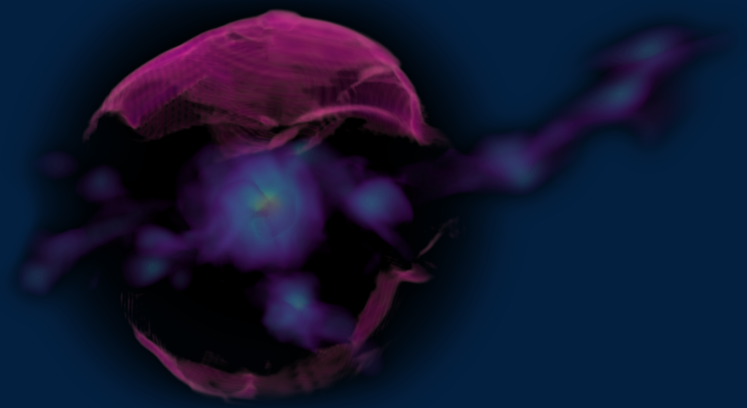
Unravelling the complexity of cosmic flows: strong gradients, shock waves and turbulence



David Vallés Pérez

Directors: **Susana Planelles Mira, Vicent Quilis Quilis**

Programa de Doctorat en Física – 3126
Departament d'Astronomia i Astrofísica
Universitat de València

TESI DOCTORAL
Abril 2024

VNIVERSITAT DE VALÈNCIA

Unravelling the complexity of cosmic flows: strong gradients, shock waves and turbulence

David Vallés Pérez

Programa de Doctorat en Física – 3126
Departament d'Astronomia i Astrofísica
Universitat de València

TESI DOCTORAL

Directors: **Susana Planelles Mira, Vicent Quilis Quilis**

Abril 2024

DECLARACIÓ

Dra. SUSANA PLANELLES MIRA, Professora Titular de la Universitat de València,

i

Dr. VICENT QUILIS QUILIS, Catedràtic de la Universitat de València,

CERTIFIQUEN

Que la present memòria, titulada ***Unravelling the complexity of cosmic flows: strong gradients, shock waves and turbulence***, ha sigut realitzada sota la seua direcció en el Departament d'Astronomia i Astrofísica de la Universitat de València per David Vallés Pérez, i constituix la seua Tesi Doctoral per a optar al títol de Doctor per la Universitat de València una vegada cursats els estudis en el Doctorat en Física.

I, perquè en quede constància en compliment de la legislació vigent, signem el present certificat en Burjassot, a 22 d'abril de 2024.

Signat: Susana Planelles Mira

Signat: Vicent Quilis Quilis

Portada: renderització tridimensional mostrant la densitat de matèria bariònica al voltant d'un cúmul de galàxies de $\sim 3 \times 10^{14}\, M_\odot$ en una simulació cosmològica.
Més enllà d'una forma central aproximadament esfèrica, permeada per un parell de filaments còsmics que alimenten el cúmul i diverses subestructures, es representa en color magenta la superfície dels xocs externs d'acreció del cúmul.

Cover: three-dimensional render showing the baryonic matter density around a $\sim 3 \times 10^{14}\, M_\odot$ galaxy cluster from a cosmological simulation.

Beyond a roughly spherical shape, permeated by a couple of cosmic filaments that feed the cluster and several substructures, the external accretion shock shell of the cluster is represented in magenta.

Give me matter, and I will construct a world out of it.
—Immanuel Kant (preface to *Universal Natural History and Theory of the Heavens*, 1755)

Yay, space!
—Barbie (Greta Gerwig's film, 2023)

Now, my own suspicion is that the Universe is not only queerer than we suppose, but queerer than we *can* suppose...
—J.B.S. Haldane (*Possible Worlds*, 1927)[1]

[1] A different connotation of the word "queer" may have been intended here.

Agradecimientos

Los *Acknowledgements*, qué movida, ¿no? Da vértigo pararse a pensar en todas las personas que, de una forma u otra, en lo académico, en lo laboral o en lo personal, han sido condición necesaria para que el trabajo de estos casi cuatro años haya llegado a buen puerto y quede plasmado en esta Tesis. Ahí va mi mejor intento de agradeceros a todes. Espero no dejarme a nadie, pero si fuera el caso, espero que no me lo tengáis en cuenta: sois muches... y *estic molt major*.

El primer lugar de esta larga lista lo ocupan, como no podía ser de otro modo, mis directorxs, Susana y Vicent. Gracias a ambos por haberme guiado durante estos años, por estar siempre disponibles para comentar cualquier asunto y, a la vez, por haberme dado la independencia para desarrollarlo y confiar en mi trabajo. Pero, por encima de todo, gracias a lxs dos por vuestra calidez, por preocuparos siempre para que tuviera las mejores condiciones para trabajar, pero también por que dejara de trabajar cuando tocaba.

A Susana la conocí en 2015, durante mi primer año del Grado, como mi profesora de problemas de *Álgebra y Geometría I* y *II*: recuerdo que, en un determinado momento, un –otrora joven– David, que había leído sobre las *movidas* que se investigaban en el Departament d'Astronomia i Astrofísica, le preguntó qué era eso de la *Cosmología Computacional*[1]. Susana volvió a entrar al aula 4112 de la Facultat de Física, donde nos daba clase, y nos explicó a todes –¡de modo que unes boquiabiertes estudiantes de primer curso lo pudiéramos entender!– cómo se podía simular la formación de galaxias y otras estructuras

These Acknowledgements are deliberately tetralingual.

[1] Hoy en día, cuando me lo preguntan a mí, suelo decir que es como lo de echar las cartas, pero con el ordenador. Por suerte, la respuesta de Susana no fue en esta dirección.

en un ordenador. Gracias, Susana, no solo por haber sido la primera persona que me habló de la Cosmología Computacional, sino por estar siempre dispuesta a comentar o discutir cualquier cosa, aun cuando tenías muchísimas cosas más que atender; y, también, por tu empatía y tu gran capacidad para escuchar y resolver problemas, tanto científicos como metacientíficos. En definitiva, gracias por tu generosidad y tu cariño.

A Vicent no lo conocí hasta empezar mi Trabajo de Fin de Grado y la Beca de Colaboración, en 2019, cuando ambos me propusisteis deducir las ecuaciones de la hidrodinámica para el crecimiento de las perturbaciones en un Universo en expansión a partir de primeros principios, y aprender sobre las técnicas que usamos para resolverlas. Te tengo que agradecer cómo, desde entonces, siempre has tenido la puerta de tu despacho abierta para discutir cualquier detalle, por mínimo que sea, y por bajar unos peldaños al barro con las ecuaciones, con el código o hasta con las *flags* de compilación cuando ha sido necesario. Muchas gracias, Vicent, no solo por tu dedicación, sino por continuamente orientar mi trabajo con amplitud de miras, por saber retarme para que nos propusiéramos objetivos más ambiciosos y, cómo no, también por aguantar ese afán discutidor que me caracteriza –algunas veces, tal vez, acertado y, otras, puede que no tanto. En especial, gracias por tu confianza en mí –tanto en la investigación, como en la docencia–, por tu visión de conjunto y por tus «*Do not panic!*»/«*¡No paniquees!*».

Among the people that have made this Thesis possible, I am indebted to the members of the Thesis Committee, who have kindly agreed to read this work through in a short time and submit their comments. Many thanks to Veronica Biffi, Stefano Borgani, José María Martí, Elena Rasia, Patricia Sánchez-Blázquez, and Franco Vazza for their time and effort.

During this PhD, I have also enjoyed the opportunity to visit other institutions for a few months and learn from world-reknown experts in the field of hydrodynamical simulations of galaxy clusters.

Grazie, Franco Vazza, per avermi accolto tra Maggio e Luglio 2022 presso il Dipartimento di Fisica e Astronomia dell'Università di Bologna. È stato un vero onore poter discutere con te e il tuo gruppo di ricerca (Annalisa, Matteo, Paola, e tanti altri), e specialmente poter imparare tantissimo da voi per quello che riguarda turbolenza e campo magnetico in ammassi di galassie. Inoltre, grazie ai miei compagni d'ufficio, Max e Fabrizio, per avermi fatto sentire a casa per quei due mesi; ed a il mio amico Francesco per farmi da guida a Roma e Tivoli per un fine settimana.[2]

[2]Ho imparato l'italiano solo dopo avervi visitato a Bologna. Dovrete scusarmi!

Thanks[3], Klaus Dolag, for your kindness during my two months, from May to July 2023, at the Universitäts-Sternwarte München. I really much appreciate your willingness to receive me and to suggest me to adapt our VORTEX code to particle-based simulations, as well as your patience to discuss with me the details of the algorithm over a freshly brewed coffee at your office. In this sense, I also want to warmly thank Tirso and Freddy for all the interesting discussions, and for becoming my beta-testers for the code, as well as the rest of the members of the group (Stephan, Elena, Ludwig, Ildar, Laura, etc.) for receiving me as one of them.

I have also been lucky to meet so many awesome people –both in the professional and in the personal sense– during the conferences I've attended to and talks I have given. Thanks to all of you for making me feel welcome and for the interesting discussions we have had: Stefano Borgani, Ana Contreras, Weiguang Cui, Michela Esposito, Violeta González-Pérez, Meghan Gray, Max Grönke, Roan Haggar, Alexander Knebe, Rüdiger Pakmor, Frazer Pearce, Chris Power, Patricia Sánchez, Volker Springel, Alexander Vazdekis, Gustavo Yepes, and so many others that I am sure I might be forgetting. Tampoco me puedo olvidar del resto de mis compañerxs de grupo, *pre* y *post*docs: Óscar, Mónica, Marco y María, con quienes he podido colaborar en distintos proyectos.

En honor a la verdad, aunque también *por imperativo legal*[4], tengo que agradecer al Vicerrectorat d'Investigació por financiar mi trabajo durante estos años a través de un contrato *Atracció de talent.* Nunca es *tarde* si la dicha es *buena.*

Però una Tesi no és, només, treball científic. Hi ha moltíssimes altres persones que han estat presents perquè aquesta Tesi haja arribat a bon port. En primer lloc, no podria oblidar-me del nostre personal de Secretaria. No exagere quan dic que, al DAA, tenim la millor Secretaria de la UV, i no només per la seua professionalitat, sinó també per la seua calidesa. Gràcies, Maneel, Arancha, Feli (en els primers moments d'aquesta Tesi) i Ana (en els darrers), per fer-nos la vida més fàcil ajudant-nos amb totes les paperasses –que no en són poques– i totes les burocràcies a què havíem d'enfrontar-nos –algunes, de vegades, tan absurdes. I, també, gràcies a totes vosaltres per recordar-nos que hi ha molta vida al Departament més enllà d'allò que ocorre dins dels nostres despatxos: gràcies per les Xocolatades, per les felicitacions i les decoracions nadalenques, pels bombons per als qui llegim els correus fins el final... i per tota la resta de detalls que fan que el DAA siga un lloc més acollidor.

[3] I am not daring with German for this paragraph!

[4] Dejaremos para otra ocasión, tal vez, los comentarios sobre las AAPP...

Muchas gracias a todes les compañeres, más *junior* y más *senior*, con quienes he compartido conversaciones, organización de *Journal Clubs*, *cafeses*, comidas, *Olegaris* y *Viñedos*, o cualquier otra ocasión. Gracias, Raquel, Beatrice, Davide, Miquel Lluís, Fabrizio, Sergio, y tantes más, por ser tan bueña compañía en el sector *junior* del departamento. Moltes gràcies, també, a Manel, Chema i Joan, pels esmorzars en Borbotó, pels dinars en la Llotgeta i per les paelles en l'horta, a més de per les converses científiques i no tan científiques que hem pogut gaudir. Y, también, a tantos de nuestros *senior* con quien he podido interactuar de uno u otro modo durante estos años: gracias a Miguel Ángel, Pablo, Toni, Iván, Juan Antonio, Álex, y tantos más.

No podría olvidarme en estas líneas de *Cuca*, nuestra *Cuca*, que nos acompañó diariamente durante los cafés vespertinos en la entrada del edificio en los primeros años de esta Tesis. Espero que, allá donde estés, te acuerdes de nosotros.

Tampoco es posible obviar a tantas personas que te preparan, no solo en contenidos, sino también en herramientas y actitudes, antes de llegar al doctorado. De mis años del grado, docentes que dejaron especial huella fueron Chema (por su infinito rigor en sus clases de Álgebra, amén de sus preciosas pizarras), María Jesús Hernández (a quien, además de introducirme por primera vez a la física experimental y, por tanto, a todo ese arte de tomar y tratar datos, le tengo que agradecer su constante compromiso con la diversidad en estos ambientes –a veces tan poco diversos– que habitamos, un tema en el que he podido trabajar de cerca con ella), Salvador Mafé (quien me introdujo a una de las materias más interesantes que he estudiado, la Física Estadística) y, en las propias palabras de este último, su *logaritmo*, José Antonio Manzanares (quien cambió mi manera de ver la Termodinámica, no sin cierto sufrimiento mediante –que, echando la vista atrás, queda justificado), Benito Gimeno (con su enorme capacidad para transmitirnos la Electrodinámica Clásica), així com Miquel Portilla i Juan Antonio Morales (que em van introduir a la Geometria Diferencial i a la Relativitat General). Tampoc puc oblidar-me de dos professors del batxillerat, d'aquests que et canvien la manera de vore les coses: Vicent, que em va donar la primera aproximació rigorosa a les Matemàtiques; i Rafa, qui em va ensenyar a apreciar la Filosofia, la Lògica Formal i l'art del *pensar per pensar*.

Junto a ellxs, me he encontrado a compañeres extraordinaries durante mis años de estudio, que se han convertido en verdaderes amigues. Pienso en Ivan, Jorge y Jorge, Mar, Nacho, Carmen, con quienes, pese a que cada uno ha tomado un camino bien distinto con el paso de los años, todavía tengo la suerte de encontrarme y compartir experiencias; así como a tantes otres: Marina, Cris, Álvaro, Bernardo, y un largo etcétera. Mención especial merece Antonio, quien,

pese a ser del IFIC, ha sido mi compañero de tantas comidas, espacio seguro para quejas de todo tipo (nunca viperinas, porque somos unas chicas encantadoras), así como conciertos durante estos años. També vull tindre unes paraules per a les persones amb qui, fa ja uns anys, vam fundar *Espectre Visible*, un col·lectiu per a tractar de visibilitzar i sensibilitzar a la Facultat de Física sobre la diversitat afectivosexual i de gènere en ambients sovint tan poc diversos com ho són els científics. Gràcies a Carles, Héctor, Anna, Marina, Carmen, Tono, Víctor, y a mi *Confidente de Crisis Académicas*, Antonio, por todo lo que hemos hecho y aprendido juntes. Més enllà d'aquests entorns tan propers a la Facultat de Física –i és que, de vegades, ens costa ixir d'ací–, aquestes línies quedarien incompletes si no esmentara a les meues amigues *del poble*, *de sempre*, les quals m'han hagut d'aguantar durant tots aquests anys de transició des d'adolescent fins a físic desquiciat: Alejandro, Alí*ss*ia, Isabel, Ivan, Lluis, Lourdes, Marc i Marc, Maria i Maria, Miquel, Rita... De vegades, m'heu vist el pèl menys del que m'agradaria, però gràcies per traure'm de casa de tant en tant i recordar-me que també hi ha vida més enllà de la Física.

Jose y Miquel, Miquel y Jose... En palabras de Gandalf el Blanco, *no diré «no lloréis», pues no todas las lágrimas son amargas.* Qué fácil es escribir unas líneas para vosotros, y qué difícil es escribir *las* líneas que os merecéis. Gracias a los dos por hacer del 3.02 un lugar tan especial. Lo voy a intentar lo mejor que pueda, pero nada de lo que diga podrá estar a la altura de estos tres años que hemos pasado juntos en el 3.02. Hala, ya he hecho el *Disclaimer* y, parafraseando a Noemí Argüelles en Paquita Salas, «*¡que me registren, porque yo ya lo dije!*».

Jose, querido *YoU*, apreciado *Horse*, u.t.c. JSOE. Este mundo completamente delirante y ajeno a la realidad que hemos creado en el 3.02, con su *quokkisme* y sus *«Avall la revolució!»*[5], es tan tuyo como mío. Podría ser un síntoma de que ambos hemos perdido completamente el juicio durante nuestros respectivos Doctorados, pero yo prefiero pensar que sencillamente es indicativo de la buena sintonía que hemos vivido en el despacho estos años. Eres una persona brillante y me siento afortunado de que hayamos podido trabajar en proyectos juntos. Sin duda, nos has demostrado ser capaz de conseguir cada meta que te has propuesto, de modo que no me cabe duda de que, en los próximos años, nos seguirás dando más alegrías, desde el ESAC o desde donde tengas a bien sorprendernos.

Miquel, estimat *Mikos*, benvolgut *Maiq.* Encara que vam començar junts el Grau, no va ser fins al Màster i, especialment, el primer any de Doctorat,

[5] Puedo entender que el lector ajeno a este contexto esté pensando: «¿Qué dice este *petardo*?». Prometo que todos estos disparates tienen su contexto, pero lo dejaremos también para otra ocasión.

arrel de compartir despatx, que ens vam començar a conéixer bé. Qui m'anava a dir que, darrere d'ixe *heteríssim* que havia vist de lluny s'amagava una de les persones més carismàtiques i carinyoses que he conegut mai, un *golden* recobert de *tremendo maikisme*. Gràcies per tot el que hem compartit: les *turres* amb les cridades i la cara de concentració, les *ràdios* sobre quasi qualsevol tema que ens han amenitzat els dinars, el màrqueting de les teues habilitats culinàries[6]... Et desitge tota la sort del món en l'etapa postdoctoral que tu també comences prompte (qui seran els afortunats i afortunades d'experimentar el *maikisme* ara?) i, qui sap, si ens retrobarem en algun moment al DAA. O a *Rottnest Island*, si fa falta.

Als dos, gràcies per la vostra companyia diària i per la vostra estima. M'alegre d'haver conegut durant aquests anys també a les vostres parelles, Ana i Marina, que són dos persones extraordinàries que vos complementen a la perfecció. En definitiva, *quod ergo 3.02 coniunxit, homo non separet.*

Por último, quiero referirme a mi familia, porque sin cada unx de ellxs, esta Tesis no hubiera sido en absoluto posible. A mis padres, gracias por apoyarme en mis estudios y darme todo lo que he necesitado en mi afán por seguir esta carrera que puede tener momentos tan precarios, por todos los «*posa-li més arròs al xiquet, que no veus que està estudiant?*», por las visitas durante las estancias... A Iván y a Maria, el *germà* y la *cunyà*, muchas gracias por todas las veces que me habéis recibido allá donde habéis estado: en Madrid, Cambridge, (¡lástima que no nos pudimos ver en Luxemburgo!)... Gracias a ambos por estar siempre disponibles como mis *data analyst* y *experta en cosas gráficas* de guardia, como espero yo haber sido un buen *físico y matemático-wannabe* de guardia. Tampoco puedo, por supuesto, dejar de mencionar aquí a Bimba por su apoyo incondicional, aunque ella no tenga ni idea.

A Josep, ni sis pàgines més d'aquesta secció farien justícia a tot el que t'he d'agrair. Gràcies pel teu recolzament incondicional, per la teua infinita paciència quan les hores de treball són més llargues del que ens agradaria, pels viatges a l'aeroport i per les visites durant les estades, per la teua capacitat per fer-me riure quan més ho necessite i per entendre'm i per suportar-me en el *mapatxisme* i en el *totutepisme*. Gràcies, Josep, per ser, a més de la meua parella, el meu millor amic, per ser el meu millor company de viatge. Gràcies per tot el que hem compartit i per tot el que ens queda per compartir. Gràcies per tot el que ens queda per viure, i per tot el que ens queda per estimar-nos.

A todes les que estáis aquí mencionades, y a les que he olvidado, creedme cuando os digo que habéis sido condición necesaria para haber llegado hasta aquí

[6] Ben justificat, una volta hem tastat la teua aclamada *petxuga* al curri...

de esta manera. En la línea que me caracteriza (Vicent diría: «*¡tío, es que eres Shakespeariano, mira que te gusta la prosa!*»), no he sido capaz de cerrar estos *Agradecimientos* en menos de seis páginas. Claro que, quienes me conozcan, tampoco se estarán llevando gran sorpresa... A todos, todas y todes,

Gracias de corazón,
David Vallés Pérez

Abstract

The formation and evolution of cosmic structures proceed through the gravitational collapse of primordial density fluctuations, which is a non-linear process that results in the emergence of the cosmic web and a rich hierarchy of structures. Besides gravity, many more phenomena associated with the physics of baryons are at play during this evolution and shape the physical and observational properties of galaxies, galaxy clusters and the environments they inhabit. The processes of cosmological structure formation are complex, in the sense of very non-linear, and involve a wide range of physical scales, making numerical simulations an indispensable tool for their understanding. The scientific goals of this Thesis have been mainly oriented towards the numerical study of the assembly of galaxy clusters, in the first place, and cosmic voids, in the second place, as two complementary ends of the range of cosmic structures. In galaxy clusters, the stirring of the gas due to accretion and mergers, together with other processes, generates two important non-linear hydrodynamic phenomena: shock waves and turbulence, which have a consequential impact on the dynamical, thermodynamical and observational properties of clusters. Regarding cosmic voids, although much of their interest comes from their role as pristine environments for galaxy evolution due to their mostly outflowing velocity field, the situation can be more complex in a cosmological context. The main results of this Thesis can be summarised in four closely intertwined main lines. The first line of work has been the development of numerical tools for the analysis of cosmological simulations of different types. The most salient outputs have been VORTEX, a code for performing a Helmholtz-Hodge decomposition of multiresolution velocity fields,

and a new version of ASOHF, a spherical overdensity halo and galaxy finder. Both codes have been publicly released. The second line concerns the assembly of galaxy clusters, comprising two different works exploring the quantitative study of their accretion histories and the determination of the assembly state of dark matter haloes from a set of observables. A third axis of this Thesis is the study of the previously mentioned non-linear hydrodynamic phenomena, including the presence of turbulent motions in the intracluster medium in connection with its assembly history and the properties of cosmological accretion shock waves. Finally, the fourth line of work has been devoted to the study of the velocity field in cosmic voids, including the development of a new method for their identification in cosmological simulations. This Thesis is laid out as a compendium of publications, and the results presented here have been published in 6 papers in peer-reviewed journals.

Resum

Abstract. La formació i evolució d'estructures còsmiques avança a través del col·lapse gravitacional de fluctuacions de densitat primordials, que és un procés no-lineal que dona lloc a l'aparició de la teranyina còsmica i una rica jerarquia d'estructures. A més de la gravetat, molts altres fenòmens associats a la física dels barions estan presents durant aquesta evolució i configuren les propietats físiques i observacionals de galàxies, cúmuls de galàxies i els entorns que habiten. Els processos de formació d'estructures cosmològiques són complexos, en el sentit de ser altament no-lineals, i involucren una àmplia gamma d'escales físiques, fent que les simulacions numèriques esdevinguen una eina indispensable per a la seua comprensió. Els objectius científics d'aquesta Tesi s'han orientat principalment cap a l'estudi numèric de l'assemblatge de cúmuls de galàxies, en primer lloc, i buits còsmics, en segon lloc, com a dos extrems complementaris de l'abast d'estructures còsmiques. En els cúmuls de galàxies, l'agitació del gas a causa de l'acreció i les fusions, juntament amb altres processos, genera dos importants fenòmens hidrodinàmics no-lineals: ones de xoc i turbulència, que tenen un impacte decisiu en les propietats dinàmiques, termodinàmiques i observacionals dels cúmuls. Pel que fa als buits còsmics, malgrat que molt del seu interés prové del seu paper com a entorns pristins per a l'evolució galàctica a causa del seu camp de velocitat majoritàriament sortint, la situació és més complexa en un context cosmològic. Els resultats principals d'aquesta Tesi es poden resumir en quatre línies principals, estretament relacionades. La primera

This chapter, a summary of the Thesis in Spanish or Catalan, is included to comply with PhD Thesis regulations by Universitat de València.

línia de treball ha estat el desenvolupament d'eines numèriques per a l'anàlisi de simulacions cosmològiques de diferents tipus. Els resultats més destacats han sigut VORTEX, un codi per realitzar una descomposició de Helmholtz-Hodge de camps de velocitat multiresolució, i una nova versió d'ASOHF, un buscador d'halos i galàxies de sobredensitat esfèrica. Ambdós codis s'han posat a disposició de la comunitat científica. La segona línia tracta de l'assemblatge de cúmuls de galàxies, amb dos treballs diferents que exploren l'estudi quantitatiu de les seues històries d'acreció i la determinació de l'estat d'assemblatge dels halos de matèria fosca a partir d'un conjunt d'observables. Un tercer eix d'aquesta Tesi és l'estudi dels fenòmens hidrodinàmics no-lineals esmentats prèviament, incloent-hi la presència de moviments turbulents en el medi intracúmul en connexió amb la seua història d'assemblatge, i les propietats de les ones de xoc d'acreció cosmològiques. Finalment, la quarta línia de treball s'ha dedicat a l'estudi del camp de velocitat en buits còsmics, inclòs el desenvolupament d'un nou mètode per a la seua identificació en simulacions cosmològiques. Aquesta Tesi es presenta com un compendi de publicacions. Els resultats ací detallats s'han publicat en 6 articles en revistes arbitrades.

Introducció

El treball portat a terme durant el curs d'aquesta Tesi Doctoral, i reportat en aquesta memòria, s'incardina dins del camp de l'Astrofísica Computacional i, més específicament, la Cosmologia Numèrica. En particular, els principals objectes científics d'interés han sigut els cúmuls de galàxies i els buits còsmics, que representen dos extrems oposats dins del rang d'estructures generades pel col·lapse gravitacional a l'Univers.

Els cúmuls de galàxies (en endavant, també *cúmuls*) van ser descoberts fa poc menys de dos-cents cinquanta anys pels astrònoms Charles Messier i William Herschel, amb distintes motivacions. Tanmateix, no va ser fins ben entrat el segle XX, amb el desenvolupament de la Cosmologia moderna, que la seua comprensió astrofísica va començar. Un segle després, els cúmuls han esdevingut un objecte d'estudi essencial, no només en Astrofísica i Cosmologia, sinó també en altres disciplines com la Física de Partícules, la Física de Plasmes, etc.

Pel que fa als buits còsmics (o *buits*), el seu descobriment és més recent, donat que va haver d'esperar a l'arribada dels primers grans catàlegs espectroscòpics de galàxies, a partir de la dècada dels setanta i vuitanta.

Malgrat l'interés que ha generat des de fa segles, la formació d'estructures cosmològiques és, per tant, un camp relativament jove. Açò es deu, en part, a

la impossibilitat d'accedir directament als objectes d'estudi o recrear les seues condicions en un laboratori, al contrari que en altres camps de la Física. Per això, el desenvolupament d'aquest camp ha anat de la mà del desenvolupament de l'Astrofísica Observacional, per una banda, i del naixement de l'Astrofísica Computacional, per l'altra, com a eina fonamental per tal de conciliar les prediccions teòriques de la Física fonamental amb les observacions.

Des de les primeres simulacions numèriques cosmològiques, als anys 70 i 80 del segle passat, el ràpid desenvolupament tecnològic i l'esforç conjunt en Astrofísica i Cosmologia ens ha permés obtenir un coneixement profund sobre la formació i evolució de les galàxies, els cúmuls, els seus entorns i els fenòmens físics que els governen. Així, en les últimes dècades, teoria, simulacions i observacions han convergit en establir el paradigma jeràrquic com l'escenari estàndard de la formació d'estructures còsmiques, i el model ΛCDM com el model cosmològic més exitós fins a la data, a pesar dels seus problemes oberts.

Al mateix temps, treballs pioners amb simulacions hidrodinàmiques numèriques durant les últimes tres dècades han revelat la presència de característiques complexes associades a la física dels barions, com turbulència, ones de xoc (també anomenades, directament, *xocs*), o camps magnètics, que tenen un impacte significatiu en les propietats observables dels cúmuls a pesar del rol energèticament subdominant dels barions (la *matèria ordinària*) en el context cosmològic. Aquests fenòmens, l'estudi dels quals està associat a la no-linealitat de les equacions que regixen la dinàmica dels fluids i l'agrupament gravitatori, remarquen novament l'important rol de les simulacions numèriques en l'estudi de la formació i l'evolució d'estructures còsmiques.

Objectius

El principal objectiu, en termes generals, d'aquesta Tesi Doctoral ha sigut el d'explorar, quantificar i contribuir a ampliar el nostre coneixement al voltant de diversos aspectes associats amb el procés de formació d'estructures còsmiques, que resulta en l'emergència de complexos *fluxos còsmics* de matèria fosca i bariònica, tant laminars i globals com turbulents, que donen títol a aquesta Tesi. Aquest objectiu general pot concretar-se en els següents objectius més particulars:

- La caracterització de la història de formació, evolució i assemblatge dels cúmuls i grups de galàxies (la seua intensitat, propietats qualitatives, evolució, la seua relació amb propietats locals i globals del mitjà intracúmul

[ICM], etc.). En aquesta mateixa direcció, la caracterització del complex estat dinàmic dels cúmuls a partir d'una sèrie d'observables també ha sigut un tema d'interés durant aquesta Tesi.

- L'estudi al voltant de la presència, la naturalesa i l'evolució de la turbulència en el medi intracúmul, així com la seua relació amb la història d'assemblatge dels cúmuls.

- La investigació sobre les propietats i la localització dels xocs en cúmuls de galàxies, especialment pel que fa als xocs d'acreció a gran escala associats al col·lapse gravitacional de les inhomogeneïtats primordials, i la relació d'aquests fenòmens amb les propietats de l'objecte que els alberga.

- L'estudi dels fluxos de matèria bariònica al voltant dels buits i, especialment, la natura del camp de velocitats en aquestes regions i les possibles conseqüències per a les propietats del gas difús que hi residix.

- El desenvolupament de noves eines numèriques, així com l'optimització de les ja existents, per tal de dur a terme l'anàlisi de simulacions cosmològiques. De manera constant durant aquesta Tesi, s'ha treballat per estendre les eines desenvolupades per a poder ser aplicades a codis de simulacions de distinta natura. Així mateix, els codis desenvolupats en el marc d'aquesta Tesi Doctoral han sigut sempre fets públics i de lliure accés per a la comunitat científica.

Marc teòric

La Part I de la Tesi fa una revisió bibliogràfica del marc teòric i numèric en què s'inscriu el treball d'aquest Doctorat. En les següents seccions, es descriu breument el contingut de cada capítol de manera general.

Formació d'estructures còsmiques

Com a culminació d'un segle de treball al voltant de la física de la formació d'estructures baix el marc de la relativitat general (GR), el model ΛCDM ha emergit com el model estàndard de la cosmologia moderna. Aquest model, que descriu l'Univers com un sistema dinàmic en expansió, considera una barreja de matèria fosca i energia fosca, a més de matèria ordinària, i s'ajusta a les observacions cosmològiques amb una precisió sense precedents. No obstant això, el model ΛCDM no està exempt de problemes: més enllà de l'absència d'una

descripció satisfactòria de la naturalesa de la matèria i l'energia fosca, el model presenta problemes a xicoteta escala.

Baix el paradigma ΛCDM, la mètrica de l'espai-temps ve descrita per un model de Friedmann-Lemaître-Robertson-Walker (FLRW), que descriu un Univers homogeni i isòtrop en expansió. La història d'expansió d'un univers FLRW queda determinada una vegada especificada la composició, en termes energètics, del mateix, per mitjà de les equacions de Friedmann. En el model ΛCDM, el contingut energètic de l'Univers està dominat per l'energia fosca (una densitat d'energia constant amb pressió negativa que es manifesta en l'expansió accelerada de l'Univers) i la matèria fosca (una forma de matèria no-bariònica que no interacciona electromagnèticament). Només el $\sim 5\%$ restant correspon a matèria ordinària. Malgrat el seu rol energèticament subdominant, és possible que alguns dels problemes oberts del model ΛCDM troben la seua solució en la física dels barions. Per exemple, la retroalimentació energètica associada a la formació d'estrelles i l'activitat dels nuclis actius de galàxies (AGN) podrien explicar l'absència de distribucions de densitat amb pics.

Pertorbacions sobre un Univers FLRW

Els models FLRW descriuen universos homogenis i isòtrops. La formació d'estructures, en tant que inhomogeneïtats en l'Univers, s'estudia com l'evolució de pertorbacions (de gran magnitud i, per tant, no-lineals), sobre aquest *fons cosmològic*. Per tal que les estructures puguen formar-se, és necessari que existiren pertorbacions en instants molt primerencs, que s'amplificarien posteriorment per acció de la gravetat. Els models inflacionaris, que descriuen una fase d'expansió accelerada de l'Univers en els seus primers instants, poden explicar l'origen d'aquestes pertorbacions primigènies i quantificar-ne les seues propietats.

Encara que existixen tècniques analítiques per a estudiar l'evolució de les pertorbacions (per exemple, el model lineal, o el model *top-hat*), la seua naturalesa no-lineal fa que siga necessari recórrer a mètodes numèrics per a estudiar-les. En aquest sentit, les simulacions numèriques cosmològiques s'han convertit en una eina fonamental per a l'estudi de la formació d'estructures còsmiques. Aquestes simulacions, que consistixen en la resolució numèrica de les equacions de la dinàmica de la matèria fosca i la matèria ordinària, es descriuen en la propera secció.

Cúmuls de galàxies

Els cúmuls de galàxies són les estructures gravitacionalment lligades més massives de l'Univers, amb masses entre els $10^{14} M_\odot$ i $10^{15} M_\odot$ i grandàries típiques de $\sim 5\,\mathrm{Mpc}$. La formació d'aquests objectes està entre els fenòmens més energètics de l'Univers, involucrant energies de fins a $10^{65}\,\mathrm{erg}$.

La composició dels cúmuls és aproximadament representativa de la composició global de l'Univers, amb fraccions de gas f_{gas} similars o només lleugerament per baix de la còsmica. Açò, junt amb la seua singular posició en el cim de la jerarquia d'estructures còsmiques, els convertix en objectes d'especial interés per a la Cosmologia.

Per altra banda, els cúmuls també esdevenen excel·lents laboratoris per a l'estudi, no només dels processos de formació de galàxies en entorns densos, sinó també la física de l'ICM, que representa la major part de la massa bariònica del cúmul i correspon a un plasma quasi completament ionitzat, amb temperatures de fins a $10^8\,\mathrm{K}$ i densitats de fins a $10^{-1}\,\mathrm{cm}^{-3}$, parcialment en equilibri amb el camp gravitatori del cúmul.

Els cúmuls poden ser detectats i observats a través de, pràcticament, tot l'espectre electromagnètic, des de les ones de ràdio fins als raigs gamma. Les tres principals finestres per a estudiar els cúmuls, això no obstant, corresponen a les microones, els raigs X i la llum visible. En la banda òptica, els cúmuls poden observar-se a causa de la llum emesa per les galàxies que en formen part, i aquest canal correspon precisament als primers descobriments d'aquests objectes. Addicionalment, els cúmuls també poden ser estudiats en la banda òptica per la deflexió que produïxen de la llum de galàxies més llunyanes, fenomen conegut com a *lent gravitacional*.

En la banda de les microones, els cúmuls són observats a través de la distorsió que provoquen sobre la radiació de fons de microones (CMB) que travessa l'ICM, en un efecte conegut com a Sunyaev-Zeldovich. Finalment, en la banda dels raigs X, els cúmuls poden ser estudiats a través de l'emissió tèrmica de l'ICM que, a causa de la seua alta temperatura, emet de manera molt potent radiació de frenat (*bremsstrahlung*) en aquest rang d'energies.

Buits còsmics

Els buits còsmics són les regions més grans i menys denses de l'Univers, amb grandàries característiques de fins a 100 Mpc i densitats de matèria al voltant d'una cinquena part la densitat mitjana de l'Univers. Aquests objectes, que evolucionen expandint-se i buidant-se a mesura que la matèria circumdant col-

lapsa per a formar parets, filaments i cúmuls, són de gran interés, tant per les seues aplicacions cosmològiques, com per a l'estudi de l'evolució de galàxies en entorns poc densos.

Simulacions numèriques cosmològiques

Per tal de simular els processos associats amb la formació i evolució d'estructures còsmiques, és necessari plantejar-los com un problema de valor inicial. És a dir, és necessari descriure quines són les equacions que regixen l'evolució del sistema, i quines són les condicions inicials per tal de començar aquesta evolució.

A causa de la seua distinta naturalesa, la matèria fosca i la matèria ordinària són tractades de manera distinta en les simulacions cosmològiques.

Dinàmica de la matèria fosca

La matèria fosca freda (i.e., no relativista) es pot descriure com un fluid sense col·lisions (és a dir, on les seues partícules no interaccionen apreciablement entre sí), que es mou sota l'acció de la gravetat. Encara que existixen alternatives, la tècnica més emprada per a estudiar l'evolució de la matèria fosca és la de les simulacions d'N-cossos, on es mostreja l'espai de fases (posicions i velocitats) de la matèria fosca amb un conjunt discret d'N partícules que evolucionen baix l'acció del camp gravitatori global (generat per la mateixa matèria fosca, però també pels barions).

La diferència entre els distints mètodes per a evolucionar la matèria fosca residix en les tècniques emprades per a calcular el camp gravitatori. Idealment, la força sobre cada partícula es calcularia sumant sobre les $N-1$ partícules restants. Desafortunadament, aquest mètode, on el nombre d'operacions a realitzar escala com $\mathcal{O}(N^2)$, és prohibitiu per a grans simulacions. Per això, s'han desenvolupat tècniques més eficients per a calcular aproximacions precises al camp gravitatori. D'entre elles, un dels enfocaments més adoptats és el mètode *Partícula-Malla* (PM), que calcula el potencial gravitatori com la solució de l'equació de Poisson en un espai periòdic discretitzat en una malla uniforme. Aquest mètode, que escala com $\mathcal{O}(N \log N)$, produïx bons resultats per a escales superiors a la resolució de la malla, però no és capaç de resoldre la força entre partícules que n'estan més pròximes. Per això, altres mètodes anomenats híbrids, com els mètodes d'arbre-PM, o el *Partícula-Partícula/Partícula-Malla* ($\mathrm{P^3M}$), combinen el mètode PM amb mètodes més precisos però computacionalment costosos per a calcular la força entre partícules pròximes.

Dinàmica de la matèria ordinària

Al contrari que la matèria fosca, els barions són una component material altament col·lisional, que es pot descriure com un fluid. Així, la dinàmica de la matèria ordinària es pot formular mitjançant les equacions de la hidrodinàmica, que descriuen l'evolució de la densitat, la velocitat i la pressió (o l'energia interna) del fluid. Aquestes equacions són altament no-lineals, implicant que s'han de resoldre numèricament en virtualment qualsevol situació d'interés. Més enllà, la seua no-linealitat implica el desenvolupament de discontinuïtats, o de propietats a escales molt menudes, demandant per tant tècniques complexes i resolucions molt altes per a ser tractades correctament.

Les equacions de la hidrodinàmica poden discretitzar-se de diverses maneres, segons s'escriguen en forma euleriana (les propietats del fluid s'estudien, en funció del temps, en una posició determinada) o lagrangiana (s'estudien les propietats d'un element de fluid, a mesura que aquest evoluciona i es desplaça pel domini).

En el primer cas, les tècniques eulerianes permeten estudiar, mitjançant mètodes de volums finits, l'evolució de les quantitats integrades en volums fixos resolent equacions en derivades parcials. Aquestes tècniques tenen certs avantatges, com la possibilitat de tractar de manera molt precisa les discontinuïtats (amb els mètodes de captura de xocs d'alta resolució, HRSC) i de conservar explícitament certes quantitats (com la massa, el moment lineal, l'energia, etc.). Malgrat això, la discretització en volums fixos implica costos computacionals molt alts per a assolir resolucions altes. Aquest problema pot ser parcialment alleugerat amb la tècnica de refinament adaptatiu de la malla (AMR), on determinades regions són escollides d'acord amb distints criteris per a ser discretitzades amb una resolució més alta que la resta del domini, de manera recursiva i dinàmica (és a dir, el refinament canvia amb el temps).

En el segon cas, les tècniques lagrangianes com la hidrodinàmica de partícules suavitzades (SPH) discretitzen el fluid en una sèrie de partícules de massa donada, que es mouen d'acord amb unes equacions diferencials ordinàries. Les quantitats del fluid en una posició donada es calculen mitjançant una mitjana entre les partícules més pròximes a aquesta posició. Aquesta tècnica, a causa de la seua discretització en termes de massa, en lloc de volum, permet obtenir resolucions molt més altes que els mètodes eulerians en les regions d'alta densitat, que són usualment les de major interés. No obstant això, el seu tractament de les regions de baixa densitat és, per construcció, pobre. Addicionalment, aquestes tècniques tenen majors problemes per a resoldre discontinuïtats o gradients molt forts.

Finalment, existixen mètodes mixts, que utilitzen una formulació híbrida euleriana-lagrangiana. Aquests mètodes permeten combinar l'adaptivitat dels mètodes lagrangians, amb la descripció precisa de discontinuïtats dels mètodes eulerians, i han guanyat popularitat en els últims anys.

Física addicional de la matèria ordinària

Amb l'objectiu de reproduir fidelment les condicions termodinàmiques del gas en els cúmuls, o per a descriure els diversos processos astrofísics que condicionen la formació de galàxies, les equacions bàsiques de la hidrodinàmica han de ser complementades amb processos addicionals, típicament implementats mitjançant parametritzacions fenomenològiques.

Aquests processos inclouen processos de refredament del gas, que depenen de les condicions químiques del medi (la composició o *metal·licitat*), a banda de paràmetres com la densitat i la temperatura, i permeten que el gas radie part de la seua energia interna i puga generar estructures fredes i denses capaces de formar estrelles. La formació estel·lar s'implementa freqüentment mitjançant la formació de partícules no-col·lisionals (com la matèria fosca) a partir de gas fred i dens, d'acord amb parametritzacions compatibles amb les observacions. Aquestes estrelles, que no corresponen a objectes particulars sinó a poblacions completes, poden posteriorment retornar part de la seua massa (enriquida químicament) i energia al mitjà, a través de processos de *retroalimentació energètica*.

Addicionalment, altres processos usualment inclosos en les simulacions són la formació de forats negres i la seua retroalimentació energètica deguda a l'acreció de gas, la presència de camps magnètics (que implica una revisió completa de les tècniques hidrodinàmiques), rajos còsmics, camps de radiació, viscositat, conducció tèrmica, o modelització de la pols galàctica, entre d'altres.

Condicions inicials

Les condicions inicials per a les simulacions cosmològiques es dividixen, típicament, en dos tipus. D'una banda, trobem simulacions de grans volums que tracten de reproduir les propietats a gran escala de l'Univers. En aquest cas, hi ha mètodes per a generar camps de densitat inicials (a molt alt redshift, $z \sim 100$) amb propietats estadístiques equivalents a les observades. Per altra banda, existixen tècniques per a restringir aquestes condicions inicials de tal manera que reproduïsquen, per exemple, un cúmul de galàxies massiu en el centre del domini computacional o, fins i tot, l'entorn de l'Univers Local.

Fenòmens no-lineals i no-tèrmics en fluxos cosmològics

Part dels resultats presentats en aquesta Tesi estan relacionats amb l'estudi de fenòmens associats a la no-linealitat de les equacions de la hidrodinàmica. Aquests inclouen, per exemple, turbulència i ones de xoc.

La turbulència està present en fluxos cosmològics a escales molt diverses, des de l'interior d'estrelles fins als cúmuls de galàxies. De manera general, la turbulència pot definir-se com a la propietat d'un fluid altament irregular, tant en espai com en temps, que es caracteritza per la transferència d'energia entre diverses escales. Pel que fa a cúmuls de galàxies, els moviments turbulents poden ser produïts per diversos processos, incloent-hi fusions de cúmuls de galàxies i acreció de gas, moviment de galàxies a través de l'ICM, retroalimentació energètica d'AGN, entre d'altres.

Pel que fa als xocs, els fluxos cosmològics de gas originats pel col·lapse gravitacional entren freqüentment en el règim supersònic, quan la velocitat del gas caient al pou de potencial de l'estructura supera la velocitat local del so. Açò implica l'aparició de xocs, que es manifesten com a superfícies de discontinuïtat en certes propietats del fluid (pressió, temperatura, densitat, etc.). En el context dels cúmuls de galàxies, els xocs tenen un paper fonamental en la termalització de l'ICM fins a les temperatures observades.

A més del seu paper en la termalització de l'ICM, els xocs també poden convertir una fracció significativa de l'energia cinètica involucrada en el col·lapse dels cúmuls en acceleració de partícules a energies relativistes (rajos còsmics), mitjançant mecanismes que estan, al seu torn, associats amb la presència d'emissió difusa en ones de ràdio. De la mateixa manera, els moviments turbulents al cúmul també poden accelerar rajos còsmics en regions extenses, a més de contribuir a l'amplificació dels camps magnètics o d'aportar una part significativa del suport de l'ICM contra la gravetat.

Metodologia

Més enllà de comparacions amb resultats d'altres codis de simulacions (per exemple, els continguts en les Publicacions A5 i A6), la major part dels resultats reportats en aquesta Tesi han sigut obtinguts a partir de l'anàlisi de simulacions portades a terme amb MASCLET (Mesh Adaptive Code for CosmologicaL structurE evoluTion), del qual m'he convertit, durant el curs d'aquesta Tesi, en un usuari i desenvolupador.

Simulacions cosmològiques amb MASCLET

MASCLET és un codi hidrodinàmic eulerià, dissenyat fonamentalment per a aplicacions cosmològiques, acoblat a una implementació d'N-cossos i un solucionador de la gravetat basat en el mètode PM. Per a ser capaços de resoldre els alts rangs dinàmics requerits per les simulacions cosmològiques, a la volta que es mantinga un cost computacional adequat, MASCLET utilitza un esquema AMR, per tal d'incrementar recursivament la resolució temporal i espacial en les zones d'interés, que poden ser escollides d'acord amb diversos criteris segons l'aplicació.

En particular, quant a la component col·lisional (matèria bariònica), MASCLET implementa mètodes HRSC sobre les equacions de la hidrodinàmica en un marc cosmològic en expansió escrites en forma conservativa. Quant als algorismes particulars, encara que MASCLET implementa diverses possibilitats per a cada procediment, la configuració més habitual utilitzada en aquesta Tesi implica una reconstrucció parabòlica (PPM), un solucionador de Riemann HLLE, i un esquema de Runge-Kutta de segon ordre (predictor-corrector) per a l'evolució temporal. Per a la descripció precisa dels mètodes, dirigim el lector interessat a la Sec. 4.

Quant a la component no-col·lisional (matèria fosca i altres espècies, com estrelles o forats negres), MASCLET implementa un solucionador de la gravetat basat en el mètode PM. Les partícules són assignades a la malla computacional mitjançant un esquema *Triangular Shaped Cloud* (TSC), que fa que les forces resultants siguen contínues i diferenciables. Per a la malla base (sense refinaments), l'equació de Poisson es resol mitjançant una transformada de Fourier ràpida (FFT). Finalment, les partícules són avançades en el temps emprant un esquema Runge-Kutta de segon ordre, amb un pas intermedi.

Naturalment, en presència de refinament de la malla, aquests solucionadors bàsics han de modificar-se per tal de tindre en compte les condicions de frontera imposades pels dominis menys resolts que els envolten. Açò implica modificacions tant per a l'esquema hidrodinàmic, com per a l'avançament de les partícules i el solucionador de la gravetat, que són descrits en major detall en la Sec. 4.2.

A més de gravetat i hidrodinàmica adiabàtica, MASCLET implementa diversos processos astrofísics addicionals, d'entre els esmentats anteriorment, com diversos mecanismes de refredament del gas, formació d'estrelles i forats negres, enriquiment químic, així com retroalimentació energètica de supernoves i forats negres.

A més, encara que en el marc d'aquesta Tesi no s'ha publicat encara cap resultat referent a les simulacions amb camps magnètics, existix una versió de MASCLET adaptada per a solucionar les equacions de la magnetohidrodinàmica en un Univers en expansió.

Resultats, conclusions i perspectives

La Part II d'aquesta Tesi conté un resum dels resultats obtinguts durant el curs d'aquest Doctorat. D'acord amb les regulacions de l'*Escola de Doctorat* per al format de *Tesi per compilació de publicacions*, els articles, que contenen la major part del treball d'aquest Doctorat, s'han inclòs a l'Apèndix A. En les següents seccions, es descriu breument el contingut de cada un dels quatre capítols en què s'ha dividit el treball d'aquesta Tesi.

Eines numèriques

Una part significativa del treball durant aquest Doctorat s'ha dedicat al desenvolupament de noves tècniques numèriques per a l'anàlisi de simulacions cosmològiques, així com a la renovació i optimització d'altres ja existents. De manera constant, els codis desenvolupats en el marc d'aquesta Tesi han sigut generalitzats, de manera que han esdevingut eines que poden ser aplicades, no només a simulacions AMR, sinó a qualsevol simulació, i han sigut fets públics i de lliure accés per a la comunitat científica.

Un algorisme per a la descomposició de Helmholtz-Hodge en simulacions cosmològiques

Aquesta secció es basa en Vallés-Pérez, Planelles, and Quilis (2021b)[7], i el text complet de l'article es pot trobar a l'Apèndix A2. La implementació de l'algorisme presentat en aquesta publicació, VORTEX, és de lliure accés.[8]

Context. L'estudi dels camps de velocitat turbulents (així com moltes altres disciplines, on també puga ser rellevant) sovint requerix la descomposició d'un camp vectorial en les seues components compressiva (o sense rotacional) i solenoidal (o sense divergència). Açò és un problema matemàtic ben estudiat, conegut com la descomposició de Helmholtz-Hodge (HHD). L'HHD, donades

[7] D. Vallés-Pérez, S. Planelles, and V. Quilis. "Unravelling cosmic velocity flows: a Helmholtz-Hodge decomposition algorithm for cosmological simulations." In: *Computer Physics Communications* 263, 107892 (June 2021). DOI: 10.1016/j.cpc.2021.107892

[8] https://github.com/dvallesp/vortex.

condicions de frontera adients, pot ser realitzada d'una manera relativament senzilla quan el camp vectorial està mostrejat en una malla regular. No obstant això, en el cas de malles no uniformes, com les utilitzades en les simulacions d'aquesta Tesi, l'HHD esdevé un problema substancialment més complex.

Objectius. El desenvolupament d'un algorisme computacionalment eficient per a realitzar l'HHD sobre els camps de velocitat en simulacions AMR basades en blocs, com les utilitzades en aquesta Tesi.

Mètodes. L'algorisme proposat en aquesta publicació es basa en la formulació de l'HHD com un conjunt d'equacions en derivades parcials el·líptiques, cadascuna d'elles formalment equivalent a l'equació de Poisson. L'algorisme, per tant, empra les tècniques estàndard per a resoldre aquestes equacions (en particular, les mateixes que MASCLET), és a dir, una combinació de FFT per a la malla base (periòdica) i solucionadors iteratius per als nivells de refinament. L'algorisme està implementat en FORTRAN i paral·lelitzat d'acord amb les directrius estàndard d'OPENMP.

Resultats. El codi ha sigut validat amb una sèrie de tests idealitzats i d'altres més complexos. La descomposició es porta a terme amb errors típics per baix de l'1% al percentil 95% en els tests idealitzats, mentre que en aquells més complexos, on s'inclouen camps de velocitat amb fluctuacions en quasi tres ordres de magnitud en escala, els errors medians són de l'ordre de l'1%.

Conclusions. El procediment presentat en aquest treball proporciona un algorisme eficient i raonablement precís per a portar a terme l'HHD, i ha sigut aplicat a estudis posteriors, tant dins com fora d'aquesta Tesi Doctoral. A més, el codi pot ser adaptat per a altres tipus de simulacions, incloent-hi les simulacions basades en partícules.

Adaptació de l'algorisme per a simulacions basades en partícules. Aquests resultats corresponen al treball desenvolupat durant una estada a l'Universtäts-Sternwarte de Munic, de maig a juliol del 2023, sota la supervisió del Prof. Klaus Dolag.

Amb l'objectiu de generalitzar els mètodes desenvolupats en el marc d'aquest Doctorat, el codi VORTEX s'ha adaptat per a ser capaç d'analitzar simulacions basades en partícules o de malla mòbil. Amb l'esperit de preservar tot el possible del codi VORTEX original, la implementació d'aquesta versió, anomenada VORTEX-P, es basa en la interpolació del camp de velocitats des de les partícules a una jerarquia de malles ad-hoc.

Aquesta implementació s'ha validat amb una sèrie de tests similars als del codi original, i s'ha aplicat a simulacions SPH i de massa finita sense malla (MFM), amb i sense viscositat física, per tal de demostrar les capacitats d'aquest codi. Aquests resultats corresponen a un treball que ha sigut enviat i recentment acceptat a la revista *Computer Physics Communications*.

Una revisió profunda del codi ASOHF

Aquesta secció es basa en Vallés-Pérez, Planelles, and Quilis (2022)[9], i el text complet de l'article es pot trobar a l'Apèndix A5. La implementació de l'algorisme presentat en aquesta publicació, ASOHF, és de lliure accés.[10]

Context. L'anàlisi de simulacions cosmològiques requerix, sovint, la identificació de les estructures que s'hi formen. En aquest context, els halos de matèria fosca són els blocs de construcció de l'estructura a gran escala, generant els pous de potencial gravitacional que desencadenen el col·lapse de les galàxies i cúmuls de galàxies observats. Encara que un halo de matèria fosca és, en essència, una estructura localment sobredensa i lligada gravitacionalment, no existix una definició operativa única i s'han desenvolupat diverses tècniques per a la seua identificació. A més, la tendència creixent en les capacitats computacionals ha permés que les simulacions cresquen en mida i resolució, de manera que l'anàlisi de les simulacions esdevé un problema computacional en si mateix.

Objectius. La revisió del codi ASOHF, basat en el mètode de sobredensitat esfèrica, amb l'objectiu de: (i) millorar el seu rendiment paral·lel i la seua capacitat per a tractar simulacions molt grans, (ii) millorar la seua capacitat per a identificar subestructura, (iii) identificar galàxies dins dels halos de matèria fosca.

Mètodes. ASOHF es basa en el paradigma de la sobredensitat esfèrica, juntament amb una interpolació de densitat multiresolució per a detectar subestructura. Les principals novetats d'aquesta revisió inclouen la nova definició de subestructura, els esquemes de recentrat, la capacitat per a identificar galàxies, la descomposició del domini, un nou procediment de construcció de l'arbre de fusions, entre d'altres. L'algorisme està implementat en FORTRAN i paral·lelitzat d'acord amb les directrius estàndard d'OPENMP.

[9] D. Vallés-Pérez, S. Planelles, and V. Quilis. "The halo-finding problem revisited: a deep revision of the ASOHF code." In: *Astron. Astrophys.* 664, A42 (Aug. 2022), A42. DOI: 10.1051/0004-6361/202243712

[10] https://github.com/dvallesp/ASOHF.

Resultats. El codi ha sigut validat amb una sèrie de tests idealitzats amb l'objectiu de verificar el rendiment dels diferents processos crítics de l'algorisme de detecció d'halos. A més, el seu rendiment ha sigut comparat amb altres cercadors d'halos disponibles públicament, mostrant un bon acord i un rendiment notable en la detecció de subestructura.

Conclusions. La nova versió d'ASOHF produïx resultats notables en termes de detecció d'halos, subhalos i galàxies, rendiment paral·lel i cost computacional reduït.

ASOHF com a buscador de galàxies: aplicació a les simulacions DIANOGA. Com a verificació de les capacitats d'ASOHF per a identificar galàxies, s'ha aplicat aquest codi a una de les regions de DIANOGA. Els catàlegs de galàxies obtinguts s'han comparat amb els proporcionats per SUBFIND, mostrant un bon acord entre ambdós, tant pel que fa a la coincidència entre galàxies, com pel que fa a les seues propietats.

Formació de cúmuls de galàxies

L'assemblatge dels cúmuls de galàxies determina moltes de les seues propietats, des de les seues poblacions estel·lars fins als estats termodinàmics i la cinemàtica de l'ICM. Aquesta secció resumix els treballs realitzats en el marc d'aquesta Tesi Doctoral amb l'objectiu d'explorar la descripció de l'assemblatge dels cúmuls de galàxies, i el seu impacte en les propietats d'aquests objectes, a partir de l'estudi de les taxes d'acreció i els indicadors d'estat dinàmic.

Sobre la història d'acreció dels cúmuls de galàxies

Aquesta secció es basa en Vallés-Pérez, Planelles, and Quilis (2020)[11], i el text complet de l'article es pot trobar a l'Apèndix A1.

Context. Els cúmuls de galàxies creixen a través de l'acreció altament anisotròpica de matèria de la teranyina còsmica circumdant, així com a través de fusions amb altres estructures de mida similar. Treballs anteriors han confirmat la relació entre l'acreció i diverses propietats de l'ICM, com els perfils de diverses quantitats termodinàmiques, la posició dels cúmuls respecte a les relacions d'escala, o l'el·lipticitat de l'ICM.

[11] D. Vallés-Pérez, S. Planelles, and V. Quilis. "On the accretion history of galaxy clusters: temporal and spatial distribution." In: *Mon. Not. R. Astron. Soc.* 499.2 (Dec. 2020), pp. 2303–2318. DOI: 10.1093/mnras/staa3035

Objectius. La caracterització de l'acreció de matèria d'una petita mostra de cúmuls de galàxies a partir d'una simulació numèrica, amb l'objectiu d'examinar diferents indicadors de la taxa d'acreció, així com l'impacte de l'acreció en les propietats internes del cúmul, i la distribució espacial dels fluxos d'acreció.

Mètodes. S'han comparat diversos indicadors de la taxa d'acreció (MAR), incloent-hi $\Gamma_{200\mathrm{m}}$ calculat a partir de l'arbre de fusions complet dels halos, α_{200m} calculat a partir del perfil de velocitat, i el flux de massa integrat a través del radi virial. $\Gamma_{200\mathrm{m}}$ s'ha comparat, a més, amb la presència d'esdeveniments de fusió i les densitats circumdants al cúmul. Finalment, la distribució espacial del flux d'acreció s'ha estimat mitjançant un enfocament pseudolagrangià sobre les dades eulerianes del gas de la simulació, i s'ha analitzat utilitzant anàlisi multipolar.

Resultats. Els diferents indicadors de la MAR estan feblement correlacionats entre ells, destacant la dificultat de trobar bones mesures del creixement del cúmul associades a l'arbitrarietat en la definició del seu límit. Tant la MAR bariònica com la total es correlacionen amb la presència d'esdeveniments de fusió, però, quan es mesuren a R_{200m}, es mantenen altes durant molt de temps després dels episodis de fusió. L'enfocament novell per a estudiar la distribució angular del flux d'acreció remarca el comportament altament anisotròpic del flux de matèria en la frontera virial, i la seua anàlisi multipolar mostra diferències significatives en algunes propietats termodinàmiques com l'entropia del gas acretat.

Conclusions. En l'advent de pròximes campanyes observacionals que seran capaces d'observar les regions externes dels cúmuls, la caracterització de les taxes d'acreció i la seua comparació amb els observables del cúmul és d'importància cabdal. Aquest treball exploratori exemplifica alguns enfocaments temptatius, però es requerix treball amb mostres més grans per a traure conclusions robustes.

L'estat d'assemblatge dels halos de matèria fosca a través de la història còsmica

Aquesta secció es basa en Vallés-Pérez et al. (2023)[12], i el text complet de l'article es pot trobar a l'Apèndix A6.

Context. L'estat dinàmic i les característiques morfològiques de les galàxies i cúmuls de galàxies, així com les dels seus homòlegs de matèria fosca, estan

[12] D. Vallés-Pérez et al. "On the choice of the most suitable indicator for the assembly state of dark matter haloes through cosmic time." In: *Mon. Not. R. Astron. Soc.* 519.4 (Mar. 2023), pp. 6111–6125. DOI: 10.1093/mnras/stad059

estretament relacionades amb la seua història d'assemblatge al llarg d'un ampli interval de redshifts. Per tant, aquestes característiques poden contenir informació crucial sobre la formació i l'evolució d'aquestes estructures còsmiques.

Objectius. Aquest treball pretén examinar críticament com l'estat d'assemblatge d'aquestes estructures, entés com la presència de fusions recents o períodes d'acreció forta, pot ser determinat a partir d'una sèrie de propietats a un instant donat.

Mètodes. Utilitzant la combinació de la desviació del centre, el quocient virial, la velocitat radial mitjana, la dispersió del perfil de densitat i l'el·lipticitat dels halos de matèria fosca extrets d'una simulació de volum moderat, estudiem com els llindars d'aquests paràmetres, així com els seus pesos relatius per a definir un indicador combinat, han d'evolucionar amb el redshift per a correlacionar-se el màxim possible amb l'activitat de fusió i d'acreció inferida a partir dels arbres de fusions.

Resultats. La classificació resultant, que involucra una categoria *totalment relaxada*, una *no relaxada*, i una intermèdia (*marginalment relaxada*), es correlaciona fortament amb l'activitat de fusió extreta dels arbres de fusions, així com amb les taxes d'acreció. Açò s'ha comprovat també amb dades d'una simulació independent, del projecte CAMELS.

Conclusions. Una classificació de l'estat dinàmic centrada en la història d'assemblatge (presència de fusions i taxes d'acreció) ha de tindre en compte que diferents indicadors poden proporcionar més informació que altres en un temps còsmic particular. De la mateixa manera, els llindars que s'han d'aplicar a aquests indicadors per a produir la millor classificació també semblen dependre del redshift.

Turbulència i xocs durant l'evolució de cúmuls de galàxies

Tal com s'ha descrit al Marc Teòric, els fluxos turbulents i les ones de xoc són dos fenòmens associats a la no-linealitat de les equacions de la hidrodinàmica que tenen conseqüències molt importants per a la física de l'ICM, així com per a les seues propietats observables. En aquesta secció, es resumixen els resultats de dos estudis relacionats amb aquests fenòmens.

Turbulència, enstròfia i helicitat a l'ICM

Aquesta secció es basa en Vallés-Pérez, Planelles, and Quilis (2021a)[13], i el text complet de l'article es pot trobar a l'Apèndix A3. Els mètodes desenvolupats per a realitzar una descomposició de Reynolds sobre camps de velocitat definits en una malla AMR basada en blocs han sigut inclosos en el codi públic VORTEX.[14]

Context. La turbulència és un fenomen present en molts fluxos astrofísics i, en particular, en els cúmuls de galàxies. No obstant això, és particularment difícil de modelitzar numèricament, a causa de la seua naturalesa intrínsecament multiescala, que requerix altes resolucions. A més, treballs anteriors han mostrat que la turbulència està estretament relacionada amb la història de formació de l'ICM.

Objectius. Explorar la connexió entre la generació i la dissipació de turbulència en l'ICM amb la història d'assemblatge (tant de matèria fosca com de barions) dels cúmuls de galàxies.

Mètodes. El codi públic VORTEX s'ha ampliat per a incloure la possibilitat de realitzar una descomposició de Reynolds (component laminar i turbulenta) abans de l'HHD. Açò, aplicat als dos cúmuls de galàxies millor resolts de la simulació, permet extraure estadístiques globals (per exemple, funcions d'estructura compressives i solenoidals) i estadístiques locals (per exemple, enstròfia) dels fluxos turbulents.

Resultats. Encara que les funcions d'estructura mostren una dependència important amb la distància al centre del cúmul, açò es deu, en gran mesura, a la resolució no constant. L'estudi de l'evolució temporal de les funcions d'estructura de segon ordre a escala fixa revela una clara correlació amb les taxes d'acreció i els períodes de fusió. Seguint l'evolució temporal i la distribució espacial i en espai de fases de l'enstròfia, un indicador de moviments solenoidals, els nostres resultats confirmen un escenari proposat prèviament on els moviments solenoidals que omplin el volum són generats, primer, per baroclinicitat en els xocs més externs que envolten el cúmul, i després són amplificats per compressió i canalitzats avall dels xocs per estirament de vòrtexs.

Conclusions. Encara que la mesura de quantitats relacionades amb el flux turbulent en dades AMR, on diferents regions han sigut evolucionades amb diferents resolucions en diferents temps, està lluny de ser senzilla, els resultats reportats

[13] D. Vallés-Pérez, S. Planelles, and V. Quilis. "Troubled cosmic flows: turbulence, enstrophy, and helicity from the assembly history of the intracluster medium." In: *Mon. Not. R. Astron. Soc.* 504.1 (June 2021), pp. 510–527. DOI: 10.1093/mnras/stab880

[14] https://github.com/dvallesp/vortex.

en aquest treball mostren que és possible extraure informació significativa sobre el flux turbulent en l'ICM amb aquestes simulacions.

Xocs d'acreció com una ferramenta per a mesurar la massa dels cúmuls de galàxies

Aquesta secció es basa en Vallés-Pérez, Quilis, and Planelles (2024, in press)[15], que ha sigut acceptat per a la seua publicació en la revista *Nature Astronomy*. El manuscrit original (previ a revisió, d'acord amb la llicència de l'editorial) s'ha inclòs a la Secció 8.2.

Context. Els xocs d'acreció cosmològics creats durant la formació dels cúmuls de galàxies són un fenomen ubic en tot l'Univers. Aquests xocs i les seues característiques estan íntimament relacionats amb l'energia gravitatòria posada en joc durant la formació dels cúmuls de galàxies.

Objectius. Estudiar les relacions d'escala que lliguen la massa dels cúmuls de galàxies amb el radi i la intensitat dels seus xocs d'acreció, i explorar la possibilitat d'utilitzar aquestes relacions per a mesurar la massa dels cúmuls de galàxies.

Mètodes. Emprem una mostra de cúmuls de galàxies i grups de galàxies extrets d'una simulació cosmològica euleriana de mida i resolució moderades ($100\,h^{-1}\,\mathrm{Mpc}$ de costat, resolució pic de $\Delta x \sim 9\,\mathrm{kpc}$). Els xocs s'identifiquen en la simulació utilitzant un algorisme de cerca de xocs basat en la separació de coordenades i els salts de temperatura, i les superfícies dels xocs d'acreció més externs es caracteritzen a partir d'ací amb una sèrie de criteris heurístics.

Resultats. Demostrem que els objectes de la nostra mostra es troben en un plànol dins de l'espai tridimensional de la massa total del cúmul, el radi del xoc i el nombre de Mach (una mesura de la intensitat del xoc).

Conclusions. Utilitzant aquesta relació, i considerant que observacions futures seran capaces de mesurar els radis i les intensitats dels xocs d'acreció, plantegem la idea que el contingut de matèria fosca dels cúmuls de galàxies podria mesurar-se indirectament amb un error d'aproximadament el 30% al nivell de confiança d'1σ. Aquest procediment seria un mètode nou i independent per a mesurar la massa de matèria fosca en les estructures còsmiques, i una nova manera de posar a prova el paradigma ΛCDM.

[15] D. Vallés-Pérez, V. Quilis, and S. Planelles. "Cosmic accretion shocks as a tool to measure the dark matter mass of galaxy clusters." In: *Nat. Astron.* (2024, in press). DOI: 10.1038/s41550-024-02303-x

Buits còsmics

Aquesta secció es basa en Vallés-Pérez, Quilis, and Planelles (2021)[16], i el text complet de l'article es pot trobar a l'Apèndix A4.

Context. Els buits còsmics són regions de baixa densitat que ocupen la major part del volum de l'Univers i emergixen en regions que contenen fluctuacions primordials de densitat negatives, que posteriorment s'expandixen mentre la matèria al seu voltant col·lapsa per a formar parets, filaments i halos.

Objectius. Estudiar la naturalesa dels fluxos de matèria a través de les fronteres dels buits, per a comprovar si el seu camp de velocitat és purament sortint o sorgixen patrons més complexos.

Mètodes. A partir dels resultats d'una simulació cosmològica especialment dissenyada per a descriure amb detall les regions que formaran buits i els seus voltants, s'extrau una mostra de buits definits com els majors el·lipsoides possibles al voltant de mínims de densitat en expansió, possiblement envoltats de gradients de densitat forts. Els fluxos de massa de gas a través de la frontera dels buits s'estimen en post-processament utilitzant un enfocament pseudolagrangià.

Resultats. Contràriament a l'expectativa comuna, al voltant del 10% de la massa de gas en els buits a $z = 0$ ha sigut acretada des de regions sobredenses, arribant a fraccions més grans per a alguns buits (per exemple, $\sim 35\%$ en el percentil 84). A més, seguint els fluxos d'entrada de matèria fosca, una fracció significativa de la massa que entra als buits roman en ells durant llargs períodes de temps, arribant a regions centrals del buit.

Conclusions. Els resultats reportats en aquest treball suggerixen que, si els buits es defineixen a partir del camp de densitat com les regions més grans possibles al voltant dels mínims de densitat, llavors no és possible afirmar que el seu camp de velocitat és purament sortint, fins i tot per als més grans. L'existència d'entrada de matèria als buits, una fracció d'ella procedent de gas que ha habitat regions més denses i que posteriorment ha sigut deslligat, pot tindre conseqüències importants per al paradigma de la formació de galàxies dins dels buits.

Conclusions i perspectives

Aquesta Tesi Doctoral presenta resultats referents a la recerca teòrica i numèrica al voltant de la natura, les propietats i l'evolució de diversos aspectes relacionats

[16] D. Vallés-Pérez, V. Quilis, and S. Planelles. "Void Replenishment: How Voids Accrete Matter Over Cosmic History." In: *Astrophys. J. Lett.* 920.1, L2 (Oct. 2021), p. L2. DOI: 10.3847/2041-8213/ac2816

amb els fluxos còsmics, amb especial atenció als cúmuls de galàxies i els buits còsmics com a principals objectes d'estudi. El treball reportat en aquesta Tesi, principalment contingut en les sis publicacions que constituïxen l'Apèndix A, pot ser emmarcat generalment en el camp de la Cosmologia Numèrica o Computacional. Més enllà dels treballs científics publicats, aquest Doctorat també ha suposat un esforç important en el desenvolupament de noves eines numèriques, que han sigut fetes públiques i de lliure accés per a la comunitat científica.

En cadascun dels capítols de la Part II es descriuen les línies de continuació del treball desenvolupat durant aquest Doctorat, tant les més immediates com aquelles a mitjà i llarg termini. A manera de resum breu, les eines numèriques dissenyades i desenvolupades durant aquesta Tesi, així com el codi cosmològic MASCLET, del qual he esdevingut un usuari i contribuïdor, són codis en constant desenvolupament i optimització. En el vessant computacional, les principals línies de continuació inclouen noves estratègies de paral·lelització per tal de fer els codis més eficients i escalables, d'acord amb la tendència actual en computació d'alt rendiment, així com una major interoperabilitat amb codis externs.

Pel que fa al vessant científic, un nou conjunt de simulacions que ja està en preparació ens permetrà estendre i ampliar els resultats presentats en aquesta Tesi i les seues publicacions associades. Conjuntament amb les versions públiques d'ASOHF i VORTEX per a codis basats en partícules, les futures extensions a simulacions majors i més resoltes dels resultats i de les tècniques desenvolupades en aquesta tesi ens permetran explorar amb gran detall la connexió entre l'assemblatge dels cúmuls i la presència de característiques no-lineals (turbulència, xocs, etc.) en els cúmuls de galàxies, així com la seua relació amb observables.

Publications

This PhD Thesis is based on the following publications (listed in chronological order):

- **Vallés-Pérez, D.**, Planelles, S., & Quilis, V. (2020).
  On the Accretion History of Galaxy Clusters: temporal and spatial distribution.
  Monthly Notices of the Royal Astronomical Society, 499(2), 2303-2318.

- **Vallés-Pérez, D.**, Planelles, S., & Quilis, V. (2021).
  Unravelling cosmic velocity flows: a Helmholtz-Hodge decomposition algorithm for cosmological simulations.
  Computer Physics Communications, 263, 107892.

- **Vallés-Pérez, D.**, Planelles, S., & Quilis, V. (2021).
  Troubled cosmic flows: turbulence, enstrophy and helicity from the assembly history of the intracluster medium.
  Monthly Notices of the Royal Astronomical Society, 504(1), 510-527.

- **Vallés-Pérez, D.**, Quilis, V., & Planelles, S. (2021).
  Void Replenishment: How Voids Accrete Matter Over Cosmic History.
  The Astrophysical Journal Letters, 920(1), L2.

- **Vallés-Pérez, D.**, Planelles, S., & Quilis, V. (2022).
  The halo-finding problem revisited: a deep revision of the ASOHF code.
  Astronomy & Astrophysics, 664, A42.

- **Vallés-Pérez, D.**, Planelles, S., Monllor-Berbegal, Ó., & Quilis, V. (2023). *On the choice of the most suitable indicator for the assembly state of dark matter haloes through cosmic time.* Monthly Notices of the Royal Astronomical Society, 519(4), 6111–6125.

The following correspond to work undertaken during this PhD and contained in this Thesis, but were not accepted/published at the time of submission. In chronological order of their submission date:

- **Vallés-Pérez, D.**, Quilis, V., & Planelles, S. (2024) *Cosmic accretion shocks as a tool to measure the dark matter mass of galaxy clusters.* Nature Astronomy, in press (doi:10.1038/s41550-024-02303-x).

- **Vallés-Pérez, D.**, Planelles, S., Quilis, V., Groth, F., Marín-Gilabert, T, & Dolag, K. (2024) *VORTEX-P: A Helmholtz-Hodge+Reynolds decomposition algorithm for particle-based simulations.* Computer Physics Communications, in press.

During the course of the PhD, I have also participated in the following co-authored publications (listed in chronological order):

- Monllor-Berbegal, Ó., **Vallés-Pérez, D.**, Planelles, S., & Quilis, V. (2024) *The imprints of galaxy cluster internal dynamics on the Sunyaev-Zeldovich effect.* Astronomy & Astrophysics, 686, A243.

- López-Miralles, J., Perucho, M., **Vallés-Pérez, D.**, et al. *Unravelling the relativistic magnetised jet dynamics of the microquasar SS 433.* In preparation.

Contents

List of Figures

List of Tables

Tables within the publications

Nomenclature

Acronyms / Abbreviations

AGN	active galactic nuclei
AMR	adaptive mesh refinement
BAO	baryon acoustic oscillations
BCG	brightest cluster galaxy
BH	black hole
CDM	cold dark matter
CIC	cloud-in-cell
CMB	cosmic microwave background
CM	centre of mass
DM	dark matter
EdS	Einstein-de Sitter
FFT	fast Fourier transform
FLRW	Friedmann-Lemâitre-Robertson-Walker
FoF	friends-of-friends

GR	General Relativity
HHD	Helmholtz-Hodge decomposition
HRSC	high-resolution shock-capturing
ICL	intracluster light
ICM	intracluster medium
kSZ	kinetic Sunyaev-Zeldovich
ΛCDM	Λ-cold dark matter
LSS	large-scale structure
MAH	mass assembly history
MAR	mass accretion rate
MFM	meshless finite mass
MHD	magnetohydrodynamics
NFW	Navarro-Frenk-White (density profile)
NGP	nearest grid point
ODE	Ordinary Differential Equation
PDE	Partial Differential Equation
OMP	Open Multi-Processing
PM	particle-mesh
PPM	piecewise-parabolic method
SN(e)	supernova(e)
SPH	smoothed-particle hydrodynamics
SOR	successive overrelaxation
SO	spherical overdensity
SZ	Sunyaev-Zeldovich
TSC	triangular-shaped cloud
tSZ	thermal Sunyaev-Zeldovich

Part I

Background

CHAPTER 1

Introduction

Contents

1.1 Motivation

The work undertaken during the course of this PhD Thesis, and reported in this dissertation, is generally framed within the field of Computational Astrophysics and, more specifically, Numerical Cosmology. In particular, the main scientific objects of study have been galaxy clusters and voids, representing two opposite ends within the range of structures generated by gravitational collapse.

Galaxy clusters (from here on, used interchangeably with *clusters*) were discovered almost 240 years ago, with different motivations, by astronomers Charles Messier and William Herschel. The latter described, already back in 1785, one of the most widely-studied clusters today (Herschel 1785):

> *"[...] that remarkable collection of many hundreds of nebulae which are to be seen in what I have called the nebulous stratum of Coma Berenices."*

However, it was not until a few decades into the 20th century, with the advent of modern Cosmology, that the actual astrophysical comprehension of these objects commenced. A century later, galaxy clusters have become an essential object of study, not only for Cosmology and Astrophysics, but also for more fundamental fields of research, such as Particle Physics, Plasma Physics, etc.

Regarding cosmic voids, their discovery has been more recent and had to wait until the era of the first galaxy surveys (Gregory and Thompson 1978, Jõeveer, Einasto, and Tago 1978). Their interest has steadily grown ever since, because of their potential as cosmological probes and their role in galaxy evolution. Altogether, galaxy clusters and voids represent the two opposite ends of a complex network of structures, known as the cosmic web, containing a hierarchy of haloes, filaments, sheets, and voids.

It is therefore evident that, despite the fact that it has triggered interest since long times ago, cosmological structure formation in its modern form is a relatively young field of research. This is in part due to the obvious fact that, in contrast to many other fields of Physics, it is not possible to directly access the objects of study or to recreate their conditions in a laboratory. Hence, the advancement of this field has inevitably been tightly linked to, first, the development of Observational Astrophysics and, second, the birth of Computational Astrophysics as a fundamental tool to bridge the gap between fundamental physics and observations.

The first simulations of gravitational collapse were performed by Holmberg (1941), who ingeniously devised a method to use the intensity of a set of lightbulbs to compute the gravitational force (since they both decay with distance as $1/r^2$). The earliest computational numerical integrations, however, had to wait until the advent of the first electronic computers (von Hoerner 1960, Aarseth 1963). Around a decade later, Press and Schechter (1974) performed the first truly cosmological simulations of gravitational collapse, while Larson (1969), Peebles (1970) and White (1976) carried out some of the first simulations with hydrodynamics in this context. Since then, in the 50 years of maturity of the field, cosmological simulations have grown in sophistication and realism, and have become an indispensable tool for the study of the cosmic web and its different elements.

The joint effort in Cosmology and Astrophysics, aided by the rapid technological development of the last decades and the consequent ever-growing computational capacities, has allowed to garner a deep understanding on the formation and evolution of galaxies, galaxy clusters, their environments, and the

physical phenomena associated to them. In the last decades, theory, simulations and observations have converged to establish the hierarchical paradigm as the standard scenario for the formation of cosmic structures, and the ΛCDM model has been confirmed as the most successful cosmological model to date, despite their remaining open problems.

Together with this, pioneer works with numerical simulations have revealed, during the past 25 years, the presence of complex features associated to the physics of baryons, such as turbulence, shock waves, or magnetic fields, which have been shown to have a significant impact on the observable properties of galaxy clusters despite the energetically subdominant role of baryons in the cosmic picture. These phenomena, whose study is associated to the non-linearity of the equations governing gravitational clustering and fluid dynamics, highlight once again the importance of numerical simulation in the research around the formation and evolution of cosmic structures.

1.2 Thesis goals and aims

The main, general aim of this Thesis is to explore, quantify and contribute to deepening our knowledge on several aspects associated to the assembly of cosmic structure, which results in the emergence of complex *cosmic flows* of dark and baryonic matter, both bulk and turbulent, that give title to this dissertation. This is attained, primarily, through the design and analysis of cosmological simulations, which are the main tool used in this Thesis. This general objective can be specified in the following, more particular goals:

- The characterisation of the assembly history of galaxy clusters and groups, the intensity of accretion, its qualitative features, evolution, and their relation with several global and local properties of the intracluster medium (ICM). In this same direction, the characterisation of the complex and multifaceted assembly and dynamical state of galaxy groups and clusters through a set of observables has also been a topic of interest during this Thesis.

- The study of the presence, the nature, the role, and the evolution of turbulent motions filling the ICM, and their relation with the assembly history of galaxy clusters referenced in the previous point.

- The investigation on the properties and the location of shocks surrounding galaxy clusters, especially regarding the large-scale accretion shocks associated with the gravitational collapse of the primordial inhomogeneities, and

the relation of these features with the properties of the objects harbouring them.

- The study of the gas flows around voids and, especially, the nature of the velocity field in these regions and its possible consequences for the properties of the diffuse gas in voids.
- The development of new numerical tools, as well as the optimisation of existing ones, to allow the analysis of cosmological simulations. A constant line of work throughout this Thesis has been the extension of the developed tools to enable them to be applied to the analysis of other simulation codes of different nature, as well as the public release of any code that has been mainly developed during the course of the Thesis.

1.3 Outline of the Thesis

This Thesis is lay out as a compendium of publications, which comprise the main body of work conducted during the duration of the PhD. To comply with the requirements imposed by the *Escola de Doctorat* of the Universitat de València, all publications are deferred to Appendix A. The main body of the Thesis is structured as follows:

- Part I (*Background*) contains the theoretical and bibliographic review of the topics covered in the Thesis. It is divided in four chapters:
  - Chapter 2 overviews the basic physics around the formation of cosmic structures, including a review of the ΛCDM model, the growth of cosmological perturbations and, in particular, the formation of galaxy clusters and voids.
  - Chapter 3 describes the basic mathematical modelling and numerical tools for addressing the problem of cosmological structure formation as an initial-value problem.
  - In Chapter 4, the main features of MASCLET, the code employed to perform most of the simulations used in this Thesis, are described.
  - Chapter 5 gives an overview of some phenomena associated to the non-linearity of the evolution of the baryonic component of the Universe, that are particularly relevant for the study of cosmic flows.
- Part II (*Results*) contains a brief summary of each publication, together with some details on unpublished results. It is divided in four chapters:

  - Chapter 6 covers the works related to the development of numerical tools for the analysis of cosmological simulations.
  - Chapter 7 covers the works related to the study of the assembly history of galaxy groups and clusters.
  - Chapter 8 describes the results of the study of turbulence and accretion shocks during the formation of galaxy groups and clusters.
  - The results obtained from the study of cosmic flows around voids are explained in Chapter 9.

- Part III (*Conclusions*) highlights the main conclusions of this Thesis and lines out some future perspectives. It is lay out in a single chapter, Chapter 10.

- Part IV corresponds to the Appendices.

  - Appendix A contains the publications that constitute the main body of work of this PhD Thesis, in chronological order of their publication dates, according to the regulations of the *Escola de Doctorat* of the Universitat de València.
  - Appendix B contains a brief summary of the publications where I have participated as a co-author, outside the main body of work of this Thesis, together with a short description of my contribution to each of them.

CHAPTER 2

Cosmological structure formation

Contents

The formation and the evolution of the cosmological structures studied through this PhD Thesis are intimately interwoven with the evolution of the Universe as a whole. In this Chapter, I give a brief overview of the physics of cosmological structure formation. This includes an introduction to the ΛCDM model (Sec. 2.1) and to the paradigm of the growth of perturbations over a cosmological background (Sec. 2.2), as well as a review of the main properties, from the physical and the observational points of view, of galaxy clusters (Sec. 2.3) and cosmic voids (Sec. 2.4).

2.1 The ΛCDM model

The standard cosmological model, also dubbed the Λ-cold dark matter (ΛCDM) or *concordance* model, has been the culmination of a century of work around the physics of structure formation under the framework of General Relativity (GR), and it has turned to be remarkably successful in describing the evolution of the

large-scale structure of the Universe. Nevertheless, several fundamental and practical issues still remain open, especially when dealing with the smallest scales (length scales below $\sim 1\,\mathrm{Mpc}$, and masses below $\sim 10^{11}\,\mathrm{M}_\odot$; see Hamilton 2014, Bullock and Boylan-Kolchin 2017, for reviews). In the following paragraphs, its fundaments, the observations supporting the model, and its open issues are briefly overviewed.

2.1.1 The Friedman-Lemâitre-Robertson-Walker metric

The ΛCDM cosmological paradigm is a particular case of a larger class of models —the Friedman-Lemâitre-Robertson-Walker (FLRW) models; see Peebles (1993) for a textbook introduction—, which stem from two major assumptions:

- Gravity is the main force driving the evolution of the Universe (on its cosmological scales), and it can be well described by Einstein's field equations of GR.
- When smoothed on sufficiently large scales (at least, above a few 100 Mpc; Yadav, Bagla, and Khandai 2010), the Universe can be modelled as a homogeneous and isotropic system (i.e., the Cosmological principle, ultimately stemming from the Copernican principle which states the absence of privileged observers).

Under these assumptions, and when written in the appropriate coordinate system (the so-called *comoving*, spherical coordinates), the FLRW metric is usually given by the arc element ($\mathrm{d}s$):

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + a^2(t)\left[\frac{\mathrm{d}r^2}{1-kr^2} + r^2\mathrm{d}\Omega^2\right], \tag{2.1}$$

where c is the speed of light, t is the time measured by the comoving observer, r and $\Omega(\theta,\phi)$ are the usual radial and unit sphere coordinates, k is the scalar curvature of the FLRW metric and $a(t)$ is the scale factor. The scalar curvature k can take the values of -1, 0 or 1, corresponding to open, flat and closed universes, respectively. The scale factor $a(t)$ encodes the expansion history of the Universe and, in the case of flat ($k = 0$) universes, it can be normalised so that $a(t_0) \equiv a_0 = 1$, where t_0 is the present time. Hence, a distance $D_\mathrm{c} = \int \mathrm{d}r$ will be referred to as a *comoving* distance, while $D_\mathrm{p} = \int a(t)\mathrm{d}r$ will be a *physical* or *proper* distance. A particularly useful quantity is the Hubble parameter, $H(t) \equiv \dot{a}/a$, which measures the rate of expansion of the

Universe. Its value at the present time, or the *Hubble constant*, is often given as $H_0 \equiv 100\,h\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, where $h \sim 0.7$ is the *dimensionless Hubble parameter*.

2.1.2 The energetic budget of the Universe

The description of a FLRW model is complete once its parameters, k and $a(t)$, are specified by solving Einstein's field equations, written in tensor form as $\mathbf{G} + \Lambda\mathbf{g} = \frac{8\pi G}{c^4}\mathbf{T}$. Here, $\mathbf{g}$ is the metric tensor, $\mathbf{G}$ is the Einstein tensor (a second-order differential concomitant of the metric tensor), and $\mathbf{T}$ is the energy-momentum tensor. Assuming $\mathbf{T}$ to be given by a perfect fluid, Einstein equations reduce to the so-called Friedman equations:

$$\left(\frac{\dot{a}(t)}{a(t)}\right)^2 = \frac{8\pi G}{3}\rho_\mathrm{B}(t) - \frac{kc^2}{a(t)^2} + \frac{\Lambda}{3}, \tag{2.2}$$

$$\dot{\rho}_\mathrm{B}(t) + 3H(t)\left(\rho_\mathrm{B}(t) + \frac{p_\mathrm{B}(t)}{c^2}\right) = 0, \tag{2.3}$$

where $\rho_\mathrm{B}(t)$ is the density of the background (i.e., homogeneous) Universe and $p_\mathrm{B}(t)$ its pressure. Generally, the composition of the Universe is assumed to be a mixture of several components, each of them with a given equation of state $p_\mathrm{B} = w\rho_\mathrm{B}c^2$, where w is the equation of state parameter. Therefore, the composition of the Universe can be specified by the equation of state parameter of each component, together with the values of their densities at one time (typically, at the present time, ρ_{X0}). The latter are often given in terms of the dimensionless density parameters $\Omega_{X0} \equiv \rho_{X0}/\rho_{\mathrm{c}0}$, where $\rho_{\mathrm{c}0} = 3H_0^2/8\pi G$ is the critical density (at $z = 0$), or the value of ρ_B that would make a universe with $\Lambda = 0$ flat ($k = 0$).

In the ΛCDM paradigm (Blumenthal et al. 1984, Deruelle and Uzan 2018), the main constituents of the Universe are dark matter (DM; a non-electromagnetically interacting form of gravitational matter whose fundamental nature is yet unknown) and baryonic (or *ordinary*) matter, accounting for $\Omega_\mathrm{DM} \sim 0.25$ and $\Omega_\mathrm{b} \sim 0.05$, both of which are pressureless ($w = 0$) due to being non-relativistic species. Incidentally, the cosmological constant term can be interpreted as an energy density, $\rho_\Lambda \equiv \Lambda/8\pi G$, with negative pressure ($w = -1$) so that its density remains constant according to Eq. (2.3). This elusive form of *dark energy* makes up to $\Omega_\Lambda \sim 0.7$ of the total energy content of the Universe, making it consistent with flatness ($k = 0$).

2.1.3 Observational probes supporting the ΛCDM model

The predictions of ΛCDM have been tested against a wide variety of observations, which range from the cosmic microwave background (CMB) to the large-scale structure (LSS) of the Universe. Besides serving as tests for the model, many of these probes are able to provide complementary and stringent constraints on the values of the free parameters of the model. In the following paragraphs, some of the most relevant probes are briefly described.

Expansion of the Universe. Naturally, the first probe in this category corresponds to the one by Hubble (1929), who interpreted the fact that galaxies are redshifted proportionally to their distance to the observer as a (relativistic) Doppler effect from the recession of galaxies. This is only made compatible with the cosmological principle if the Universe is globally expanding. The rate of this cosmic expansion at the present time, after decades of long-standing debate, has been constrained to yield around $h \sim 0.7$. More recently, the observations of distant type-Ia supernovae (SN), dimmer than expected, have been used to obtain precise measurements of the expansion rate at late times, suggesting an acceleration of the expansion rate and providing compelling evidence for the necessity of dark energy (Riess et al. 1998, Perlmutter et al. 1998, Perlmutter, Turner, and White 1999).
Incidentally, the expansion of the Universe provides a useful manner to measure distances and times. The cosmological redshift, z, is defined as a dimensionless measure of time, since it can be shown to be related to the expansion factor by

$$a(t) = \frac{1}{1+z}, \tag{2.4}$$

and the expansion factor itself, $a(t)$, is a monotonous function of cosmic time for flat and open universes.

Cosmic microwave background. The CMB, which originated at the epoch of recombination at around $z \simeq 1100$, was predicted by Gamow (1948), and only serendipitously discovered two decades later by Penzias and Wilson (1965) and promptly interpreted by Dicke et al. (1965). Despite its extreme isotropy (with relative fluctuations only in the order of $\sim 10^{-5}$), subsequent missions (e.g., COBE, Smoot et al. 1992; WMAP, Spergel et al. 2003; Planck, Planck Collaboration et al. 2020) have been able to measure the CMB temperature with increasing precision and angular resolution. The spectral study of the primary temperature fluctuations of the CMB (i.e., the ones that were imprinted on

the surface of last scattering, including the primordial fluctuations [see below, Sec. 2.2] and the baryonic acoustic oscillations [BAO]; Hu and Dodelson 2002) provides stringent constraints on the cosmological parameters (see, for example, Komatsu et al. 2011, Planck Collaboration et al. 2020), becoming one of the main cornerstones of the cosmological model.

Abundance of light elements. The production of light atomic nuclei other than ^{1}H during the first minutes after the *Big-Bang* is extremely sensitive to the thermal history of the Universe and, in particular, to the baryonic matter density. Big-Bang nucleosynthesis (Alpher, Bethe, and Gamow 1948) allows to predict the abundances of ^{2}H, ^{3}He, ^{4}He and ^{7}Li, providing an independent probe to Ω_b and serving as a fundamental test to the cosmological model.

Large-scale structure. The LSS of the Universe and, in particular, the clustering of matter at $\gtrsim$ Mpc scales, allows placing additional constraints on several cosmological parameters. For instance, large-scale galaxy surveys, such as CfA (de Lapparent, Geller, and Huchra 1986), SDSS (York et al. 2000), 2dF (Colless et al. 2001) or, more recently, DEEP2 (Newman et al. 2013) and GAMA (Liske et al. 2015), can be used to constrain the autocorrelation power spectra, or the *mass function* (i.e., the mass distribution of gravitationally-bound objects), which are particularly sensitive on $\Omega_m \equiv \Omega_{\rm DM} + \Omega_b$ and σ_8 (see Sec. 2.2 for its definition). Additionally, X-ray or Sunyaev-Zeldovich (SZ) observations can be used to measure the baryon fraction in galaxy clusters, providing a probe for Ω_b.

Cluster and galaxy dynamics. Perhaps the earliest example of cosmology with galaxy clusters dates back to Zwicky's (1933) inference of the presence of DM from the study of the velocity dispersion of galaxies in the Coma cluster. Although it did not have a large initial impact, the interest in DM resurged a few decades later, after the observations of Rubin, Ford, and Thonnard (1980) on the rotation curve of spiral galaxies. As of today, observations of the hot, X-ray emitting gas in galaxy clusters (Böhringer and Werner 2010) and gravitational lensing (Bartelmann 2010) are customarily used to probe the distribution of DM in clusters, providing complementary probes supporting and/or constraining the cosmological model. Recently, it has been shown that, even with a single galaxy, it is possible to extract information about the cosmological model (mainly, in the form of constraints on Ω_m; see Villaescusa-Navarro et al. 2022).

As an example of the information garnered by different probes, Fig. 2.1 shows the constraints that BAOs, type-Ia SNe, and the CMB impose on the

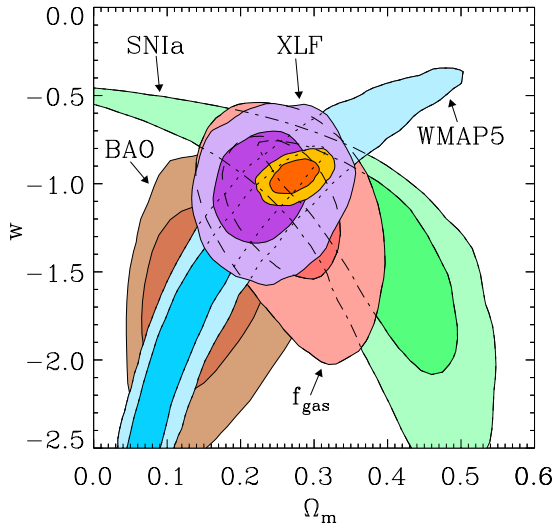


Figure 2.1: Example of constraining power of different probes on the equation of state of dark energy (w) and the matter density parameter (Ω_m). Differently coloured regions correspond to BAOs, type Ia SNe, CMB (from WMAP data), gas mass fraction in galaxy clusters ($f_{\rm gas}$), and X-ray luminosity function (XLF). The yellow contour represents the confidence region combining the five probes. Figure reproduced from Mantz et al. (2010) with permission.

joint distribution of the dark energy equation of state (w) and matter density parameter (Ω_m).

2.1.4 Open issues of the ΛCDM model

Besides the lack of a fundamental explanation for the nature of dark matter and dark energy, the ΛCDM model still has several open issues regarding structure formation, especially at the smallest scales. Some of these issues are briefly described in the following paragraphs.

The core-cusp problem. Predictions from DM-only simulations within the ΛCDM framework point at a cuspy density profile for DM haloes, where density scales as $\rho(r) \propto r^{-\gamma}$ with $\gamma \approx 1$ at small radii (Navarro, Frenk, and White 1997, Navarro et al. 2010). However, rotation curves of low-mass galaxies tend to prefer a cored density profile, i.e., $\gamma \approx 0$ (Walker and Peñarrubia 2011).

The missing satellites' problem. High-resolution cosmological simulations predict that, rather than being monolithic objects, dark matter haloes contain a rich amount of substructure at all scales. In a Milky Way-like halo, for instance, thousands of subhaloes with masses above $10^7 M_\odot$ would be expected from the

results of DM simulations (Springel et al. 2008). However, only around 50 satellites have been observed around the Milky Way (Drlica-Wagner et al. 2015).

The too-big-to-fail problem. One possible solution to the missing satellite problem could be related to the fact that most of the missing satellites correspond to faint, yet undetected galaxies, in such a way that the observed Milky Way satellites correspond to the largest subhaloes. The most massive subhaloes in DM-only simulations are expected to be the most concentrated ones, and therefore the most easily detectable. However, the measured central mass of the Milky Way satellites are systematically lower than that of the most massive DM haloes in high-resolution simulations, suggesting that these high-mass haloes would be failing at forming galaxies (Boylan-Kolchin, Bullock, and Kaplinghat 2011).

These and other problems (e.g., the planes of satellites, Kroupa, Theis, and Boily 2005; or the phenomenology of dark, e.g. Simon and Geha 2007, relic, e.g. Quilis and Trujillo 2013, and ultradiffuse, e.g. Román and Trujillo 2017, galaxies) have been the subject of intense debate in the last decades. It may be possible that the solution to these problems is found in the modification of the cosmological model, such as the introduction of self-interacting DM (Moore et al. 2000, Tulin and Yu 2018), or the modification of the gravitational force at small scales (Famaey and McGaugh 2012). Nevertheless, there is still room for many of these conundrums to still find an answer lying within the ΛCDM model. For instance, baryonic feedback from SNe and active galactic nuclei (AGN) activity may be capable of erasing DM cusps (Mashchenko, Wadsley, and Couchman 2008), thus motivating the importance of deepening the understanding of the physics of baryons.

2.2 Growth of perturbations over a cosmological background

In Sec. 2.1, a homogeneous universe of mean values has been introduced to discuss its global properties and evolution. This section succinctly discusses the evolution of inhomogeneities over this background, which ultimately lead to the formation of the structures populating the Universe.

2.2.1 The seed of density perturbations

In order for structures to gravitationally collapse, there must be some inhomogeneities already present in the density field at high redshift. While the origin of these *primordial* inhomogeneities is not fully understood, the most widely accepted mechanisms place the origin of these perturbations in the inflationary epoch. Cosmic inflation, a period of exponential expansion driven by a constant density scalar field, which took place at very early cosmic times ($t \lesssim 10^{-32}\,\mathrm{s}$), was initially proposed to solve several open problems in Cosmology, such as the horizon problem and the flatness problem (Guth 1981, Linde 1982) or the lack of magnetic monopoles (Guth and Weinberg 1981). Incidentally, this model also predicts the spatial magnification of quantum fluctuations up to cosmic scales, thus providing a mechanism for the emergence of tiny density fluctuations that would constitute the seed for structure formation (see also García-Bellido 1999, Mukhanov 2005 for reviews).

Due to the weakness of its interaction with baryonic matter, DM was able to decouple from the primordial plasma at early times and start its collapse, forming potential wells. Baryonic matter, instead, would remain coupled to photons, preventing its collapse, until the epoch of recombination. This difference in the behaviour of DM and baryons prior to recombination is imprinted in the CMB via several effects, such as the BAOs or the non-integrated Sachs-Wolfe effect.

Inflationary models predict the power spectrum of primordial fluctuations to be self-similar, i.e., given by a power law (Guth and Pi 1982),

$$P(k) = Ak^{n_s}, \tag{2.5}$$

where A is the amplitude of the power spectrum and $n_s \sim 1$ is its spectral index. Instead of A, it is frequent to use the variance of the primordial density field in spheres of $8\,h^{-1}\,\mathrm{Mpc}$, σ_8, as a measure of the amplitude of the power spectrum. The baryonic effects mentioned above are often parametrised by a transfer function (Eisenstein and Hu 1998; see also Fig. 2.2), $T(k)$, in such a way that the power spectrum at the epoch of recombination can be given by

$$P(k) = A|T(k)|^2 k^{n_s}. \tag{2.6}$$

2.2.2 The linear model for the evolution of perturbations

Even though the behaviour of density perturbations is generally not analytically solvable (see below, Sec. 2.2.4 and Sec. 3), some knowledge can be gained

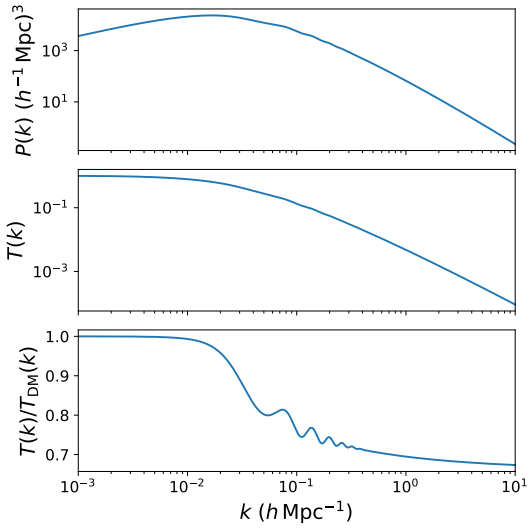


Figure 2.2: *Top panel:* Power spectrum for the primordial density fluctuations (linearly extrapolated at $z = 0$), generated using the Colossus library (Diemer 2018) and assuming a Planck Collaboration et al. (2020) cosmology. *Middle panel:* Transfer function, $T(k)$, for the same cosmology. *Bottom panel:* Ratio between the transfer function shown above, and the one for an equivalent cosmology with only cold dark matter and no baryons.

from the analytic study of the linear regime, i.e., when deviations from the homogeneous background are small in magnitude. It is useful to define the density contrast,

$$\delta(\mathbf{x}, t) \equiv \frac{\rho(\mathbf{x}, t) - \rho_{\mathrm{B}}(t)}{\rho_{\mathrm{B}}(t)}, \tag{2.7}$$

which, according to the linearised continuity and Euler equations for a self-gravitating, pressureless fluid (see Sec. 3.1.2 for a more thorough discussion of these equations), can be shown to evolve following (Peebles 1980):

$$\frac{\partial^2 \delta}{\partial t^2} + 2H\frac{\partial \delta}{\partial t} = 4\pi G \rho_{\mathrm{B}} \delta, \tag{2.8}$$

which admits a separable solution in the form

$$\delta(\mathbf{x}, t) = A(\mathbf{x}) D_+(t) + B(\mathbf{x}) D_-(t), \tag{2.9}$$

where $D_{+(-)}(t)$ are the growing (decaying) linear growth factors, and $A(\mathbf{x})$ and $B(\mathbf{x})$ are set by the initial conditions. In what follows, I shall adopt an Einstein-de Sitter (EdS; flat, $k = 0$, and without cosmological constant, $\Omega_\Lambda = 0$,

hence $\Omega_m = 1$) cosmology for simplicity, although equivalent results can also be obtained for ΛCDM (often not admitting closed forms and requiring numerical integration). In an EdS cosmology, the linear growth factors are given by $D_+(t) \propto t^{2/3} \propto a(t)$, $D_-(t) \propto t^{-1}$. Therefore, once the decaying mode can be neglected, the density contrast grows as $\delta \propto a(t)$, i.e., linear overdensities (underdensities) collapse (expand) with the same rate as cosmic expansion in a matter-dominated universe.

The linear analysis based on the determination of the growth factor for a given cosmology, although only valid for $|\delta| \ll 1$, provides important insight on the evolution of inhomogeneities in different cosmologies. For instance, in a universe dominated by dark energy ($\Omega_\Lambda \sim 1$), the growing mode is constant and, thus, the collapse of structures is suppressed with respect to the matter-dominated one, consequently preventing the formation of structures at late times.

2.2.3 The spherical collapse model

One of the few situations where the exact solution for the evolution of a non-linear perturbation can be obtained is the collapse of a spherically-symmetric, homogeneous overdensity (usually called *top-hat*, for its flat shape) in a matter-dominated Universe. In this case, the evolution can be studied as an isolated FLRW Universe, and it is possible to obtain analytic solutions. The spherical top-hat collapse model (Gunn and Gott 1972, Bertschinger 1985) describes the evolution of an initially homogeneous overdensity $\delta_i \ll 1$, with initial radius R at time t_i, in an EdS cosmology. The motion of each spherical shell can be solved under the assumption that shells do not cross (i.e., if $r_1(t_i) > r_2(t_i)$, then $r_1(t) > r_2(t)\ \forall t$).

The overdensity initially grows in size due to cosmic expansion but, provided $\Omega_m(t_i)(1 + \delta_i) > 1$, the perturbation will reach a maximum radius $R_{\rm ta}$ (the so-called *turn-around radius*) at $t = t_{\rm ta}$, after which it will detach from the background expansion and collapse. Using energetic considerations, the collapse time $t = t_{\rm vir}$ is defined as the moment when the virial theorem, $U + 2K = 0$ is satisfied, being U and K, respectively, the gravitational and kinetic energies of the perturbation. In an EdS Universe, the collapse time is given by

$$t_{\rm vir} = t_i \frac{9\pi}{10} \sqrt{\frac{3}{5\delta_i^3(1+\delta_i)}}, \tag{2.10}$$

from which the density predicted by the linear model ($\delta \propto t^{2/3}$) at collapse can be computed:

$$\delta_{\rm vir} = \delta_+(t_{\rm vir}) = \frac{3}{5}\left(\frac{3\pi}{2}\right)^{2/3} \approx 1.686 \tag{2.11}$$

Note, however, that this universal value for the linear-theory extrapolation of the density at collapse time already falls outside the domain of validity of the linear theory. Conversely, the non-linear overdensity ($\Delta \equiv 1 + \delta$) at the time of collapse is found to be

$$\Delta_{\rm vir} = 18\pi^2 \approx 178. \tag{2.12}$$

Compared to $\delta_{\rm vir} \approx 1.69$, this gives a sense of the rapid evolution of overdensities once the non-linear evolution onsets. For a more detailed discussion on the spherical collapse model, the reader is referred to, e.g., Planelles, Schleicher, and Bykov (2015) and Bovy (2023), and references therein.

2.2.4 Beyond the linear model

Although the linear model can be very illuminating and provides a useful description of the evolution of density perturbations at very large scales, which have not yet entered non-linearity, it does not suffice to describe the complex evolution of cosmic structures. Similarly, also the spherical collapse model has important limitations: under the assumption of no shell-crossing, the collapse would continue so that, at $t = 2t_{\rm ta}$, all shells would end up at $r = 0$. This unphysical conclusion is prevented by, for instance, deviations from spherical symmetry causing shell-crossing, where non-linear interactions between shells prevent this collapse.

Notwithstanding the fact that other analytic or semi-analytic models, such as the Zeldovich (1970) approximation or higher-order perturbation theory (Buchert 1992), are customarily used to study the evolution of inhomogeneities in the Universe, especially at high redshift or at large scales (where the perturbations remain linear or mildly non-linear), the scientific community soon resorted to cosmological, N-Body simulations to study the evolution of cosmic density fields in the non-linear regime (e.g., Press and Schechter 1974; see Sec. 3 for a more detailed discussion on simulations, both N-Body and hydrodynamic).

Complemented with data from observations and simulations, our current understanding of the formation of cosmic structures indicates that structures form:

- Anisotropically, i.e., galaxies (hence, haloes) are not distributed homogeneously in space in scales of $\sim 10\,$Mpc, but they form topologically distinct

regions (Bond, Kofman, and Pogosyan 1996). Matter tends to accumulate in two-dimensional structures called sheets, which surround large regions of low matter density (voids). In turn, sheets are not homogeneous, but they are instead traversed by filaments, which intersect at nodes, where the most massive structures (clusters) reside (Bhavsar and Ling 1988, Cautun et al. 2014).

- Hierarchically, i.e., smaller and less massive structures form at earlier times, and subsequently merge to form larger structures (White and Rees 1978). This is also known as the *bottom-up* scenario. Incidentally, this also implies the presence of substructure within cosmic structures: haloes may contain a population of smaller haloes, called sub-haloes (and so on, recursively), which are the remnants of smaller structures that have been accreted by the main halo (e.g., Springel et al. 2008). Similarly, voids are permeated by a tenuous network of filaments, haloes and sub-voids (Sheth and van de Weygaert 2004).

2.2.5 Dark matter haloes

Within the paradigm described above for the non-linear emergence of the cosmic web, gravitationally-bound spheroidal structures of DM –the so-called *dark matter haloes*– are the basic building blocks of the LSS of the Universe, providing the gravitational potential wells that support the collapse of virtually all observed structures between the scales of dwarf galaxies and galaxy clusters. In this framework, the distribution, internal structure, evolutionary history and the surrounding environment of DM haloes are key ingredients to understand the formation and evolution of galaxies and galaxy clusters. Below, I briefly review some basic properties of DM haloes, while the interested reader is referred to, e.g., Zavala and Frenk (2019) for a thorough review.

Definition of a DM halo. Within the spherical collapse model for an EdS universe, the overdensity of a halo in virial equilibrium is $\Delta_{\rm vir,EdS} = 18\pi^2$ (Eq. 2.12). However, in a fully cosmological environment, this has to be taken with caution. In the first place, the dynamical environment the halos inhabit, and their lack of spherical symmetry, makes the definition of a halo somewhat arbitrary. Generally, there is not a clear boundary for a DM halo[1] and this

[1]Nevertheless, there are more physically-motivated definitions of the extent of a DM halo. For instance, the *splashback* radius has been proposed as a physical boundary of haloes (Diemer and Kravtsov 2014), which separates the regions were DM particles are infalling from the regions where they are orbiting the potential centre. However, even with more

brings an unavoidable arbitrariness in its definition. However, it is still common to define the extent of a halo as the spherical region encompassing a density Δ_m times the background (matter) density, or Δ_c times the critical density. While common choices for Δ_m and Δ_c are 200 or 500, a widely-used option in the literature is to choose the virial overdensity (Bryan and Norman 1998),

$$\Delta_{\rm c,vir} = 18\pi^2 + 82x - 39x^2, \qquad \text{where } x \equiv \Omega_m(z) - 1. \tag{2.13}$$

Yet another definition of a halo can be obtained through the friends-of-friends (FoF) definition, which groups together all particles (i.e., DM particles in an N-Body simulation [see Sec. 3.1.1]; or any kinematic tracer, e.g., galaxies) separated by less than a given *linking length*, l, which is customarily defined in terms of the mean particle density $\bar{n}$ as $l = b\bar{n}^{-1/3}$, $b \sim 0.2$ being a dimensionless parameter that approximately sets the overdensity of the resulting haloes in $\Delta \sim b^{-3}$.

Mass function. The mass function of DM haloes is defined as the number density of haloes per unit mass and unit volume, $n(M, z)$. It can alternatively be defined as an anti-cumulative quantity, i.e., the number density of haloes with mass above M at redshift z, in which case I shall denote it $n(> M, z)$ within this Thesis. Numerical simulations roughly predict (Frenk et al. 1988, Crain et al. 2009)

$$n(M, z) \propto M^{-1.9}, \tag{2.14}$$

that is, massive clusters are highly suppressed, in consistency with a bottom-up formation scenario. The normalisation constant is a function of redshift and, possibly, other factors related to the environment (i.e., higher density environments yield a higher normalisation). Mass functions can be reasonably predicted from theoretical grounds using Press and Schechter (1974) or excursion set (Bond et al. 1991) formalisms, and are essential ingredients for cluster cosmology.

Internal structure: density profile. In the past few decades, a large amount of effort was devoted to the study of the internal structure of DM haloes, which in general terms presents a negative logarithmic slope that gets steeper with radius. Navarro, Frenk, and White (1997) found that, across a large range in

physically-motivated definitions, the lack of spherical symmetry and the interactions between different elements of the cosmic web do not leave room for ideal definitions.

halo masses, the spherically-averaged density profile of DM haloes in dynamical equilibrium can be well described by a functional form that is independent of the initial condition of the overdensity, and depends essentially on two parameters — namely, a scale density, ρ_s, and a scale radius, r_s — which can be related to the total mass and concentration of the halo. This functional form, which is widely known as the *Navarro-Frenk-White* (NFW) profile, is given by

$$\rho(r) = \frac{\rho_s}{\frac{r}{r_s}\left(1+\frac{r}{r_s}\right)^2}. \tag{2.15}$$

A fundamental quantity derived from the NFW profile is its concentration, $c = R_{\rm vir}/r_s$. Alternatively, concentration can be defined at any other overdensity as $c_\Delta = R_\Delta/r_s$. Concentration is itself a function of mass and redshift (Dutton and Macciò 2014).

Shape. While, in part for simplicity, haloes are often described by spherical boundaries and spherically-averaged profiles, the shape of DM haloes is —even after removing the substructure— generally triaxial, with a slight preference towards prolateness. The shape of DM haloes is the result of several factors, including the shape of the initial overdensity (Bardeen et al. 1986), and the anisotropic environment, which can affect the shape of the halo through tidal torques (Dubinski and Carlberg 1991, Jing and Suto 2002).

Substructure. Besides a smooth component, DM haloes are populated by large numbers of substructures —the so-called DM subhaloes. In a bottom-up scenario, subhaloes correspond to the remnants of haloes accreted onto the main host, most of which end up stripped due to processes like tidal stripping (Read et al. 2006) or dynamical friction (White 1976). The presence of substructure in a halo has been widely used as an indicator of its past merging history (West and Bothun 1990).

2.3 Galaxy clusters

Galaxy clusters stand in the apex of the hierarchy of gravitationally bound objects, being the latest structures to have collapsed and virialised by $z \sim 0$ (or are still in the process of doing so). Having masses in the order of $10^{14} - 10^{15}\,M_\odot$, and extents of $\sim 5\,$Mpc, the processes associated to their formation are amongst the most energetic ones in the Universe, involving up to $10^{65}\,$erg over time scales of several Gyr. In this section, I briefly review some of the most relevant

properties of galaxy clusters, while the interested reader is referred to, e.g., Kravtsov and Borgani (2012) for a more thorough review.

2.3.1 Physical properties of galaxy clusters and their evolution

As the most massive collapsed structures in the Universe, the composition of galaxy clusters, which is customarily expressed in terms of their mass fractions,

$$f_{\rm DM} = \frac{M_{\rm DM}}{M_{\rm tot}}, \qquad f_{\rm gas} = \frac{M_{\rm gas}}{M_{\rm tot}}, \qquad f_* = \frac{M_*}{M_{\rm tot}}, \tag{2.16}$$

where '$*$' stands for cold baryons (stars), is approximately representative of the cosmic matter budget. Even though these objects were initially discovered from the emission of their stellar component, it only corresponds to a few per cents of the total mass of the cluster ($f_* \sim 0.01 - 0.05$). The baryon fraction ($f_{\rm b} = f_{\rm gas} + f_* \sim 0.14 - 0.16$) is found by cosmological simulations to be consistent or just slightly below the cosmic value (Planelles et al. 2013). The remaining mass being in the form of DM ($f_{\rm DM} \sim 0.85$), much of what has been mentioned in the previous section about the properties of DM haloes applies to galaxy clusters themselves.

Since these objects are still in the process of assembling by $z \simeq 0$, its dynamical state has drawn considerable attention in the recent decades. There are numerous ways of assessing the dynamical states of galaxy clusters, both in three-dimensional simulation data (Shaw et al. 2006, Neto et al. 2007, Cui et al. 2017) which can additionally track the evolution of these objects (e.g., Nelson et al. 2019), and in two-dimensional, projected observations (Mohr, Fabricant, and Geller 1993, Buote and Tsai 1995; see also Rasia, Meneghetti, and Ettori 2013 for a review); and these classifications do not necessarily match each other. A proper selection of relaxed clusters is a fundamental step for treating clusters as cosmological probes (Mantz et al. 2015).

Besides their cosmological role, galaxy clusters also constitute excellent laboratories to study, not only galaxy formation processes in dense environments (Gunn and Gott 1972), but also the physics of the bulk of baryonic matter in clusters, the hot intracluster medium (ICM). The ICM is an almost fully ionised plasma with typical temperatures of $T \sim 10^7\,{\rm K}$ and densities in the range $n \sim 10^{-5} - 10^{-1}\,{\rm cm}^{-3}$, which is nearly in equilibrium within the gravitational potential well of the cluster. The chemical and thermodynamic properties of this medium are of utmost importance for the study of the physical process at play during galaxy and galaxy cluster assembly. Besides adiabatic hydrodynamics,

galaxy clusters also harbour non-thermal processes contributing to a significant amount of their energetic budget. For instance, large-scale shock waves located at the peripheries of clusters ($r \sim [2-3]\,R_{\rm vir}$), together with weaker, smaller-scale internal shocks, have a significant role in the thermalisation of the infalling gas and the heating of the ICM up to the observed temperatures (Quilis et al. 1998, Miniati et al. 2000), while residual, turbulent motions fill a large fraction of the ICM volume and contribute significantly to its support against gravity (Vazza et al. 2009, Biffi et al. 2016). Non-thermal processes in the ICM and their observational implications are reviewed in Chapter 5.

While, so far, observations of galaxy clusters are restricted to their central regions ($\lesssim R_{\rm 500c}$), future telescopic facilities along the electromagnetic spectrum will be able to shed light on the physics of their outer regions (Walker et al. 2019), where the physics of accretion leaves its imprint on the ICM. Due to them being dynamically active and low-density regions, cluster outskirts exhibit a plethora of non-equilibrium phenomena, deviations from spherical symmetry, etc., such as gas clumping (Nagai and Lau 2011), intense bulk and turbulent motions (Lau, Kravtsov, and Nagai 2009, Vazza et al. 2009), or non-equilibrium conditions for electrons and ions (Avestruz et al. 2015).

2.3.1.1 Galaxy cluster formation and the thermodynamics of the ICM

Gravity is the main force driving the collapse of galaxy clusters, and therefore it is responsible, not only of setting the properties of the DM halo, but also the thermodynamic quantities of the ICM, which in turn determine its observational properties. Diffuse gas infalling from very low-density regions into the cluster soon achieves supersonic velocities ($\sim 1000\,{\rm km/s}$), producing the large-scale, strong ($\mathcal{M} \sim 100$) accretion shocks that are already predicted by the spherically-symmetric collapse model of Bertschinger (1985). In a three-dimensional, cosmological context, the picture gets more complex due to the presence of an anisotropic environment (e.g., filaments) and internal shocks due to mergers and the passage of substructures. Both these classes of shocks have an essential role in determining the thermal properties of the ICM (Skillman et al. 2008). They are also crucial to understand the velocity structure of the ICM, since gas transitions from ordered, laminar infall at high clustercentric distances to develop vortical and small-scale motions inside the accretion shock boundaries (e.g., Vazza et al. 2009).

However, the low levels of entropy in the cores of clusters, together with the very presence of galaxies forming out of cold gas, indicate that additional

processes able to cool the gas must be at play during cluster evolution (in this regard, see, e.g., the review by Voit, Kay, and Bryan 2005). Incidentally, the inclusion of cooling introduces a physical scale and breaks the otherwise self-similar evolution of the ICM (Voit and Bryan 2001). At the same time, several mechanisms must be present to overcome an excessive cooling of the gas and to regulate star formation, such as the injection of energy from SNe and AGN activity (Springel and Hernquist 2003, McNamara and Nulsen 2007). Besides these mechanisms, also thermal conduction (Dolag et al. 2004), physical viscosity (Sijacki and Springel 2006), and cosmic rays (Ruszkowski, Yang, and Reynolds 2017), amongst others, are expected to play a significant role in shaping the thermodynamic properties of the ICM, especially in the outskirts.

The complexity and diversity of the physical processes acting during cluster formation and evolution can only be self-consistently addressed with the aid of numerical simulations, which are discussed in Sec. 3.

2.3.2 Observational properties of galaxy clusters

Clusters of galaxies can be detected and observed throughout the electromagnetic spectrum, from radio wavelengths to high-energy gamma-rays, but the three main windows correspond to the microwave, the optical and the soft X-ray regions of the spectrum. In the following paragraphs, I briefly review the main observational properties of galaxy clusters in these three windows, while the interested reader is referred to, e.g., Voit, Kay, and Bryan (2005) for a classical review on the observable properties of galaxy clusters through the electromagnetic spectrum. Additionally, Sec. 5.4 briefly discusses some aspects of clusters in radio wavelengths. In Fig. 2.3, a multiwavelength view of the Coma cluster is displayed.

2.3.2.1 Observations in the optical band

The historical discovery of galaxy clusters corresponds to this band of the spectrum, with the observations of Charles Messier and William Herschel of the Virgo and Coma clusters, respectively (see Biviano 2000 for a historical perspective to the optical study of galaxy clusters). A large leap forward in the optical identification and characterisation of optical clusters was performed by Abell (1958) and Abell, Corwin, and Olowin (1989), who systematised the identification and classification of galaxy clusters with a well-defined heuristic procedure. Still today, modern optical cluster identification techniques are based on similar principles to the ones of Abell (Lumsden et al. 1992, Dalton et al.

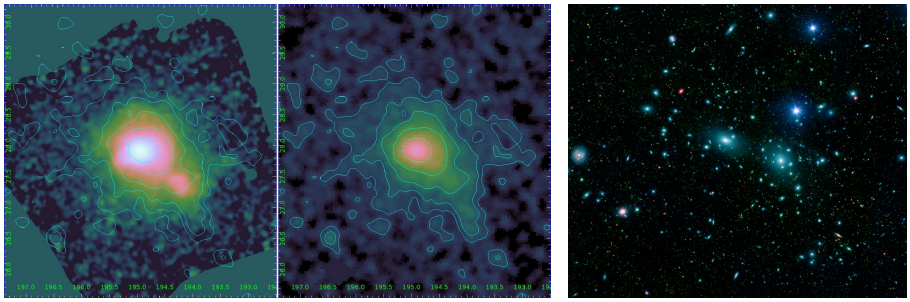


Figure 2.3: Region around the Coma cluster ($z = 0.023$) seen at different wavelengths. The left-hand side, middle and right-hand side panels correspond, respectively, to X-ray (SRG/eROSITA, $0.4 - 2\,\mathrm{keV}$), SZ (Planck) and optical + infrared (SDSS+Sptizer). The X-ray and SZ panels cover a region of $(4.4\,\mathrm{deg})^2$, and have been reproduced from Churazov et al. (2021) with permission (CC-BY license). The contours in both panels correspond to the SZ signal. The optical image covers a smaller, $41.7 \times 36.2\,\mathrm{arcmin}^2$ region. Image credit (public domain): NASA/JPL-Caltech/L. Jenkins (GSFC).

1997, Gladders and Yee 2000; cf. Euclid Collaboration et al. 2019 for a discussion on the detection of clusters with the very recent Euclid mission).

Given a list of member candidates, the status of the apparent concentration of galaxies as an actual cluster has to be confirmed, since projection effects can complicate their characterisation. This is done with a two-fold motivation: confirming the status of true members of the candidate galaxies, and measuring the mass of the cluster. In the optical band, there are at least three complementary ways of doing so:

Optical richness. Defined as a categorical measure (from 1, the poorest, to 5, the richest) of the number of galaxies in the cluster above a magnitude limit, it serves as a proxy for the total optical luminosity of the cluster and, in turn, for its mass (as long as light traces mass).

Galactic kinematics. It can be used together with dynamical considerations to estimate the gravitational mass of the cluster, together with the boundedness state of each constituent. Optical spectroscopy allows to measure the radial velocities, i.e. the velocities along the line of sight, v_{los}, of each member, from which the one-dimensional velocity dispersion, σ_{1D}, or even the whole (radial) velocity distribution function can be estimated. Using the virial theorem, as in the original approach by Zwicky (1933), the mass can be estimated. Alternatively, more sophisticated methods are used nowadays to estimate the mass profile of the cluster, such as the caustic technique (Diaferio and Geller 1997), or the Jeans equations (Mamon, Biviano, and Boué 2013).

Gravitational lensing of background sources. The lensing of background objects (most typically, galaxies) by the deep gravitational potential well of clusters can produce multiple images of the lensed galaxy (in the *strong lensing* regime), or a mild shear distortion of the background galaxies (in the *weak lensing* regime). Originally suggested by Zwicky (1937), the latter can be used to constrain the distribution of mass in the cluster (Tyson, Valdes, and Wenk 1990; see also Umetsu 2020 for a recent review).

The optical component of galaxy clusters does not consist exclusively of galaxies. In the recent years, the intracluster light (ICL), a diffuse component not bound to any galaxy, but only to the cluster's potential well, has triggered considerable attention (e.g., Contini et al. 2014, Mihos et al. 2017), in part because it has been suggested to trace total mass (Montes and Trujillo 2019) and the assembly history of the cluster (Presotto et al. 2014).

2.3.2.2 Observations in X-ray

The first detections of X-ray emission outside our Galaxy were reported by Byram, Chubb, and Friedman (1966) and Bradt et al. (1967) in the direction of the radio-galaxy M87, the most dominant galaxy of the Virgo cluster. Promptly, X-ray photons were also detected in the directions of Perseus (Fritz et al. 1971) and Coma clusters (Gursky et al. 1971), suggesting that galaxy clusters are generally X-ray sources (Cavaliere, Gursky, and Tucker 1971) and already attributing it primarily to bremsstrahlung emission (Felten et al. 1966).

X-ray emission from clusters appears as a consequence of the inefficiency of star formation, i.e., of the fact that only a small fraction of baryons ($\sim 10\%$) is in the form of stars. Diffuse baryons in clusters get compressed in the cluster potential well and shock-heated to temperatures of $T \sim GM_{\rm vir} m_{\rm p}/(k_{\rm B} R_{\rm vir}) \sim 10^{7-8}\,{\rm K}$, thus becoming powerful X-ray emitters. The X-ray spectrum (see Rosati, Borgani, and Norman 2002, Böhringer and Werner 2010 for classical reviews) is mainly contributed by two components, which I shall briefly review below:

The thermal bremsstrahlung continuum. It is especially dominant at high temperatures ($k_{\rm B}T \gtrsim 2\,{\rm keV}$), and is produced by the scattering of free electrons off ions. The spectral energy distribution of this component is described by the emissivity (Gronenschild and Mewe 1978)

$$\epsilon(\nu, T) \propto T^{-1/2} n_e n_i Z^2 g_{\rm ff}(\nu, T) e^{-\frac{h\nu}{k_{\rm B}T}}, \tag{2.17}$$

where $n_{e(i)}$ is the electron (ion) number density, Z is the effective charge of the ion, $g_{\rm ff} \sim 1$ is the gaunt factor that accounts for the quantum nature of the scattering process, and ν, T are, respectively, the photon frequency and the plasma temperature. The spectrum presents a sharp cut-off at high frequencies due to the exponential term, which can be used to infer the temperature of the emitting plasma as long as it is within the energy window of the X-ray telescope.

When considering bolometric X-ray luminosities, it is frequent to parametrise the temperature dependence inside a *cooling function*, $\Lambda(T) \propto T^{1/2}$, in such a way that the bolometric emissivity is given by $n_e n_i \Lambda(T)$. The total X-ray luminosity of a cluster is then given by the integral of the emissivity over the cluster volume, which is most usually expressed in terms of the plasma density using $n_e \sim n_i \sim \rho_{\rm gas}/(\mu m_{\rm p})$, where $\rho_{\rm gas}$ is the gas density and $\mu m_{\rm p}$ is the mean mass per particle. The bolometric X-ray luminosity of a cluster is then given by

$$L_{\rm X} = \iiint \left(\frac{\rho_{\rm gas}}{\mu m_{\rm p}}\right)^2 \Lambda(T) {\rm d}V, \tag{2.18}$$

yielding typical values of $L_{\rm X} \sim 10^{43} - 10^{45}\,{\rm erg/s}$.

Emission lines. Especially relevant at low temperatures ($k_{\rm B}T \lesssim 2\,{\rm keV}$; Raymond and Smith 1977), they allow the measurement of chemical abundances of metals in the ICM (see, e.g., Mernier and Biffi 2022 for a recent review on the chemical enrichment of the ICM). Early works (e.g., Renzini 1997) established that clusters had typical metallicities of one third of the solar value, $Z_{\rm ICM} \sim Z_\odot/3$, while subsequent studies, aided by the improvement of the spectroscopic capabilities of X-ray telescopes, have focused on the distribution of metals in the ICM (De Grandi and Molendi 2001, Gatuzz et al. 2023).

2.3.2.3 Observations in microwaves

Soon after the discovery of the CMB, Weymann (1966) predicted the distortions of its spectrum by inverse Compton scattering with high-energy electrons, which shifts photons to higher energies. Sunyaev and Zeldovich (1970) applied this to the hot gas in clusters, predicting the existence of a spectral distortion of the CMB, now known as the Sunyaev-Zeldovich (SZ) effect. First claims of the detection of such effect were only reported two decades later (Birkinshaw 1990), while modern facilities such as Planck (Planck Collaboration et al. 2011) or the South Pole Telescope (Bleem et al. 2015) have led to the identification of hundreds of galaxy clusters.

The effect described above corresponds to the so-called thermal Sunyaev-Zeldovich (tSZ) effect, which is due to the thermal motion of electrons in the ICM. Its magnitude is proportional to the integrated electron pressure, $P_\mathrm{e} = n_\mathrm{e} k_\mathrm{B} T_\mathrm{e}$, along the line of sight, and is usually parametrised by the Compton-y parameter, defined as (see, e.g., Mroczkowski et al. 2019 for a review)

$$y_\mathrm{tSZ} = \frac{\sigma_T}{m_\mathrm{e} c^2} \int P_\mathrm{e} \mathrm{d}l, \tag{2.19}$$

where σ_T is the Thomson scattering cross section. Since the ICM is optically thin, this is a small distortion ($y_\mathrm{tSZ} \ll 1$). Bulk motions of the cluster, and even its motion as a whole, produce an additional contribution, known as the kinetic Sunyaev-Zeldovich (kSZ) effect, which can be interpreted as a relativistic Doppler boosting (Rephaeli 1995). Its intensity is proportional to the line of sight velocity, and can be parametrised with an analogue parameter, y_kSZ.

$$y_\mathrm{kSZ} = \frac{\sigma_T}{c} \int n_e \mathbf{v} \cdot \mathrm{d}\mathbf{l} \tag{2.20}$$

Only recently, the direct detection of kSZ effect from stacked observations has been reported (e.g., Tanimura, Zaroubi, and Aghanim 2021). A property which makes the SZ effect especially interesting is the fact that, consisting on a scattering of CMB photons off the ICM electrons, its intensity is essentially redshift-independent, making it a promising tool for the study of high-redshift systems (e.g., Jones 1998).

2.3.3 The self-similar model

While the actual formation and evolution of the properties of galaxy clusters is complex and demands numerical simulations (Sec. 3), the self-similar model (Kaiser 1986; see also Borgani and Kravtsov 2011 for a comprehensive review) provides a useful baseline to compare with observations and simulations. Within this model, the properties of galaxy clusters are determined by the driving mechanism in their evolution, i.e., gravity. The predictions of this model are most usually presented in the form of scaling relations, which are briefly reviewed below. In their simplest form, scaling relations connect any observable measured within an overdensity aperture Δ, X_Δ, with the corresponding mass (M_Δ) and redshift z. For thorough reviews on the topic of scaling relations, the reader is addressed to Giodini et al. (2013) and Lovisari and Maughan (2022).

As seen in Sec. 2.2.5, DM haloes (and hence, clusters) can be defined based on an arbitrary overdensity with respect to the background density ($\rho = \Delta_m \rho_\mathrm{B}(z)$)

or the critical density ($\rho = \Delta_c \rho_c(z)$). The scaling relations can be formulated for any of these, but care must be taken with their different evolution: while $\rho_B(z) = \rho_{B0}(1+z)^3$, the evolution of $\rho_c(z)$ is related to that of $H(z)$. It is customary to express $H(z) = H_0 E(z)$, where

$$E(z) = \sqrt{\Omega_m (1+z)^3 + \Omega_\Lambda}, \tag{2.21}$$

for a universe with no curvature nor radiation, and where the Ω_X written here correspond to their values at $z = 0$. With this definition, $\rho_c(z) = \rho_{c0} E(z)^2$. Then, the mass enclosed within an overdensity Δ_m is given by

$$M_{\Delta_m} \propto \Delta_m R^3_{\Delta_m}, \tag{2.22}$$

while for a Δ_c overdensity

$$M_{\Delta_c} \propto \Delta_c \frac{E(z)^2}{(1+z)^3} R^3_{\Delta_c}, \tag{2.23}$$

where all radii are expressed in comoving coordinates. From this, assuming gravity to be the driving mechanism of the evolution of the system, one would expect the internal (thermal) energy per particle to be proportional to its binding energy, i.e., $k_B T \propto GM(\mu m_p)/R$. This implies that the temperature of the ICM should scale with mass and redshift as

$$T_{\Delta_m} \propto \Delta_m^{1/3} (1+z) M^{2/3}_{\Delta_m}, \tag{2.24}$$

$$T_{\Delta_c} \propto \Delta_c^{1/3} E(z)^{2/3} M^{2/3}_{\Delta_c}. \tag{2.25}$$

Similarly, the ICM entropy, defined as $K \equiv k_B T n_e^{-2/3}$, should scale as

$$K_{\Delta_m} \propto \Delta_m^{-1/3} (1+z)^{-1} M^{2/3}_{\Delta_m}, \tag{2.26}$$

$$K_{\Delta_c} \propto \Delta_c^{-1/3} E(z)^{-2/3} M^{2/3}_{\Delta_c}. \tag{2.27}$$

Equivalent relations can be written for the X-ray bolometric luminosity, L_X, gas pressure, etc.; as well as for parameters derived from SZ observations, such as the integrated Compton-y parameter, Y_{SZ}. These relations can be observationally used to infer masses from integrated quantities of the cluster.

2.3.3.1 Breaking of self-similarity

While the self-similar model has been extremely successful in predicting average properties of galaxy clusters, some of its predictions have been found to be at variance with observations, indicating the presence of additional processes on top of gravity.

For instance, Reiprich and Böhringer (2002) reported an $L_X - M$ relation steeper than the self-similar prediction (with a logarithmic slope of 4/3). Similarly, the $T - M$ relation has also been reported to be steeper than expected (Markevitch et al. 1998, Arnaud and Evrard 1999), together with a non-constant gas mass fraction across masses (Sanderson et al. 2003). Additionally, the scatter around some of these relations (e.g., the $L_X - M$ relation) can be important, and this has been associated to the presence of cool cores[2] (Fabian et al. 1994). While excising the core is a common practice in X-ray studies (Markevitch et al. 1998), departures from self-similarity still remain present.

Subsequent studies using hydrodynamical cosmological simulations have been able to link departures from self-similarity to structure formation processes. Just to cite a few, Planelles and Quilis (2009) show how clusters move along the scaling relations planes as a result of mergers. In Lau et al. (2015), the shape of the ICM thermodynamic profiles is studied in relation to accretion rates, demonstrating how accretion impacts the outskirts of clusters. Also in this direction, Chen et al. (2019) explore the connection between the residuals with respect to the $T - M$ relation and accretion rates as well as the ellipticity of the ICM.

Besides the purely hydrodynamical and gravitational processes, the introduction of radiative cooling, together with heating from SNe and AGN feedback, are also thought to be major contributors to the breaking of self-similarity in the ICM (Borgani et al. 2004, McCarthy et al. 2004, McCarthy et al. 2008).

2.4 Cosmic voids

For as long as galaxy surveys have been compiled (e.g., Einasto, Joeveer, and Saar 1980), voids have been known as a characteristic feature of the LSS of the Universe, filling up most of its volume (Cautun et al. 2014) and constituting the largest structures in the cosmic web, with typical sizes up to $\sim 100\,\mathrm{Mpc}$ as in the case of the colossal Boötes void (Kirshner et al. 1981). In this section, I

[2] In relaxed systems, dense cores that are capable of undergoing efficient radiative cooling may develop. Cool gas gets removed from the X-ray emitting phase but, perhaps paradoxically, these objects exhibit an enhanced central surface brightness due to its high density

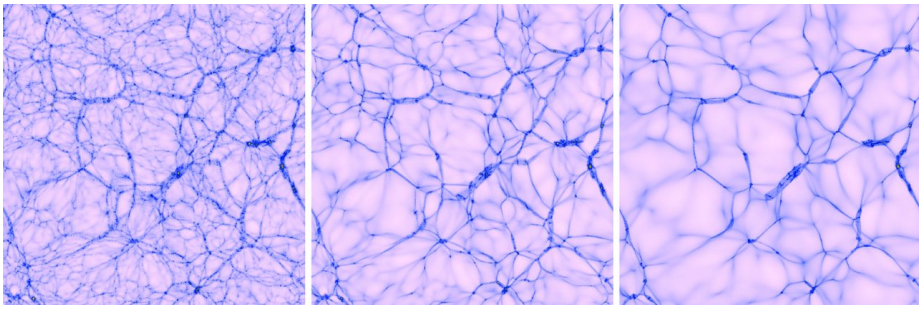


Figure 2.4: Density field across a thin slice of a cosmological simulation at $z = 0$, where the initial conditions have been smoothed at different scales: no smoothing (left-hand side panel), smoothed at scales of $2\,h^{-1}$ Mpc (middle panel), and $4\,h^{-1}$ Mpc (right-hand side panel); showing the emergence of a hierarchy of voids and subvoids. Reproduced from Aragon-Calvo and Szalay (2013) with permission.

briefly review some of the most relevant properties of cosmic voids, while the interested reader is referred to, e.g., van de Weygaert and Platen (2011) for a more thorough review.

Voids emerge from the evolution of primordial underdensities which, according to the linear theory (see Sec. 2.2.2), get emptier with cosmic time, as matter around them collapses to form walls, filaments and haloes. In this sense, cosmic voids have triggered considerable attention in the last decades, for their potential as cosmological probes (e.g., Lavaux and Wandelt 2012, Nadathur et al. 2019, Pisani et al. 2019). Essentially, their lower mean density allows treating isolated voids, by means of the Tolman-Bondi approach (Tolman 1937, Bondi 1947), as regions with super-Hubble expansion (e.g., Baushev 2021) where structure formation proceeds at a slower pace.

Nevertheless, voids are not monolithic regions of negative density contrast, but instead contain tenuous filaments and walls, subvoids and haloes, in the light of the hierarchical structure formation scenario (Sheth and van de Weygaert 2004). To exemplify this, Fig. 2.4 shows slices of the density field from a cosmological simulation smoothed at different scales, demonstrating the hierarchical nature of these structures. The pristine, low-density environment of voids constitutes an ideal laboratory to study the formation and evolution of galaxies, where interactions are extremely infrequent (Hahn et al. 2007, Kreckel et al. 2011, Ricciardelli et al. 2014).

As already revealed by models of isolated voids, as they evolve, matter inside them gets evacuated by the repulsive peculiar gravitational field, which is stronger in the central regions and gradually decreases going outwards, as density also increases. Consequently, matter in central regions moves outwards faster than

matter in its outskirts, leading to an accumulation of matter in the surrounding of the void, usually denominated as the void's *shell*. A characteristic timescale in void evolution is that of shell-crossing, when inner shells of matter overtake the outer ones. Analytic models show that, by this time, voids reach a self-similar, slower, expansion phase (Sheth and van de Weygaert 2004), achieving non-linear density contrasts of $\delta_{\rm v} \sim -0.8$.

In this picture, also confirmed by three-dimensional simulations (Ricciardelli, Quilis, and Planelles 2013, Aragon-Calvo and Szalay 2013), the velocity field in voids is dominated by its nearly-radial expansion, especially for larger voids that are not contaminated by the shear of larger-scale structure. Regarding their shapes, voids are generally triaxial, with a slight preference towards prolateness. Even though analytical models predict that aspherical voids become more spherical as they undergo expansion (Icke 1984), simulations show that they do not get more spherical as evolution progresses due to their anisotropic environment (Shandarin et al. 2006). As a matter of fact, the aforementioned reference also shows how voids are not only aspherical in the sense of being triaxial, but also because of their intricate, porous shapes. Lastly, several groups (Hamaus, Sutter, and Wandelt 2014, Ricciardelli et al. 2014) have studied the universality of void density profiles, coming up with functional forms capturing their shape and evolution.

All this said, the very definition of a void is still matter of debate. Different identification techniques have been proposed, ranging from finding the largest spheres around density minima or connecting underdense cells, to performing a watershed transform on the density field or studying the dynamics of test particles; and either using the DM particle distribution from simulations, or halo/galaxy catalogues. The extent to which different void finding strategies converge has been thoroughly studied, for example, in Colberg et al. (2008), revealing important differences in basic properties of the void distribution, such as the sizes of the largest void regions and their shapes.

CHAPTER 3

Numerical simulations of cosmological structure evolution

Contents

As introduced in Chapter 2, cosmological structure formation is a complex process, involving the interaction of many different phenomena (e.g., gravity, hydrodynamics, galaxy formation physics, etc.) over a wide range of scales. The complexity of this intrinsically multiscale and multiphysics problem render numerical simulations as, perhaps, the most viable option for tackling it.

For a classical review of numerical methods and simulation techniques for cosmological simulations, the reader is referred to Dolag et al. (2008). The recent review by Vogelsberger et al. (2020) offers an updated view on the topic, including a thorough discussion on the additional physics associated to galaxy formation. In this Chapter, I discuss how the evolution of cosmic inhomogeneities is posed as an initial-value problem (Sec. 3.1), and the numerical techniques that are most often involved in its solution (Sec. 3.2).

3.1 Cosmological structure evolution as an initial value problem

In this Section, the mathematical framework regarding the evolution of cosmic inhomogeneities is introduced. In the following, I shall use comoving coordinates in the description of the dynamics of the different components. Throughout this text, the comoving position is denoted by $\mathbf{x}$, and it is related to the physical position vector, $\mathbf{r}$, by $\mathbf{r} = a(t)\mathbf{x}$. Henceforth, the explicit time dependence of any function related to cosmology is dropped, unless in the cases where it may trigger confusion. The physical velocity, $\mathbf{u} \equiv \mathrm{d}\mathbf{r}/\mathrm{d}t$, is related to the comoving velocity, $\mathbf{v} \equiv a\mathrm{d}\mathbf{x}/\mathrm{d}t$, by $\mathbf{u} = \mathbf{v} + \dot{a}\mathbf{x}$, where the second term corresponds to the Hubble flow. The differential operator ∇ always makes reference to derivatives with respect to the components of the comoving position, unless explicitly written as $\nabla_{\mathbf{r}}$.

The dynamics of any material component can be fully described, in the continuum limit, by its distribution function, $f(\mathbf{x}, \mathbf{p}, t)$, specifying its phase space density at a given time t. The distribution function is defined such that $f(\mathbf{x}, \mathbf{p}, t)\mathrm{d}^3\mathbf{x}\mathrm{d}^3\mathbf{p}$ is the number of particles in the phase space volume $\mathrm{d}^3\mathbf{x}\mathrm{d}^3\mathbf{p}$ around the point $(\mathbf{x}, \mathbf{p})$. Once f is known, the value of any given field $Q(\mathbf{x})$ is obtained by integrating the distribution function over momenta,

$$Q(\mathbf{x}, t) = \frac{\int q(\mathbf{x}, \mathbf{p}) f(\mathbf{x}, \mathbf{p}, t)\mathrm{d}^3\mathbf{p}}{\int f(\mathbf{x}, \mathbf{p}, t)\mathrm{d}^3\mathbf{p}}. \tag{3.1}$$

The evolution of the distribution function is governed by the Boltzmann equation, which can be written in comoving coordinates as (see, for instance, Padmanabhan 2002 and Winther 2021 for introductory reviews)

$$\frac{\partial f}{\partial t} + \frac{\mathbf{v}}{a} \cdot \nabla f + \dot{\mathbf{p}} \cdot \nabla_{\mathbf{p}} f = \left(\frac{\delta f}{\delta t}\right)_{\mathrm{coll}}, \tag{3.2}$$

where $\dot{\mathbf{p}}$ is the rate of change of momentum due to external force fields (e.g., gravity, Hubble flow), and $(\delta f/\delta t)_{\mathrm{coll}}$ is the collision term, which accounts for the changes in the distribution function due to particle interactions. In most real applications, this equation is numerically intractable due to its high-dimensionality, and approximate approaches are sought for. Generally, most cosmological simulations use Newtonian dynamics for the description of the perturbations over the FLRW background, since Newtonian and GR predictions match in the linear approximation for the matter-dominated case (Padmanabhan

2002), and the velocities and gravitational fields at these scales are far from the relativistic regime (see, e.g., the discussion by Vogelsberger et al. 2020).

Owing to their fundamentally different natures, DM and baryons receive distinct treatments. In Sec. 3.1.1, the dynamics of DM are introduced, including the description of the gravitational field, which couples the different material components. Sec. 3.1.2, instead, discusses the dynamics of baryonic matter. Given a recipe for the evolution, the mathematical statement of the problem is complete once the initial conditions are established. These are covered in Sec. 3.1.3.

3.1.1 Dynamics of dark matter

Cold dark matter (CDM) can be described as a *collisionless* fluid, i.e., a fluid where self-interactions can be neglected. This implies dropping the right-hand side term in Eq. (3.2), $(\delta f/\delta t)_{\rm coll} = 0$. The resulting equation is also known as the *Vlasov* equation, or the *collisionless Boltzmann* equation.

3.1.1.1 The N-Body approach

While there are alternative approaches (see the next heading), the most widely-used approach for addressing the dynamics of collisionless systems is the N-Body method, in which the phase space density of DM is sampled by a set of N tracer particles, each of them characterised by a point-like phase space position ($\mathbf{x}_{\rm i}$, $\mathbf{v}_{\rm i}$),

$$f(\mathbf{x},\mathbf{p},t) \sim \sum_{i=1}^{N} m_i \delta^{(3)}(\mathbf{x}-\mathbf{x}_{\rm i})\delta^{(3)}(\mathbf{p}-\mathbf{p}_{\rm i}). \tag{3.3}$$

Here, $\delta^{(3)}$ represents a three-dimensional Dirac delta. Each individual tracer is evolved in time according to the Newtonian equations of motion, which are a set of $6N$ coupled ordinary differential equations (ODEs) that, written the comoving frame, have the form

$$\dot{\mathbf{x}}_{\mathbf{i}} = \frac{1}{a}\mathbf{v}_{\mathbf{i}}, \tag{3.4}$$

$$\dot{\mathbf{v}}_{\mathbf{i}} = -\frac{1}{a}\nabla\phi - H\mathbf{v}_{\mathbf{i}}, \tag{3.5}$$

where ϕ is the *peculiar* Newtonian gravitational potential, related to the usual gravitational potential (Φ) by

$$\phi = \Phi + \frac{1}{2} a\ddot{a}x^2 \tag{3.6}$$

and generated by the total density contrast (accounting for all matter density, i.e., including baryons), $\delta_{\rm T}$, through the comoving Poisson equation,

$$\nabla^2\phi = 4\pi G a^2 \rho_{\rm B}\delta_{\rm T} = \frac{3}{2}H^2 a^2 \delta_{\rm T}. \tag{3.7}$$

Updating particle positions and velocities in this way can be shown to be equivalent to solving the original Vlasov equation using the *method of characteristics*. An N-Body method can also be seen as a Monte-Carlo approach for solving the Vlasov equation. As such, this method is affected by sampling, or *shot* noise, that is reduced with increased resolution. For a thorough discussion on N-Body simulations for solving the gravitational dynamics of collisionless systems, see also Dehnen and Read (2011). In Sec. 3.2.1, the techniques to tackle Eqs. (3.4) and (3.5) are discussed in detail.

3.1.1.2 Alternative approaches

Despite the hegemony of N-Body methods to solve the gravitational dynamics of DM and other collisionless species, alternative methods have been developed and tested to try to overcome some of its shortcomings (e.g., the ones associated to discreteness noise; Splinter et al. 1998). These include, for instance, the reformulation of the problem as a Schrödinger-Poisson system (Widrow and Kaiser 1993, Schaller et al. 2014), Lagrangian tessellation methods (Hahn, Abel, and Kaehler 2013), or the direct integration of the distribution function from Vlasov's equation (Yoshikawa, Yoshida, and Umemura 2013, Colombi et al. 2015).

3.1.2 Dynamics of baryons

Contrary to DM, baryons are a highly collisional material component, therefore undergoing qualitatively different dynamics than what has been reviewed before for DM. At early times, the baryonic mass in the Universe is in the form of a gas mainly composed of hydrogen and helium. Some of this gas is able to cool efficiently and form bound objects, such as stars, that detach from the global dynamics of the gaseous component and behave effectively as a collisionless species. Nevertheless, at least $\sim 80\%$ of the Universe's baryons are still forming a diffuse gas and, hence, the bulk of baryonic matter can be described with a hydrodynamic approach.

In the Eulerian specification of the flow, any quantity regarding the fluid is mathematically described as a function of (comoving) position and time: e.g., its density, $\rho(\mathbf{x}, t)$; its velocity, $\mathbf{v}(\mathbf{x}, t)$; its pressure, $P(\mathbf{x}, t)$; etc. This is in contrast to the Lagrangian specification of the flow, where instead of focusing on the properties of the fluid at a fixed position, the properties of a given fluid element are tracked as it moves through space. In this case, it is customary to describe the properties of the fluid parcel in terms of its initial position, $\mathbf{q}$. Through this text[1], I shall denote the Eulerian time derivative as $\frac{\partial}{\partial t}$, while the Lagrangian, *total* or *material* derivative is $\frac{\mathrm{d}}{\mathrm{d}t} \equiv \frac{\partial}{\partial t} + \frac{\partial \mathbf{x}}{\partial t} \cdot \nabla = \frac{\partial}{\partial t} + \frac{\mathbf{v}}{a} \cdot \nabla$.

The set of partial derivative equations (PDEs) governing the evolution of a collisional, inviscid fluid in an expanding frame, under the action of a peculiar gravitational potential, ϕ, are (e.g., Peebles 1980):

$$\frac{\partial \delta}{\partial t} + \frac{1}{a}\nabla \cdot [(1+\delta)\mathbf{v}] = 0, \tag{3.8}$$

$$\frac{\partial \mathbf{v}}{\partial t} + \frac{1}{a}(\mathbf{v} \cdot \nabla)\mathbf{v} + H\mathbf{v} = -\frac{1}{a}\nabla\phi - \frac{1}{\rho a}\nabla P, \tag{3.9}$$

$$\frac{\partial E}{\partial t} + \frac{1}{a}\nabla \cdot [(E+P)\mathbf{v}] = -3H(E+P) - H\rho v^2 - \frac{\rho \mathbf{v}}{a} \cdot \nabla\phi, \tag{3.10}$$

where ρ is the gas density, $\delta \equiv \rho_{\mathrm{gas}}/\rho_{\mathrm{B}} - 1$ is the gas density contrast, and $E \equiv \rho u + 1/2\rho v^2$ is the *total* energy density, contributed by internal and kinetic comoving energy densities, u being the specific internal energy. In qualitative terms, Eq. (3.8), the *continuity* equation, expresses mass conservation; Eq. (3.9) is the law of motion of a gas parcel, or the *Euler* equation, corresponding to the conservation of linear momentum; and Eq. (3.10), the *energy* equation, expresses conservation of energy and contains the first law of Thermodynamics.

Once the background cosmology is solved (i.e., $a(t)$, $H(t)$ are known functions), and provided the total peculiar gravitational potential is solved for using Eq. (3.7), the system of equations (3.8-3.10) presents 6 scalar unknowns (namely, δ, the three Cartesian components of $\mathbf{v}$, E and P) and only 5 scalar equations. The system is only closed after choosing an equation of state for the gas, $P = P(\delta, \mathbf{v}, E)$ in the most general case. The most common option is to choose the equation of state of an ideal gas,

$$P = (\gamma - 1)\rho u, \tag{3.11}$$

[1]Also following the notation of, e.g., Landau and Lifshitz (1987). Note that other authors choose to notate the total (Lagrangian) time derivative as $\frac{\mathrm{D}}{\mathrm{D}t}$.

where γ is the adiabatic exponent ($\gamma = 5/3$ for a monatomic ideal gas).

As discussed by, e.g., Dolag et al. (2008), the equations of hydrodynamics in an expanding frame under the action of gravity are highly non-linear, as a result of the non-linearity of gravitationally-driven clustering in the Universe. On the one hand, the collapse of overdense regions by gravitational instability triggers supersonic motions leading to the appearance of strong shock discontinuities. On the other hand, the non-linearity of the equations requires a sizeable dynamical range to be resolved, from the LSS sizes of $(10-100)$ Mpc to the scales of galaxy formation, below the kpc. In this context, complex turbulent fluid flows of high Reynolds number develop within collapsing structures. These two phenomena are described in greater detail in Chapter 5.

Finally, it is worth mentioning that Eqs. (3.8-3.10) can be alternatively formulated in terms of the Lagrangian derivative, or in integral form for an arbitrary Lagrangian-Eulerian formulation, and this gives rise to the different methods for discretising and solving the hydrodynamic equations, as discussed below in Sec. 3.2.2. With the definition of the Lagrangian derivative, the equations of hydrodynamics become (Dolag et al. 2008):

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\frac{\rho}{a}\nabla\cdot\mathbf{v} - 3H\rho, \tag{3.12}$$

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -\frac{1}{a}\nabla\phi - \frac{1}{\rho a}\nabla P - H\mathbf{v}, \tag{3.13}$$

$$\frac{\mathrm{d}(\rho u)}{\mathrm{d}t} = -(\rho u + P)\left(\frac{1}{a}\nabla\cdot\mathbf{v} + 3H\right) \tag{3.14}$$

where u is the specific internal energy, i.e., $u \equiv E/\rho - v^2/2$. Finally, the equations can also be written in integral form in an arbitrary Eulerian-Lagrangian formulation, where the integration volume $V(t)$ can evolve arbitrarily in time. Dropping the expansion terms ($a = 1$, $H = 0$) and the gravitational force for ease of notation, the equations can be written as (Vogelsberger et al. 2020)

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint_{V(t)} \rho \mathrm{d}V = -\oiint_{\partial V(t)} \rho(\mathbf{v}-\mathbf{w}_\mathrm{S})\cdot\mathrm{d}\mathbf{S}, \tag{3.15}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint_{V(t)} \rho\mathbf{v} \mathrm{d}V = -\oiint_{\partial V(t)} \rho\mathbf{v}(\mathbf{v}-\mathbf{w}_\mathrm{S})\cdot\mathrm{d}\mathbf{S} - \oiint_{\partial V(t)} P\mathrm{d}\mathbf{S}, \tag{3.16}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint_{V(t)} \rho e \mathrm{d}V = -\oiint_{\partial V(t)} \rho e(\mathbf{v}-\mathbf{w}_\mathrm{S})\cdot\mathrm{d}\mathbf{S} - \oiint_{\partial V(t)} P\mathbf{v}\cdot\mathrm{d}\mathbf{S}, \tag{3.17}$$

where $e \equiv u + v^2/2 = E/\rho$ is the total specific energy, and $\mathbf{w}_\mathrm{S}$ is the velocity of the surface ∂V.

The equations of hydrodynamics can be modified to include the presence of magnetic fields, giving rise to the equations of magnetohydrodynamics (MHD). Since magnetic fields have not been the main object of the results presented within this PhD Thesis, the equations of MHD are not discussed here for conciseness, and only briefly mentioned in Chapter 4 for completeness.

3.1.2.1 Other phenomena beyond adiabatic hydrodynamics

With the aim of faithfully reproducing the thermodynamic conditions of the gas within galaxy clusters, or in order to describe the various astrophysical processes that shape galaxy formation, the equations above need to be complemented with additional physics that are most usually modelled with subgrid or subresolution prescriptions, since they occur at scales below the ones directly resolved by simulations. In the following, the main processes that are usually included in cosmological simulations are briefly described, with special focus on those relevant for galaxy clusters.

Gas cooling. Internal energy of the gas is dissipated through a variety of processes, including inverse Compton scattering, electron-nuclei recombination, collisional excitation, and free-free emission. These terms are usually parametrised through a *cooling function*, Λ, in such a way that the rate of change of specific internal energy due to cooling can be written as

$$\left(\frac{\mathrm{d}u}{\mathrm{d}t}\right)_\mathrm{cooling} = -\Lambda/\rho. \tag{3.18}$$

The cooling function is often tabulated in terms of the abundances of each species (H, HI, He, HeI, HeII, etc.) and temperature (e.g., Sutherland and Dopita 1993), and is usually computed assuming ionisation equilibrium and an optically thin gas in order to derive the elemental abundances (Katz, Weinberg, and Hernquist 1996, Theuns et al. 1998). Simulations aimed to reproduce galaxy formation in more detail track the individual abundances of each species implementing chemical reaction networks that are solved during the evolution of the simulation (e.g., Tornatore et al. 2007; see also Smith et al. 2017). An example of the cooling rates in the first case (where they are parametrised only as a function of gas density) is shown in Fig. 3.1.

Although it is not a source of cooling, gas can interact with a background of high-energy (ultraviolet) photons, generated by high-redshift quasars, heating

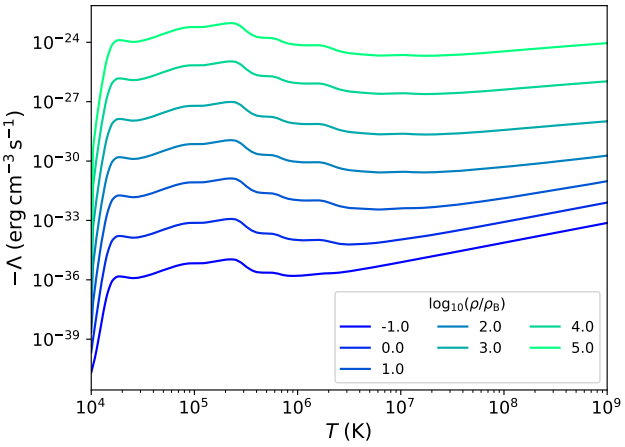


Figure 3.1: Cooling rate ($-\Lambda$) in cgs units shown as a function of temperature for several densities (expressed in units of the background density at $z = 0$), for a gas with metallicity $Z \simeq Z_\odot/3$.

the gas in a process known as reionisation. This ionising background is time-dependent because of the evolution of the quasar population, and can be treated similarly to cooling, by including its contribution in the cooling function. The most widely-used approach is to include a parametrisation, such as the one in Haardt and Madau (1996; see also the discussion in Theuns et al. 1998), as a heating term in the energy equation.

Star formation. In cosmological simulations, which are generally far from resolving the scale of molecular clouds giving rise to star formation, it is represented by the creation of collisionless particles (stars) out of cold, dense gas. These stellar particles represent single-age, single-metallicity stellar populations, with an underlying initial mass function. The rate of star formation is usually computed according to a prescription consistent with the Kennicutt-Schmidt law (Schmidt 1959, Kennicutt 1998),

$$\frac{\mathrm{d}M_*}{\mathrm{d}t} = \varepsilon \frac{M_{\mathrm{gas}}}{\tau_{\mathrm{ff}}}, \tag{3.19}$$

where M_* is the stellar mass, M_{gas} is the gas mass, $\tau_{\mathrm{ff}} \propto (G\rho)^{-1/2}$ is the gravitational free-fall time, and $\varepsilon \sim 0.01 - 0.1$ is a free parameter related to the efficiency of star formation (Springel and Hernquist 2003), or the fraction of gas mass converted to stars per free-fall time. When implementing star formation, Eq. (3.19) is not applied to all gas elements, but only to those satisfying a set

of criteria. Although there are many prescriptions, common choices include a minimum density and a maximum temperature (Springel and Hernquist 2003, Hopkins et al. 2014, Schaye et al. 2015), a criterion based on Jeans length (Hopkins et al. 2018) or on the divergence of the gas velocity (Stinson et al. 2006), gravitational boundedness (Stinson et al. 2006, Semenov, Kravtsov, and Gnedin 2017), or a restriction of star formation exclusively to the molecular gas phase (Hopkins et al. 2014, Hopkins et al. 2018).

Once formed, stellar particles evolve as a collisionless component under the action of the gravitational potential, and may yield part of their mass back to the gaseous phase, enriching the intergalactic medium with metals using parametrisations derived from the study of type-II SNe or, more recently, winds from stars in the asymptotic giant branch (Wiersma, Schaye, and Smith 2009) and type-Ia SNe (Vogelsberger et al. 2013), amongst others.

Stellar feedback. Stellar particles inject energy and momentum back into their surrounding medium, typically parametrising the effect of SNe (with a typical value of $\sim 10^{51}$ erg per event), to generate a self-regulated feedback loop that prevents excessive cooling of the gas and star formation. The main difference amongst models lies on the particular mechanism by which energy and momentum are deposited into the gaseous phase. Broadly, energy can be injected in the form of kinetic energy or internal (thermal) energy.

Feedback prescriptions where energy is exclusively deposited in the form of internal energy of the gas are ineffective, because in this high-density gas resolved at the (relatively) low resolution of cosmological simulations, the energy is quickly radiated away. Possible remedies for this problem rely on artificially disabling cooling for some tens of Myr for the gas affected by feedback (Stinson et al. 2006), or on the implementation of stochastic modes of heating the gas (Dalla Vecchia and Schaye 2012). By injecting the energy in kinetic form, the immediate loss of the feedback energy is prevented, because it first needs to thermalize (by shocks or viscous dissipation).

Naturally, there are even more sophisticated stellar feedback schemes, that take into account other feedback channels (e.g., young, massive stars; e.g., Agertz, Teyssier, and Moore 2011, Stinson et al. 2013), or that track the multiphase nature of baryons in star-forming regions (Murante et al. 2010).

Black hole (BH) formation and AGN feedback. Supermassive black holes are present in the centres of many galaxies, from massive (Gehren et al. 1984) to dwarf ones (Reines et al. 2011). Since the actual seeds of these BHs and

the detailed accretion processes fall many orders of magnitude below the scales resolved by cosmological simulations, BHs are usually seeded in DM haloes with a minimum DM, gas or stellar mass (see, e.g., the discussion by Wang et al. 2019), or directly in regions of high stellar density fulfilling several conditions (as is the case of MASCLET; see Sec. 4.3.3).

BHs then act as sink particles that accrete surrounding gas based on some prescription for their accretion rate. Most often, the accretion rate is set by the Bondi-Hoyle model (Bondi 1952, Hoyle and Lyttleton 1941),

$$\dot{M}_{\rm BH} = \frac{4\pi G^2 M_{\rm BH}^2 \rho}{(c_{\rm s}^2 + v_{\rm BH}^2)^{3/2}}, \tag{3.20}$$

where $M_{\rm BH}$ is the BH mass and $\dot{M}_{\rm BH}$ its accretion rate, ρ is the gas density in the vicinity of the BH, $c_{\rm s}$ is the sound speed, and $v_{\rm BH}$ is the BH velocity relative to the gas. The accretion rate is usually superiorly limited by the Eddington accretion rate, which indicates the limit where the outward radiation pressure would counteract the gravitational force and stop the accretion process,

$$\dot{M}_{\rm Edd} = \frac{4\pi G M_{\rm BH} m_{\rm p}}{\epsilon \sigma_{\rm T} c} \tag{3.21}$$

where $m_{\rm p}$ is the proton mass, $\epsilon \sim 0.1$ is the efficiency in turning gravitational energy into thermal energy, and $\sigma_{\rm T}$ is the Thomson cross-section.

Intensely accreting BH particles may release energy back to the surrounding gas aiming to mimic the phenomenology of AGN (see, e.g., Krolik 1999 for a thorough review), serving to regulate the growth of the BH and the star formation activity in the host galaxy. Most commonly, AGN feedback is divided in two broad modes: quasar mode (at high accretion rates, energy is released through photon emission, that may deposit energy and momentum in the gas through radiation pressure; Di Matteo, Springel, and Hernquist 2005) and radio mode (at low accretion rates, through the launching of highly-collimated, relativistic jets, which in turn may inflate X-ray bubbles; Sijacki et al. 2007).

Other processes. Besides the aforementioned processes and feedback mechanisms, present-day cosmological and galaxy formation simulations have implemented additional physical effects. These include, but are not restricted to, cosmic rays (able to drive galactic outflows, Pakmor et al. 2016; and heat the ICM, Ruszkowski, Yang, and Reynolds 2017; see Pfrommer et al. 2017 and Böss et al. 2023 for recent implementations), radiation hydrodynamics to study reionisation (e.g., Gnedin and Kaurov 2014, Rosdahl et al. 2018), physical

viscosity (Sijacki and Springel 2006, ZuHone et al. 2015), thermal conduction (Ruszkowski et al. 2011, Arth et al. 2014, Kannan et al. 2016) or dust modelling (McKinnon, Torrey, and Vogelsberger 2016, Vogelsberger et al. 2019).

3.1.3 Initial conditions

The initial conditions for cosmological simulations are most often obtained by performing a realisation of the power spectrum at high redshift $z \sim 50 - 1000$, $P(k)$, introduced in Sec. 2.2.1. In an unconstrained realisation of the primordial Gaussian random field, where the complex phases are independent, the real and imaginary part of each Fourier mode, $\delta_{\mathbf{k}}$, are drawn from two standardised Gaussian distributions and scaled by $\sqrt{P(|\mathbf{k}|)}$ (Bardeen et al. 1986). This is normally done by the Box and Muller (1958) method, i.e.,

$$\Re(\delta_{\mathbf{k}}) = \sqrt{P(|\mathbf{k}|)}\,\sqrt{-2\log\xi_1}\cos(2\pi\xi_2), \tag{3.22}$$

$$\Im(\delta_{\mathbf{k}}) = \sqrt{P(|\mathbf{k}|)}\,\sqrt{-2\log\xi_1}\sin(2\pi\xi_2), \tag{3.23}$$

where ξ_1 and ξ_2 are two independent random numbers drawn from a uniform distribution in the interval $]0, 1]$. The inverse Fourier transform of $\delta_{\mathbf{k}}$ yields the initial overdensity field in real space $\delta(\mathbf{x})$. In order to get initial velocities, as well as the initial displacements for particles, it is customary to define the perturbation potential, Ψ, as the solution of the elliptic equation

$$\nabla^2\Psi = \delta, \tag{3.24}$$

whose solution can be easily written in Fourier space as

$$\Psi_{\mathbf{k}} = -\frac{\delta_{\mathbf{k}}}{k^2}. \tag{3.25}$$

Particles are placed in a glass-like arrangement (Baugh, Gaztanaga, and Efstathiou 1995), with Lagrangian coordinates $\mathbf{q}$, and then displaced to their initial positions using Zeldovich (1970) theory,

$$\mathbf{x} = \mathbf{q} - D_+(z)\nabla_{\mathbf{q}}\Psi(\mathbf{q}), \tag{3.26}$$

where $D_+(z)$ is the linear growth factor, defined in Sec. 2.2.2. Similarly, the initial velocities can be found from the above equation, since $\mathbf{v} = a\dot{\mathbf{x}}$, thus yielding

$$\mathbf{v} = -a(z)\dot{D}_+(z)\nabla_{\mathbf{q}}\Psi(\mathbf{q}). \tag{3.27}$$

For more thorough reviews on the Zeldovich approximation and on the generation of initial conditions, see, respectively, Shandarin and Zeldovich (1989) and Efstathiou et al. (1985). Alternatively, second-order Lagrangian perturbation theory can also be used to set the initial perturbations (Bertschinger 2001, Jenkins 2010, Garrison et al. 2016).

While unconstrained initial conditions are useful to simulate large volumes with statistical properties akin to the Universe at large scales, they may not be the optimal solution for the study of individual objects. To study individual objects (e.g., a galaxy cluster) while still retaining the large-scale environment, which can affect the object of study through the action of its tidal field, it is frequent to use the *zoom-in* technique (Tormen, Bouchet, and White 1997, Klypin et al. 2001). In this approach, the region of interest is selected and sampled at high resolution. This can be done by either degrading the resolution outside the volume of interest, even using a nested set of refinements (Miniati 2014), or by adding higher frequency modes to lower-resolution initial conditions (see, e.g., the discussion in Jenkins 2010, its section 3).

Finally, it is also possible to generate realisations of the primordial Gaussian random field subject to constraints, using for example the relatively simple algorithm by Hoffman and Ribak (1991). With this algorithm, it is possible to produce constrained simulations of massive objects without the need of simulating the large volumes required to expect at least one of them, due to their scarcity, as it is done in several articles in this Thesis (see Sec. A1 and A3). Alternatively, this algorithm has also been used in combination with observational data to produce constrained simulations of the local Universe (e.g., Gottloeber, Hoffman, and Yepes 2010).

3.2 Numerical techniques

The mathematical modelling for the evolution of cosmic inhomogeneities, reviewed in Sec. 3.1, implies, in its most basic form, the resolution of a system of five non-linear hyperbolic PDEs (Eqs. 3.8-3.10), a set of $6N$ coupled ODEs (Eqs. 3.4 and 3.5), coupled together through the gravitational force, which is obtained as the solution of the Poisson equation (Eq. 3.7), an elliptic PDE. The inclusion of additional physics, such as the ones discussed in Sec. 3.1.2.1, may add source terms to the hydrodynamic equations, or even introduce new equations, and the

particularities depend on each mechanism and each implementation. Here, I give an overview of the most standard numerical techniques for the basic system described above: Sec. 3.2.1 discusses the techniques for solving gravity and the dynamics of DM, while Sec. 3.2.2 reviews the techniques for addressing the hydrodynamic equations.

3.2.1 Overview of methods for N-Body simulations

The modelling of collisionless species is most often done through N-Body approaches, as discussed in Sec. 3.1.1. Here, the two main elements of N-Body simulations, i.e., the force computation (Sec. 3.2.1.1) and the time integration and advancement of the particles (Sec. 3.2.1.2), are discussed. It is worth mentioning that the discussion on the force computation is also relevant for the hydrodynamics part, but it is included here for the sake of the argument.

3.2.1.1 Force computation

Given a distribution of N particles with masses $\{m_{\rm i}\}_{i=1}^{N}$ in positions $\{\mathbf{x}_{\rm i}\}_{i=1}^{N}$, any N-Body code needs to be capable of computing the gravitational force acting on each particle. This can be done by either directly summing the forces, or by solving for $\phi(\mathbf{x})$ and finite-differencing it. The most widely-used alternatives are discussed below.

Direct summation. The (conceptually) simplest method consists on the obtention of the gravitational force acting on a particle by explicitly summing the contributions of the remaining $N-1$ particles. That is to say,

$$\mathbf{F}_{\rm i} = -Gm_{\rm i}\sum_{j\neq i} m_j \frac{\mathbf{x}_{\rm i}-\mathbf{x}_{\rm j}}{\left(|\mathbf{x}_{\rm i}-\mathbf{x}_{\rm j}|^2+\epsilon^2\right)^{3/2}}. \tag{3.28}$$

which, in the case $\epsilon = 0$, would represent the exact Newtonian force acting on the particle. Since N-Body particles do not represent actual physical particles, but rather a Monte-Carlo sampling of the underlying density field, they should not be regarded as point-like particles. Hence, the divergence of the force amongst arbitrarily close pairs of particles is prevented by the inclusion of the gravitational softening length, ϵ, which weakens gravity on scales $\lesssim \epsilon$ and reduces spurious two-body interactions. The softening length is usually set globally to a fraction (broadly, $0.01-0.1$) of the mean interparticle separation.

Within this approach, the algorithmic complexity of an N-Body force computation is $\mathcal{O}(N^2)$, what renders it prohibitive for large numbers of particles (large

volumes and/or high resolution). Nevertheless, direct summation simulations are still used, either by using dedicated hardware (Ito et al. 1990, Makino et al. 1997, Kawai et al. 2000, Makino et al. 2003) or by performing the calculations in graphic processing units (GPUs; e.g., Schive et al. 2008, Maureira-Fredes and Amaro-Seoane 2018).

Tree methods. The computational load of direct-sum methods can be alleviated by simplifying the contributions of distant particles and substituting them by a low-order multipolar expansion, as if they were produced by a distant "macro-particle". There are different ways to perform this substitution but, perhaps, the most widely used approach is the tree algorithm by Barnes and Hut (1986). In this algorithm, the whole particle distribution is covered by a cube (the *root* node), which is split in its eight octants. This subdivision is performed recursively, each time halving the side length, until each leaf of the tree contains only one particle. In order to compute forces, the algorithm needs to 'walk' the tree, starting by its root, and only entering a node if necessary to comply with a pre-defined precision criterion.

In this way, the algorithm can achieve algorithmic complexity as good as $\mathcal{O}(N \log N)$ for a homogeneous particle distribution, although its performance is worsened when strong inhomogeneities develop. Tree methods for solving gravity have been extensively used (e.g., Dehnen 2000), especially in combination with other methods (see below).

Particle-mesh (PM). The PM method simplifies the computation of the force to $\mathcal{O}(N \log N)$ by computing it on a grid and considering a finite-difference version of the Poisson equation (Eq. 3.7). There are three main steps for the algorithm:

1. **Density assignment.** Density is computed on an $N_{\rm g}^3$ grid with cell resolution Δx, where it is common to choose $N_{\rm g}^3 \sim N$. In doing so, the density of a cell, $\rho_{\rm g}$, can be written as

$$\rho_{\rm g} = \sum_{i=1}^{N} m_{\rm i} W(\mathbf{x}_{\rm i} - \mathbf{x}_{\rm g}), \tag{3.29}$$

   where W is a smoothing kernel, that determines the particular scheme for splitting the mass of the particle amongst the neighbouring cells. The three most common choices are:

- Nearest Grid Point (NGP), where the mass of a particle is assigned to the cell containing it. That is to say, the particle is treated as point-like. The resulting forces are discontinuous at cell boundaries, representing the main drawback of this approach.
- Cloud-in-Cell (CIC), where the mass of a particle is split amongst the eight immediate neighbouring cells, according to the fraction of the particle's volume contained in each cell. That is to say, the particle is treated as a cubical cloud with uniform density. In this way, forces are continuous, but not differentiable at cell interfaces.
- Triangular-Shaped Cloud (TSC), where the mass of a particle is split amongst the 3^3 cells around the particle using a quadratic weighting function that ensures that forces are continuous and differentiable at cell interfaces.

For the particular functional form of the smoothing kernels and a more thorough discussion, the reader is addressed to Hockney and Eastwood (1981). Note that, by iterating on the particles, any of these density assignment methods is only $\mathcal{O}(N)$.

2. **Solution of Poisson's equation in Fourier space.** From the density contrast evaluated on the grid, δ_{ijk}, it is possible to solve the second-order centred finite-difference version of Poisson's equation (Eq. 3.7),

$$\begin{aligned}\phi_{i+1,j,k} + \phi_{i-1,j,k} + \phi_{i,j+1,k} + \phi_{i,j-1,k} + \phi_{i,j,k+1} + \phi_{i,j,k-1} - \\ -6\phi_{ijk} = 4\pi G a^2 (\Delta x)^2 \delta_{ijk}.\end{aligned} \tag{3.30}$$

This equation can be solved in Fourier space, where it admits a Green function of the form:

$$\tilde{G}_{lmn} = -\frac{(\Delta x/2)^2}{\sin^2\left(\frac{\pi l}{N_g}\right) + \sin^2\left(\frac{\pi m}{N_g}\right) + \sin^2\left(\frac{\pi n}{N_g}\right)}, \tag{3.31}$$

being l, m, n the indices of the Fourier modes. The solution of Eq. (3.30) is then given by

$$\tilde{\phi}_{lmn} = 4\pi G a^2 \tilde{G}_{lmn} \tilde{\delta}_{lmn}. \tag{3.32}$$

where $\tilde{\delta}_{lmn}$ are the discrete Fourier transform components of δ_{ijk}, and the solution potential is obtained by taking the inverse discrete Fourier

transform on $\tilde{\phi}_{lmn}$. Both these Fourier transforms can be performed by the Fast Fourier Transform (FFT; Cooley and Tukey 1965) algorithm, which requires $\mathcal{O}(N \log N)$ operations.

3. **Computation of the forces.** Finally, forces (per unit mass) are obtained by finite-differencing the potential, e.g.,

$$f_{ijk}^{(x)} = -\frac{\phi_{i+1,j,k} - \phi_{i-1,j,k}}{2\Delta x}, \tag{3.33}$$

for a second-order centred scheme (although higher-order schemes, involving longer stencils, can also be used; see for example Dolag et al. 2008). Forces on the grid can be interpolated back to particles using the same scheme as in the density assignment step.

Note that, strictly, the PM technique can only be applied in this version to periodic boxes since, otherwise, the FFT is artificially imposing periodicity. Additionally, the main limitation of the PM method lies on the fact that the force resolution is fixed by Δx and, once structures start to cluster strongly, many particles may end up within an individual cell and the forces amongst them are unresolved. This can be –at least partially– overcome by using the Adaptive-Mesh Refinement (AMR) technique, which is introduced in more detail in the section about hydrodynamics (Sec. 3.2.2) and within the discussion of the MASCLET code (Sec. 4). Essentially, the grid can be refined in regions of interest (e.g., where finer force resolution is required). In these regions, the potential can be initialised to the coarser, inaccurate solution, and then Eq. (3.30) can be iterated to solve for ϕ_{ijk} in the refined region. This iteration can be performed by different methods (see, e.g., Tomida and Stone 2023 for a recent discussion on different approaches), but generally have a cost $\mathcal{O}(N^{4/3}) - \mathcal{O}(N^{5/3})$. An alternative with better scaling properties are the multigrid methods (e.g., as implemented in the AMR code Athena++ by Tomida and Stone 2023), which bring the scaling back to $\mathcal{O}(N \log N)$.

Other methods. Finally, there exist several methods that combine some of the above, by using a computationally-cheaper method for long-range forces, and a more accurate solver for short-range ones. Normally, the splitting is performed in Fourier space, where the potential is written as

$$\phi_{\mathbf{k}} = \phi_{\mathbf{k}}^{\text{long}} + \phi_{\mathbf{k}}^{\text{short}}. \tag{3.34}$$

With ϵ being the length scale separating the long-range and short-range scales, the long-range potential is obtained from the PM solution as

$$\phi_{\mathbf{k}}^{\text{long}} = \phi_{\mathbf{k}} e^{-k^2\epsilon^2}, \tag{3.35}$$

and the short range potential can be specified in real space as a direct sum with a long-range cut-off (Bagla 2002),

$$\phi^{\text{short}}(\mathbf{x}) = \sum_i \frac{-Gm_i}{|\mathbf{x}-\mathbf{x_i}|} \text{erfc}\left(\frac{|\mathbf{x}-\mathbf{x_i}|}{2\epsilon}\right). \tag{3.36}$$

The short-range force can be computed by different methods. In the TreePM method (Xu 1995, Bode, Ostriker, and Xu 2000, Bagla 2002), it is computed with a tree algorithm. Over time, it has become one of the most widely-used methods for solving gravity in cosmological simulations, being implemented in some of the most widely used public codes (e.g., GADGET, Springel 2005, Arepo, Springel 2010). Alternatively, the short-range force can be computed by direct summation, giving rise to the *Particle-Particle/Particle-Mesh* method (P^3M; Hockney and Eastwood 1981).

3.2.1.2 Time integration

Once the gravitational forces are known, the ODEs in Eqs. (3.4) and (3.5) can be solved by a variety of methods. The simplest of them, although seldom used, is the first-order accurate, explicit Euler method. Writing the system of $6N$ differential equations in vector form as $\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}, t)$, Euler's method corresponds to performing the iteration:

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \mathbf{f}(\mathbf{y}_n, t_n)\Delta t \tag{3.37}$$

This simple, explicit method assumes constant derivatives over the integration interval, Δt. This can be overcome by the usage of *implicit* methods. The equivalent first-order implicit method would be

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \mathbf{f}(\mathbf{y}_{n+1}, t_{n+1})\Delta t, \tag{3.38}$$

its main disadvantage being the fact that $\mathbf{y}_{n+1}$ appears also at the right-hand side of the equation and, in general, requires numerical inversion. As a compromise, there are explicit rules that make use of information of the evolved timestep, such as the *predictor-corrector* integration, where a first, Euler-like step is performed to obtain $\mathbf{y}_{n+1}^{\text{pred}}$, and then this prediction is used to compute the corrected value,

$$\mathbf{y}_{n+1}^{\rm corr} = \mathbf{y}_n + \frac{1}{2}\left[\mathbf{f}(\mathbf{y}_n, t_n) + \mathbf{f}(\mathbf{y}_{n+1}^{\rm pred}, t_{n+1})\right]\Delta t. \tag{3.39}$$

In this case, the integration is second-order accurate, at the cost of requiring two force computations instead of just one. In general, all these methods can be seen as particular cases of the Runge-Kutta methods (Runge 1895, Kutta 1901), which use Newton-Cotes formulae to evaluate the integral $\int_t^{t+\Delta t} \mathbf{f}(\mathbf{y}(t'), t')\mathrm{d}t'$ to the desired order of accuracy using several evaluations of $\mathbf{f}$.

Yet another possibility, given that the system to be integrated is second-order ($\ddot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$), is the *leap-frog* method, where the evaluation of velocities and positions are shifted by half a timestep. In the kick-drift-kick implementation, the iteration is:

$$\begin{aligned} \mathbf{v}_{n+1/2} &= \mathbf{v}_n + \mathbf{f}(\mathbf{x}_n, t_n)\frac{\Delta t}{2}, \\ \mathbf{x}_{n+1} &= \mathbf{x}_n + \frac{1}{a}\mathbf{v}_{n+1/2}\Delta t, \\ \mathbf{v}_{n+1} &= \mathbf{v}_{n+1/2} + \mathbf{f}(\mathbf{x}_{n+1})\frac{\Delta t}{2}. \end{aligned} \tag{3.40}$$

While this method is still second-order accurate and requires two force evaluations as well, it is widely employed in N-Body integrations due to its stability, related to its symplectic properties (see the discussion in Springel 2005).

Finally, a particularly relevant issue is choosing the timestep, Δt. In a general N-Body simulation, one can associate a timescale from the maximum acceleration in the previous timestep ($\max|\mathbf{a}|$), as

$$\Delta t = \alpha\sqrt{\frac{\epsilon}{\max|\mathbf{a}|}}, \tag{3.41}$$

where $\alpha \sim 0.1$ is a dimensionless, tolerance parameter, and ϵ is a characteristic length scale, or the cell size in PM methods. In such a way, no particle can change significantly its velocity over a timestep. Alternatively, especially in implementations using PM, it is also frequent to use a Courant-like condition, where the timestep is fixed by the cell-crossing time of the fastest particle,

$$\Delta t = \alpha\frac{\Delta x}{\max|\mathbf{v}|}. \tag{3.42}$$

For a thorough discussion and comparison of different time-stepping options, the reader is referred to Power et al. (2003).

3.2.2 Overview of methods for hydrodynamic simulations

Based on the different formulations of the equations of hydrodynamics (Eqs. 3.8-3.10 in Eulerian form; Eqs. 3.12-3.14 in Lagrangian form; Eqs. 3.15-3.17 in Eulerian-Lagrangian form), there are different numerical techniques to solve them. This section briefly introduces the main three formulations, without aiming to be exhaustive but rather to give a general overview of the different approaches and highlighting some differences amongst them.

3.2.2.1 Eulerian HRSC methods

The equations of hydrodynamics in Eulerian form can be rewritten as a hyperbolic system of conservation laws (e.g., Quilis, Ibañez, and Sáez 1993, Ryu et al. 1993, Quilis, Ibañez, and Sáez 1994),

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} + \frac{\partial \mathbf{G}}{\partial y} + \frac{\partial \mathbf{H}}{\partial z} = \mathbf{S}, \tag{3.43}$$

where $\mathbf{U} = \{\delta, (1+\delta)v_x, (1+\delta)v_y, (1+\delta)v_z, E\}$ is the vector of conserved quantities[2], $\mathbf{F}$, $\mathbf{G}$ and $\mathbf{H}$ are the fluxes, and $\mathbf{S}$ are the source terms. These equations, formally resemblant to the continuity equation, imply that in the absence of sources, $\mathbf{S} = \mathbf{0}$, the volume integral of $\mathbf{U}$ would be conserved, granting these codes with excellent conservation properties.

While early implementations discretised these equations using central differences (Cen 1992), these methods are unable to treat regions with discontinuities and need to explicitly include artificial viscosity. Modern approaches, based on high-resolution shock-capturing (HRSC) methods, are essentially higher-order versions of the method by Godunov and Bohachevsky (1959). Here, it is worth keeping in mind that fluid quantities are defined as cell averages, $\bar{u} \equiv \int_{x_{n-1/2}}^{x_{n+1/2}} f_{u_n}(x)\mathrm{d}x$, where $f_{u_n}(x)$ represents the underlying continuous distribution of the variable. The general procedure for evolving the cell-averaged values according to these schemes can be summarised in three main steps (for the sake of simplicity, in the one-dimensional case):

1. Reconstruction of the quantities at cell boundaries. In order to compute the fluxes at cell interfaces while being able to treat strong discontinuities and shocks, HRSC methods compute the values at the left-hand side and the right-hand side of the interface. That is to say, in order to compute the flux of the variable u from cell i to cell $i+1$, one needs $u^L_{i+1/2}$ and $u^R_{i+1/2}$, where L and

[2]Conversely, $\mathbf{V} = \{\delta, v_x, v_y, v_z, \varepsilon\}$ are the *primitive* variables. Here, the specific internal energy, ε, can be interchanged with pressure, P.

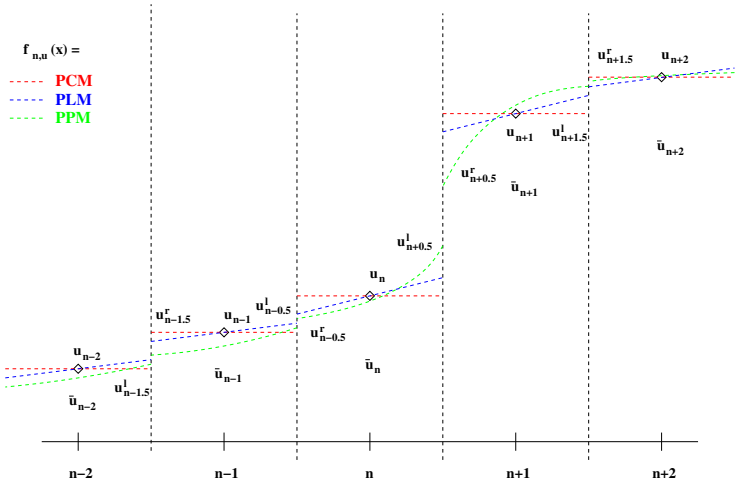


Figure 3.2: Reconstruction of the cell-averaged quantities at cell interfaces. Squares represent the cell-averaged values, while vertical dashed lines mark the cell interfaces. Red, blue and green lines correspond to piecewise constant, piecewise linear and piecewise parabolic reconstructions. Figure reproduced from Dolag et al. (2008, CC-BY-NC permission)

.

R stand for left- and right-hand side of the interface. This is usually done by piecewise polynomials, that are set by the values in the neighbouring cells. In increasing order of accuracy, basic choices are the piecewise constant method, the piecewise linear method (Colella and Glaz 1985) and the piecewise parabolic method (PPM; Colella and Woodward 1984). An example of the reconstruction of the values at cell interfaces is shown in Fig. 3.2.

The reconstruction procedures need to add several constraints to, for instance, avoid oscillations (i.e., the development of spurious extrema). For instance, in linear reconstructions it is customary to introduce slope limiters (e.g., the total variation diminishing schemes, Harten and Hyman 1983). In parabolic reconstruction, additional controls are placed to guarantee the monotonicity of the reconstruction (e.g., Colella and Woodward 1984). Finally, it is worth mentioning that, although the reconstruction can be performed either on primitive or conserved variables, the former has been argued to have more robust properties (Martí and Müller 1996), such as the avoidance of the occurrence of unphysical values.

2. Solution of the Riemann problem: computation of the fluxes. The evolution of an initial discontinuity that separates two constant states, i.e., a Riemann problem, has been thoroughly studied since many decades ago (Courant

and Friedrichs 1948). A plethora of methods to give either an exact solution (involving iterations) or an approximate solution (with a reduced computational cost, hence, most broadly used) to this problem have been developed over the years. Notable examples are the Roe (1981) solver, which solves exactly a linearisation of the system; the HLLE solver (Harten and Hyman 1983, Einfeldt 1988), which simplifies the structure of the system and only considers the fastest and slowest propagation velocities; and many variations of the latter (e.g., HLLC, Toro, Spruce, and Speares 1994).

While a detailed description of Riemann solvers exceeds the scope of this Chapter, a comprehensive description can be found in the textbooks by LeVeque (1992) or Toro (2009). In any case, from the left and right values reconstructed at the interfaces, $u^L_{i+1/2}$ and $u^R_{i+1/2}$, any particular Riemann solver will yield the fluxes at the interfaces, $F_{i+1/2}$, which are the necessary ingredient for time-stepping in the next paragraph. This description can be naturally extended to three spatial dimensions, by solving the Riemann problem along each direction independently.

3. Time-stepping. After solving the Riemann problem and obtaining all the fluxes at the interfaces, the cell-averaged values of the conserved variables can be evolved in time with an ODE integrator. Exemplifying with a first-order, explicit Euler step,

$$\begin{aligned}\mathbf{U}_{ijk}(t+\Delta t) = \mathbf{U}_{ijk}(t) - \Delta t \left(\frac{\mathbf{F}_{i+1/2,j,k} - \mathbf{F}_{i-1/2,j,k}}{\Delta x} + \right. \\ \left. + \frac{\mathbf{G}_{i,j+1/2,k} - \mathbf{G}_{i,j-1/2,k}}{\Delta y} + \frac{\mathbf{H}_{i,j,k+1/2} - \mathbf{H}_{i,j,k-1/2}}{\Delta z} \right) + \Delta t \mathbf{S}_{ijk}.\end{aligned} \tag{3.44}$$

In practice, higher order schemes granting better conservation and stability properties are used, like the third-order Runge-Kutta of Shu and Osher (1988). This integration is in principle stable if the timestep is chosen according to the Courant, Friedrichs, and Lewy (CFL; 1928) condition,

$$\Delta t \leq \frac{a(t)\Delta x}{\max |c_s + v|}, \tag{3.45}$$

where c_s is the sound speed and v is the velocity. In practice, the timestep is usually chosen as a fraction of the maximum allowed by the CFL condition, $\Delta t = \alpha \Delta t_{\rm CFL}$, with $\alpha \sim 0.1 - 0.6$.

Generally speaking, Eulerian methods are well-suited to describe both high-density and low-density regions, and handle accurately shocks, discontinuities

and strong gradients, without the need of explicitly introducing artificial viscosity. However, they involve a large computational cost, severely limiting resolution and consequently the dynamical range that can be resolved in a simulation. This is why the AMR technique (Berger and Oliger 1984, Berger and Colella 1989) has become a standard feature in Eulerian codes for cosmological structure evolution (Teyssier 2002, Quilis 2004, Bryan et al. 2014). Even though implementations are diverse, in general AMR allows to dynamically refine regions of interest and solve them with enhanced temporal and spatial resolution. For more details on the AMR implementation in MASCLET, see Sec. 4.2.

3.2.2.2 Lagrangian SPH methods

The most widely spread implementation for solving the Lagrangian form of the equations of hydrodynamics are the so-called smoothed particle hydrodynamics methods (SPH; Lucy 1977, Gingold and Monaghan 1977). In this approach, the fluid is discretised in finite mass elements (particles), and the properties of the fluid at a given position are obtained by smoothing the properties of the neighbouring particles. That is to say, the value of a given field $A(\mathbf{x})$ is computed from the value of this magnitude amongst the neighbouring particles, A_i, as

$$A(\mathbf{x}) = \sum_{j=1}^{\mathrm{N_{ngh}}} \frac{m_j}{\rho_j} A_j W(\mathbf{x} - \mathbf{x_j}, h_j), \tag{3.46}$$

where N_{ngh} is the number of neighbours considered for the computation, m_j/ρ_j represents the volume associated to the particle in position $\mathbf{x_j}$ and $W(\cdot, h_j)$ is a smoothing kernel with smoothing length h_j, which is usually determined from the radius of a sphere containing N_{ngh} neighbours around the particle. In this way, also derivatives can be computed by summing over the neighbouring particles, since

$$\frac{\partial A}{\partial x^\alpha} = \sum_{j=1}^{\mathrm{N_{ngh}}} \frac{m_j}{\rho_j} A_j \frac{\partial}{\partial x^\alpha} W(\mathbf{x} - \mathbf{x_j}, h_j). \tag{3.47}$$

where x^α is an arbitrary component of the vector $\mathbf{x}$. Using the above relations at the position of the i-th particle, $\mathbf{x_i}$, and using properties of the derivatives to symmetrise the expressions above, the Lagrangian equations of hydrodynamics (Eqs. 3.12-3.14), in the absence of gravity and expansion (for the sake of simplicity), are discretised as:

$$\frac{\mathrm{d}\mathbf{v_i}}{\mathrm{d}t} = -\sum_{j=1}^{\mathrm{N_{ngh}}} m_j \left(\frac{P_i}{\rho_i^2} + \frac{P_j}{\rho_j^2} + \Pi_{ij} \right) \nabla_i W(\mathbf{x_i} - \mathbf{x_j}, h_i), \tag{3.48}$$

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} = \frac{1}{2} \sum_{j=1}^{\mathrm{N_{ngh}}} m_j \left(\frac{P_i}{\rho_i^2} + \frac{P_j}{\rho_j^2} + \Pi_{ij} \right) (\mathbf{v_j} - \mathbf{v_i}) \nabla_i W(\mathbf{x_i} - \mathbf{x_j}, h_i). \tag{3.49}$$

Here, the term Π_{ij} is the artificial viscosity, introduced by hand to deal with shocks and discontinuities (Monaghan and Gingold 1983, Balsara 1995, Monaghan 1997). Since artificial viscosity damps the signal at small scales, it is desirable to keep its effects as contained as possible. Most modern implementations follow the idea of Morris and Monaghan (1997), where each particle carries its own artificial viscosity, which gets activated when undergoing shocks and decays outside them.

The number of neighbours, N_{ngh}, as well as the order of the kernel, are chosen as a compromise between resolution, accuracy and computational cost. Many improvements over this baseline methods do exist, such as the entropy-conserving formulation (Springel and Hernquist 2002), or schemes with reduced artificial viscosity (Beck et al. 2016). While Lagrangian methods achieve better resolution in high-density regions than Eulerian methods with equivalent computational cost, their treatment of low-density regions is poor by construction. Additionally, due to artificial viscosity, these techniques have a degradation of their resolution in shocks, and may have problems in resolving some instabilities (Agertz et al. 2007). For a thorough review on SPH methods, the reader is referred to, e.g., Price (2012).

3.2.2.3 Hybrid methods

The arbitrary Eulerian-Lagrangian formulation of the equations (Eqs. 3.15-3.17) can be applied to construct mesh codes where the grid is allowed to move freely. Such is the case of Arepo (Springel 2010), where a set of mesh-generating points that move with the fluid are used to define the grid through a Voronoi tessellation. In this way, the grid follows the fluid flow, continuously deforming and adapting to the local properties of the fluid, but while keeping the advantages of Eulerian methods. In the last decade, mesh-free codes based on an arbitrary Eulerian-Lagrangian formulation have also triggered considerable interest (e.g., Hopkins 2015, Groth et al. 2023).

CHAPTER 4

Cosmological simulations with MASCLET

Contents

BESIDES COMPARISONS WITH RESULTS from other simulation codes (for instance, the ones contained within Secs. 6 and 8.2, and papers A5 and A6), the majority of the results reported within this Thesis correspond to the analysis of simulations carried out with MASCLET (Mesh Adaptive Code for CosmologicaL structurE evoluTion; Quilis 2004, Quilis, Martí, and Planelles 2020), which combines a Eulerian description of the collisional component with a PM scheme for DM and other collisionless species, both implemented on top of an AMR scheme. In this Chapter, the main features of the code, together with its recent developments during the last years, are presented. In particular, in Sec. 4.1, the fix-grid hydrodynamics, N-Body and gravity solvers are described. Their AMR implementation is discussed in Sec. 4.2, while the particularities of the additional physics included in MASCLET are presented in Sec. 4.3, together with some comments on the analysis tools specifically created for MASCLET in Sec. 4.4.

4.1 The monolithic solvers

Here, I describe the basic hydrodynamics (Sec. 4.1.1), N-Body (Sec. 4.1.2) and gravity (Sec. 4.1.3) solvers implemented in MASCLET for the base grid[1], which also constitute the basic ingredients for the following AMR implementation.

4.1.1 The hydro solver

For addressing the evolution of the collisional component, MASCLET makes use of Eulerian, HRSC methods, which have been introduced in some detail in Sec. 3.2.2.1. Although MASCLET implements several reconstruction mechanisms, Riemann solvers, and time integrators, this section is restricted to the most usual configuration.

Reconstruction. Most frequently and, in particular, for all applications within this PhD Thesis, MASCLET uses the PPM reconstruction scheme (Colella and Woodward 1984), following the prescriptions and parameters of Martí and Müller (1996), on the set of primitive variables ($\mathbf{V} = \{\delta, v_x, v_y, v_z, u\}$) in order to get the values at left-hand side and right-hand side of each interface. Reconstructing the specific internal energy, u, instead of the total energy density, E, prevents the former to be dominated by numerical error in kinetically-dominated flows ($\rho u \ll E$).

Besides the standard PPM reconstruction (which includes several controls to avoid non-monotonic reconstructions and detect shocks, in which cases the reconstruction order is downgraded to constant), the order of the reconstruction is additionally lowered to linear or constant in some situations:

- For the borders of the grid (i.e., for the ghost cells used to enforce the boundary conditions, $i = 0$ and $i = N_x + 1$), the reconstruction is zeroth-order, while for cells $i = 1$ and $i = N_x$ it is first-order.
- In low-density regions ($\delta < -0.99$), the reconstruction is also zeroth-order, since these cases are easily contaminated from arithmetic error stemming from the conversion between primitive and conserved variables ($1 + \delta \ll 1$).

Riemann solver. The standard Riemann solver in MASCLET is HLLE (Harten and Hyman 1983, Einfeldt 1988), which provides a conceptually simple,

[1]Here, *base grid* makes reference to the uniform grid (usually cubic) covering the whole domain with constant resolution, in contrast to the *refinement patches* that will be introduced in Sec. 4.2.

robust and computationally cheap solution, even though it is known to produce excessive numerical diffusion in the case of contact discontinuities (Toro, Spruce, and Speares 1994). According to this method, the prescription for computing the flux at the x-interface between cell i and cell $i+1$, $\mathbf{F}^{\rm HLLE}_{i+1/2}$, is the following[2]:

1. Compute the maximum and minimum signal velocities at the interface[3], $\mathcal{S}^{\rm max}_{i+1/2}$ and $\mathcal{S}^{\rm min}_{i+1/2}$, respectively, as:

$$\mathcal{S}^{\rm max}_{i+1/2} = \max\left[0, \frac{v^R_{x,i+1/2} + c^R_{s,i+1/2}}{a}, \frac{v^L_{x,i+1/2} + c^L_{s,i+1/2}}{a}\right] \tag{4.1}$$

$$\mathcal{S}^{\rm min}_{i+1/2} = \min\left[0, \frac{v^R_{x,i+1/2} - c^R_{s,i+1/2}}{a}, \frac{v^L_{x,i+1/2} - c^L_{s,i+1/2}}{a}\right] \tag{4.2}$$

   where a is the ΛCDM expansion factor and c_s is the sound speed. Hence, note that $\mathcal{S}^{\rm min}_{i+1/2} < 0 < \mathcal{S}^{\rm max}_{i+1/2}$ corresponds to subsonic motion of the gas, while $\mathcal{S}^{\rm min}_{i+1/2} \geq 0$ or $\mathcal{S}^{\rm max}_{i+1/2} \leq 0$ corresponds to supersonic motion.

2. Compute the left and right fluxes, $\mathbf{F}^L_{i+1/2} = \mathbf{F}(\mathbf{U}^L_{i+1/2})$ and $\mathbf{F}^R_{i+1/2} = \mathbf{F}(\mathbf{U}^R_{i+1/2})$, where $\mathbf{U}^L_{i+1/2}$ and $\mathbf{U}^R_{i+1/2}$ are the left and right states at the interface, respectively.

3. Compute the HLLE flux as:

$$\mathbf{F}^{\rm HLLE}_{i+1/2} = \frac{\mathcal{S}^{\rm max}_{i+1/2}\mathbf{F}^L_{i+1/2} - \mathcal{S}^{\rm min}_{i+1/2}\mathbf{F}^R_{i+1/2} + \mathcal{S}^{\rm max}_{i+1/2}\mathcal{S}^{\rm min}_{i+1/2}(\mathbf{U}^R - \mathbf{U}^L)}{\mathcal{S}^{\rm max}_{i+1/2} - \mathcal{S}^{\rm min}_{i+1/2}} \tag{4.3}$$

   Note that, according to this expression, in the trivial case of supersonic motions ($\mathcal{S}^{\rm min}_{i+1/2} \geq 0$ or $\mathcal{S}^{\rm max}_{i+1/2} \leq 0$), the HLLE flux will be determined by the left or right flux alone, respectively, since information is unable to propagate upstream of the shock.

Time integration. Although a third-order Runge-Kutta scheme is also available in MASCLET, the standard time integrator is a second-order Runge-Kutta, or Heun's method, which is a predictor-corrector scheme (i.e., the second-order case described by Shu and Osher 1988). Explicitly, the time integration is as follows:

[2]The three-dimensional case is addressed analogously, as discussed in Sec. 3.2.2.1. Here, the process is discussed in just one spatial dimension for conciseness.

[3]Through this discussion, I shall use calligraphic $\mathcal{S}$ for the signal velocities, to avoid confusion with the source vector, $\mathbf{S}$.

$$\begin{aligned}
\mathbf{U}^{\rm pred}_{i,t+\Delta t} &= \mathbf{U}_{i,t} + \Delta t \left[-\frac{\mathbf{F}_{i+1/2} - \mathbf{F}_{i-1/2}}{\Delta x} + \mathbf{S}_i \right]_{\mathbf{U}_{i,t}} , \\
\mathbf{U}_{i,t+\Delta t} &= \frac{1}{2}\mathbf{U}_{i,t} + \frac{1}{2}\mathbf{U}^{\rm pred}_{i,t+\Delta t} + \frac{\Delta t}{2} \left[-\frac{\mathbf{F}_{i+1/2} - \mathbf{F}_{i-1/2}}{\Delta x} + \mathbf{S}_i \right]_{\mathbf{U}^{\rm pred}_{i,t+\Delta t}} .
\end{aligned} \tag{4.4}$$

Naturally, in the three-dimensional case, the reconstruction and Riemann solver are performed separately in each Cartesian direction, and the time advancement is only performed at the end of this process, adding also the **G** and **H** terms of Eq. (3.43) (see Eq. 3.44 for the three-dimensional, first-order accurate evolution rule). It is worth noting that the source term due to gravity is assumed to be frozen during the substeps of the Runge-Kutta, i.e., for the computation of $\mathbf{S}_i|_{\mathbf{U}^{\rm pred}_{i,t+\Delta t}}$, the gravitational acceleration is the same as the one in $\mathbf{S}_i|_{\mathbf{U}_i}$. This avoids the necessity of having to couple the substeps within the integrators of the hydrodynamics and gravity solvers.

4.1.2 The N-Body solver

The N-Body solver in MASCLET is an implementation of the PM scheme (see Sec. 3.2.1.1, where this technique has already been introduced in detail). The description in this section restricts to the particularities of the DM density assignment onto the grid and the time integration, while the solution for the gravitational potential is discussed in Sec. 4.1.3.

Density assignment. MASCLET uses a TSC scheme for density assignment, i.e., a given particle will contribute to the cell containing it, and its $3^3 - 1$ immediate neighbours. The mass assignment function is quadratic and is separable in the three Cartesian directions, in such a way that the mass on a given cell (with indices i, j, k) is

$$m_{ijk} = \sum_{n=1}^{N_{\rm part}} m_n \, W\left(\frac{|x_n - x_i^g|}{\Delta x}\right) W\left(\frac{|y_n - y_j^g|}{\Delta y}\right) W\left(\frac{|z_n - z_k^g|}{\Delta z}\right), \tag{4.5}$$

where m_n is the mass of the n-th particle, x_n, y_n, z_n are the particle coordinates and x_i^g, y_j^g, z_k^g are the coordinates of the cell centre. W is the one-dimensional mass assignment function which, for TSC, is defined as

$$W(d) = \begin{cases} \frac{3}{4} - d^2 & \text{if } d < \frac{1}{2} \\ \frac{1}{2}\left(\frac{3}{2} - d\right)^2 & \text{if } \frac{1}{2} \leq d < \frac{3}{2} \\ 0 & \text{otherwise} \end{cases} \tag{4.6}$$

In a uniform grid, the contribution of a particle is only non-zero for the 3^3 mesh points closest to the particle, and it is normalised so that $\sum_{i,j,k}^{3} W_i W_j W_k = 1$.

Time integration. N-Body particles are advanced in time using a second-order accurate, Runge-Kutta integrator involving an intermediate step at $t^{n+1/2} = t^n + \Delta t/2$. Explicitly, the temporal discretisation of Eqs. (3.4) and (3.5) is as follows:

1. The intermediate step, $n + 1/2$, is computed as:

$$\mathbf{x}^{n+1/2} = \mathbf{x}^n + \frac{\Delta t}{2}\frac{\mathbf{v}^n}{a^n}, \tag{4.7}$$

$$\mathbf{v}^{n+1/2} = \mathbf{v}^n - \frac{\Delta t}{2}\left(\frac{\nabla\phi^n}{a^n} + H^n\mathbf{v}^n\right), \tag{4.8}$$

   where superindices correspond to the timestep, and $a^n \equiv a(t^n)$, $H^n \equiv H(t^n)$ are the expansion factor and Hubble parameter at time t^n, respectively.

2. Then, the final step, $n + 1$, is computed as:

$$\mathbf{x}^{n+1} = \mathbf{x}^n + \Delta t\frac{\mathbf{v}^{n+1/2}}{a^{n+1/2}}, \tag{4.9}$$

$$\mathbf{v}^{n+1} = \mathbf{v}^n - \Delta t\left(\frac{\nabla\phi^{n+1/2}}{a^{n+1/2}} + H^{n+1/2}\mathbf{v}^{n+1/2}\right). \tag{4.10}$$

   Here, the gravitational force at $t^{n+1/2}$ is computed by linear extrapolation of ϕ from the values at t^n and t^{n-1}, thus avoiding the necessity of solving Poisson's equation during the substep.

In both substeps, the gravitational field is computed by taking second-order accurate, centred differences from ϕ in the 2^3 cells closest to the particle position, and linearly interpolating it back to the particle position.

4.1.3 The gravity solver

Gravity at the base level (without any mesh refinement) follows what has been described in general within Sec. 3.2.1.1, implying that, by design, periodic

boundaries are always assumed. For the sake of a fast, parallel implementation of the FFT, the library FFTW (Frigo and Johnson 2005) is used.

4.1.4 Time-stepping

To preserve the stability of the integration, several constraints on the timestep are imposed. These include the following (each of them depending on a free parameter, η_i):

- Particle cell-crossing time, to prevent particles from moving more than a given fraction of the cell side length in a single timestep (since the code is using a PM approach). If $v_{\max}$ is the maximum absolute value amongst the Cartesian components of all particles' peculiar velocities, the associated timestep is
$$\Delta t_1 = \eta_1 \frac{a\Delta x}{v_{\max}}. \tag{4.11}$$

- Dynamical time, to prevent a fast, unresolved collapse of a cold (pressureless) overdensity. If $\delta_{\mathrm{T}}^{\max}$ is the maximum value of the total density contrast (including gas, DM, as well as stars and BH, if present), the associated timestep is then:
$$\Delta t_2 = \eta_2 \sqrt{\frac{3\pi^2}{4G\rho_B(z)(1+\delta_T^{\max})}}. \tag{4.12}$$

- Courant time, as discussed in Sec. 3.2.2.1. In the three-dimensional case, it is customary to compute the most stringent one-dimensional signal velocity,
$$v_{\max} = \max_{ijk} \left[\max(|v_x|, |v_y|, |v_z|) + c_s\right], \tag{4.13}$$
and then define the timestep as
$$\Delta t_3 = \eta_3 \frac{a\Delta x}{\sqrt{3} v_{\max}}, \tag{4.14}$$
where the $\sqrt{3}$ factor arises as a conservative estimate of the maximum signal velocity in the three-dimensional case given the one-dimensional maximum.

- The timestep associated to the expansion factor is set so that $a(t)$ does not change by more than a given fraction, η_4, in a single timestep:
$$\Delta t_4 = \eta_4 \frac{a}{\dot{a}} \equiv \frac{\eta_4}{H(t)}. \tag{4.15}$$

- If cooling is present, the timestep is also constrained by the cooling time, which is defined as the ratio between the thermal energy density and the cooling rate, $\tau_{\rm cool} = \rho u/|\Lambda|$. The timestep is then defined as

$$\Delta t_5 = \eta_5 \min_{ijk} \tau_{\rm cool}. \tag{4.16}$$

- Stability condition: to prevent sudden changes in the timestep that may lead to the onset of instabilities, the timestep is additionally constrained to increase less than a maximum factor, η_6, in a single timestep:

$$\Delta t_6 = (1 + \eta_6)\Delta t_{\rm old}. \tag{4.17}$$

Typically, all the η_i parameters are set to some value smaller than 1, although the particularities may change depending upon the application. In some cases, some of the criteria can be disregarded: for instance, in simulations with gas cooling but no efficient mechanisms to offset it, the criteria Δt_2 and Δt_5 may yield prohibitively small integration times as result of the strong dynamics of a negligible region (in terms of volume or mass), and are often dropped provided they are no cause of numerical instabilities. With the suitable set of timestep constraints selected, once within the evolution, the most restrictive of the timescales is chosen at each integration step, i.e.,

$$\Delta t = \min\left[\Delta t_1, \Delta t_2, \Delta t_3, \Delta t_4, \Delta t_5, \Delta t_6\right]. \tag{4.18}$$

4.2 The Adaptive-Mesh Refinement (AMR) strategy

To overcome the lack of resolution inherent to Eulerian techniques, MASCLET implements a block-based AMR strategy, in the manner of Berger and Colella (1989). This is done in such a way that, each global timestep, the code picks the regions in the original domain that must be refined according to some criteria (Sec. 4.2.1) and maps them with a set of finer grids (the *refinement patches*). This is done a number of times in a hierarchical manner achieving high dynamical ranges. The equations from the coupled evolution of DM and baryons are then solved with increased spatial and temporal resolution (Sec. 4.2.2), with some modifications to the base solvers that need to be taken into account in the presence of mesh refinement (Secs. 4.2.3 to 4.2.5).

4.2.1 Creation of the mesh

Before beginning each global timestep of the simulation, the whole grid hierarchy is rebuilt to keep tracking the regions that need to be resolved with enhanced resolution. This is always done hierarchically, from a first level of refinement, $\ell = 1$, to a maximum level, $\ell = n_\ell$. Although AMR implementations in the literature vary, in MASCLET all refinement patches halve the cell size with respect to their *progenitor* patches. That is to say, the resolution at level ℓ is $\Delta x_\ell = \Delta x/2^\ell$.

Even though the criteria may change depending on the application, for typical cosmological structure formation simulations focusing on high-density regions, these criteria include:

- Pseudo-Lagrangian criterion based on baryonic density: a cell at level ℓ is flagged as refinable if its baryonic overdensity, Δ_B, exceeds $\Delta_\ell^\mathrm{thr} = f_\mathrm{B} 8^{\ell-1}$, where f_B is the cosmic baryon fraction. In practice, a value smaller than 8 can be used for the basis of the exponential, to favour the refinement of high density regions. Additionally, in simulations with BH formation and AGN feedback, the regions around BH particles are flagged as refinable at all levels.

- Pseudo-Lagrangian criterion based on DM density: similarly to the former, a cell is flagged as refinable if its DM overdensity, Δ_DM, exceeds $\Delta_\ell^\mathrm{thr} = f_\mathrm{DM} 8^{\ell-1}$, where f_DM is the cosmic DM fraction.

- Jeans length criterion: if the Jeans length, $\lambda_\mathrm{J} = \sqrt{15 k_\mathrm{B} T/(4\pi G m_\mathrm{part}\rho)}$, is unresolved, $\lambda_\mathrm{J} < 4\Delta x$, the cell is flagged as refinable.

- If the cell is physically collapsing, i.e., $\nabla_\mathbf{r} \cdot \mathbf{u} = \frac{1}{a}\nabla \cdot \mathbf{v} + 3H(t) < 0$, it is flagged as refinable.

- Cells at a given level, ℓ, containing particles that were within a refinement level $\ell' > \ell$ in the initial conditions, are flagged as refinable[4].

Once the whole set of refinable cells at level ℓ is obtained, the mesh creation routine generates the patches at level $\ell + 1$ as follows (exemplified in Fig. 4.1):

[4]In MASCLET, the AMR strategy is also customarily used at the level of the initial conditions. This is done by performing a low-resolution simulation, from which the regions of interest at $z = 0$ are flagged. The DM particles within these regions at $z = 0$ are traced back to $z = z_\mathrm{ini}$, and the initial conditions can be regenerated with a number of refinement levels covering the Lagrangian surroundings of the regions of interest. DM particles belonging to a level-ℓ patch at $z = z_\mathrm{ini}$ also get divided in 8^ℓ particles.

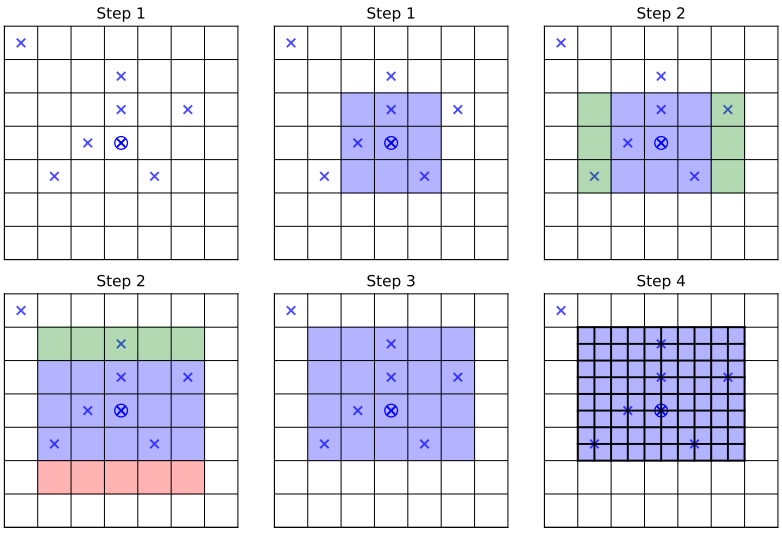


Figure 4.1: Schematic representation of the mesh creation strategy in MASCLET. Crosses represent refinable cells, while the circled cross represents the cell maximising the refinability condition. A 3^2 coarse cells patch is considered around this cell, and it is subsequently enlarged in each direction if new refinable cells are included during step 2. Step 3 shows the conclusion of the growing procedure, if no more refinable cells are to be added. In step 4, the patch has been mapped with twice the resolution of its parent patch.

1. The cell maximising the refinement condition (in the order specified above) is chosen, and extended to a cubic, 5^3-coarse cells region around it.

2. Iteratively, the patch is extended along each of its six faces, only if at least one refinable cell is included by the extension.

3. The iteration is stopped once the largest of the dimensions of the patch reaches a maximum value, $N_{\text{patch}}^{\text{max}}$. This is done because, due to of reasons of efficient memory layout, the AMR fields are stored in four-dimensional arrays of fix size $(N_{\text{patch}}^{\text{max}})^3 \times N_{\text{patches}}$.

4. The patch is accepted if the smallest of its dimensions has a minimum number of cells, $N_{\text{patch}}^{\text{min}}$. Otherwise, the patch is rejected, and the considered cell is removed from the list of refinable cells.

5. The new, *child* patch is mapped with twice the resolution of its parent patch, and the parent cells are removed from the list of refinable cells.

The process is repeated until all refinable cells have been either refined or discarded. Once the whole set of patches at level $\ell + 1$ has been generated,

the values of its conserved variables are obtained. For each cell, there are two possibilities:

- If the cell was already refined at level $\ell+1$ in the previous timestep, the value of the variable is obtained by copying the value from the corresponding cell at the end of the previous timestep.

- If the cell was not already refined at level $\ell+1$, then the value of the variable is obtained by trilinear interpolation from the values of the highest-available level of refinement, ℓ', with $\ell' \leq \ell$.

By this procedure, and unlike other AMR implementations based on octrees[5], the patches at level $\ell + 1$ may overlap with each other. While each patch works as an independent computational domain during the evolution, subject to the boundary conditions imposed by the coarser levels, overlapping regions need to be synchronised after each AMR step of the hydrodynamics solver (see Sec. 4.2.2 for a description of the time-stepping in the presence of AMR), so that any quantity has a unique value at a given position and time, and different, overlapping patches do not present divergent solutions due to their possibly different boundary conditions. Likewise, to maintain the conservation properties of the scheme, when the patches at level $\ell + 1$ reach the time of patches at level ℓ, the values of the conserved variables at level $\ell + 1$ are averaged to update the values at level ℓ. The mesh creation strategy also needs to take into account the overlaps to avoid refining the same region twice.

4.2.2 Time-stepping in the presence of mesh refinement

In its current version, MASCLET employs the same time-stepping scheme for DM and hydrodynamics when AMR is activated. Once the base grid has been advanced in time from t to $t + \Delta t$ following the procedure described in Sec. 4.1, the refinement levels need to be evolved. Patches at level ℓ will require 2^ℓ steps, each of $\Delta t_\ell = \Delta t/2^\ell$.

At any given moment, both for the hydro solver and for the gravity solver (see Secs. 4.2.3 and 4.2.5 below), it may be necessary to establish the boundary conditions of the patch from the information available at coarser levels. Since finer levels do more timesteps than coarser ones, the information of the coarser levels is generally not directly available at the times of the steps of the finer

[5] In an octree-AMR implementation, the refinement is performed in a cell-wise basis, instead of in patches, as described here for MASCLET.

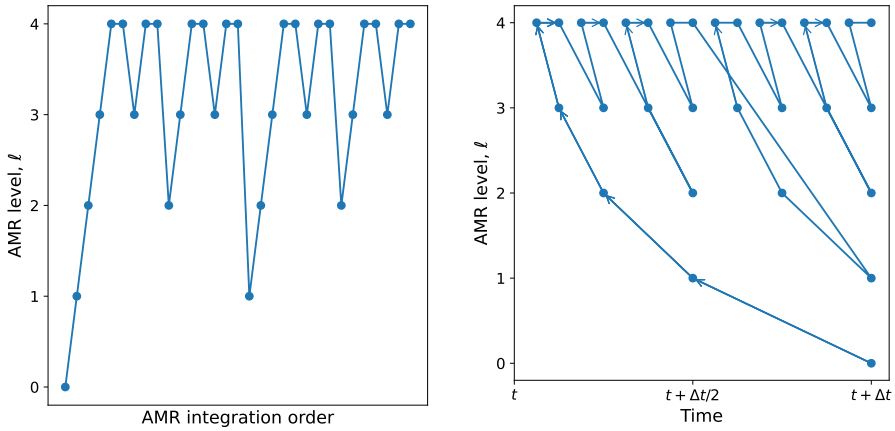


Figure 4.2: Schematic representation of the time-stepping scheme in MASCLET AMR implementation. Left-hand side panel shows the heuristic order in which the different levels are integrated. Right-hand side panel shows the time advancement of the different levels, with the arrows indicating the order of the integration.

levels, and is obtained by linear interpolation from the values at the immediately previous and next steps of the coarser level.

Therefore, the coarser levels always need to be evolved before the finer ones. This is achieved in MASCLET by a W-like cycle, which is represented graphically in Fig. 4.2. Besides this graphical representation, it might be useful to explicitly describe the process for a simple case ($n_\ell = 2$). In this case, the time-stepping scheme is as follows:

1. The base grid is evolved from t to $t + \Delta t$.

2. The level $\ell = 1$ is evolved from t to $t + \Delta t/2$, and its boundary conditions are set by the base grid at t.

3. The level $\ell = 2$ is evolved from t to $t + \Delta t/4$, and its boundary conditions are set by the $\ell = 1$ level at t.

4. The level $\ell = 2$ is evolved from $t + \Delta t/4$ to $t + \Delta t/2$, and its boundary conditions are set by interpolation of the $\ell = 1$ level at t and $t + \Delta t/2$. At this moment, level $\ell = 2$ has caught up with level $\ell = 1$, and hence the coarse level can be updated with the average of the fine level.

5. The level $\ell = 1$ is evolved from $t + \Delta t/2$ to $t + \Delta t$, and its boundary conditions are set by interpolation of the base grid at t and $t + \Delta t$. At this

moment, level $\ell = 1$ has caught up with the base grid, and hence the base grid can be updated with the average of the $\ell = 1$ level.

6. The level $\ell = 2$ is evolved from $t + \Delta t/2$ to $t + 3\Delta t/4$, and its boundary conditions are set by the $\ell = 1$ level at $t + \Delta t/2$.

7. The level $\ell = 2$ is evolved from $t + 3\Delta t/4$ to $t + \Delta t$, and its boundary conditions are set by interpolation of the $\ell = 1$ level at $t + \Delta t/2$ and $t + \Delta t$. Again, $\ell = 1$ patches can be updated with the average of the $\ell = 2$ level.

It is worth to mention that, in MASCLET's AMR implementation, a given cell at level ℓ can be refined, not by a child patch, but by the child patch of any other patch at the same level ℓ which overlaps it (a *nephew* patch). In this case, the fine-to-coarse averaging procedure also needs to update these values accordingly. Similarly, the necessary boundary conditions for a given patch may not be available within its parent patch (e.g., because the child reaches the edge of its parent), but in a sibling of the parent patch. These aspects increase the complexity of block-based AMR codes, as well as their parallelisation, beyond that of octree codes. However, the block-based approach brings several advantages, such as a smaller fraction of the domain corresponding to fine-coarse interfaces, where corrections have to be made (see below, Sec. 4.2.3) and numerical instabilities can arise.

An additional complexity brought up by the AMR scheme is that, due to the time synchronisation scheme imposed, it is possible that the Courant condition (or any other of the time-stepping constraints) is violated during the AMR steps, since it is computed just once per global timestep. This normally demands the use of more conservative values for the η_i parameters in the timestep constraints (see Sec. 4.1.4).

4.2.3 Hydrodynamics on an adaptive mesh

The hydrodynamics solver on the AMR refinement patches is equivalent to the one on the base grid, relying on PPM reconstruction, HLLE Riemann solver and the two-step, predictor-corrector Runge-Kutta time integrator. However, a few details relating to the treatment of boundaries are worth to be mentioned:

Boundary conditions and ghost cells. Let a given refinement patch at level ℓ have $N_x^{\rm patch}$ cells along the x direction, which correspond to the domain that the mesh-creation strategy has determined that needs to be refined. However, in order to perform the PPM reconstruction around the i-th cell, the shock

detector needs the values of some quantities (in particular, pressure) in the cells $i-3$ and $i+3$. This requires three *ghost* zones along each direction ($i=-2,-1,0,N_x+1,N_x+2,N_x+3$), which are not evolved but are used only for the reconstruction.

These cells are obtained, if possible, by copying the value from a patch at the same level of refinement ℓ. If the corresponding region is not refined, the value is linearly interpolated in space and time from the values of the highest level $\ell' < \ell$ possible.

Conservation properties at the fine-coarse interfaces. The conservation properties of HRSC methods, as described in Sec. 3.2.2.1, rely on the fact that whatever quantity is subtracted from the cell at the left-hand side of an interface corresponds exactly to what is added to the cell at its right-hand side. Since the fluxes are non-linear functions of the set of primitive variables, fluxes computed at different resolutions do not necessarily match, posing a serious shortcoming due to the explicit violation of the conservativity of the method.

This problem is mitigated by the flux correction, or *refluxing*, scheme of Berger and Colella (1989). In simple terms, the basic idea, applied to the interface between cells at level ℓ and non-refined cells at level $\ell-1$ located at their right (e.g., the situation in the rightmost column of the bottom, right panel in Fig. 4.1), is to compute the difference between the *fine* fluxes that have been used to integrate the cells of the ℓ-level grid at the left of the interface, and the coarse fluxes that have been used to integrate the cells of the $(\ell-1)$-level grid at the right of the interface, and ascribe the resulting difference in the conserved quantities to the coarse cell as a manual correction after the evolution.

Focusing on a particular $\ell-1$ cell, the flux mismatch would be computed as

$$\delta\mathbf{F} = -\mathbf{F}^{\rm coarse} + \frac{1}{8}\sum_{\substack{\text{2 AMR}\\ \text{steps}}}\sum_{\substack{\text{4 neigh-}\\ \text{bouring cells}}}\mathbf{F}^{\rm fine}, \tag{4.19}$$

that is, the difference between the coarse flux, and the average of the fine flux over the 2 AMR timesteps and the 4 fine cells touching the coarse cell. This difference is then added to the coarse cell, so that it gets updated as

$$\mathbf{U} \leftarrow \mathbf{U} + \delta\mathbf{F}\frac{\Delta t}{\Delta x}. \tag{4.20}$$

In practical applications, even though this procedure where a cell is modified 'by hand' guarantees conservativity exactly, it has some drawbacks. First, it can introduce instabilities due to the modification of the cell conserved values,

especially in regions of low density or kinetically dominated. Additionally, it implies the necessity of saving the fine and the coarse fluxes at the boundaries of the refinement patches. Last, there is not such a simple solution for performing a correction in the case of a boundary between a cell at level ℓ and a cell at level $\ell' < \ell - 1$.

4.2.4 N-Body on an adaptive mesh

Particles in MASCLET AMR simulations get assigned a level, which corresponds to the level of refinement of the highest-resolution patch the particle has ever been embedded in. That is to say, if a particle enters a higher-resolution region, it gets promoted to a higher level, but it never gets demoted. Particles get evolved following the same scheme as described in Sec. 4.2.2, i.e., a particle at level ℓ performs 2^ℓ AMR timesteps of size $\Delta t/2^\ell$ during a global timestep of the simulation.

The integration process for each AMR step is similar to the one described for the monolithic N-Body solver, but the gravitational fields are computed from finite-differences of the potential at the highest resolution available. If the level of the most well-resolved patch containing the particle of level ℓ is $\ell_{\rm patch} = \ell$, the process is straightforward. If $\ell_{\rm patch} < \ell$, the force at the timestep of the particle is computed by linear interpolation in space and time. By construction, the case $\ell_{\rm patch} > \ell$ cannot happen, since the particle would have then been promoted to a higher level.

4.2.5 Gravity on an adaptive mesh

In order to solve Poisson's equation for the gravitational potential within the refinement patches, iterative solvers have to be used, as mentioned in Sec. 3.2.1.1. In the case of MASCLET, a successive over-relaxation (SOR; see, for instance, Press et al. 1992) method is applied on the discretised version of Poisson's equation.

The potential in the interior of the patch is initialised by the solution of the previous AMR timestep, which is usually a close approximation to the updated gravitational potential. The boundary conditions are imposed in a ghost zone around the patch, and are obtained by linear interpolation from the parent patch. Subsequently, the value of the potential can be iteratively updated in a chessboard pattern as:

$$\phi^{\rm new}_{i,j,k} = \omega\phi^{*}_{i,j,k} + (1-\omega)\phi^{\rm old}_{i,j,k}, \tag{4.21}$$

where

$$\phi^*_{i,j,k} = \frac{1}{6}\left[\phi^{\rm old}_{i+1,j,k} + \phi^{\rm old}_{i-1,j,k} + \phi^{\rm old}_{i,j+1,k} + \phi^{\rm old}_{i,j-1,k} + \phi^{\rm old}_{i,j,k+1} + \phi^{\rm old}_{i,j,k-1} - (\Delta x_\ell)^2 f_{i,j,k}\right], \tag{4.22}$$

being $f_{i,j,k}$ the value of Poisson's equation source, i.e., $f_{i,j,k} = 4\pi G a^2 \rho_B \delta_{T,i,j,k}$ (see Eq. 3.7), and $1 < \omega < 2$ is the overrelaxation parameter, which is set according to the Chebyshev acceleration procedure (see Press et al. 1992). Having the potential at the previous timestep as a guess for the solution, this process usually converges in a few iterations. Nevertheless, in order to save computing time, in MASCLET, Poisson's equation at a given level ℓ of the AMR hierarchy is only solved when level ℓ catches up with level $\ell - 1$. Otherwise, the forces are linearly extrapolated in time. In this way, the computational cost is cut by half, and the boundary conditions can be imposed trivially, since the patches at level ℓ and their parent patches are at the same time. Shall the boundary conditions need to be interpolated from an even coarser level, $\ell' < \ell - 1$, then linear interpolation in space and time is used.

4.3 Other ingredients

Besides gravity and adiabatic hydrodynamics, MASCLET includes several additional physics modules, such as cooling, star formation and a basic chemical enrichment model, supernova feedback, AGN feedback, etc. These, together with a glimpse to the MHD version of MASCLET, are briefly described in the following pages.

4.3.1 Gas cooling

MASCLET implements cooling as an additional source term to be added to the energy evolution equation (Eq. 3.10). The cooling function is tabulated as a function of gas temperature, density, and metallicity, and linearly interpolated in logarithmic space. The abundances of the different chemical species (H, HI, He, HeI, HeII, as well as e^-) are obtained by iteratively solving the equations of collisional ionisation equilibrium for a given temperature, following Katz, Weinberg, and Hernquist (1996). From these values, the cooling terms due to bremsstrahlung and Compton are obtained using the tables of Sutherland and Dopita (1993). These abundances are also used to compute the heating rates

due to the UV background, following Haardt and Madau (1996) and Theuns et al. (1998).

When high resolutions are involved, the cooling term can be rather strong and, in particular, the cooling time, $\tau_{\rm cool} = \rho u/\Lambda$ of a cell can be several orders of magnitude below the Courant time and the rest of timescales. To prevent numerical problems with such stiff source terms, the cooling rate Λ is always limited superiorly by $\rho u/\Delta t_\ell$. While other possibilities exist to deal with these terms (e.g., implicit methods or subcycling; Hernquist and Katz 1989, Zhu, Smith, and Hernquist 2017), in this way an excessive cooling that would produce unphysical results is prevented without strong alterations of the time-stepping scheme.

4.3.2 Star formation, chemical enrichment and supernova (SN) feedback

In MASCLET, star formation, as well as all its associated feedback mechanisms, is performed only each global timestep. For a cell to be candidate to form stars, it must fulfil several conditions, including an upper threshold on temperature ($\sim 2 \times 10^4\,{\rm K}$), a metallicity-dependent lower threshold on gas density (with typical values around $\sim 1\,m_p\,{\rm cm}^{-3}$; Schaye 2004), a cooling time shorter than its dynamical time ($\tau_{\rm cool} < \tau_{\rm dyn}$), negative velocity divergence ($\nabla \cdot \mathbf{v} < 0$), or being Jeans-unstable ($m_{\rm cell} > M_{\rm Jeans} = c_s^3/\sqrt{G^3\rho}$, with c_s the local sound speed).

From this cell, a maximum fraction of the gas mass ($m_{\rm SFR} \sim 0.25 m_{\rm gas}$) is allowed to convert into stars. The real amount of stellar mass that is formed is determined by a star formation efficiency, $\epsilon_{\rm SFR} \sim 0.02$, which sets the typical fraction of gas mass that is able to form stars per dynamical time $\tau_{\rm dyn}$. Since the typical mass of a stellar particle (m_s, usually similar to the smallest DM mass in the simulation) is typically much smaller than $m_{\rm SFR}$, many stellar particles would be generated if all the gas mass was to be converted to stars. Hence, the efficiency is implemented stochastically, by generating $\lfloor m_{\rm SFR}/m_s \rfloor$ random numbers, $r \in [0,1]$, and only accepting the particle if $r < \epsilon_{\rm SFR}\Delta t/\tau_{\rm dyn}$. Spawned stellar particles are then associated a random position within the cell, and their initial velocity gets interpolated from the underlying gas velocity field.

For a given time, $\Delta t_{\rm active} \sim 10$ Myr, stellar particles are considered 'active', in the sense that they produce feedback in the form of mass and energy release. In particular, stars progressively yield a fraction, $\beta \sim 0.1$, of their mass back to the gaseous phase, with a fixed metal yield (~ 0.02). Thermal feedback from type-II SNe is distributed in a cloud around the cell containing the particle, and

its magnitude is set assuming a 10^{51} erg release per SN event and a Salpeter (1955) initial mass function (see Dalla Vecchia and Schaye 2012), together with some efficiency $\epsilon_{\rm SN} \sim 0.1$.

All values quoted here with a '$\sim$' symbol are just to be taken as informative guesses, while the actual values used in galaxy-formation simulations need to be carefully tuned to match several observables, such as the cosmic star formation rate history or the stellar mass function, amongst others (see, e.g., Oh et al. 2020).

4.3.3 Active galactic nuclei (AGN) feedback

While the simulations analysed within this Thesis do not include AGN feedback, it is succinctly described here for completeness. Similarly to star formation and SN feedback, BHs in MASCLET are only formed in global timesteps. BHs are collisionless particles that form out of gas mass in regions of high stellar and gaseous density ($\rho_* \gtrsim 1\, m_p\, {\rm cm}^{-3}$, $\rho_{\rm gas} \gtrsim 10\, m_p\, {\rm cm}^{-3}$), which are additionally required to have a minimum velocity dispersion ($\sim 100\, {\rm km\, s}^{-1}$) to restrict BH seeding to gravitationally-bound objects. This mechanism overcomes the necessity of performing an on-the-fly halo finding, as it is customarily done in many codes (e.g., Springel, Di Matteo, and Hernquist 2005, Sijacki et al. 2007). In MASCLET, BHs are typically seeded with a fix mass, $\sim 10^5\, {\rm M}_\odot$.

Subsequently, BHs can accrete gas from their surroundings, with a Bondi-Hoyle accretion rate capped by the Eddington accretion rate (as described in Sec. 3.1.2.1). The implementation of thermal feedback from AGN, as well as additional interactions such as the drag force experimented by the BH particle, fall beyond the scope of this revision, but generally follow the ideas of Booth and Schaye (2009), Teyssier et al. (2011), and Chapon, Mayer, and Teyssier (2013).

4.3.4 Magnetohydrodynamics

Rather than being a relatively simple addition on top of the basic solvers, as in the case of the previous feedback mechanisms, the incorporation of magnetic fields to the basic hydrodynamic system, even in the context of *ideal magnetohydrodynamics* (MHD), as done in MASCLET, implies significant changes on top of the basic scheme presented in Secs. 4.1.1 and 4.2.3. The magnetic field, $\mathbf{B}$, gets introduced within the fluxes and sources of the system of conservation laws (Eq. 3.43), and evolves according to the induction equation, which is stated in terms of the comoving magnetic field, $\mathbf{B}' \equiv a^2\mathbf{B}$, as

$$\frac{\partial \mathbf{B}'}{\partial t} - \frac{1}{a}\nabla \times (\mathbf{v} \times \mathbf{B}') = \mathbf{0}. \tag{4.23}$$

This needs to be complemented with a method to ensure the divergence-free condition of the magnetic field, $\nabla \cdot \mathbf{B}' = 0$, which is not guaranteed by the evolution equation. This is achieved in MASCLET by the divergence-cleaning algorithm of Dedner et al. (2002), which consists in the addition of a new scalar field, ψ, which acts as a Lagrange multiplier on the induction equation enforcing the divergence-free condition. For more details on the particular implementation in MASCLET, see the original paper by Quilis, Martí, and Planelles (2020).

4.3.5 Tracer particles

During the course of this PhD Thesis, *ghost* or *tracer* particles have been added as a new feature of MASCLET, to overcome the inherent limitation of Eulerian methods to trace a specific Lagrangian fluid element. This can be especially useful in the study of the formation history of different objects. In their current implementation, ghost particles are seeded at a given redshift, $z_{\rm ini}^{\rm ghost}$, by randomly sampling the gas mass distribution. These particles get then advected with the gas velocity field on global timesteps. The limitation of performing this advection on global timesteps is imposed for the sake of keeping the modularity of this feature, in such a way that tracer particles do not need to be intertwined with the more complex AMR machinery. This implies that the integration of the tracer particles' trajectories cannot use the information of arbitrarily fine AMR levels, since this procedure would be numerically unstable. Instead, a maximum level of refinement (typically $\ell = 2$ or 3) can be used for interpolating the velocity field on tracer particles, which is usually enough to study accretion flows onto galaxy clusters, but may not suffice to extend this study to, e.g., galaxies.

4.4 Analysis codes

Alongside the development of MASCLET, a series of analysis codes have been developed to post-process the simulation outputs. Many of these codes have been initially envisioned to be directly coupled to the outputs of MASCLET. As a part of this PhD Thesis, some of them have been evolved into fully-standalone, public tools that can be applied on the outcomes of any cosmological simulation (see Sec. 6, as well as Appendices A2 and A5). Amongst the codes either developed, optimised or extensively-used during the course of this

PhD Thesis, are a spherical-overdensity DM halo and galaxy finder (ASOHF; Planelles and Quilis 2010, Vallés-Pérez, Planelles, and Quilis 2022), a shock finder (Planelles and Quilis 2013), two cosmic void finders (Ricciardelli, Quilis, and Planelles 2013, Vallés-Pérez, Quilis, and Planelles 2021), and a code to perform a Helmholtz-Hodge and Reynolds decomposition on cosmic velocity fields (VORTEX, Vallés-Pérez, Planelles, and Quilis 2021b, Vallés-Pérez, Planelles, and Quilis 2021a).

CHAPTER 5

Non-linear and non-thermal phenomena in cosmological flows

Contents

Part of the results presented within this PhD Thesis (in particular, Sec. 8) are related to phenomena associated to the non-linearity of the hydrodynamic equations. Here, a succinct description of the most salient non-linear phenomena arising within the study of cosmic flows is presented. These include turbulence (Sec. 5.1) and shock waves (Sec. 5.2). In Sec. 5.3, the relevance of these phenomena on galaxy cluster formation is described, while some of their observational implications are discussed in Sec. 5.4.

5.1 Turbulence

Turbulence is present in astrophysical flows on many different scales, ranging from the interior of stars to galaxy clusters (see, e.g., Brandenburg and Nordlund 2011 for a general review on astrophysical turbulence). Regarding galaxy clusters,

turbulent motions can be driven by different processes, including galaxy cluster mergers (or, generally, intense gas accretion), galactic motions through the ICM, AGN feedback, or the sloshing of a cool core, amongst others (Brüggen and Vazza 2015, for a review).

Turbulence is customarily defined, in a loose manner, as a property of fluid flows being highly irregular, both in space and time (e.g., Landau and Lifshitz 1987; even sometimes dubbed as *chaotic* and studied within the context of dynamical systems' theory). Two fundamental qualitative aspects of hydrodynamical turbulence, as described by Spiegel (2011), are: (i) there is a continuous interchange of energy between various states of motion, to be understood as an energy transfer between motions or *modes* at different scales; and (ii) viscosity plays a central role in the dynamics of turbulence, by being the only mechanism available for the fluid to dissipate the injected energy.

Given the importance of the energy transfer between different scales, it is customary to describe turbulent velocity fields in Fourier space. Using a discrete representation for the Fourier transforms for ease of the notation (the same expressions can be stated in terms of integrals with the proper normalisation), the Fourier components of the velocity are introduced as

$$\mathbf{v}(\mathbf{x}) = \sum_{\mathbf{k}} \mathbf{v}_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{5.1}$$

where $\mathbf{k}$ is the wavevector. Averaging over a given volume V, the mean specific kinetic energy can be written in Fourier space as

$$\frac{1}{2}\overline{\mathbf{v}^2} = \frac{1}{2}\sum_{\mathbf{k}} |\mathbf{v}_{\mathbf{k}}|^2 \longrightarrow \frac{1}{2}\frac{V}{(2\pi)^3}\iiint |\mathbf{v}_{\mathbf{k}}|^2 \mathrm{d}^3\mathbf{k}, \tag{5.2}$$

where the first equality corresponds to Parseval's identity, and the last one represents to the continuum limit of the sum. From here, the turbulence spectrum is defined as the specific kinetic energy per unit volume in Fourier space (except for a factor of 1/2), i.e.,

$$P(\mathbf{k}) = \frac{V}{(2\pi)^3}|\mathbf{v}_{\mathbf{k}}|^2. \tag{5.3}$$

The three-dimensional power spectrum, $P(\mathbf{k})$, especially in the case of isotropic turbulence, is customarily simplified in the one-dimensional power spectrum, $P(k)$,

$$P(k) = \frac{V}{(2\pi)^3}\langle|\mathbf{v}_{\mathbf{k}}|^2\rangle_{\hat{\mathbf{k}}} = \langle P(\mathbf{k})\rangle_{\hat{\mathbf{k}}}, \tag{5.4}$$

where $\langle\cdot\rangle_{\hat{\mathbf{k}}}$ denotes averaging over the directions of $\mathbf{k}$, i.e., over the wavevectors such that $|\mathbf{k}| = k$.

It is also frequent to introduce, in the case of isotropic turbulence, the one-dimensional spectrum, often referred to as the *energy spectrum*, defined as

$$E(k) = 2\pi k^2 P(k), \tag{5.5}$$

implying that the mean specific kinetic energy can be written as

$$\frac{1}{2}\overline{\mathbf{v}^2} = \int_0^\infty E(k)\mathrm{d}k. \tag{5.6}$$

Hence, the energy spectrum, $E(k)$, measures the specific kinetic energy in the fluid around a scale $l = 2\pi/k$, per unit k. That is to say, $E(k)\mathrm{d}k$ is the specific kinetic energy contained in scales $[k, k + \mathrm{d}k]$.

While turbulence in real astrophysical scenarios, and in particular in galaxy clusters, can be much more complex, it is still useful to introduce a baseline model due to Kolmogorov (1941) for the description of fully-developed, isotropic, solenoidal turbulence. In an intuitive description, the model assumes that turbulent motions span a wide range in scales, from a *macroscale*, where energy is injected into the flow in the form of bulk motions, to a *microscale*, where kinetic energy is dissipated by viscous forces. The non-linear interactions between different scales are responsible for passing the energy from larger to smaller scales, as stated in the poem by Richardson (1922):

> *Big whirls have little whirls*
> *That feed on their velocity,*
> *And little whirls have lesser whirls*
> *And so on to viscosity.*
>
> – L. F. Richardson

Mathematically, the basic predictions of Kolmogorov's theory stem from two hypotheses, or *principles of similarity*:

(i) There is a range of scales, called the *inertial range*, between the macroscale and the microscale, where the spectrum (or any other related quantity) does not depend on the boundary conditions of the fluid, but only on the rate at which (specific) energy is injected into the flow, ε, and on the viscosity of the fluid, ν.

(ii) In the inviscid limit, $\nu \to 0$, these spectra and associated quantities depend exclusively on ε.

Exploiting the dimensional character of $[\varepsilon] = L^2T^{-3}$ and $[\nu] = L^2T^{-1}$, where L and T are the units of length and time, respectively, it is possible to construct length and velocity scales that depend only on these quantities. The viscous length is defined as

$$\eta = \left(\frac{\nu^3}{\varepsilon}\right)^{1/4}, \tag{5.7}$$

which indicates the scale below which viscous dissipation becomes dominant. A quantity with units of velocity would be constructed as

$$V = (\nu\varepsilon)^{1/4}. \tag{5.8}$$

From this, given that the dimensions of $E(k)$ are $L^3T^{-2} = V^2L$, the energy spectrum must be of the form

$$E(k) = \nu^{5/4}\varepsilon^{1/4} f\left[k\left(\frac{\nu^3}{\varepsilon}\right)^{1/4}\right], \tag{5.9}$$

where f is a function of the dimensionless quantity $k\eta$. Using hypothesis (ii), by which $E(k)$ is independent of ν, and assuming a power-law behaviour for f, it follows that $f \propto \nu^{-5/4}$ and, therefore, $f(x) \propto x^{-5/3}$. Hence, the Kolmogorov spectrum must be given by:

$$E(k) = C\varepsilon^{2/3}k^{-5/3}, \tag{5.10}$$

where C is a constant. This spectrum is valid for $k\eta \ll 1$, i.e., for scales much larger than the viscous scale, but still below the macroscale, where the velocity field retains the imprint of the boundary conditions and the particularities of the energy injection. The energy spectrum is thus a power-law with a slope of $-5/3$ within the inertial range.

From the scaling properties of the energy spectrum, it is possible to derive a scaling for the magnitude of solenoidal velocity fluctuations in the inertial range. By inspection of Eq. (5.6), it follows that the characteristic velocity fluctuations on a scale $l \equiv 2\pi/k$, δv_l, scale as $\delta v_l^2 \propto \int E(k)\mathrm{d}k \propto \int k^{-5/3}\mathrm{d}k \propto k^{-2/3}$, and therefore $\delta v_l \propto l^{1/3}$. It is also frequent to use the velocity structure functions of order p, defined as

$$S_p(l) = \langle|\mathbf{v}(\mathbf{x} + l\hat{\mathbf{n}}) - \mathbf{v}(\mathbf{x})|^p\rangle_{\mathbf{x},\hat{\mathbf{n}}}, \tag{5.11}$$

which contain equivalent information to the energy spectrum, and where the averaging is performed over positions $\mathbf{x}$ and directions $\hat{\mathbf{n}}$. Within the Kolmogorov theory, structure functions of order p scale as $S_p(l) \propto l^{p/3}$.

5.2 Shock waves

Cosmological flows originated by the gravitational collapse of structures often enter the supersonic regime, where the speed of the material infalling into the potential well of the structure exceeds the local speed of sound. Hydrodynamic flows are fundamentally different depending on whether they are subsonic or supersonic. One of the most distinctive features of motions in the supersonic regime is the development of shock waves, that arise as surfaces of discontinuity for various quantities of the fluid (e.g., density, pressure, velocity) that can propagate around the medium (Landau and Lifshitz 1987). Perhaps the best example of these phenomena in the context of cosmological structure formation are the large-scale accretion shocks that surround galaxy clusters (Quilis et al. 1998, Miniati et al. 2000, Ryu et al. 2003). For a review on the physics of the cosmological plasma in shocks, see Bykov, Dolag, and Durret (2008).

Some quantities must be conserved through the surfaces of discontinuity associated to shocks. These can be directly derived from the equations of hydrodynamics (Eqs. 3.8-3.10), which, in the non-cosmological case, without sources, and in one spatial dimension for the sake of simplicity, can be written as:

$$\partial_t \rho + \partial_x (\rho v) = 0, \tag{5.12}$$

$$\partial_t (\rho v) + \partial_x (\rho v^2 + P) = 0, \tag{5.13}$$

$$\partial_t E + \partial_x \left[(E + P) v \right] = 0, \tag{5.14}$$

where $\partial_t \equiv \frac{\partial}{\partial t}$ and $\partial_x \equiv \frac{\partial}{\partial x}$. Assuming the discontinuity to be located at $x = 0$ and using the rest frame of the shock surface, one can integrate the equations from $x = -\epsilon$ to $x = +\epsilon$, where $\epsilon \to 0$. This yields the *Rankine-Hugoniot* jump conditions,

$$\rho_1 v_1 = \rho_2 v_2, \tag{5.15}$$

$$\rho_1 v_1^2 + P_1 = \rho_2 v_2^2 + P_2, \tag{5.16}$$

$$\frac{1}{2}\rho_1 v_1^2 + \rho_1 u_1 + P_1 = \frac{1}{2}\rho_2 v_2^2 + \rho_2 u_2 + P_2, \tag{5.17}$$

which relate the properties in the pre-shock medium (upstream of the shock surface) and the post-shock medium (downstream of it). Notating '1' for the pre-shock medium and '2' for the post-shock medium ($v_1 > v_2$, $\rho_1 < \rho_2$), using the equation of state (Eq. 3.11) in terms of the adiabatic coefficient γ, and defining the shock Mach number as the quotient between the pre-shock velocity and the pre-shock speed of sound, $\mathcal{M} \equiv v_1/c_1$, the jump conditions can be written as (Landau and Lifshitz 1987):

$$\frac{\rho_2}{\rho_1} = \frac{v_1}{v_2} = \frac{(\gamma+1)\mathcal{M}^2}{(\gamma-1)\mathcal{M}^2+2}, \tag{5.18}$$

$$\frac{P_2}{P_1} = \frac{2\gamma\mathcal{M}^2 - (\gamma-1)}{\gamma+1}, \tag{5.19}$$

$$\frac{T_2}{T_1} = \frac{\left[(\gamma-1)\mathcal{M}^2+2\right]\left[2\gamma\mathcal{M}^2-(\gamma-1)\right]}{(\gamma+1)^2\mathcal{M}^2}. \tag{5.20}$$

The jump conditions are thus fully determined by the Mach number of the shock. The case where $\mathcal{M} \gg 1$ is usually called a *strong shock*, as it is the case of the outermost accretion shocks of galaxy clusters, where the pre-shock velocity is high and the sound speed is rather small, since the gas is still cold. In this limit, the jumps in pressure and temperature can be arbitrarily high ($T_2/T_1 \propto \mathcal{M}^2$, $P_2/P_1 \propto \mathcal{M}^2$), while the density jump saturates to $\rho_2/\rho_1 \to (\gamma+1)/(\gamma-1)$.

5.3 The role of turbulence and shock waves in structure formation

Turbulence and shocks appear as a natural consequence of the non-linear collapse of cosmic structures, and, in particular, of galaxy clusters. As the initially cold baryonic matter falls and collapses into walls, filaments and clusters, its supersonic motions induce the generation of shock waves (Ryu and Kang 1997, Quilis et al. 1998). These shocks are one of the main actors shaping the thermal history of the ICM, since they provide an essential mechanism for heating the

gas to the ICM-like temperatures around 10^7 K, from the dissipation of the kinetic energy involved in the collapse.

As studied by, e.g., Miniati et al. (2000), Miniati (2002), and Ryu et al. (2003), cosmic shocks associated to the collapse of galaxy clusters can be broadly subdivided in two categories: external, accretion shocks, which surround the cluster at distances of $\gtrsim 2R_{\rm vir}$ and emerge in regions of low density and temperature, with high Mach numbers ($\mathcal{M} \sim 50 - 1000$); and internal shocks associated to mergers or to gas flows of previously-shocked gas, with much smaller Mach numbers ($\mathcal{M} \sim 2 - 10$). The aforementioned works found that, despite their modest intensity, most of the energy dissipation in the ICM is due to these weaker, internal shocks, relegating the accretion shocks to a more subdominant role in energetic terms. Further works involving high-resolution simulations, both using HRSC and SPH methods, have confirmed these results and explored in detail several aspects of the role of shocks in the evolution of the thermodynamics of the ICM (e.g., Vazza et al. 2010, Planelles and Quilis 2013, Planelles et al. 2021).

Besides their thermodynamic role, cosmic shocks (especially the internal ones) are also able to convert a significant fraction of kinetic energy, due to incomplete plasma thermalisation, in cosmic-ray energy, i.e., in the acceleration of particles (generated at the accretion shocks, Kang, Ryu, and Jones 1996; or either ejected by radio galaxies, e.g. Vazza et al. 2023) to relativistic energies (Blandford and Ostriker 1978, Vazza et al. 2015). These mechanisms are, in turn, associated with the presence of diffuse radio emission, as discussed in Sec. 5.4.

Regarding turbulence, many different mechanisms stir the ICM and generate motions that cascade down to smaller scales in a Kolmogorov-like picture. Amongst the most important ones are those associated to the hierarchical nature of structure formation, such as galaxy cluster mergers. Extensive work with numerical simulations has shown how clusters of similar mass but in very distinct dynamical states, have extremely different fractions of their volumes hosting significant turbulent motions (Subramanian, Shukurov, and Haugen 2006, Iapichino and Niemeyer 2008, Vazza et al. 2011). Additionally, core sloshing (Markevitch, Vikhlinin, and Mazzotta 2001), galaxy motions (Faltenbacher et al. 2005, Ruszkowski and Oh 2011), or AGN outflows (Quilis, Bower, and Balogh 2001, Vazza, Roediger, and Brüggen 2012, Gaspari 2015) have been also studied in the context of turbulence generation in galaxy clusters. Furthermore, shocks themselves can also be responsible for generating small-scale solenoidal motions (Vazza et al. 2017, Wittor et al. 2017).

Once generated, turbulent motions are significant actors in the evolution of clusters. On the one hand, they contribute significantly to support the ICM against gravity, acting as a non-thermal contribution to pressure (Dolag et al. 2005, Subramanian, Shukurov, and Haugen 2006, Biffi et al. 2016, Vazza et al. 2018b). Turbulence is also able to accelerate cosmic rays in extended regions through the second-order Fermi process (Fujita, Takizawa, and Sarazin 2003, Cassano and Brunetti 2005), and to amplify magnetic fields through the small-scale dynamo mechanism (Porter, Jones, and Ryu 2015, Beresnyak and Miniati 2016, Vazza et al. 2018a).

5.4 Observational implications

Besides the interest they trigger on a more fundamental level, shocks and turbulence are intimately connected to observational properties of galaxy groups and clusters and, in particular (although not exclusively), the properties of their non-thermal emission. In the following, a brief description of some of these observational implications is presented, without the aim of being exhaustive. Some reviews that cover extensively the topics associated to non-thermal phenomena and their implications for X-ray observations, as well as radio emission in galaxy clusters, are Rephaeli et al. (2008) and Ferrari et al. (2008), respectively.

Turbulence has a direct impact on the X-ray emission from galaxy clusters, through the broadening of spectral lines (e.g., Churazov et al. 2008, Sanders et al. 2010, Sanders, Fabian, and Smith 2011; see also §3.5 of Böhringer and Werner 2010 for a review). Nevertheless, direct measurements of turbulent velocities are still rare. Before its unexpected early demise, the Hitomi satellite was able to measure velocity fluctuations of $\sim 200\,\mathrm{km\,s^{-1}}$ on scales of $\sim 60\,\mathrm{kpc}$ in the central region of the Perseus cluster (ZuHone et al. 2018). Planned X-ray missions, such as Athena, could importantly enhance the current prospects of resolving turbulent motions in clusters (Roncarelli et al. 2018).

Indirect evidence of the existence of turbulent motions can be obtained from the analysis of X-ray brightness fluctuations, associated to density fluctuations introduced by the stirring of the ICM by (compressive) turbulent motions (Schuecker et al. 2004, Churazov et al. 2012, Gaspari et al. 2014, Zhuravleva et al. 2014). This interpretation is further backed by the spatial correlation between these brightness fluctuations and the loci of diffuse radio emission (Eckert et al. 2017, Bonafede et al. 2018). Nevertheless, recent simulation work has shown how the relation between density and velocity fluctuations is not a straightforward

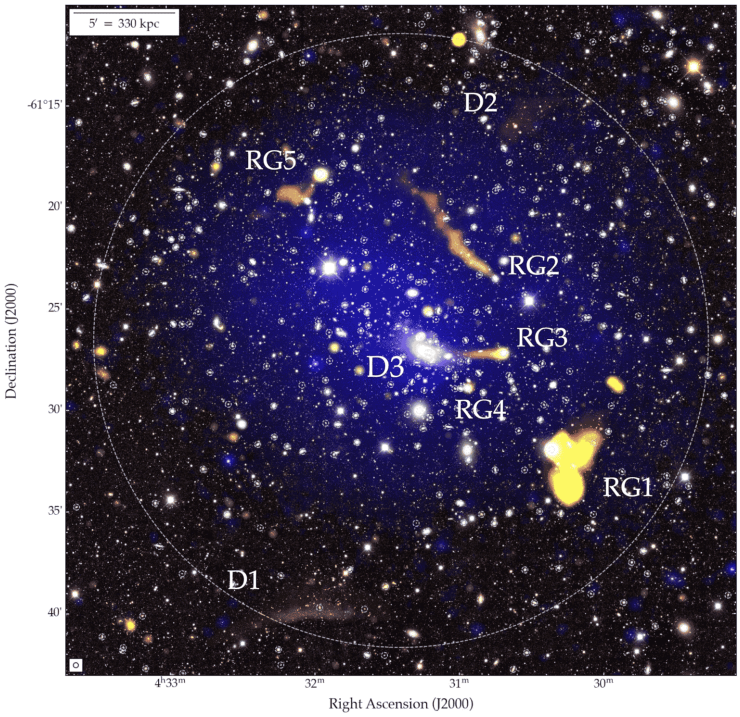


Figure 5.1: Composite image of Abell 3266: radio (ASKAP 943 MHz, red channel; and ATCA 2.1 GHz, green channel) + X-ray (XMM-Newton 0.5-2 keV, blue channel), overlaying an optical RGB image (i, r, and g bands from DES DR2). The dashed silver cicle marks a radius of 1 Mpc from the cluster centre. RG labels indicate active radio galaxies, while D labels indicate diffuse sources. For instance, D1 is a (wrong-way) radio relic. Below the discrete sources, there is a low-surface brightness radio halo. Figure reproduced from Riseley et al. (2022) with permission.

one, and it is affected by significant scatter that could hinder the possibility of measuring turbulent velocities by this method (Simonte et al. 2022).

Connected to turbulence and at the very foundation of X-ray cluster cosmology, the fact that turbulent motions act as a non-thermal pressure component against the gravitational collapse of the ICM biases the measurement of masses by the hydrostatic equilibrium assumption (Biffi et al. 2016, Vazza et al. 2018b, Angelinelli et al. 2020, Bennett and Sijacki 2022).

Radio observations of galaxy clusters have also provided compelling evidence for the presence of significant turbulent motions. For instance, the polarisation angle of the synchrotron emission of cluster galaxies gets changed by the presence of magnetic fields, in a very well-known effect known as Faraday rotation. The study of this Faraday rotation reveals the presence of small-scale, $\sim 1\,\mu$G, tangled

magnetic fields, suggesting the existence of turbulent motions at these scales (Vogt and Enßlin 2005, Vacca et al. 2010, Bonafede et al. 2010).

Closely related to the processes of acceleration of cosmic rays by shocks and turbulence mentioned in Sec. 5.3, the presence of diffuse radio emission is a clear signature of their existence and their location within in galaxy clusters. While the first detections of extended radio emission from galaxy clusters date back to Large, Mathewson, and Haslam (1959) and Willson (1970) observations of the Coma cluster, the advances in sensitivity of subsequent radio facilities have allowed to unveil a rich diversity of extended radio emission in clusters. For thorough reviews, the reader is addressed to Ferrari et al. (2008) and Feretti et al. (2012).

Very broadly, the two most salient classes are *radio haloes* and *radio relics*, which are exemplified in a composite (radio+X-ray) image of Abell 3266 in Fig. 5.1. While both classes share several properties, such as the nature of the emission and a steep radio spectrum, they mainly differ on their morphology and location:

- Radio haloes are Mpc-scale sources of diffuse radio emission, located towards the centre of galaxy clusters. They show no obvious connection with a galaxy, but instead appear to display a rather regular morphology, usually similar to that of the X-ray emission. They appear to be located at the centres of merging clusters, although not all disturbed systems exhibit a radio halo.

- Radio relics, on the other hand, are elongated ($\sim 1\,\mathrm{Mpc}$ long), arc-like sources of radio emission, located at the outskirts of their hosts, and with a clear connection to the presence of a merger shock. Normally, their major axis is perpendicular to the direction of the cluster centre.

The understanding of the formation mechanisms for both these classes of extended radio emission is still not fully closed and falls beyond the scope of this short review, but generally involves the acceleration of relativistic particles by either shocks, turbulence or a combination of both.

Part II

Results

CHAPTER 6

Numerical tools

Contents

During the course of this PhD Thesis, a significant amount of effort has been devoted to either the development of new analysis tools for cosmological simulations, or the renewal of existing ones. In all cases, these tools have been made publicly available. A substantial focus has also been placed on the generalisation of these tools so that they can be applied, not only to AMR simulations, but also to the outcomes of other simulation codes, including particle-based simulations. This is reported in the subsections of each heading.

6.1 A Helmholtz-Hodge decomposition algorithm for cosmological simulations

This Section is based on Vallés-Pérez, Planelles, and Quilis (2021b)[1], and the full paper can be found in Appendix A2. The implementation of the algorithm presented in this publication is publicly available.[2]

Overview

Context. The study of turbulent velocity fields (among many other disciplines where it may also be relevant) often requires the decomposition of a vector field in its compressive (curl-free) and its solenoidal (divergence-free) components. This is a well-known mathematical problem, known as the Helmholtz-Hodge decomposition (HHD). The HHD, provided suitable boundary conditions are given, can be performed in a reasonably simple way when data is uniformly sampled. However, in the context of non-uniform grids such as the ones used through the simulations in this Thesis, the HHD becomes a more complex problem.

Aims. The development of a computationally-efficient algorithm for performing an HHD on velocity fields defined on a block-based AMR set of grids.

Methods. The method is based on the formulation of the HHD as a set of elliptic PDEs, formally equivalent to the Poisson equation. The algorithm uses the standard machinery for solving these equations (the same as described for MASCLET in Secs. 4.1.3 and 4.2.5), i.e., a combination of FFT for the base, periodic grid and iterative solvers for the refinement patches. The code, named `vortex`, has been implemented in Fortran and parallelised according to the OpenMP (OMP) standard directives.

Results. The code is tested against a set of idealised and more complex test cases. The decomposition is performed with typical errors below 1% at the 95-percentile in the idealised tests, and around $\sim 1\%$ median errors in the complex tests where velocity fields including fluctuations on almost 3 orders of magnitude in scale are considered.

[1] D. Vallés-Pérez, S. Planelles, and V. Quilis. "Unravelling cosmic velocity flows: a Helmholtz-Hodge decomposition algorithm for cosmological simulations." In: *Computer Physics Communications* 263, 107892 (June 2021). DOI: 10.1016/j.cpc.2021.107892

[2] https://github.com/dvallesp/vortex.

Conclusions. The procedure introduced in this paper provides a reasonably-accurate, computationally-efficient method for performing an HHD, and has been applied to subsequent studies within and outside this PhD Thesis. Furthermore, it can be adapted to other types of simulations, including particle-based ones.

6.1.1 An HHD algorithm for particle-based simulations

The results reported in this section correspond to the work developed during a stay at the Universitäts-Sternwarte München, from May 2023 to July 2023, under the supervision of Prof. Klaus Dolag. The simulations shown here have been carried out by Tirso Marín-Gilabert and Frederick Groth, whom I thank for their feedback during the development of the code. These results constitute the basis for a forthcoming publication, which has already been submitted to the journal Computer Physics Communications.[3]

With the aim of generalising the methods developed within the frame of this PhD Thesis and making them publicly available, the code VORTEX has been adapted to work with particle-based (or moving-mesh) simulations. In this short section, I briefly summarise the main features of the code and show some preliminary results on idealised tests and applications to SPH and meshless finite mass (MFM) simulations with OpenGADGET. The code is publicly available.[4]

6.1.1.1 The algorithm

In the spirit of preserving as much from the original VORTEX code as possible, the algorithm relies on the interpolation of the particle-based velocity field onto an ad-hoc hierarchy of AMR grids. The elliptic equations yielding the potentials, from which the velocity components are subsequently obtained, are solved as described in the original paper (Paper A2). In the end, the results can be output either with the AMR structure, or reinterpolated back to particles. Hence, the main differences with respect to the original code restrict to:

Mesh creation. First, a cubic base grid of dimensions N_x^3 covering the whole domain of interest must be defined. As it is the case of the simulations shown below (Sec. 6.1.1.3), many SPH resimulations of a cluster involve a large domain (e.g., a 1 Gpc box), but only the $\sim$ 20 Mpc around the central cluster are

[3] D. Vallés-Pérez et al. "VORTEX-P: A Helmholtz-Hodge+Reynolds decomposition algorithm for particle-based simulations." In: *Computer Physics Communications* (2024, in press).

[4] https://github.com/dvallesp/vortex-p/.

populated with SPH particles (or mesh-generating points, in the case of MFM; hereafter, I shall refer exclusively to *particles* for simplicity). In this case, the base grid can be defined to cover a side length slightly larger than this, so that the imposition of periodic boundary conditions does not affect the resulting decomposition in the region of interest. A reasonable rule-of-thumb is to set the base grid resolution, N_x, as the closest power of 2 such that $N_x^3 \sim N_{\rm part}$, where $N_{\rm part}$ is the number of particles in the region of interest.

Once the base grid is set, the code needs to define the AMR hierarchy. The process is parallel to the one followed in the new implementation of ASOHF (see Paper A5, its §2.2.1). Essentially, at any refinement level ℓ, all cells comprising more than $n_{\rm part}^{\rm refine}$ particles (a free parameter) will be flagged as refinable. Since SPH quantities at a given position are defined as a kernel-weighted average over the nearest $N_{\rm ngh} \sim 50 - 300$ neighbours, $n_{\rm part}^{\rm refine}$ does not generally need to be set to a very restrictive (small) value. Subsequently, the mesh-creation routine loops over the refinable cells and covers them as efficiently as possible with a set of ortohedral patches with half the cell size. The free parameters for this mesh-creation process are the same ones as in ASOHF and can be found in Table A5.1.

Velocity interpolation. Velocity needs to be smoothed and assigned from the particles to the multi-resolution grid in such a way that it preserves the details in the regions of high particle density, but is sufficiently smooth so as to apply the Helmholtz decomposition theorem. In VORTEX-P, the values at each cell represent cell-centre values, and are obtained at a position $\mathbf{x}$ by an average of the particle individual velocities, $\mathbf{v}_i$, over a length $h(\mathbf{x}) = \max(l_{N_{\rm ngh}}, \Delta x)$, where $l_{N_{\rm ngh}}$ is the distance to the $N_{\rm ngh}$-th nearest neighbour, while Δx is the cell size at the AMR level considered:

$$\mathbf{v}(\mathbf{x}) = \frac{\sum_{i\in{\rm ngh}} \mathbf{v}_i \, W(|\mathbf{x}-\mathbf{x}_i|, h(\mathbf{x}))}{\sum_{i\in{\rm ngh}} W(|\mathbf{x}-\mathbf{x}_i|, h(\mathbf{x}))}. \tag{6.1}$$

Several kernels, $W(\cdot,\cdot)$, have been implemented in the code, including the cubic spline (M_4; see Monaghan and Lattanzio 1985) and Wendland's (1995) C^4 and C^6 kernels. Since this process requires locating $N_{\rm part}$ particles onto $\mathcal{O}(N_{\rm part})$ cells, a naive implementation would have computational complexity scaling as $\mathcal{O}(N_{\rm part}^2)$, prohibitive for practical applications. Instead, VORTEX-P makes use of an implementation of a k-dimensional space-partitioning tree in the library CORETRAN[5], which reduces the computational complexity to $\mathcal{O}(N \log N)$.

[5] https://github.com/leonfoks/coretran.

Since this is a computationally-expensive process, especially in low-density regions, where it might be necessary to walk the tree out to large distances to find the required number of particles, part of the computational cost of the base grid can be alleviated in the case of simulations around a massive object. In this case, it is possible for the code to automatically detect the cube within the original domain containing all particles. Outside this cube, the velocity field is only evaluated in one out of each 4^3 cells, and filled in by interpolation. A second cube can be defined inside the previous one, delimited along each direction by the 1st and 99th percentiles of the list of particle coordinates. Inside the cube, the velocity field is evaluated in all cells, while outside the cube, it can be evaluated each 2^3 cells, and filled in by interpolation. This is a reasonable process, since in these regions of low-particle density, which correspond to a large fraction of the volume, the velocity field is expected to be smooth and is only required to appropriately set the boundary conditions.

Reynolds decomposition. In Paper A3 (see Sec. 8.1), the multi-scale filter of Vazza, Roediger, and Brüggen (2012) and Vazza et al. (2017) was implemented in VORTEX with the aim of decomposing the velocity field into its bulk and turbulent components. According to this procedure, the bulk velocity is computed as the volume-weighted average over a sphere of radius $L(\mathbf{x})$. This filtering length is locally determined by an iterative procedure in which $L(\mathbf{x})$ grows until the turbulent velocity converges (indicating the outer scale of turbulence has been reached), or until a strong shock enters the sphere. This last condition is necessary because the averaging of pre-shock and post-shock velocities would prevent the convergence of the procedure and could introduce spurious measurements of the turbulent velocity. While, in the original implementation, the code is fed with the outputs of a mesh-based shock finder, in VORTEX-P there are two possibilities:

- If the simulation has been run together with an on-the-fly shock-finding scheme (e.g., Beck, Dolag, and Donnert 2016), then this quantity can be mapped onto the grid in the same way as the velocity, and the Reynolds decomposition can be performed as in the original VORTEX code.

- If this information is not available, strong shocks are detected in a computationally-cheap way by establishing a lower threshold on artificial viscosity, and an upper threshold on velocity divergence (only in SPH simulations). While this procedure is not as accurate and introduces two

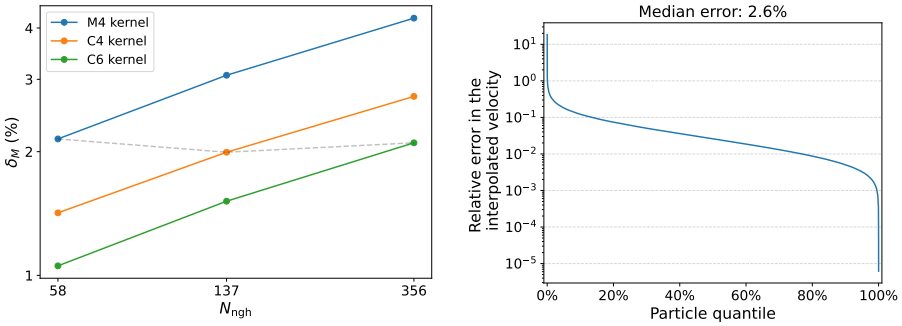


Figure 6.1: *Left-hand side panel:* Relative error on the interpolation of the mass onto the grid, for three kernels (M_4, C^4 and C^6) and three numbers of neighbours. The grey, dashed line joins three values with the same width, σ. *Right-hand side panel:* Anti-cumulative distribution function of the relative error between the original velocity of the SPH particles and the re-interpolated velocity from the grid, for the M_4 kernel with $N_{ngh} = 137$.

free dimensional parameters, it overcomes the necessity of performing a more expensive shock-finding procedure within the code.

6.1.1.2 Tests

By design, the interpolation procedure introduced above is not conservative, i.e., the sum of the interpolated values over the cells does not necessarily equal the sum of the values over the particles, due to the necessity of obtaining a smooth interpolation for cell-centred values (and not the cell-averaged ones). This is not necessarily a problem, since the quantity that is being interpolated is not a conserved quantity itself. However, it is still worth assessing the level of conservativity of such an interpolation. This is shown in the left-hand side panel of Fig. 6.1 for the interpolation of the mass, for different kernels and numbers of neighbours. The result clearly points out that, the lower the width of the kernel (higher-order and smaller number of neighbours), the more conservative the interpolation.

Another fact worth bearing in mind is that, if the results are to be interpolated back to particles, the re-interpolated values will never coincide with the original SPH particle velocities, since the former represent the smoothed velocities at the location of each particle. This is shown, in the form of an anti-cumulative distribution function of the relative errors between the original and the re-interpolated velocities, in the right-hand side panel of Fig. 6.1, for the particular case of the M_4 kernel with $N_{ngh} = 137$, with a median error of $\approx 3.2\%$.

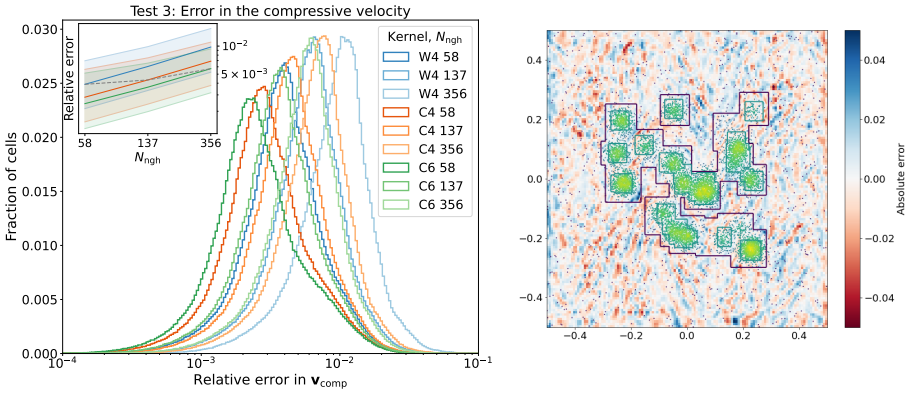


Figure 6.2: Results from Test 3 in Vallés-Pérez, Planelles, and Quilis (2021b) applied to VORTEX-P. *Left-hand side panel:* Distribution of errors in $\mathbf{v}_{\rm comp}$ for several kernels and number of neighbours. The inset shows the medians and the $(16-84)\%$ region of the distributions. *Right-hand side panel:* Thin slice of the absolute error for the case of the M_4 kernel with $N_{\rm ngh}=137$, together with the distribution of particles and the AMR grid (levels 1 and 2, in purple and blue, respectively).

The performance of VORTEX-P has been subsequently validated with the same set of tests introduced in Paper A2. Fig. 6.2 exemplifies this with Test 3 in the original VORTEX paper (see Fig. A2.4 for its results in the AMR version of VORTEX). This test introduces a combination of low spatial frequency, compressive and solenoidal, sinusoidal oscillations in a periodic box, with known decomposition. To perform these tests in a situation where mesh refining is relevant, a mock particle distribution is generated by considering a background distribution of particles in the unit cube, randomly sampled from a uniform distribution ($N_{\rm part}^{\rm bkg}=5\times10^5$ particles), and a number of blobs ($N_{\rm blobs}=100$, for a total of $N_{\rm part}^{\rm blobs}=1.5\times10^6$ particles), each generated by sampling a three-dimensional Gaussian distribution with $\sigma=0.02$ and a random displacement within the domain.

The distribution of errors is shown in the left-hand side panel of Fig. 6.2 as a function of the number of neighbours and the kernel used for the interpolation. The inset summarises the distribution of median and $(16-84)\%$ error percentiles for each setting. As a general trend, the relative errors (in the order of $\sim10^{-3}$) increase with the number of neighbours and with less compact kernels. This is reasonable, since the interpolation does not necessarily preserve the solenoidal/compressive character of the input fields and, hence, the more local the interpolation is, the lesser amount of cross-talk between these two components there will be. The right-hand side panel of Fig. 6.2 shows the error

map for the case of the M_4 kernel with $N_{\rm ngh} = 137$, together with the particle distribution and structure of the AMR grid for reference. No significant errors associated to the AMR boundaries are found.

6.1.1.3 Application

This section intends to show some applications of the VORTEX-P to actual simulation data, without the aim of being exhaustive in their discussion but just to exemplify different contexts where the code is currently being used. In particular, Fig. 6.3 compares graphically the velocity structure of an SPH (left) and an MFM (right) simulation of the same massive galaxy cluster, with the same nominal resolution, at $z = 0$. Looking at the density maps (top row), while SPH seems to resolve more gaseous substructure, MFM captures more sharply the discontinuities associated to shocks (e.g., in the centre of the panel). Subsequent rows compare the total, compressive and solenoidal velocity magnitudes, respectively. The difference between SPH and MFM is more evident when comparing the compressive velocity field, where MFM shows a sharp description of shocks, while in SPH they are smoothed out.

These differences can be quantified through the energy power spectra, $E(k)$ (see its definition in Sec. 5, Eqs. 5.5 and 5.6), which are shown in the top row of Fig. 6.5. The left hand-side panel shows the total, compressive and solenoidal spectra for SPH (solid lines) and MFM (dashed lines). Although the trends are similar, as shown in the right-hand side panel, MFM shows an overall higher normalisation (up to 70% higher) at almost all scales and for all components, but especially at intermediate scale ($k \sim [5 - 10]\, h\mathrm{Mpc}^{-1}$). This implies higher kinetic energy in MFM, perhaps as a consequence of a greatly reduced numerical dissipation.

In Fig. 6.4, a similar comparison is shown for two SPH simulations of the same cluster, without (left) and with (right) physical viscosity. Both simulations have numerical resolution 10× with respect to the ones shown above. The differences are striking even when looking at the density slices, with the simulation with physical viscosity showing a much more complex morphology of clumps and filamentary structures. The third and fourth rows show the compressive and solenoidal components of the turbulent (small-scale) velocity field, as extracted by the multi-scale filter in VORTEX-P. While in the non-viscous simulation shock surfaces are clearly apparent in the compressive turbulent velocity field, when physical viscosity is added, compressive velocities get strongly damped and the shock surfaces are smoothed out. Also, when considering the solenoidal velocity,

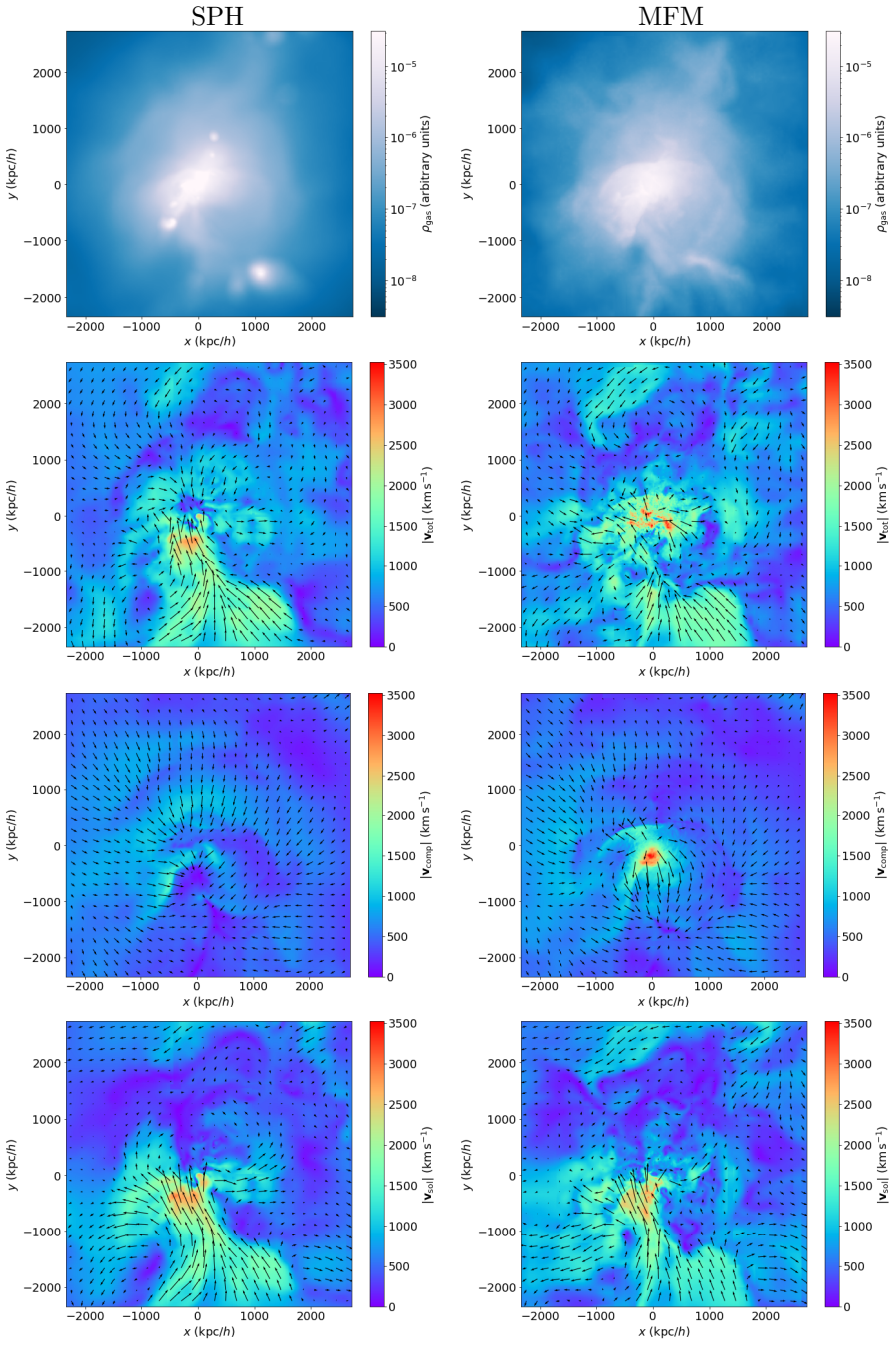


Figure 6.3: Comparison, performed with VORTEX-P, of the density and velocity structure of an SPH (left-hand side column) and an MFM (right-hand side column) simulation of the same cluster, with equivalent resolution, at $z = 0$. From top to bottom, the different rows show thin slices of gas density, total velocity, compressive velocity and solenoidal velocity, respectively. In the velocity maps, the background colour shows the magnitude, while the arrows show the directions and magnitude of the xy projection.

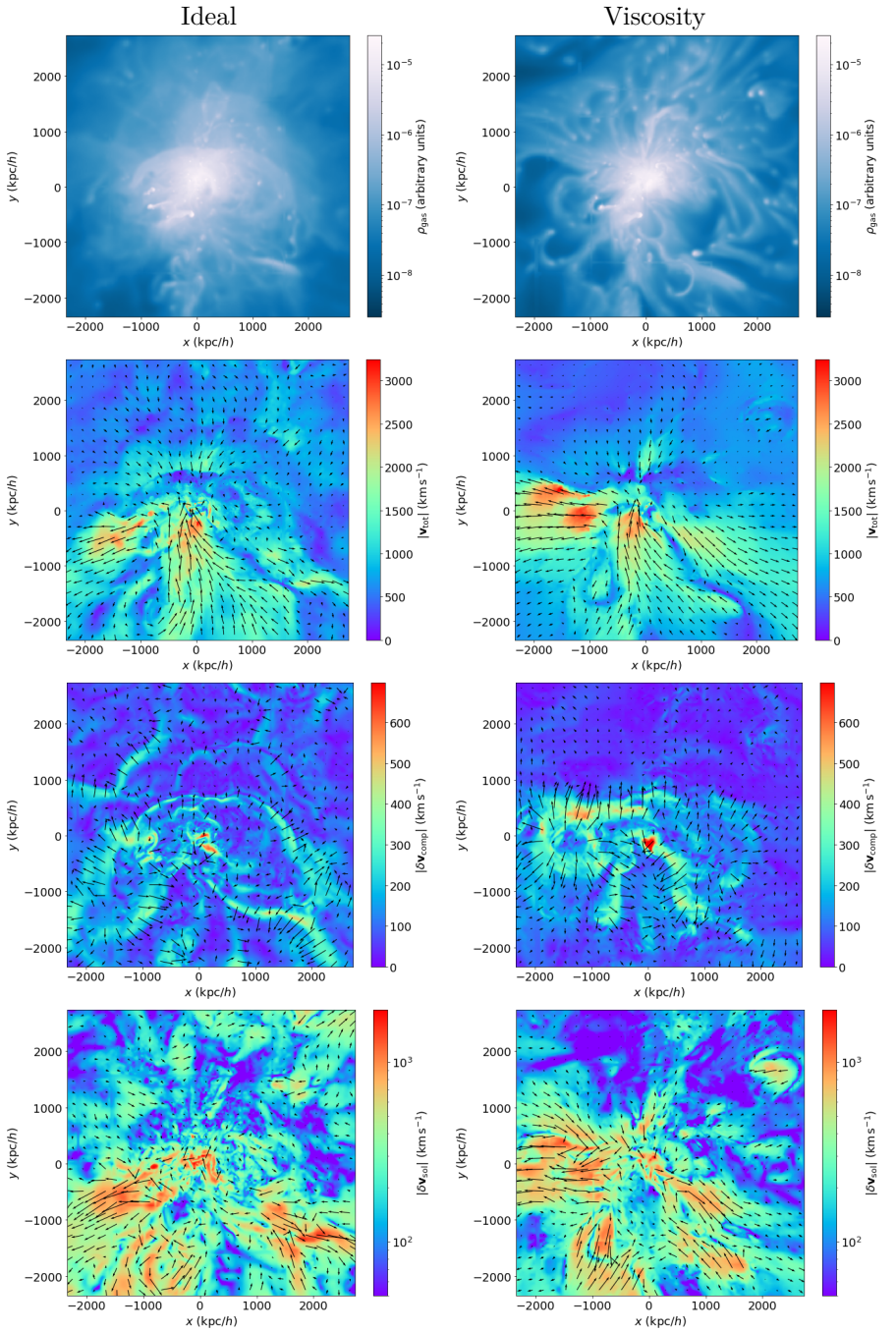


Figure 6.4: Comparison, performed with VORTEX-P, of the density and velocity structure of an SPH simulation without viscosity (left-hand side column) and with viscosity (right-hand side column) of the same cluster in Fig. 6.3, with 10 times the resolution, at $z = 0$. From top to bottom, the different rows show gas density, total velocity, compressive part of the turbulent velocity and solenoidal part of the turbulent velocity, respectively. In the velocity maps, the background colour shows the magnitude, while the arrows show the directions and magnitude of the xy projection.

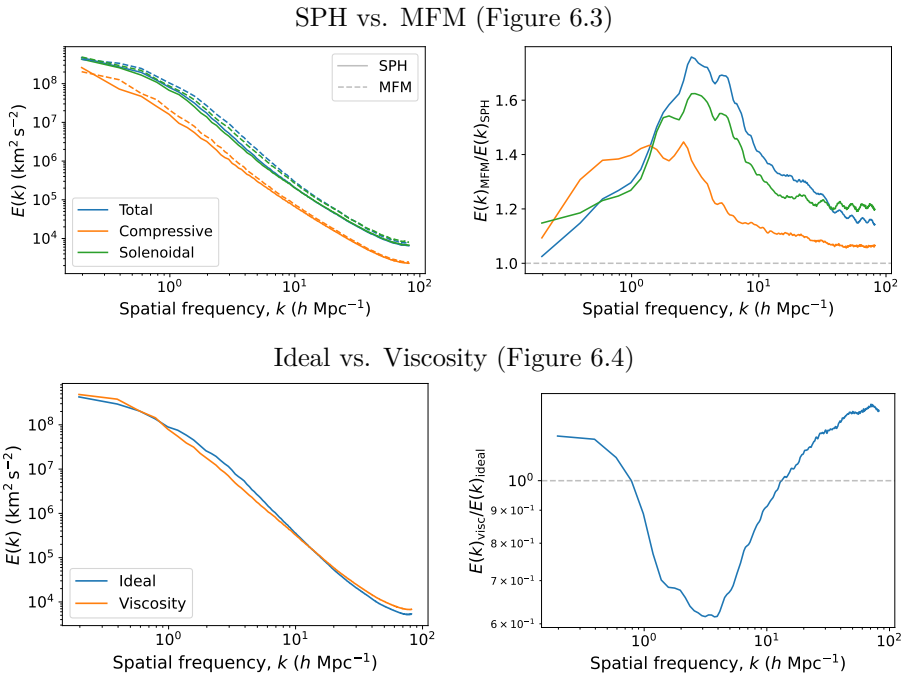


Figure 6.5: *Top panels:* Comparison of the kinetic energy spectra (for the total velocity, in blue; and for the compressive and solenoidal, in orange and green, respectively) for SPH (solid lines) and MFM (dashed lines). For better comparison, the right-hand side panels show the ratio of the MFM to SPH spectra. *Bottom panels:* Same as above, but for the ideal (solid lines) vs. viscosity (dashed lines) comparison, only for the total velocity.

it becomes apparent from the maps that viscosity is suppressing solenoidal turbulence on small scales.

This is shown more quantitatively through the energy spectra in the bottom panels of Fig. 6.5. The left-hand side panel shows the energy spectra of the total velocity for the case without viscosity (blue) and with viscosity (orange). To better highlight the differences, the right-hand side panel shows the ratio between these two spectra. Interestingly, while at medium scales ($1\,h\,\mathrm{Mpc}^{-1} \lesssim k \lesssim 10\,h\,\mathrm{Mpc}^{-1}$) viscosity is suppressing velocity fluctuations by up to $\sim 40\%$, it seems that the viscous case has higher velocity fluctuations on smaller scales ($k \gtrsim 10\,h\,\mathrm{Mpc}^{-1}$). An interpretation for this apparently counter-intuitive result could be that, while viscosity is damping velocity fluctuation on small scales, the fact that this suppresses gas mixing is generating the more extreme density fluctuations seen in the upper panels of Fig. 6.4. In turn, this may be causing the velocity field to display higher fluctuations on these small scales, as a consequence of the more complex mass distribution.

While giving a thorough physical interpretation of these results is beyond the scope of this short section, the examples presented above show the potential of VORTEX-P as a tool for analysing the velocity field in particle-based simulations.

6.2 A deep revision of the ASOHF code

This Section is based on Vallés-Pérez, Planelles, and Quilis (2022)[6], and the full paper can be found in Appendix A5. The algorithm presented in this publication is publicly available[7] and documented.[8]

Overview

Context. One of the main analyses customarily performed on cosmological simulations is the identification of structures. In this context, DM haloes are the building block of the LSS, generating the gravitational potential wells that trigger the collapse of the observed galaxies and galaxy clusters. Even when a DM halo is just a locally overdense, gravitationally bound structure, there is not a single operative definition and many halo-finding techniques have been developed in the past decades. Additionally, the ever-increasing trend in computational power has allowed simulations to grow in size and resolution, in such a way that the analysis of the simulations becomes a pressing computational problem itself.

Aims. The overhaul of the already-existing, spherical-overdensity halo finder ASOHF, aimed to: (i) improve its parallel performance and its ability of handling very large simulations, (ii) boost its capabilities of dealing with substructure, (iii) identify galaxies within DM haloes.

Methods. ASOHF is based on the spherical-overdensity paradigm, together with a multi-resolution density interpolation to detect substructure. The most salient new additions to the code include the new definition of substructure, the recentring schemes, the ability to identify galaxies, the domain decomposition strategy, a revamped merger tree procedure, among others. The algorithm is implemented in FORTRAN and parallelised according to the OMP standard directives.

[6] D. Vallés-Pérez, S. Planelles, and V. Quilis. "The halo-finding problem revisited: a deep revision of the ASOHF code." In: *Astron. Astrophys.* 664, A42 (Aug. 2022), A42. DOI: 10.1051/0004-6361/202243712

[7] https://github.com/dvallesp/ASOHF.

[8] https://asohf.github.io.

Results. The code has been tested against a battery of idealised tests aiming to verify the performance of different critical processes of the halo-finding scheme. Additionally, its performance has been compared to other publicly-available halo finders on actual simulation data, showing a generally good agreement and a remarkable performance in substructure detection.

Conclusions. The new, public version of ASOHF produces remarkable results in terms of halo, subhalo, and galaxy finding capabilities, parallel performance and a reduced computational cost.

6.2.1 ASOHF as a galaxy finder: Application to DIANOGA simulations

As a further verification of the capabilities of ASOHF as a galaxy finder, in this Section I show results from its application to the DIANOGA D1-BH2015 simulation (Planelles et al. 2014, Rasia et al. 2015)[9]. The halo-finding is restricted to the central $40\,h^{-1}\,$Mpc. Typical wall time (per snapshot) in this configuration is between $\sim 20\,$s and $\sim 3\,$min, depending on the choice of the free parameters, using 24 cores on a desktop workstation.

In the pages below, the results are briefly summarised in comparison with the galaxy catalogues extracted from SUBFIND (Springel et al. 2001, Dolag et al. 2009) outputs. To produce the galaxy catalogues from SUBFIND, all subhaloes (excluding the central one, if there is more than one subhalo per group) containing at least 15 stellar particles are considered.

For reference, the ASOHF catalogues have been obtained with the parameters (see the precise definition in the code paper, Table A5.1):

- Auxiliary mesh: $N_x = 128$, $n_\ell = 8$, $N_{\rm min}^{\rm patch} = 12$, $n_{\rm part}^{\rm refine} = 3$. Kernel length assigned by local density, with stars using an $\ell = 6$ kernel.
- Minimum number of particles for DM (sub)haloes: $n_{\rm part}^{\rm min} = 15$.
- Stellar halo finding parameters: $f_{\rm min} = 5$, $\ell_{\rm gap} = 1\,$ckpc, $f_{\rm B} = 1$, maximum (physical) radius for delimiting the extent and collecting stars of a BCG candidate, 200 kpc.

Fig. 6.6 shows a thin slice of the spatial distribution of DM and stellar particles in a $\sim 20\,$Mpc wide region around the central cluster. The panel on the left shows the galaxies detected by ASOHF as pink dots, while the corresponding

[9]We are grateful to Giuseppe Murante for granting us access to the DIANOGA simulations and their associated SUBFIND catalogues.

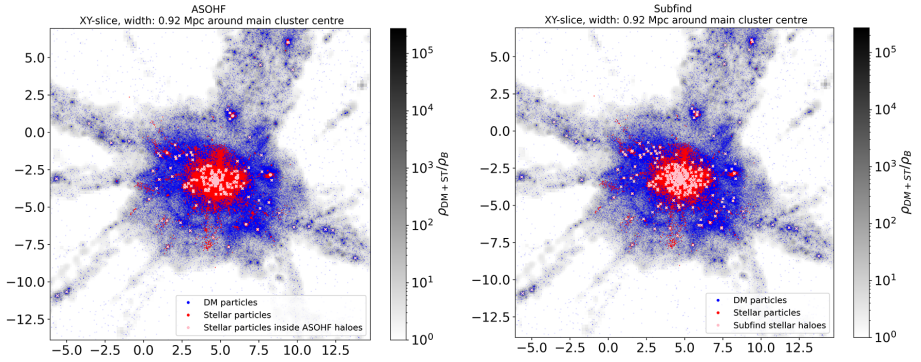


Figure 6.6: *Left panel:* In grey scale, the density of DM + stars in a box $\sim$ 20 Mpc wide and $\sim$ 1 Mpc deep around the central cluster. Blue dots represent a sampling of the DM particles, while red dots represent stellar particles. Pink dots represent the stellar particles inside the galaxies identified by ASOHF. *Right panel:* Same as the left panel, but with the haloes identified by SUBFIND. Here we show a cross for the centre of each SUBFIND subhalo.

results for SUBFIND are contained in the right-hand side panel. Overall, ASOHF identifies 1406 galaxies, while SUBFIND obtains 1859 according to the selection cuts imposed. Of these, 1227 can be matched between the two, so that SUBFIND identifies 632 haloes that ASOHF does not; while ASOHF identifies 179 that SUBFIND does not.

A summary of the properties of the galaxies identified by ASOHF is given in Fig. 6.7. Galaxies detected by ASOHF have a peaked distribution in sizes, with half-stellar mass radii ranging (1-10) kpc. Galactic stellar masses have a narrow peak between $(1-2)\times10^{10}\,M_\odot$, associated to the details of the feedback scheme used in the DIANOGA BH2015 run. In this direction, differential trends are found for $M_{1/2,*}\lessgtr 10^{10}M_\odot$ in the stellar mass vs. radii and stellar mass vs. DM mass. Note that all axes (except ellipticity) are logarithmic.

As a comparison, Fig. 6.8 shows a subset of these properties for the galaxies identified by ASOHF (orange) and by SUBFIND (blue). The most striking difference regards the radii distribution, since ASOHF delimits the extent of stellar haloes by the half-stellar mass radii, as opposed to the half-total mass radii of SUBFIND catalogues. Since stellar mass is very concentrated towards the centre of the halo, the impact on the corresponding galaxy mass is marginal. Looking at the two lower panels, SUBFIND also captures some differential behaviour below and above $\sim 10^{10}\,M_\odot$, with obvious differences with respect to ASOHF due to their different radius prescriptions (i.e., here SUBFIND reports much larger radii). One can also look at the relation between the stellar to DM mass (lower right panel). Here, the DM mass for ASOHF is the mass of the underlying DM halo

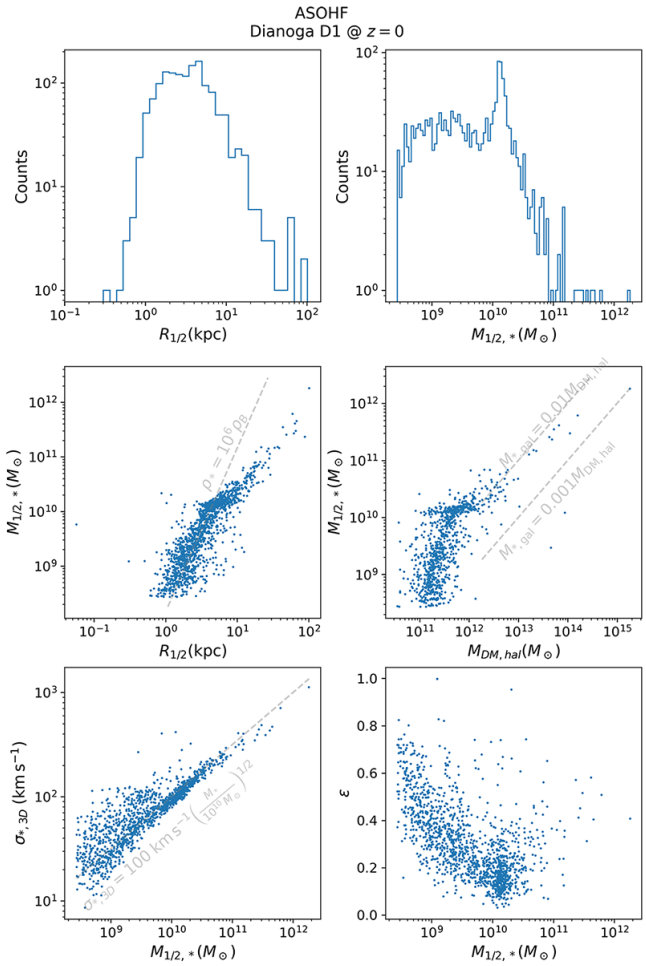


Figure 6.7: Summary of the properties of ASOHF galaxies, over the DIANOGA D1 BH2015 simulation at $z = 0$. *Top left:* distribution of half-stellar mass radii. *Top right:* distribution of half-stellar mass. *Middle left:* Half-stellar mass vs. half-stellar mass radius. *Middle right:* Stellar mass (inside $R_{1/2}$) vs. DM halo mass. *Bottom left*: 3D velocity dispersion vs. stellar mass. *Bottom right*: ellipticity vs. stellar mass.

(instead of the DM mass within the half-stellar mass radius). Thus, ASOHF tends to report higher DM halo masses.

Last, in order to validate the performance of ASOHF as a galaxy finder, it can be interesting to study the properties of the SUBFIND galaxies that are missed by ASOHF. This is shown in Fig. 6.9, where blue lines correspond to the SUBFIND catalogue, and orange lines correspond to the SUBFIND haloes that cannot be matched to any ASOHF halo.

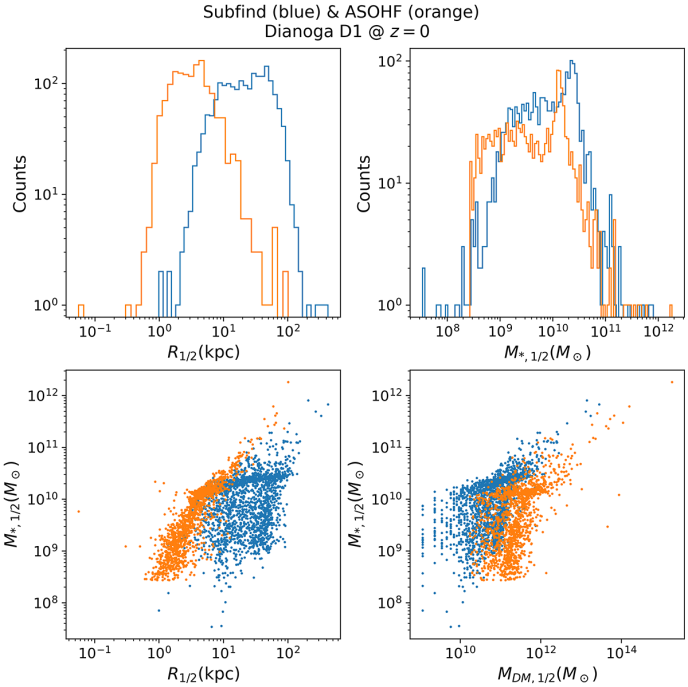


Figure 6.8: Same plots as the two upper rows of Figure 6.7, but this time comparing SUBFIND (blue) to ASOHF (orange).

Looking at the bottom panels, it is clear that the majority of haloes missed by ASOHF correspond to small haloes. Moreover, focusing on the lower right panel, it becomes obvious that the majority of haloes missed by ASOHF are DM-deficient (e.g., 55% of the missed haloes contain less than 25 DM particles, and 90% of the missed haloes contain less than 100). This limitation is inherent to ASOHF, since the identification of galaxies is built upon the detection of their underlying DM halo, even though the final galaxy may have different centre and bulk properties than the original DM halo.

6.3 Future directions

The code paper for VORTEX-P, containing a complete description of the algorithm, a thorough test set, and several applications, although not included in this Thesis, has been very recently submitted to *Computer Physics Communications*. The more complex and realistic tests, including velocity modes spanning several orders of magnitude in scales, allow to estimate the accuracy of the code when applied

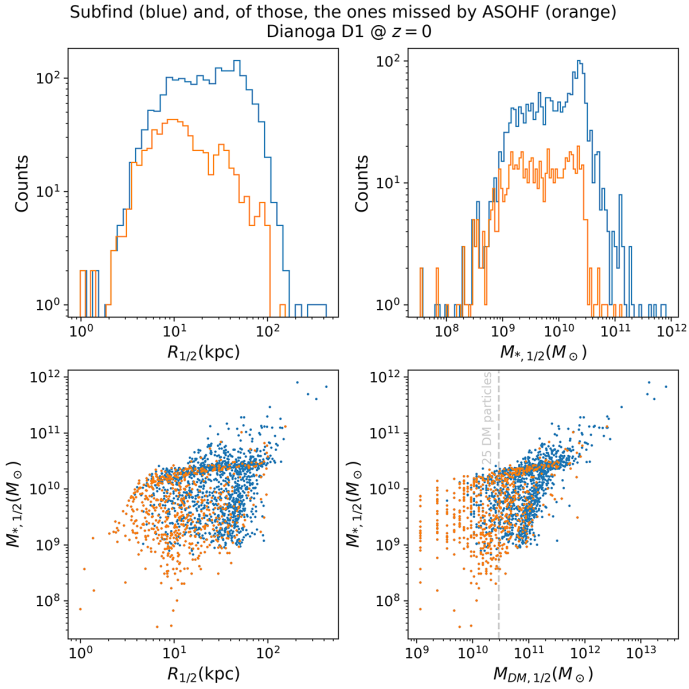


Figure 6.9: Statistics of the SUBFIND haloes that are missed by ASOHF. The two upper panels present radii and masses histograms, analogous to the previous figures, where the blue lines correspond to the SUBFIND catalogue, and the orange lines correspond to the SUBFIND haloes that cannot be matched to any ASOHF halo. Note the scale of the vertical axes in the histograms are logarithmic.

to extremely high-resolution simulations, such as the ones recently presented by Steinwandel et al. (2023), who resolve a cluster of mass $M_{\rm vir} \sim 2 \times 10^{15}\, M_\odot$ with SPH particle masses of $m_{\rm SPH} \sim 4 \times 10^5\, M_\odot$.

Naturally, applying VORTEX-P to the previously mentioned very high-resolution simulations will require a careful optimisation of many routines of the code. On the one hand, more efficient versions of the critical algorithms must be sought for in order to improve the computing time. Nevertheless, if the code is to be applied to simulations of this unprecedented size, the parallelisation of the code for shared-memory platforms (MPI) through a domain-decomposition strategy is imperative. In principle, this should not be particularly complex, since the grid hierarchy already provides a natural domain decomposition. However, the fact that patches use large buffer zones to minimise boundary effects requires boundary conditions to be sought for in patches different from their *parent* patch, complicating the communications. Additionally, load balancing could also be a critical issue. Therefore, these enhancements are mid-term goals for the code.

Also in the mid-term, both for the grid-based and the particle-based versions of VORTEX, it would be interesting to implement a new, coordinate unsplit (e.g., Skillman et al. 2008, Schaal and Springel 2015) shock-finder as an additional (optional) module for the code. This would overcome the necessity to supply the code with external, Mach number information to perform the Reynolds decomposition, or to use rough approximations as the ones currently implemented in VORTEX-P. Furthermore, this would turn VORTEX into a general framework for performing a wide range of analyses on the velocity fields of cosmological simulations.

Regarding ASOHF, even though the code has already been made use of on a variety of simulations both inside and outside our group, a simple step forward towards improving its public adoption would be the inclusion of reader routines for the most widely used simulation codes (e.g., GADGET, AREPO, RAMSES, ENZO, etc.). In the mid-to-long term, even though the domain decomposition strategy implemented in the present version of the halo finder, which is performed externally in a semi-automatic manner, provides excellent results for simulations of large volumes, it would be interesting to implement an internal, MPI-based parallelisation to deal with high-resolution simulations of small volumes, where the current domain decomposition strategy is not as efficient due to the need of large buffer zones.

Concerning the cosmological code developed and used in our group, MASCLET, a first line of work that could be addressed in the short-to-mid term is its interoperability with widely-used, public libraries for handling simulation data. Including interfaces for natively loading MASCLET outputs into, for instance, NBODYKIT (Hand et al. 2018) or YT (Turk et al. 2011) would enable a direct application of their tools for easily analysing the large-scale structure properties and to produce high-quality graphics from the simulation data, respectively. Although a few steps in this direction have already been made (i.e., an internal reader for YT has already been implemented during this PhD Thesis), full exploitation of the AMR grid (without the need of construction uniform grids from the multi-resolution data; and keeping the parallel support) is still missing.

In a longer timeframe, an important step-up for MASCLET would be its MPI parallelisation, in order to be able to efficiently run simulations in distributed-memory platforms with thousands of cores. This demands a thorough revision and rewriting of the code, due to the complexity and the dynamical nature of the AMR grid, but would ultimately enable the code to perform simulations with an unprecedented number of resolution elements. In this direction, offloading

some of the most computationally-expensive routines to GPUs would also be an important step forward, aligned with the current trend in high-performance computed towards the exascale era (Gagliardi et al. 2019). This is supported by the OpenMP 4.5 standard, which allows for the offloading of code to GPUs in a relatively simple manner by using the NVIDIA Compute Unified Device Architecture (CUDA; Nickolls et al. 2008). In the field of grid-based cosmological codes, GPU implementations have been attained, for instance, in ENZO and RAMSES (Gheller et al. 2015).

CHAPTER 7

Galaxy cluster assembly

Contents

The assembly of galaxy clusters determines many of their properties, from their stellar populations to the thermodynamical states and the kinematics of the ICM. This section summarises the works conducted within this PhD Thesis aimed to explore the description of the assembly of galaxy clusters, and its impact on the properties of these objects, from the study of accretion rates and dynamical state indicators.

7.1 On the accretion history of galaxy clusters

This Section is based on Vallés-Pérez, Planelles, and Quilis (2020)[1], and the full paper can be found in Appendix A1.

[1] D. Vallés-Pérez, S. Planelles, and V. Quilis. "On the accretion history of galaxy clusters: temporal and spatial distribution." In: *Mon. Not. R. Astron. Soc.* 499.2 (Dec. 2020), pp. 2303–2318. DOI: 10.1093/mnras/staa3035

Overview

Context. Galaxy clusters and groups grow through the highly-anisotropic accretion of matter from the surrounding cosmic web, as well as through merger events between similarly-sized structures. Past works have confirmed the relation of accretion to several ICM properties, such as the profiles of several thermodynamic quantities, the position of clusters with respect to the scaling relations, or the ellipticity of the ICM.

Aims. The characterisation of the accretion history of a small sample of galaxy clusters from a numerical simulation, aiming to examine different proxies for the accretion rate, as well as the impact of accretion on the internal properties of the cluster, and the spatial distribution of accretion flows.

Methods. Several proxies for the mass accretion rate (MAR), including $\Gamma_{200\mathrm{m}}$ computed from the full merger tree of haloes, α_{200m} computed from the velocity profile, and the integrated mass flux through the virial boundary are compared amongst themselves. $\Gamma_{200\mathrm{m}}$ is additionally compared to the presence of merging events and the densities surrounding the cluster. Finally, the spatial distribution of mass accretion flows is estimated using a pseudo-Lagrangian approach on the Eulerian gas data of the simulation, and studied through multipolar analysis.

Results. The different MAR proxies are loosely correlated amongst themselves, highlighting the difficulty in finding good measures of cluster growth associated to the arbitrariness in the definition of the cluster boundary. Both the baryonic and total MAR correlate with the presence of merging events but, when measured at R_{200m}, are kept high for a long time after the merger episodes. The novel approach to study the angular distribution of accretion flows highlights the very anisotropic behaviour of mass flows at the virial boundary, and its multipolar analysis shows significant differences in some thermodynamic properties such as the entropy of the accreted gas.

Conclusions. In the advent of surveys that will be able to probe the outskirts of clusters, the characterisation of accretion rates and its comparison with cluster observables is of paramount importance. This exploratory work exemplifies some tentative approaches, but work with larger samples is required to draw robust conclusions.

7.2 The assembly state of dark matter haloes through cosmic time

This Section is based on Vallés-Pérez et al. (2023)[2], and the full paper can be found in Appendix A6.

Overview

Context. The dynamical state and the morphological features of galaxies and galaxy clusters, as well as those of their DM counterparts, are tightly connected with their assembly history through a wide redshift interval. Hence, these features may encode crucial information about the formation and the evolution of such cosmic structures.

Aims. This work intends to critically examine how the assembly state of these structures, understood as the presence of recent mergers or periods of strong accretion, can be best assessed from a series of properties at a given redshift.

Methods. By using the combination of centre offset, virial ratio, mean radial velocity, sparsity and ellipticity of the DM haloes extracted from a moderate-volume simulation, we study how the thresholds on these parameters, as well as their relative weights to define a single, combined indicator, should evolve with redshift to best correlate with the merging and accretion activity inferred from complete merger trees.

Results. The resulting classification, involving a *totally relaxed*, an *unrelaxed*, and an intermediate (*marginally relaxed*) categories, correlates strongly with the merging activity extracted from the merger trees, as well as with the accretion rates. This has been also tested against data from a different simulation from the publicly available CAMELS project (Villaescusa-Navarro et al. 2021, Villaescusa-Navarro et al. 2023).

Conclusions. A dynamical state classification focused on the assembly history (presence of mergers and accretion rates) must take into account that different indicators may yield more insight than others at a particular redshift. Likewise, the thresholds that must be applied to these indicators to produce the best classification also appear to be redshift-dependent.

[2]D. Vallés-Pérez et al. "On the choice of the most suitable indicator for the assembly state of dark matter haloes through cosmic time." In: *Mon. Not. R. Astron. Soc.* 519.4 (Mar. 2023), pp. 6111–6125. DOI: 10.1093/mnras/stad059

7.2.1 The inclusion of substructure fraction as a dynamical state indicator

As acknowledged in the discussion contained in §2.3.6 of Paper A6, the use of the substructure fraction, $f_{\rm sub}$, as an indicator of the assembly state of DM haloes did not produce satisfactory results in the analysis presented in the paper, and was hence discarded. The mass contained in substructures within a larger halo is particularly sensitive to, both, the mass and force resolution of the simulation, and the definition of the extent of the substructures themselves. In this Section, I present results from the analysis of a new simulation, equivalent to the one used in Paper A6, but with new initial conditions using around ~ 3 times more DM particles and improvements to the N-Body solver that result on a better effective force resolution in the AMR patches, allowing a more accurate treatment of substructures. While the qualitative results (i.e., the trends of the thresholds and weights on each parameter with cosmic time) are essentially unchanged, it is still worth presenting here the new results including substructure fraction for the sake of completeness.

Fig. 7.1 shows the updated version of Figs. A6.3 (left-hand side column, showing the evolution of the thresholds on each indicator) and A6.4 (right-hand side column, showing the evolution of the corresponding weights) on Paper A6. While the fits change, they are generally consistent with the former ones within their confidence regions. Substructure fraction (bottom row) appears to be now a useful parameter to assess the assembly state of haloes at intermediate and low redshift ($z \lesssim 3$). Regarding its threshold, reaching $f_{\rm sub}^{\rm thr} \sim 0.02$ at $z \sim 0$, it must be borne in mind that the definition of substructure mass in ASOHF is much more conservative than that used in other halo finders (see the discussion in Paper A5, its §4.1).

The new fits for the parameters are given by:

$$\Delta_r^{\rm thr}(z) = 0.0716(23) + 0.0323(58)z - 0.0146(37)z^2 + 0.00192(65)z^3 \quad (7.1)$$

$$\eta^{\rm thr}(z) = 1.3411(51) + 0.162(13)z - 0.0463(84)z^2 + 0.0038(14)z^3 \quad (7.2)$$

$$\langle \tilde{v_r} \rangle^{\rm thr}(z) = 0.0858(24) \quad (7.3)$$

$$s_{200c,500c}^{\rm thr}(z) = 1.550(10) + 0.0558(87)z \quad (7.4)$$

$$\varepsilon^{\rm thr}(z) = 0.2798(27) \quad (7.5)$$

$$f_{\rm sub}^{\rm thr}(z) = 0.0178(17) - 0.0080(16)z + 0.00092(33)z^2 \quad (7.6)$$

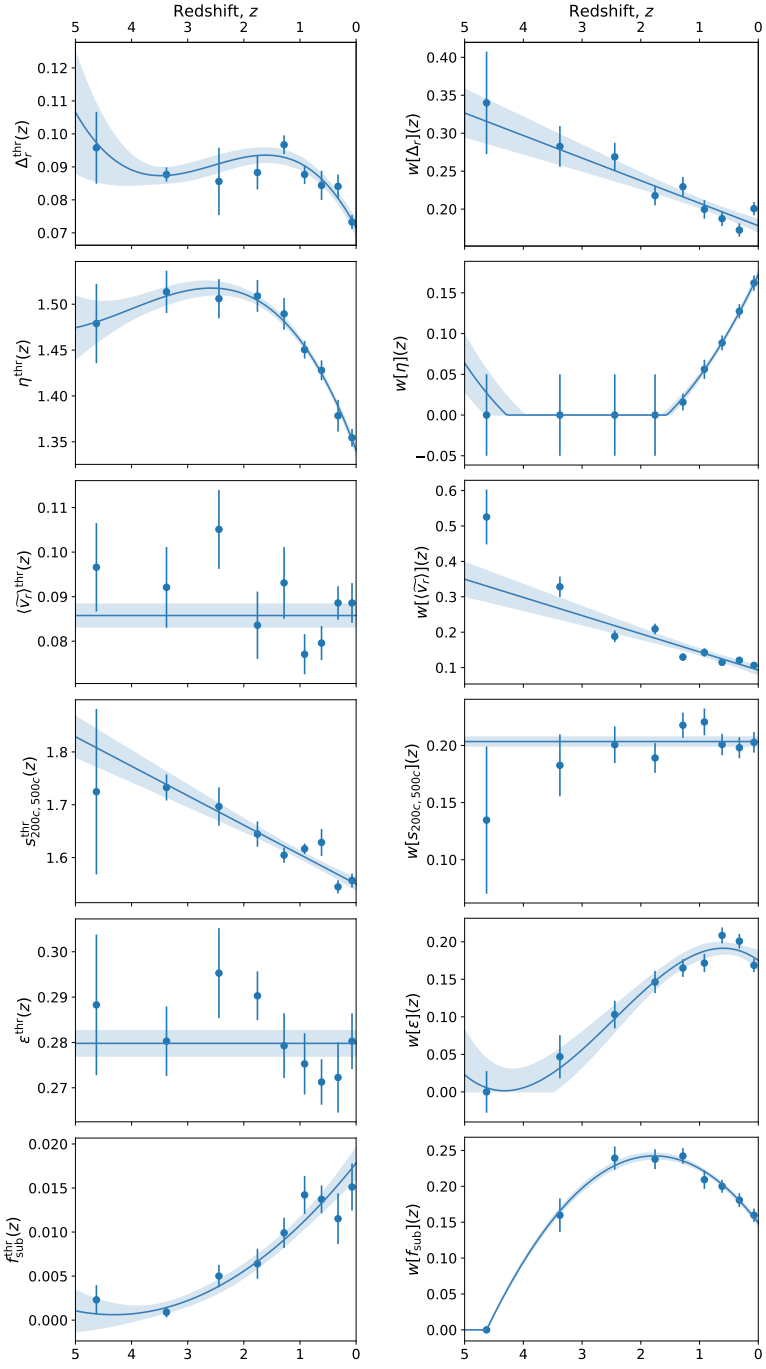


Figure 7.1: Evolution of the thresholds on the parameters (left column) and of their relative weights (right column) to provide the best assembly state classification. These figures are equivalent to Figs. A6.3 and A6.4 on Paper A6.

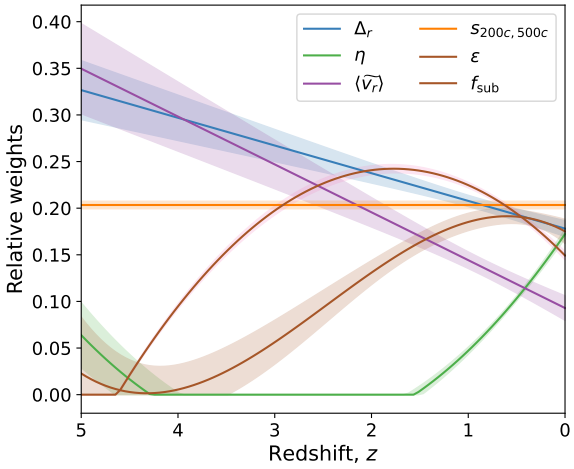


Figure 7.2: Overall fitted redshift evolution of the weights on the dynamical state indicators. This figure is equivalent to Fig. A6.5 on Paper A6.

While, for the weights:

$$w[\Delta_r](z) = 0.1781(79) + 0.0297(69)z \tag{7.7}$$

$$w[\eta](z) = 0.1722(47) - 0.1510(84)z + 0.0259(24)z^2 \tag{7.8}$$

$$w[\langle \tilde{v_r} \rangle](z) = 0.093(12) + 0.051(11)z \tag{7.9}$$

$$w[s_{200c,500c}](z) = 0.2034(38) \tag{7.10}$$

$$w[\varepsilon](z) = 0.175(12) + 0.057(33)z - 0.054(21)z^2 + 0.0074(34)z^3 \tag{7.11}$$

$$w[f_{\rm sub}](z) = 0.1492(42) + 0.1052(58)z - 0.0297(11)z^2 \tag{7.12}$$

To better visualize the performance of each indicator at different cosmic epochs, the fitted evolution of the weights is shown in a single panel in Fig. 7.2. Again, the qualitative results from the paper are kept: centre offset and mean radial velocity are still the most relevant indicators at high redshift, which gradually decline its weight at lower redshifts. At $z \sim 0$, all indicators add relevant insight about dynamical state, while substructure fraction is, as a matter of fact, the most insightful measure of assembly state at $z \sim 1 - 2$. The only mildly significant difference ($\sim 2\sigma$) with respect to the previous results appears with ellipticity at high redshift, which, in the new results, is much more irrelevant.

7.3 Future directions

The natural continuation of the studies presented within this Chapter go along the lines of extending the previous studies with new simulations, such as the one presented in Paper A6 and higher-resolution versions of it, in order to study the effects of the assembly of the cluster (galaxy cluster mergers and gas accretion) onto several cluster properties and observables using a considerably larger sample.

Therefore, in the short-term, as already started projects (albeit in a preliminary stage, so that no results are presented within this PhD Thesis), several topics that we aim to cover include:

- Impact of accretion rates and assembly states on the internal structure of galaxy clusters, by the effect on their profiles. This has been previously explored by, e.g., Lau et al. (2015), who find a noticeable dependence of the radially-averaged profiles of temperature, entropy, gas pressure and density in the outskirts of clusters with their accretion rates, using a non-radiative simulation.

- Dependence of the X-ray and SZ scaling relations (e.g., $L_X - M$, $T_X - M$, $Y_X - M$, $Y_{\rm SZ} - M$) with the accretion rates and assembly states. Several recent studies (e.g., Chen et al. 2019) point at a connection between the recent accretion rate of the cluster and the residuals with respect to the mean $T_X - M$ relation. Given that ongoing and future X-ray missions will reduce the instrumental uncertainties to the point where the scatter around the scaling relations will be dominated by the astrophysical uncertainties (Hofmann et al. 2017), a deeper understanding of these connections can be of utmost importance to improve cluster mass estimates.

- The angular distribution of gas accretion flows onto galaxy clusters and its time evolution, using the methodology introduced in Paper A1, will be extended to a sample of clusters, to study its connection with dynamical state indicators, cosmic time, and other cluster properties.

In the short-to-mid range, besides these more fundamental studies, it would also be important to explore the impact of cluster assembly on mock observations of galaxy clusters throughout the electromagnetic spectrum. To this end, the full-radiative transfer code SPEV (Mimica et al. 2009, Cuesta-Martínez, Aloy, and Mimica 2015, Mimica et al. 2016) has been already used to produce multi-wavelength mock observations out of MASCLET-simulated clusters (Planelles

et al. 2018). In this context, I have participated in a project aimed to assess the impact of the assembly state of galaxy clusters on their kSZ signal (Monllor-Berbegal et al. 2024; see App. B). In order to improve on the aforementioned works, next steps should include a more detailed treatment of the metal distribution and feedback mechanisms, the inclusion of relativistic corrections for the SZ signal (Itoh, Kohyama, and Nozawa 1998, Nozawa, Itoh, and Kohyama 2005), and the consideration of the instrumental response of current or forthcoming observational facilities.

Lastly, on a longer timeframe, we also plan to develop the necessary numerical techniques for identifying and extracting catalogues of the cosmic-web filaments out of our simulation data. The interest for this is, at least, two-fold. On the one hand, the interfaces between filaments and cluster outskirts are particularly interesting regions because of their departures from self-similarity, presence of dynamically relevant non-thermal gas motions, density inhomogeneities, accretion and merger shocks, lack of equilibrium between electrons and ions, etc. (Walker and Lau 2022, for a review). On the other hand, the study of the properties of filaments themselves can be of special interest to several fields, from galactic environments (e.g., Bonjean et al. 2018, Kuchner et al. 2022) to cosmic magnetism (Gheller et al. 2016, Banfi, Vazza, and Gheller 2021, Vernstrom et al. 2021).

CHAPTER 8

Turbulence and shocks during galaxy cluster assembly

Contents

As discussed in Chapter 5, turbulent motions and shock waves are two phenomena associated to the non-linearity of the fluid dynamics equations that have far-reaching consequences for, both, the physics of the ICM, and its observable properties. In this Chapter, two studies regarding these phenomena are presented.

8.1 Turbulence, enstrophy and helicity from the assembly history of the intracluster medium

This Section is based on Vallés-Pérez, Planelles, and Quilis (2021a)[1], and the full paper can be found in Appendix A3. The methods developed to perform a

[1] D. Vallés-Pérez, S. Planelles, and V. Quilis. "Troubled cosmic flows: turbulence, enstrophy, and helicity from the assembly history of the intracluster medium." In: *Mon. Not. R. Astron. Soc.* 504.1 (June 2021), pp. 510–527. DOI: 10.1093/mnras/stab880

Reynolds decomposition on velocity fields defined on a block-based AMR set of grids have been included in the public code VORTEX.[2]

Overview

Context. Turbulence is a pervasive phenomenon in astrophysical flows and, in particular, in galaxy clusters. However, it is particularly difficult to model numerically due to its intrinsically multiscale nature, which demands high resolutions. Additionally, previous works have shown that turbulence is tightly connected to the formation history of the ICM.

Aims. Explore the connection of the generation and dissipation of turbulence in the ICM with the assembly history (both of DM and baryons) of galaxy clusters.

Methods. The public code VORTEX has been extended to include the possibility of performing a bulk+turbulent (Reynolds) decomposition prior to the HHD. This, applied to the two most well-resolved clusters of the simulation, allows extracting global statistics (e.g., compressive and solenoidal structure functions) and local statistics (e.g., enstrophy) of the turbulent flows.

Results. While structure functions show an important dependence on cluster-centric radius and this is due, to a large extent, to the non-constant resolution, the study of the evolution of the second-order structure functions at fixed scales (comoving with $R_{\rm vir}$) still reveals a clear correlation with accretion rates and merger periods. Following the temporal evolution and the spatial and phase-space distribution of enstrophy, an indicator of solenoidal motions, our results confirm a previously proposed scenario where volume-filling solenoidal motions are generated, first, by baroclinicity at the outermost shocks surrounding the cluster, and then enhanced by compression and channelled downstream of the shocks by vortex stretching.

Conclusions. Although the measurement of quantities related to the turbulent flow in AMR data, where different regions have been evolved with different resolutions at different times, is far from straightforward, the results reported in this paper show that it may be possible to extract meaningful information about the turbulent flow in the ICM with these simulations.

[2]https://github.com/dvallesp/vortex.

8.2 Cosmic accretion shocks as a tool to measure the mass of galaxy clusters

This Section is based on Vallés-Pérez, Quilis, and Planelles (2024, in press)[3], which has been accepted for publication in the journal Nature Astronomy. The manuscript is included below, after the overview, in its original form (previous to peer review) as per licence permissions of the publisher.

Overview

Context. Cosmological accretion shocks created during the formation of galaxy clusters are a ubiquitous phenomenon all around the Universe. These shocks and their features are intimately related with the gravitational energy put into play during galaxy cluster formation.

Aims. We aim to study the scaling relations linking the mass of galaxy clusters with the radius and intensity of their accretion shocks, and to explore the possibility of using these relations to measure the mass of galaxy clusters.

Methods. We study a sample of galaxy clusters and galaxy groups extracted from a Eulerian cosmological simulation of moderate size ($100\,h^{-1}\,\mathrm{Mpc}$ on a side) and resolution (peak resolution of $\Delta x \sim 9\,\mathrm{kpc}$). Shocks are identified in the simulation using a coordinate-split shock-finding algorithm based on temperature jumps, and the outermost accretion shock surfaces are subsequently characterised based on a series of heuristic criteria.

Results. We show that objects in our sample sit in a plane within the three dimensional-space of cluster total mass, shock radius, and Mach number (a measure of shock intensity).

Conclusions. Using this relation, and considering that forthcoming new observations will be able to measure shock radii and intensities, we put forward the idea that the dark matter content of galaxy clusters could be indirectly measured with an error up to around 30 per cent at the 1σ confidence level. This procedure would be a new and independent method to measure the dark matter mass in cosmic structures, and a novel constraint to the accepted ΛCDM paradigm.

[3] D. Vallés-Pérez, V. Quilis, and S. Planelles. "Cosmic accretion shocks as a tool to measure the dark matter mass of galaxy clusters." In: *Nat. Astron.* (2024, in press). DOI: 10.1038/s41550-024-02303-x

Cosmic accretion shocks as a tool to measure the dark matter mass of galaxy clusters

David Vallés-Pérez[1,*], **Vicent Quilis**[1,2,†], **and Susana Planelles**[1,2,†]

[1]Departament d'Astronomia i Astrofísica, Universitat de València, C/Doctor Moliner, 50, Burjassot, 46100, València, Spain
[2]Observatori Astronòmic, Universitat de València, C/Doctor Moliner, 50, Burjassot, 46100, València, Spain
[*]Corresponding author: David Vallés-Pérez (david.valles-perez@uv.es),
[†]Vicent Quilis (vicent.quilis@uv.es), Susana Planelles (susana.planelles@uv.es)

ABSTRACT

Cosmological accretion shocks created during the formation of galaxy clusters are a ubiquitous phenomenon all around the Universe. These shocks, and their features, are intimately related with the gravitational energy put into play during galaxy cluster formation. Studying a sample of simulated galaxy clusters and their associated accretion shocks, we show that objects in our sample sit in a plane within the three dimensional-space of cluster total mass, shock radius, and Mach number (a measure of shock intensity). Using this relation, and considering that forthcoming new observations will be able to measure shock radii and intensities, we put forward the idea that the dark matter content of galaxy clusters could be indirectly measured with an error up to around 30 per cent at the 1σ confidence level. This procedure would be a new and independent method to measure the dark matter mass in cosmic structures, and a novel constraint to the accepted ΛCDM paradigm.

Galaxy clusters and galaxy groups sit at the top of the hierarchy of gravitationally-bound cosmological structures, as the result of a hierarchical formation history which has lasted for around 14 Gyr [1, 2, 3]. While their matter content is dominated by dark matter (DM), which accounts for roughly 85% of the gravitational mass, baryons (most of them, in the form of a diffuse, multiphase plasma; see, e.g., [4, 5, 6, 7], for reviews) also play a pivotal role in their evolution and, especially, in the modelling of their observational properties [e.g., 8, 9, 10].

The singular place of galaxy clusters makes them valuable tools for precision cosmology [11, 12, 13], yet the determination of their masses is not straightforward and can only be done by different indirect approaches, often relying on rather strong assumptions. For instance, radial profiles of some thermodynamic properties of the intracluster medium (ICM; the bulk of baryonic mass in galaxy clusters) that can be obtained from X-ray or Sunyaev-Zeldovich (SZ) observations can be used to constrain the total mass, under the assumption of sphericity and hydrostatic equilibrium [e.g., 14, 15]. Related to this, masses can also be obtained from integrated properties invoking self-similarity [16]. Both these approaches depend crucially on baryonic physics and feedback mechanisms, are affected by biases in the order of $(10-30)\%$ [14], and usually break at the group scale [e.g., 17]. Alternatively, masses can be determined by weak-lensing measurements [18], or looking at the kinematics of galaxies in the cluster (e.g., through the caustic technique, [19]), which are ultimately probes of the underlying gravitational potential. A more in-depth, recent review on current mass estimation techniques for galaxy clusters can be found in [20].

As a direct consequence of the collisional nature of the baryonic component, shock waves appear as pervading phenomena in cosmological structure formation. They play a central role in the evolution of galaxy groups and clusters, by providing the necessary non-adiabatic heating of the diffuse plasma up to the temperatures observed within cosmic structures [21, 22, 23] at the expense of removing kinetic energy from bulk flows, so that the infalling baryons can virialise within the gravitational potential wells of DM haloes. In particular, both analytic models [24, 25] and simulations [26, 27] have studied the location of the outermost accretion shocks of galaxy groups and clusters, whose position and evolution are ultimately determined by the gravitational collapse of the given structure.

The detection of these elusive features in observations has been devoted considerable attention recently, being the target of numerous surveys using new telescopic facilities. Although very preliminary, first detections of large-scale shocks have been reported by using very different observational approaches: pressure jumps using Sunyaev-Zeldovich effect observed with the Planck satellite and the South Pole Telescope (SPT) [28], γ-ray observations with the VERITAS Cherenkov array [29] or the Fermi Large Area Telescope [30], using UV absorption spectroscopy [31], or by measuring polarisation in radio observations [32].

In this work, we propose that total masses of clusters can be inferred from the physical size of the accretion shock shell and

its intensity, and calibrate these relations across a broad redshift interval using high-resolution cosmological simulations. We refer the reader to the Methods section and the Supplementary Material for technical details, while we summarise our main results below.

Results

We exemplify and further discuss our procedure for detecting the accretion shock shell and its equivalent radius, $R_{\rm sh}$, and strength, quantified through the average Mach number, $\mathscr{M}_{\rm sh}$, in the Methods section and in Supplementary Section A. Below, we discuss our results regarding the calibration of the multi-dimensional scaling relation, its evolution, and its scatter.

Multivariate relation between total mass, shock intensity and shock radii

Following the approach described in more detail in the Methods section, we use our simulation data to fit a linearised relation $\log_{10} M(<2R_{\rm vir}) = f(\log_{10} R_{\rm sh}, \mathscr{M}_{\rm sh})$. The choice of measuring the mass in spherical apertures of $2R_{\rm vir}$ is motivated by the fact that accretion shocks are located most of the times between $2R_{\rm vir}$ and $3R_{\rm vir}$ (see Supplementary Section B for a statistical summary of our sample), and hence this radius is a reasonable proxy for the mass producing the collapse of the overdensity and driving the shock evolution. However, we note that using sensibly smaller aperture radii leads to similarly accurate results (see Supplementary Section C for the equivalent results using $M(<R_{\rm vir})$).

Our best-fit relation for the sample at $z=0$ can be represented by the functional form

$$\log_{10}\frac{M(<2R_{\rm vir})}{M_\odot} = 12.760 + 1.910\log_{10}\frac{R_{\rm sh}}{\rm Mpc} + 0.0117\mathscr{M}_{\rm sh}, \qquad (\text{at } z=0) \tag{1}$$

and is presented graphically in the upper panel of Fig. 1. The mean scatter around this relation is $\sigma_{\log M}=0.134$, implying 1σ errors of $^{+36\%}_{-27\%}$ on the total mass measured within $2R_{\rm vir}$ apertures. The relation marginalised over shock intensities (the mass-shock radius relation, which we will hereon refer to as *2d relation*, in contrast to the complete 3d relation), has an intrinsic scatter of $\sigma_{2d,\log M}=0.177$ ($^{+50\%}_{-34\%}$ errors) at $z=0$, suggesting that the inclusion of information about shock intensity is crucial to recover more precise mass estimates. The slope of the 2d relation is marginally consistent with a self-similar scaling ($M(<2R_{\rm vir})\propto R_{\rm sh}^3$) within a 1σ interval (see Supplementary Section D for more details on the 2d relation).

The lower panel of Fig. 1 contains the same information at $z=1$, which is best-fitted by the relation,

$$\log_{10}\frac{M(<2R_{\rm vir})}{M_\odot} = 12.440 + 1.122\log_{10}\frac{R_{\rm sh}}{\rm cMpc} + 0.0225\mathscr{M}_{\rm sh}, \qquad (\text{at } z=1) \tag{2}$$

whose intrinsic scatter, $\sigma_{\log M}=0.144$, implies typical errors of $^{+39\%}_{-28\%}$ on the estimation of the mass enclosed by two virial radii. Also in this case, the addition of information about shock intensity decreases the error figure importantly, from a 2d value of $\sigma_{2d,\log M}=0.209$ ($^{+62\%}_{-38\%}$ errors).

Additional information about the fits in equations (1) and (2) can be found in Supplementary Section E. We have also checked that the mass accretion rate, when determined at the virial radius, does not correlate significantly with the position of the shock (see Supplementary Section F), and hence including it does not improve our results.

Evolution of the relation and fitting formulae

We have studied the evolution of our three-dimensional scaling relation with cosmic time, for the redshift interval $0\leq z\lesssim 1.5$. The resulting evolution can be fitted by a generalisation of the relation in equations (1) and (2), as

$$\log_{10}\frac{M(<2R_{\rm vir})}{M_\odot} = \gamma(z) + \alpha(z)\log_{10}\frac{R_{\rm sh}}{\rm cMpc} + \beta(z)\mathscr{M}_{\rm sh}, \qquad (\text{at } 0\leq z\lesssim 1.5) \tag{3}$$

where radii are expressed in comoving coordinates and the functions $\alpha(z)$, $\beta(z)$ and $\gamma(z)$ encapsulate the redshift evolution, which can be well-fitted by the polynomial forms below:

$$\alpha(z) = 1.523 + 0.208z - 0.386z^2 \tag{4}$$

$$\beta(z) = 0.01721 - 0.00374z + 0.00704z^2 \tag{5}$$

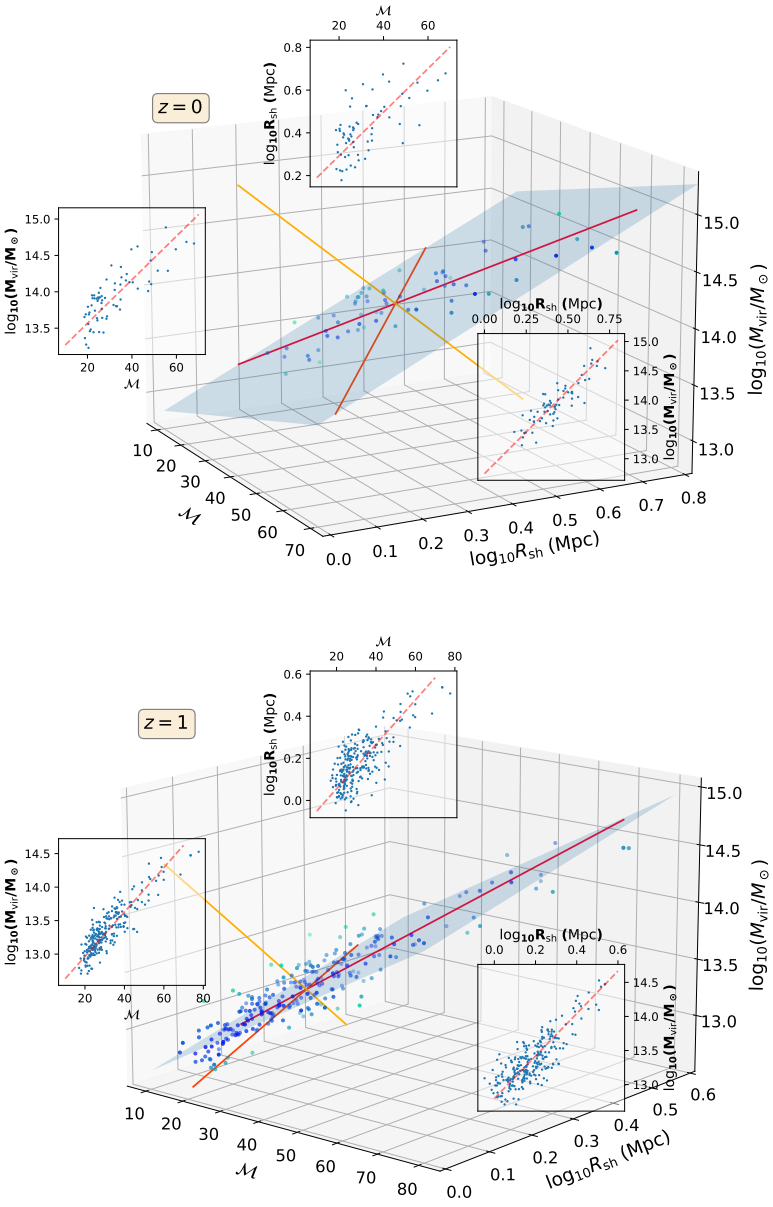


Figure 1. Best-fit relations $\log_{10} M(< 2R_{\rm vir}) = f(\log_{10} R_{\rm sh}, \mathscr{M}_{\rm sh})$, for $z = 0$ (upper panel) and for $z = 1$ (lower panel). In both panels, the information is presented as follows. Dots represent each individual cluster or group used for the fit, with their colour encoding the distance to the plane, darker being closer. The blue plane represents the best-fitting relation, which is spanned by the red and orange directions (representing the first and second principal components of the standardised data, respectively). The yellow line represents the third principal component. The plane is defined by setting this component to 0 (its mean value). The insets present the three marginal distributions, with the dashed, red line representing the best-fit linear relation to the marginal distribution.

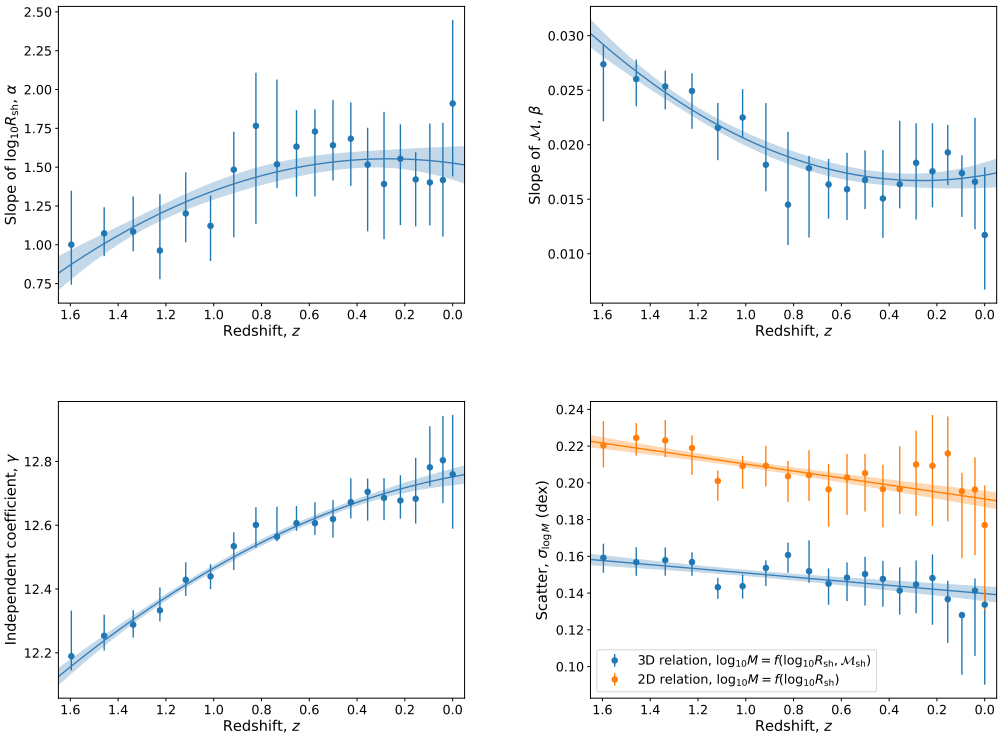


Figure 2. Scaling relation evolution summary. In all panels, dots represent the determinations at each redshift, with the bars accounting for $(16-84)\%$ percentiles obtained through bootstrap resampling. Solid lines with their shaded regions represent the best polynomial fits accounting for the evolution, with their 1σ confidence region. *Top left:* evolution of the coefficient of the shock radius. *Top right:* evolution of the coefficient of Mach number. *Bottom left:* evolution of the independent term. *Bottom right:* evolution of the scatter (in dex) around the three-dimensional scaling relation (blue line), and around the bidimensional scaling relation (orange line).

$$\gamma(z) = 12.75 - 0.143z - 0.144z^2 \tag{6}$$

The results for $\alpha(z)$, $\beta(z)$ and $\gamma(z)$ are presented graphically in the upper left, upper right, and lower left panels of Fig. 2, respectively. The lower right panel contains instead the evolution of the scatter around the fitted relation, evaluated over the same sample, $\sigma_{\log M}$. Here the blue line refers to the scatter on the three-dimensional relation, which obeys a slightly decreasing trend (from $\sigma_{\log M} \lesssim 0.16$ dex at $z \simeq 1.5$ to $\sigma_{\log M} \lesssim 0.14$ dex at $z \simeq 0$). As a comparison, the orange line presents the corresponding evolution for the bidimensional relation, i.e., the one not using information on shock intensity, confirming the behaviour observed in the examples in the previous section: adding information about shock intensity decreases the scatter across the whole redshift interval considered, and appears to be crucial in recovering precise mass estimates. We refer the reader to Supplementary Section G for more details on the evolution of the fit results. The corresponding evolution, for $M(< R_{\rm vir})$, is also presented in Supplementary Section C.

Assessment of the relation and its scatter

The study of the residuals and the scatter around our best-fit relations can be used as a calibration of the goodness of the fit and a prediction on the error intrinsic to the relation. At the same time, it is useful to assess the validity of the underlying

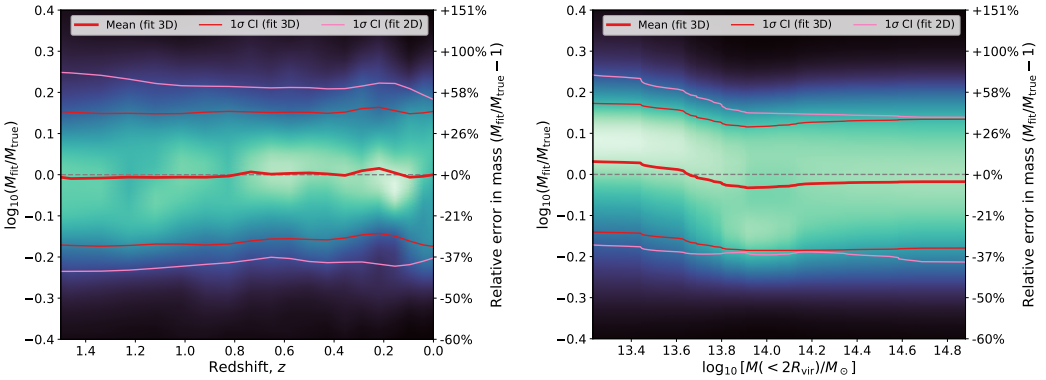


Figure 3. Distribution of the masses recovered by our best-fit relation (equation 3, with the fits given in equations 4-6) to assess its intrinsic scatter. In both cases, the vertical axis presents a conditional probability density estimation of the logarithmic error (in dex; shown as a percent error on the right of each panel), for a given redshift (left-hand panel) or for a given mass (right-hand panel; at $z = 0$). That is to say, each vertical line depicts, colour-coded (the brightest, the highest), the probability distribution of the residuals, for a given redshift or mass. The thick, red line presents the mean value, while thin, red lines enclose the $(16-84)\%$ confidence region. Similarly, the thin, pink line contains the $(16-84)\%$ confidence region around the bidimensional, $\log_{10} M(<2R_{\rm vir}) = f(\log_{10} R_{\rm sh})$ relation.

assumptions and discussing the limitations of the model.

In the left panel of Fig. 3, we represent the evolution with cosmic time (decreasing redshift towards the right) of the scatter around our relation. For each value on the x-axis (i.e., for each redshift), the colour along the y-axis represents the distribution of the residuals with respect to our relation, the brightest being the most probable. Our relation is hence unbiased with redshift, in the sense that the mean value of $\log_{10} M_{\rm fit}/M_{\rm true}$ averages to 0 at all times. Here, red [pink] thin lines enclose the $(16-84)\%$ confidence region for the complete $\log_{10} M(<2R_{\rm vir}) = f(\log_{10} R_{\rm sh}, \mathscr{M}_{\rm sh})$ [bidimensional, $\log_{10} M(<2R_{\rm vir}) = f(\log_{10} R_{\rm sh})$] relation, showing again how the inclusion of shock intensity information makes the predictions considerably more precise.

At $z = 0$, the residuals with respect to our scaling relation are presented as a function of the true mass in the right-hand side panel of Fig. 3. The mean of the probability distribution of $\log_{10} M_{\rm fit}/M_{\rm true}$ (thick, red line) presents a noticeable, although rather small ($\lesssim 0.03$ dex, corresponding to a $\lesssim 7\%$ bias), trend with mass. Note, however, that the mode of the distribution of $\log_{10} M_{\rm fit}/M_{\rm true}$ peaks at around 0.07 dex for $\log_{10} M(<2R_{\rm vir}) \lesssim 13.8$, meaning that the most likely bias on the mass determination of a single low-mass group would be $\approx +17\%$. This behaviour, which is consistent with redshift (see Supplementary Section H), is a natural consequence of having assumed a linearised model (equation 3). While this model captures to a reasonably high extent the trends of the data, the fact that this behaviour with mass is consistent accross cosmic times, and not just a result of a statistical fluctuation, would imply that it can be easily corrected. We defer this possibility for future works using simulations with more statistics.

Discussion

By using cosmological simulations, which track the coupled evolution of baryons and DM through the formation history of galaxy groups and clusters, we have studied the locations, intensities, and evolution of the outermost accretion shocks linked to these structures, where the collisional nature of baryonic matter enters into play and produces distinctive features on their structural, dynamical, thermodynamical and, potentially, observable profiles.

Our main results show that location and intensity of these shocks, together with the total mass of the hosting structures, lie on a two-dimensional plane. We have calibrated this relation to provide an empirical, fitting formula for obtaining the total (dark and baryonic) mass of galaxy groups and clusters within large aperture radii of $2R_{\rm vir}$, valid in the broad redshift interval $0 \leq z \lesssim 1.5$.

From the numerical standpoint, future simulation works with enhanced statistics should go in the direction of confirming this relation with varying physical models and simulation codes, as well as deepening our understanding on possible residual biases. From the physical perspective, while the impact of star formation physics and associated feedback mechanisms are

unlikely to make any noticeable difference on the fate of external shocks (see, for example, [6]), other physical processes, such as physical viscosity or thermal conduction, might be worth considering when describing cluster outskirts [7, 33].

The determination of galaxy cluster masses is a long-standing problem that has been studied extensively and intensively in the literature. Several methods and observational strategies have been designed to tackle this issue, each of them having its own limitations and intrinsic biases. For instance, using galaxy kinematics, [34] quote a 35% scatter between caustic and true masses. Studies based on the hydrostatic equilibrium assumption estimate masses that are affected from a bias from the breach of this condition, which usually lies in the range $(10-30)\%$, as reported by different works [35, 14, 36], while other factors (e.g., gas clumping, temperature inhomogeneities, etc.) may contribute to the scatter with similar magnitude [37]. Regarding weak-lensing masses, the uncertainties introduced by the contamination by background, unlensed galaxies may reach for errors up to 40% [38], to be added to uncertainties due to the mass modelling in the order of $\sim 10\%$ [39].

In this context, we put forward the idea that, provided that new strategies and observational facilities will detect and characterise these very large scale shocks, the results in the present work would allow to measure galaxy clusters masses within large apertures with 1σ errors in the order of $\sim 30\%$. Given these error figures, this novel method would become a new and independent manner to measure galaxy cluster masses that would be fully complementary, and with similar degree of uncertainty, to the previously mentioned procedures.

As a result, our work would lead to an independent method to indirectly constrain the dark matter content in cosmological structures, and hence the currently accepted ΛCDM paradigm.

Methods

In this section, we describe the simulation where the results have been extracted from, the algorithm employed to detect and characterise shocks in the simulation outputs, the sample of galaxy groups and clusters extracted from the simulation, the method devised to identify the outer accretion shock of each object from our sample and, finally, the procedure we follow to perform the fits shown in this article.

The simulation

In this work, we analyse the outcomes of a ΛCDM simulation of a $L^3 = (100\,h^{-1}\,\mathrm{Mpc})^3$ cubic volume, run with the adaptive-mesh refinement (AMR; 40) code `MASCLET` [41]. The evolution of the collisionless dynamics of dark matter is addressed by means of a multilevel and multispecies particle-mesh (PM; 42), which takes advantage of the AMR grid to increase the force resolution of a monolithic PM as long as the mesh is refined. For the baryonic component, we make use of Eulerian, high-resolution shock-capturing techniques (higher-order versions of the method by [43]), based on the piecewise parabolic method (PPM; 44, 45). These methods provide a faithful description of flow discontinuities and, therefore, are especially well-suited to handle shock waves without the need of introducing explicit artificial viscosity, and accurately describing the energy conversion processes (see, e.g., the discussion in [46]).

The simulation assumes periodic boundary conditions and a flat cosmology with fiducial values for the cosmological parameters, which are consistent with the latest Planck Collaboration reported values [47]: matter density parameter $\Omega_\mathrm{m} = 0.31$, baryonic density parameter $\Omega_\mathrm{b} = 0.048$, Hubble dimensionless parameter $h = 0.678$. The initial conditions correspond to a realisation of a Gaussian random field with spectrum $P(k) = Ak^{n_s}T(k)$, whose spectral index is $n_s = 0.96$ and the amplitude yields a normalisation $\sigma_8 = 0.82$ on $8\,h^{-1}$ Mpc scales. $T(k)$ is a CDM transfer function at $z = 1000$ [48]. These conditions are evolved, using the Zeldovich approximation [1], up to $z_\mathrm{ini} = 100$, where the evolution with `MASCLET` starts. In order to refine the initial conditions taking advantage of the AMR scheme, we perform a first evolution on a $N_x^3 = 256^3$ cubic grid, from whose results at $z_\mathrm{fin} = 0$ we choose the Lagrangian regions which will be mapped with enhanced resolution (using three nested levels of initial conditions) back at z_ini, with a final best DM mass resolution of $1.48 \times 10^7\,M_\odot$ and a total of around 392 million DM particles.

During the evolution, we dynamically and recursively refine regions based on a pseudo-Lagrangian criterion (local DM or baryonic density), as well as other criteria based on converging flows, Jeans length, or the presence of DM particles from refined regions in the initial conditions. With $n_\ell = 6$ refinement levels and $\Delta x_{\ell+1}/\Delta x_\ell = 1/2$, we achieve a peak resolution of $\Delta x_6 \simeq 9\,\mathrm{kpc}$. Besides gravity and hydrodynamics, the simulation includes cooling [49] and a parametrization of heating from an extragalactic UV background [50], but does not include star formation nor other feedback mechanism. While feedback from supernovae and active galactic nuclei may have an important impact on the distribution of weak, internal shocks, they are most likely irrelevant for the study of accretion shocks [see, e.g., 51].

The shock finder

In post-processing, shock waves in each snapshot of the simulation are detected and characterised making use of the shock finder presented by [52], which exploits the whole multi-resolution information of the outputs of our AMR code. The basic steps can be summarised as follows:

1. Tentative shocked cells are flagged as those with a converging gas flow ($\nabla\cdot\vec{v} < 0$) and aligned temperature and entropy gradients ($\nabla T\cdot\nabla S > 0$).
2. For each tentative shocked cell, we move to its left and to its right along the x axis (equivalently, we repeat the process with the y and z axes), until we reach a non-shocked cell. We shall refer to these cells with the subindex 'pre' and 'post' (standing for preshock [or upstream of the shock] and postshock [or downstream of the shock], respectively), where $T_{\mathrm{post},x} > T_{\mathrm{pre},x}$. For a cell to be shocked, we also require that the gas densities verify $\rho_{\mathrm{post},x} > \rho_{\mathrm{pre},x}$ in consistency with the Rankine-Hugoniot jump conditions [e.g., 53].
3. The one-dimensional Mach number, $\mathscr{M}_x$, is computed from the Rankine-Hugoniot relation,

$$\frac{T_{\mathrm{post},x}}{T_{\mathrm{pre},x}} = \frac{(5\mathscr{M}_x^2-1)(\mathscr{M}_x^2+3)}{16\mathscr{M}_x^2}, \tag{7}$$

whose inversion yields

$$\mathscr{M}_x^2 = \frac{8q-7+4\sqrt{4q^2-7q+4}}{5}, \quad \text{with } q \equiv \frac{T_{\mathrm{post},x}}{T_{\mathrm{pre},x}} > 1. \tag{8}$$

4. The three-dimensional Mach number is obtained when combining the three directions as $\mathscr{M} = \left(\mathscr{M}_x^2 + \mathscr{M}_y^2 + \mathscr{M}_z^2\right)^{1/2}$.

Thus, the shock finder corresponds to a coordinate-splitting strategy [23, 46], in which projection effects and numerical artefacts due to the Cartesian geometry are minimised by the averaging over the three spatial directions. This artificially imposed geometry is overcome by other family of shock finders, often dubbed *coordinate-unsplit*, which use the direction of the local temperature gradient to identify the pre- and post-shock states [54, 55]. While splitting and unsplit algorithms differ significantly when dealing with weak, internal shocks, remarkable convergence is found for strong, external shocks [54, their figure 4]. Finally, while there are other possibilities to solve for the Mach number than the temperature jump, such as the velocity jump, [46, 56] report generally comparable results for both methods.

Halo catalogues

For each snapshot of the simulation, we have extracted halo catalogues using the public halo finder `ASOHF` [57, 58, 59], which uses the spherical overdensity definition [60] to delimit the extent of haloes around density peaks, together with a number of particularities regarding unbinding mechanisms and handling of substructures. The extent of haloes is defined in our case by the virial radius, $R_{\rm vir}$, which is the radius enclosing an overdensity $\Delta_c = \langle\rho\rangle/\rho_{\rm c} = 18\pi^2 + 82x - 39x^2$, where $x = \Omega_m(z) - 1$ and $\Omega_m(z)$ being the matter density parameter at a redshift z [61].

At $z = 0$, the simulated volume contains 31 well-resolved clusters with masses above $10^{14}\,M_\odot$ (the largest of them having a mass of $5.7 \times 10^{14}\,M_\odot$), as well as 358 groups with DM masses exceeding $10^{13}\,M_\odot$.

Merger trees

Haloes amongst different snapshots are connected using the auxiliary `mtree.py` code of the `ASOHF` package, which identifies all objects contributing to a particular halo in a posterior snapshot. In particular, this code also determines the main progenitor of each halo at the immediately previous snapshot, using a criterion based on the most-bound particles. For more details, we refer the interested reader to the original paper of the halo finder [59]. Using the main branch we can compute the accretion rate proxy Γ_Δ [62],

$$\Gamma_\Delta = \frac{\Delta \log M_\Delta}{\Delta \log a}, \tag{9}$$

where $a = (1+z)^{-1}$ is the scale factor of the FLRW metric, and M_Δ is any spherical-overdensity mass. While in [63] we characterised the instantaneous accretion rates using Savitzky-Golay [64] filters, here we have chosen to compute the accretion rate over the last dynamical time, given that the main effect we are studying in this work, i.e. the propagation of the external accretion shock, is the result of the continued accretion history over the object's history, rather than a quantity linked to the instantaneous growth of the halo at a particular overdensity.

Dynamical state classification

At any given time, we split our sample of galaxy groups and clusters in a totally relaxed, a marginally relaxed and a disturbed class, following the methodology introduced by [65], who calibrated redshift-dependent thresholds and weights on several parameters, such as centre offset (Δ_r), virial ratio (η), mean radial velocity ($\langle\tilde{v}_r\rangle$), sparsity ($s_{200c,500c}$) and three-dimensional ellipticity ε of the dark matter halo, so as to correlate with the presence of mergers and/or strong accretion periods.

By $z = 0$, out of the 31 clusters, 5 (16%) are classified as totally relaxed, 15 (58%) as marginally relaxed, and the rest (11, a 35%) are disturbed.

Accretion shock characterisation

While above we have described the procedure to identify shock waves throughout the computational domain, here we describe the procedure by which we identify the outermost accretion shock of each group or cluster, and how we characterise it.

We start by computing directional profiles of the Mach number, using $N_\theta \times N_\phi = 50 \times 50$ bins in solid angle, equally spaced in $\cos\theta$ and ϕ. Alongside each angular direction, we take the directional profile of the Mach number, $\mathscr{M}$, from $r_{\rm min} = R_{\rm vir}$ to $r_{\rm max} = 5R_{\rm vir}$, using logarithmically spaced bins with $\Delta r = 0.01$ dex. For these regards, unshocked cells are considered to have $\mathscr{M} = 0$. The directional profiles are taken using linear interpolation from the data at the highest resolution (coarser than Δr) available at each point.

For each directional profile, $\mathscr{M}_{\theta,\phi}(r)$, we consider it crosses the accretion shock surface at a radial distance $r = R_{\rm sh}(\theta,\phi)$ if this distance corresponds to the largest local maximum of $\mathscr{M}_{\theta,\phi}(r)$ in the interval $[r_{\rm min}, r_{\rm max}]$, and its value exceeds a given threshold on the Mach number to be regarded as a strong shock, which we have set at $\mathscr{M}_{\rm thr}^{\rm strong} = 10$. Naturally, it may happen that these conditions are not met for a particular angular bin, in which case we consider that there is no crossing with the

accretion shock shell in this direction. Objects where $R_{\rm sh}(\theta,\phi)$ is not identified in at least 75% of the directions are purged from the sample.

In Suppl. Fig. 1, we show several visual impressions of the accretion shock shells recovered by this process, in order to demonstrate the robustness of the procedure. Besides the characterisation of $R_{\rm sh}(\theta,\phi)$, we also extract additional information, which we describe below:

Accretion shock equivalent radius. In order to characterise the size of the accretion shock shell, we aim to obtain a single equivalent, or *effective* radius, $R_{\rm sh}$. Near the interface between the accretion shock of the halo (where matter is infalling smoothly) and a filament, the radius may get increased (see Suppl. Fig. 1, e.g. its left panel), causing the distribution of values of $R_{\rm sh}(\theta,\phi)$ to be right-skewed. Therefore, the mean may be strongly biased high by these directions, even if they correspond to a small fraction of the solid angle around the halo. While the median is a more robust indicator, it is also sensitive to the presence of extreme values, especially when the number of directions ($\leq N_\theta N_\phi = 2500$) is not very high. The mode, or the most probable value of the distribution of $R_{\rm sh}(\theta,\phi)$, however, is insensitive to this skewness.

In order to determine the mode of the values of $R_{\rm sh}(\theta,\phi)$, which is a continuous variable, we estimate its probability density function (PDF), $f(R)$, by means of a kernel density estimation using Gaussian kernels (see Suppl. Fig. 2), whose bandwidth is fixed according to Scott's rule [66]. The shock radius is then set to the absolute maximum of the resulting PDF, $R_{\rm sh} = \operatorname{argmax} f(R)$.

Accretion shock Mach number. From the accretion shock shell, we can obtain the distribution of Mach number across it in a straightforward way, by setting $\mathscr{M}_{\rm sh}(\theta,\phi) = \mathscr{M}_{\theta,\phi}\left(r = R_{\rm sh}(\theta,\phi)\right)$. In order to obtain the characteristic strength of the shock, we define the effective Mach number as the average of $\mathscr{M}_{\rm sh}(\theta,\phi)$ over the shock shell.

Fits for the multi-dimensional scaling relations

We describe here the procedure we follow to fit the multi-dimensional (linear) scaling relations that we show in Fig. 1. We may exemplify it here with a bivariate relation, $Z(X,Y)$, although the procedure is totally analogous for relations of higher or lower number of independent variables.

Outlier removal

Even though the variable Z may be strongly correlated to X and Y, the presence of outliers can significantly bias the fitting of linear relations and, thus, hinder our ability to recover the underlying relation. These outliers can correspond to a variety of effects, ranging from underresolved objects, to complex configurations where the outer shock detection algorithm has not reached a satisfactory solution, or either it is strongly disrupted. Therefore, before proceeding with the fits, we clean the sample from data points which depart significantly from the underlying trend.

Since Z will most often be the mass, and this variable is not uniformly distributed (but instead there is an important preponderance of low-mass objects with respect to high-mass ones) care must be taken in this procedure so as not to get rid of the objects in the high-mass range.

To this end, we compute the conditional probability distribution of Z given (X,Y), by

$$\rho(Z|X,Y) = \frac{\rho(X,Y,Z)}{\rho(X,Y)}, \tag{10}$$

where $\rho(X,Y)$ is the two-dimensional probability density function of X and Y (marginalised over Z), and $\rho(X,Y,Z)$ is the complete, three-dimensional probability density function of X, Y and Z. Both of them are estimated via a Gaussian kernel density estimation procedure, in a similar manner to what we have described before for the accretion shock location procedure. Then, we choose to remove the 5% most unlikely data points (the ones with smallest value of $\rho(Z|X,Y)$) from our sample to fit our scaling relations. While this threshold is arbitrary, in our experiments it seems sufficient to prune the values that are visually far away from the general trend.

Filtering out of objects with a high uncertainty in $R_{\rm sh}$. In some cases, as discussed in Supplementary Section A (see the right panel of Suppl. Fig. 1), the distribution of values of $R_{\rm sh}(\theta,\phi)$ is not monomodal, thus difficulting the determination of the equivalent radius, $R_{\rm sh}$. Aiming to filter out these cases to prevent these uncertainties from propagating to our results, we have constructed an ad-hoc criterion to flag them. Given a kernel-density estimate (KDE) distribution $f(R)$ of the values of $R_{\rm sh}(\theta,\phi)$, we consider it to be multimodal (and thus we discard it) if there is at least one local minimum of f that simultaneously fulfils:

- Its value is below 0.8 times the minimum of the relative maxima that surround it.
- Both these maxima are above 0.5 times the absolute maximum of f.

Strictly, f would be multimodal if there it has at least one relative minimum. The first condition prevents that objects are discarded due to a relative minimum between two maxima that is not significant, and may be present due to sampling noise. The second condition prevents that we discard objects that exhibit fluctuations in the tails of the KDE, which do not introduce confusion in our determination of $R_{\rm sh}$.

Fit using principal component analysis

We describe here the process we follow in order to find the best plane (or, equivalently, the line or the hyperplane, for fewer or higher dimensions) fitting our data. Although a straightforward possibility would be to use ordinary least squares (OLS) to fit $Z(X,Y)$ to the desired functional form, we do not pursue this procedure here. The main reason behind this is the fact that, by minimising the residuals between Z_i and $Z(X_i,Y_i)$, the symmetry amongst variables is broken, i.e., the result of a fit $Z(X,Y)$ would be different from a fit $X(Y,Z)$ or $Y(X,Z)$. Aiming to find a general relation between these variables, we find no reason to break this symmetry, and perform instead a total least squares fit (TLS), where the distance from the points to the plane is minimised (instead of just the Z-axis distance, as it is done with OLS). Naturally, in doing these fits, we work with standardised variables, since otherwise the distances we made reference to, and hence our resulting fit, would depend on the unit system,

$$\tilde{x} = \frac{X-\mu_x}{\sigma_x}, \quad \tilde{y} = \frac{Y-\mu_y}{\sigma_y}, \quad \tilde{z} = \frac{Z-\mu_z}{\sigma_z}, \tag{11}$$

where $\mu_{x(,y,z)}$ and $\sigma_{x(,y,z)}$ are, respectively, the mean and standard deviation of $X(,Y,Z)$.

To find the plane that these data best fit to, we apply standard principal component analysis (PCA; [67]). We compute the covariance matrix,

$$\Sigma_{ij} = \sum_{k=1}^{N} \tilde{x}^i_{(k)} \tilde{x}^j_{(k)}, \qquad \text{where } \{\tilde{x}^i_{(k)}\}_{i=1}^3 = \{\tilde{x}_{(k)}, \tilde{y}_{(k)}, \tilde{z}_{(k)}\} \tag{12}$$

which is symmetric and hence diagonalisable with real eigenvalues and orthogonal eigenvectors. Therefore, there exists a basis in which the correlation matrix is diagonal, implying that a change of basis would make the components uncorrelated. These are the so-called principal components ($\{PC_i\}_{i=1}^3$). The eigenvalue associated to each principal component represents, if properly normalised, the fraction of the variance explained by this component (λ_i). Therefore, if we sort the principal components in non-increasing order of their eigenvalues ($\lambda_1 \geq \lambda_2 \geq \lambda_3$), then the best-fitting plane will be the one spanned by PC_1 and PC_2, i.e., it will be orthogonal to the third eigenvector. It must be borne in mind that this orthogonality is not preserved when going back to the original variables from the standardised ones. Therefore, if we call $\vec{v}_3$ the eigenvector associated to PC_3, the best-fitting plane would then be

$$v_3^{(1)}\tilde{x} + v_3^{(2)}\tilde{y} + v_3^{(3)}\tilde{z} = 0, \tag{13}$$

which, converted back to the original variables would read:

$$Z = \frac{\sigma_z}{v_3^{(3)}} \left[\left(\frac{v_3^{(1)}\mu_x}{\sigma_x} + \frac{v_3^{(2)}\mu_y}{\sigma_y} + \frac{v_3^{(3)}\mu_z}{\sigma_z} \right) - \frac{v_3^{(1)}}{\sigma_x} X - \frac{v_3^{(2)}}{\sigma_y} Y \right]. \tag{14}$$

This methodology can be extended in a straightforward way to fit a $(N-1)$-dimensional hyperplane to N-dimensional data.

Fits for the evolution of the parameters

In order to produce the fits shown in Fig. 2, we have followed a similar methodology to the one described in [65]. The evolution with redshift of each of the quantities shown is fitted, by ordinary least squares weighted to the inverse variance (which had been estimated through bootstrap resampling), by polynomials of increasing degree, until the highest degree coefficient is insignificant (p-value above 0.046 or reduced chi-squared below 0.25).

Acknowledgements
This work has been supported by the Agencia Estatal de Investigación Española (AEI; grant PID2022-138855NB-C33), by the Ministerio de Ciencia e Innovación (MCIN) within the Plan de Recuperación, Transformación y Resiliencia del Gobierno de España through the project ASFAE/2022/001, with funding from European Union NextGenerationEU (PRTR-C17.I1), and by the Generalitat Valenciana (grant CIPROM/2022/49). DVP acknowledges support from Universitat de València through an Atracció de Talent fellowship. Simulations have been carried out using the supercomputer Lluís Vives at the Servei d'Informàtica of the Universitat de València.

Author contributions statement
VQ initiated the project. VQ and DVP ran the cosmological simulation. SP and DVP developed the shock and halo finders. DVP performed the data analysis and produced the figures. DVP, SP, and VQ discussed the results and wrote this paper.

Additional information
The authors declare no competing interest.

Data availability
The data underlying this article will be shared upon reasonable request to the corresponding author.

Code availability
The halo finder `ASOHF` is publicly available in `https://github.com/dvallesp/ASOHF`. The shock finder, the simulation code (`MASCLET`) and the codes for analysing the output data and producing the figures will be shared upon reasonable request to the corresponding author.

References
1. Zel'dovich, Y. B. Gravitational instability: An approximate theory for large density perturbations. *Astron. Astrophys.* **5**, 84–89 (1970).
2. Press, W. H. & Schechter, P. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *Astrophys. J.* **187**, 425–438, DOI: 10.1086/152650 (1974).
3. Gott, I., J. R. & Rees, M. J. A theory of galaxy formation and clustering. *Astron. Astrophys.* **45**, 365–376 (1975).
4. Böhringer, H. & Werner, N. X-ray spectroscopy of galaxy clusters: studying astrophysical processes in the largest celestial laboratories. *Astron. Astrophys. Rev.* **18**, 127–196, DOI: 10.1007/s00159-009-0023-3 (2010).
5. Kravtsov, A. V. & Borgani, S. Formation of Galaxy Clusters. *Annu. Rev. Astron. Astrophys.* **50**, 353–409, DOI: 10.1146/annurev-astro-081811-125502 (2012). 1205.5556.
6. Planelles, S., Schleicher, D. R. G. & Bykov, A. M. Large-Scale Structure Formation: From the First Non-linear Objects to Massive Galaxy Clusters. *Space Sci. Rev.* **188**, 93–139, DOI: 10.1007/s11214-014-0045-7 (2015). 1404.3956.
7. Walker, S. *et al.* The Physics of Galaxy Cluster Outskirts. *Space Sci. Rev.* **215**, 7, DOI: 10.1007/s11214-018-0572-8 (2019). 1810.00890.
8. Tozzi, P. & Norman, C. The Evolution of X-Ray Clusters and the Entropy of the Intracluster Medium. *Astrophys. J.* **546**, 63–84, DOI: 10.1086/318237 (2001). astro-ph/0003289.
9. Nagai, D., Kravtsov, A. V. & Vikhlinin, A. Effects of Galaxy Formation on Thermodynamics of the Intracluster Medium. *Astrophys. J.* **668**, 1–14, DOI: 10.1086/521328 (2007). astro-ph/0703661.
10. Bykov, A. M. *et al.* Structures and Components in Galaxy Clusters: Observations and Models. *Space Sci. Rev.* **188**, 141–185, DOI: 10.1007/s11214-014-0129-4 (2015). 1512.01456.
11. Allen, S. W., Evrard, A. E. & Mantz, A. B. Cosmological Parameters from Observations of Galaxy Clusters. *Annu. Rev. Astron. Astrophys.* **49**, 409–470, DOI: 10.1146/annurev-astro-081710-102514 (2011). 1103.4829.
12. Weinberg, D. H. *et al.* Observational probes of cosmic acceleration. *Phys. Rep.* **530**, 87–255, DOI: 10.1016/j.physrep.2013.05.001 (2013). 1201.2434.
13. Clerc, N. & Finoguenov, A. *X-Ray Cluster Cosmology*, 1–52 (Springer Nature Singapore, Singapore, 2022).
14. Biffi, V. *et al.* On the Nature of Hydrostatic Equilibrium in Galaxy Clusters. *Astrophys. J.* **827**, 112, DOI: 10.3847/0004-637X/827/2/112 (2016). 1606.02293.
15. Ettori, S. *et al.* Hydrostatic mass profiles in X-COP galaxy clusters. *Astron. Astrophys.* **621**, A39, DOI: 10.1051/0004-6361/201833323 (2019). 1805.00035.
16. Lovisari, L. & Maughan, B. J. *Scaling Relations of Clusters and Groups and Their Evolution*, 1–50 (Springer Nature Singapore, Singapore, 2022).

17. Giodini, S. *et al.* Scaling Relations for Galaxy Clusters: Properties and Evolution. *Space Sci. Rev.* **177**, 247–282, DOI: 10.1007/s11214-013-9994-5 (2013). 1305.3286.

18. Umetsu, K. Cluster-galaxy weak lensing. *Astron. Astrophys. Rev.* **28**, 7, DOI: 10.1007/s00159-020-00129-w (2020). 2007.00506.

19. Diaferio, A. Mass estimation in the outer regions of galaxy clusters. *Mon. Not. R. Astron. Soc.* **309**, 610–622, DOI: 10.1046/j.1365-8711.1999.02864.x (1999). astro-ph/9906331.

20. Pratt, G. W. *et al.* The Galaxy Cluster Mass Scale and Its Impact on Cosmological Constraints from the Cluster Population. *Space Sci. Rev.* **215**, 25, DOI: 10.1007/s11214-019-0591-0 (2019). 1902.10837.

21. Quilis, V., Ibáñez, J. M. & Sáez, D. On the role of shock waves in galaxy cluster evolution. *The Astrophys. J.* **502**, 518, DOI: 10.1086/305932 (1998).

22. Miniati, F. *et al.* Properties of cosmic shock waves in large-scale structure formation. *The Astrophys. J.* **542**, 608, DOI: 10.1086/317027 (2000).

23. Ryu, D., Kang, H., Hallman, E. & Jones, T. W. Cosmological Shock Waves and Their Role in the Large-Scale Structure of the Universe. *Astrophys. J.* **593**, 599–610, DOI: 10.1086/376723 (2003). astro-ph/0305164.

24. Bertschinger, E. Cosmological self-similar shock waves and galaxy formation. *Astrophys. J.* **268**, 17–29, DOI: 10.1086/160925 (1983).

25. Shi, X. Locations of accretion shocks around galaxy clusters and the ICM properties: insights from self-similar spherical collapse with arbitrary mass accretion rates. *Mon. Not. R. Astron. Soc.* **461**, 1804–1815, DOI: 10.1093/mnras/stw1418 (2016). 1603.07183.

26. Zhang, C., Zhuravleva, I., Kravtsov, A. & Churazov, E. Evolution of splashback boundaries and gaseous outskirts: insights from mergers of self-similar galaxy clusters. *Mon. Not. R. Astron. Soc.* **506**, 839–863, DOI: 10.1093/mnras/stab1546 (2021). 2103.03850.

27. Aung, H., Nagai, D. & Lau, E. T. Shock and splash: gas and dark matter halo boundaries around ΛCDM galaxy clusters. *Mon. Not. R. Astron. Soc.* **508**, 2071–2078, DOI: 10.1093/mnras/stab2598 (2021). 2012.00977.

28. Anbajagane, D. *et al.* Shocks in the stacked Sunyaev-Zel'dovich profiles of clusters II: Measurements from SPT-SZ + Planck Compton-y map. *Mon. Not. R. Astron. Soc.* **514**, 1645–1663, DOI: 10.1093/mnras/stac1376 (2022). 2111.04778.

29. Keshet, U., Kushnir, D., Loeb, A. & Waxman, E. Preliminary Evidence for a Virial Shock around the Coma Galaxy Cluster. *Astrophys. J.* **845**, 24, DOI: 10.3847/1538-4357/aa794b (2017). 1210.1574.

30. Reiss, I. & Keshet, U. Detection of virial shocks in stacked Fermi-LAT galaxy clusters. *J. Cosmol. Astropart. Phys.* **2018**, 010, DOI: 10.1088/1475-7516/2018/10/010 (2018). 1705.05376.

31. Holguin Luna, P. & Burchett, J. Localizing the accretion shock and constraining gaseous conditions in galaxy cluster outskirts with UV absorption spectroscopy. In *American Astronomical Society Meeting Abstracts*, vol. 54 of *American Astronomical Society Meeting Abstracts*, 427.02 (2022).

32. Vernstrom, T. *et al.* Polarized accretion shocks from the cosmic web. *Sci. Adv.* **9**, eade7233, DOI: 10.1126/sciadv.ade7233 (2023). 2302.08072.

33. Walker, S. & Lau, E. Cluster Outskirts and Their Connection to the Cosmic Web. In *Handbook of X-ray and Gamma-ray Astrophysics*, 13, DOI: 10.1007/978-981-16-4544-0_120-1 (Springer Living Reference Work, 2022).

34. Andreon, S., Trinchieri, G., Moretti, A. & Wang, J. Intrinsic scatter of caustic masses and hydrostatic bias: An observational study. *Astron. Astrophys.* **606**, A25, DOI: 10.1051/0004-6361/201730920 (2017). 1706.08353.

35. Lau, E. T., Kravtsov, A. V. & Nagai, D. Residual Gas Motions in the Intracluster Medium and Bias in Hydrostatic Measurements of Mass Profiles of Clusters. *Astrophys. J.* **705**, 1129–1138, DOI: 10.1088/0004-637X/705/2/1129 (2009). 0903.4895.

36. Angelinelli, M. *et al.* Turbulent pressure support and hydrostatic mass bias in the intracluster medium. *Mon. Not. R. Astron. Soc.* **495**, 864–885, DOI: 10.1093/mnras/staa975 (2020). 1905.04896.

37. Rasia, E. *et al.* Temperature Structure of the Intracluster Medium from Smoothed-particle Hydrodynamics and Adaptive-mesh Refinement Simulations. *Astrophys. J.* **791**, 96, DOI: 10.1088/0004-637X/791/2/96 (2014). 1406.4410.

38. Okabe, N. & Smith, G. P. LoCuSS: weak-lensing mass calibration of galaxy clusters. *Mon. Not. R. Astron. Soc.* **461**, 3794–3821, DOI: 10.1093/mnras/stw1539 (2016). 1507.04493.

39. Meneghetti, M. *et al.* Weighing simulated galaxy clusters using lensing and X-ray. *Astron. Astrophys.* **514**, A93, DOI: 10.1051/0004-6361/200913222 (2010). 0912.1343.

40. Berger, M. J. & Colella, P. Local Adaptive Mesh Refinement for Shock Hydrodynamics. *J. Comput. Phys.* **82**, 64–84, DOI: 10.1016/0021-9991(89)90035-1 (1989).

41. Quilis, V. A new multidimensional adaptive mesh refinement hydro + gravity cosmological code. *Mon. Not. R. Astron. Soc.* **352**, 1426–1438, DOI: 10.1111/j.1365-2966.2004.08040.x (2004). astro-ph/0405389.

42. Hockney, R. W. & Eastwood, J. W. *Computer simulation using particles* (Institute of Physics Publishing, 1988).

43. Godunov, S. K. & Bohachevsky, I. Finite difference method for numerical computation of discontinuous solutions of the equations of fluid dynamics. *Matemati@CARON@ceskij sbornik* **47(89)**, 271–306 (1959).

44. Colella, P. & Woodward, P. R. The Piecewise Parabolic Method (PPM) for Gas-Dynamical Simulations. *J. Comput. Phys.* **54**, 174–201, DOI: 10.1016/0021-9991(84)90143-8 (1984).

45. Martí, J. M. S. S. & Müller, E. Extension of the Piecewise Parabolic Method to One-Dimensional Relativistic Hydrodynamics. *J. Comput. Phys.* **123**, 1–14, DOI: 10.1006/jcph.1996.0001 (1996).

46. Vazza, F., Brunetti, G. & Gheller, C. Shock waves in Eulerian cosmological simulations: main properties and acceleration of cosmic rays. *Mon. Not. R. Astron. Soc.* **395**, 1333–1354, DOI: 10.1111/j.1365-2966.2009.14691.x (2009). 0808.0609.

47. Planck Collaboration *et al.* Planck 2018 results. VI. Cosmological parameters. *Astron. Astrophys.* **641**, A6, DOI: 10.1051/0004-6361/201833910 (2020). 1807.06209.

48. Eisenstein, D. J. & Hu, W. Baryonic Features in the Matter Transfer Function. *Astrophys. J.* **496**, 605–614, DOI: 10.1086/305424 (1998). astro-ph/9709112.

49. Sutherland, R. S. & Dopita, M. A. Cooling Functions for Low-Density Astrophysical Plasmas. *Astrophys. J. Suppl. Ser.* **88**, 253, DOI: 10.1086/191823 (1993).

50. Haardt, F. & Madau, P. Radiative Transfer in a Clumpy Universe. II. The Ultraviolet Extragalactic Background. *Astrophys. J.* **461**, 20, DOI: 10.1086/177035 (1996). astro-ph/9509093.

51. Planelles, S. *et al.* Exploring the role of cosmological shock waves in the Dianoga simulations of galaxy clusters. *Mon. Not. R. Astron. Soc.* **507**, 5703–5719, DOI: 10.1093/mnras/stab2436 (2021). 2108.09670.

52. Planelles, S. & Quilis, V. Cosmological shock waves: clues to the formation history of haloes. *Mon. Not. R. Astron. Soc.* **428**, 1643–1655, DOI: 10.1093/mnras/sts142 (2013). 1210.1369.

53. Landau, L. D. & Lifshitz, E. M. *Fluid mechanics* (Pergamon, 1959).

54. Skillman, S. W., O'Shea, B. W., Hallman, E. J., Burns, J. O. & Norman, M. L. Cosmological Shocks in Adaptive Mesh Refinement Simulations and the Acceleration of Cosmic Rays. *Astrophys. J.* **689**, 1063–1077, DOI: 10.1086/592496 (2008). 0806.1522.

55. Schaal, K. & Springel, V. Shock finding on a moving mesh - I. Shock statistics in non-radiative cosmological simulations. *Mon. Not. R. Astron. Soc.* **446**, 3992–4007, DOI: 10.1093/mnras/stu2386 (2015). 1407.4117.

56. Vazza, F. *et al.* A comparison of cosmological codes: properties of thermal gas and shock waves in large-scale structures. *Mon. Not. R. Astron. Soc.* **418**, 960–985, DOI: 10.1111/j.1365-2966.2011.19546.x (2011). 1106.2159.

57. Planelles, S. & Quilis, V. ASOHF: a new adaptive spherical overdensity halo finder. *Astron. Astrophys.* **519**, A94, DOI: 10.1051/0004-6361/201014214 (2010). 1006.3205.

58. Knebe, A. *et al.* Haloes gone MAD: The Halo-Finder Comparison Project. *Mon. Not. R. Astron. Soc.* **415**, 2293–2318, DOI: 10.1111/j.1365-2966.2011.18858.x (2011). 1104.0949.

59. Vallés-Pérez, D., Planelles, S. & Quilis, V. The halo-finding problem revisited: a deep revision of the ASOHF code. *Astron. Astrophys.* **664**, A42, DOI: 10.1051/0004-6361/202243712 (2022). 2205.02245.

60. Cole, S. & Lacey, C. The structure of dark matter haloes in hierarchical clustering models. *Mon. Not. R. Astron. Soc.* **281**, 716, DOI: 10.1093/mnras/281.2.716 (1996). astro-ph/9510147.

61. Bryan, G. L. & Norman, M. L. Statistical Properties of X-Ray Clusters: Analytic and Numerical Comparisons. *Astrophys. J.* **495**, 80–99, DOI: 10.1086/305262 (1998). astro-ph/9710107.

62. Diemer, B. & Kravtsov, A. V. Dependence of the Outer Density Profiles of Halos on Their Mass Accretion Rate. *Astrophys. J.* **789**, 1, DOI: 10.1088/0004-637X/789/1/1 (2014). 1401.1216.

63. Vallés-Pérez, D., Planelles, S. & Quilis, V. On the accretion history of galaxy clusters: temporal and spatial distribution. *Mon. Not. R. Astron. Soc.* **499**, 2303–2318, DOI: 10.1093/mnras/staa3035 (2020). 2009.13882.

64. Savitzky, A. & Golay, M. J. E. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **36**, 1627–1639, DOI: 10.1021/ac60214a047 (1964).

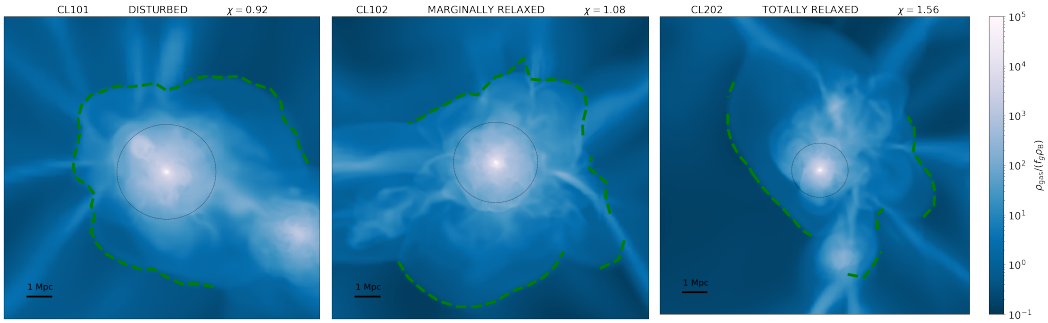
65. Vallés-Pérez, D., Planelles, S., Monllor-Berbegal, Ó. & Quilis, V. On the choice of the most suitable indicator for the assembly state of dark matter haloes through cosmic time. *Mon. Not. R. Astron. Soc.* **519**, 6111–6125, DOI: 10.1093/mnras/stad059 (2023). 2301.02253.

66. Scott, D. W. *Multivariate Density Estimation* (John Wiley & Sons, 1992).

67. Tipping, M. & Bishop, C. Mixtures of probabilistic principal component analysers. *Neural Comput.* **11**, 443–482, DOI: 10.1162/089976699300016728 (1999). Copyright of the Massachusetts Institute of Technology Press (MIT Press).

Supplementary material

A Visual impression of the accretion shock identification procedure



Suppl. Fig. 1. Visual impression of the shock shell identification results. The three panels correspond to three different galaxy clusters at $z = 0$, one of each dynamical state category. The colourmaps show thin (~ 36 kpc) slices of the gas densities (in units of the cosmic mean gas density) through the cluster centres. The virial radius is indicated by the black, dotted circle. The green, dashed lines indicate the shock shell, in the directions where it can be located. The line in the bottom left corner shows a 1 Mpc ruler for visual reference.

With the aim of showcasing the typical appearance of our detected accretion shocks, as well as presenting a visual test of the trustworthiness of our detection process, we present in Suppl. Fig. 1 three panels corresponding to three galaxy clusters at $z = 0$, each of them corresponding to a different dynamical state at the present time according to our classification. In these panels we overplot the the virial sphere (black, dotted line) and the shell corresponding to the accretion shock in the directions it has been detected (green, dashed line).

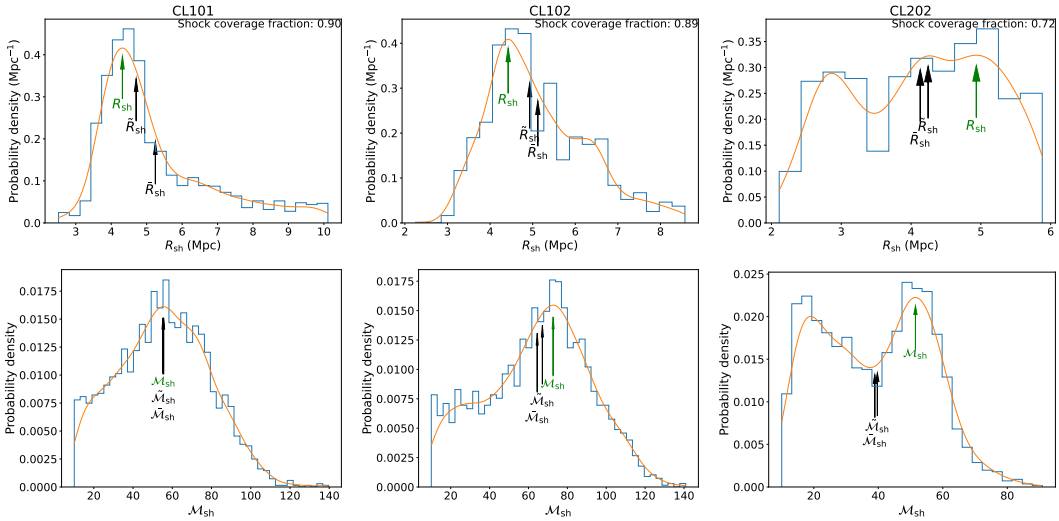
The left panel corresponds to our most massive galaxy cluster, `CL101`, which is disturbed by $z \simeq 0$ due to a recent merger with mass ratio 1:5 at $z \simeq 0.19$, that has proceeded from the bottom-right of the slice. Also in this direction, another massive structure is approaching the cluster. Thus, the accretion shock shell is disrupted in this direction, and correspondingly not detected by our procedure. In the remaining directions, the algorithm delimits correctly the location of the outer accretion shock, which can be seen as a jump in the underlying density map.

The middle panel shows a slice through `CL102`, which is a marginally relaxed cluster which suffered its last merger, with a mass ratio of $\sim$1:9, at $z \simeq 0.3$. In this case, the shock shell is rather asymmetric, mainly due to its anisotropic environment, dominated by several filaments, and to several minor mergers at $z \simeq 0.5$. In this case, it can be seen how our shock shell identification procedure naturally excludes the directions of the filaments, where the accretion shock is not present.

Finally, the right panel of Suppl. Fig. 1 corresponds to a totally relaxed cluster, `CL202`. This object suffered a minor merger (mass ratio of 1:8) at $z \simeq 0.3$, but unlike the example in the middle panel, where the merger was almost head-on, in this case the merger had a larger impact parameter. This may justify why in this case the cluster has had time to fall back to relaxation, when assessed at the virial volume, while outside the virial radius the gas density shows a more disturbed morphology, with a very anisotropic accretion shell. This cluster is on the verge of a merger with another structure, proceeding from the lower part of the slice. It is worth mentioning how, in this situation, our algorithm is picking the accretion shock boundary of the infalling structure as part of the accretion shell of the main cluster. Generally, when two haloes are merging and their accretion shocks join, it is not trivial to establish since when they must be regarded as a single object for the purpose of identifying its joint outer shock. However, our accretion shock equivalent radius definition, based on the mode of the PDF of the angular distribution of distances to the shock shell, unlike other statistics such as the arithmetic mean, gets minimally biased by these events (since the mode is robust to the presence of tails in the distribution).

Given that, in what follows, we aim to represent these shock shells with a single value of a so-called *equivalent* radius, in Suppl. Fig. 2 we comment on this issue by presenting the distribution of radial distances to the accretion shock shell, $R_{\rm sh}(\theta, \phi)$ (top row), and the distribution of Mach numbers through the accretion shock shell, $\mathscr{M}_{\rm sh}(\theta, \phi)$ (bottom row), for the same three clusters displayed in Suppl. Fig. 1.

In most cases (such as `CL101` or `CL102`; left and central figures), the distribution of shock radii is monomodal and right-skewed, implying that the mean and, to a lesser extent, the median get dragged towards larger values by the tails of the distribution. In these cases, however, the peak of the PDF is well-defined and provides a robust measurement of the shock radius, in the sense that it does not depend on the particular value of $r_{\rm max}$ (which has a large influence on the right tail of



Suppl. Fig. 2. Examples on the determination of R_{sh} and $\mathscr{M}_{sh}$, on the same clusters as Suppl. Fig. 1. The top rows present the distribution of values of $R_{sh}(\theta,\phi)$, with a histogram (blue line) and with its corresponding Gaussian kernel density estimation (orange line). The green arrow indicates the most frequent value (i.e., the peak of the probability density function, or mode), which we associate to the shock radius (R_{sh}), while the two black arrows show the mean ($\bar{R}_{sh}$) and median ($\tilde{R}_{sh}$) values, for comparison. The bottom row contains the same information regarding the distribution of $\mathscr{M}_{sh}(\theta,\phi)$. The upper panels also inform about the shock coverage fraction, i.e., the fraction of the solid angle around the object in which we identify the shock cell.

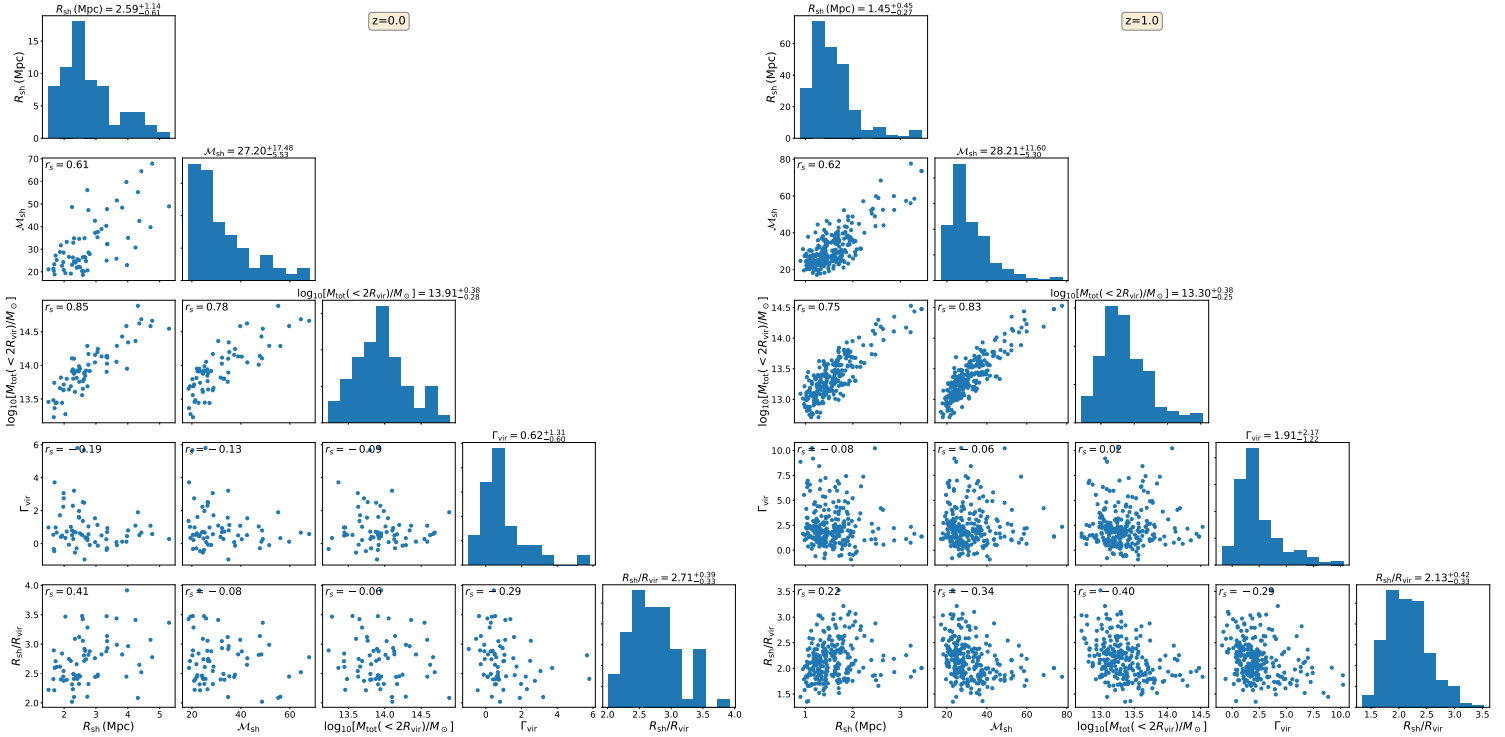
the distribution). In the vast majority of these cases, also the distribution of $\mathscr{M}_{sh}(\theta,\phi)$ is monomodal, with typically small differences between the mean, median and mode of the distribution.

On the other hand, in some cases the distribution of values of $R_{sh}(\theta,\phi)$ might be multimodal. Such is the case of `CL202` (right column), where three peaks are captured by the kernel density estimate. Additionally, in this case, a smaller fraction of the solid angle around the cluster is found to contain the shock shell, indicating a disrupted morphology, most likely associated to the presence of the infalling object in the lower direction (see right panel of Suppl. Fig. 1). Also the Mach number distribution appears to be bimodal in this case. In cases like this, though infrequent, the characterisation is not so straightforward. This is why we have cleaned our sample from multimodal distributions of $R_{sh}(\theta,\phi)$ (see the Methods section).

It is interesting to highlight how shock morphologies, at least from the qualitative point of view, do not seem to be significantly correlated to our usual measures of dynamical state. This must not be surprising, since these measures are associated to a particular aperture (i.e., R_{vir}) much smaller than the typical volume enclosed by the shock shells (a factor $2-3$ in radius). Indeed, we may find clusters that are dynamically disturbed, when looking at the virial volume, that exhibit rather spherical accretion shocks (e.g., `CL101`; see Suppl. Fig. 1); or highly relaxed clusters with more complex-shaped outer boundaries (such as `CL202`).

B Statistical summary of the sample

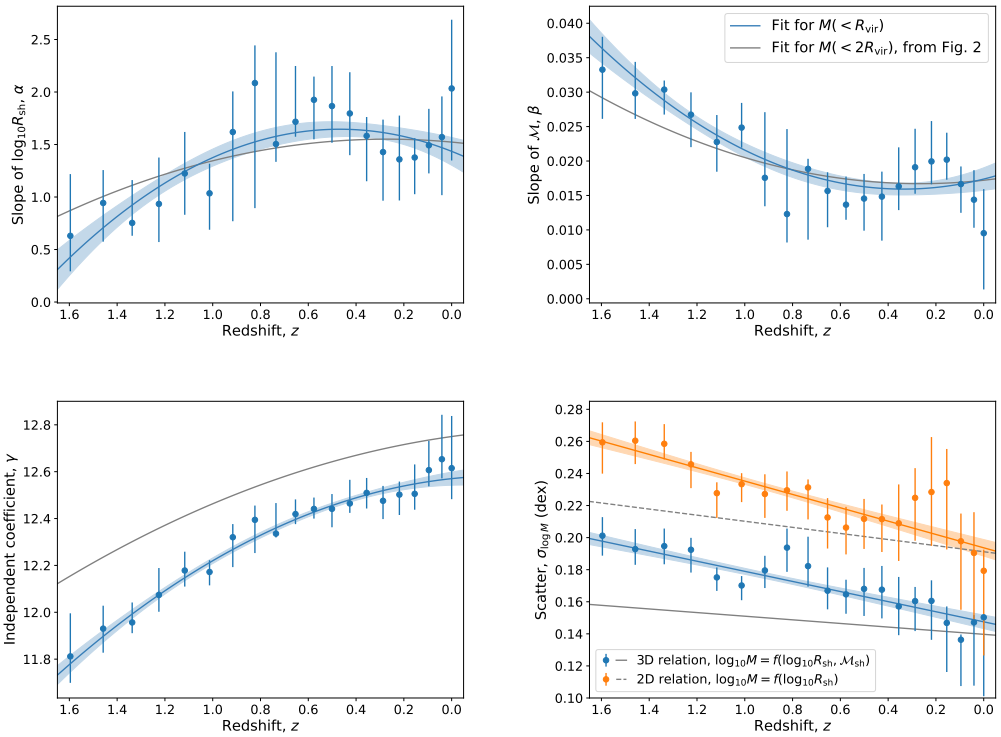
Suppl. Fig. 3 presents a summary, in statistical terms, of our sample, both at $z=0$ and at $z=1$ (left-hand side and right-hand side panels, respectively). For each cosmic epoch, we present a corner-like plot, including the univariate distributions of shock radius R_{sh}, mean Mach number $\mathscr{M}_{sh}$, total mass within $2R_{vir}$, accretion rate Γ_{vir}, and R_{sh}/R_{vir}; and scatter plots between each pair of variables to illustrate their correlations or lack thereof.



Suppl. Fig. 3. Statistical summary of our galaxy clusters and galaxy groups sample. The left-hand side panel corresponds to the sample at $z = 0$, comprising 67 objects. The right-hand side panel contains the same information for $z = 1$, when the sample is composed of 249 objects. The plots over the principal diagonal contain histograms of the corresponding variable (shock equivalent radius, mean Mach number, total mass within $2R_{\rm vir}$, accretion rate $\Gamma_{\rm vir}$, and shock radius in units of the virial radius). Above each of these panels we inform about the median and the scatter (quantified through the distance to the first and third quartiles) of the variable. The off-diagonal panels contain scatter plots for each pair of variables. Written on each of these plots is the Spearman correlation coefficient (r_s) between the corresponding pair of variables.

C Fits for the mass measured in an aperture of $R_{\rm vir}$

While in the results that we report in the main article we consistently use $2R_{\rm vir}$ as our mass definition, and this is the mass that we have proposed to infer using shock sizes and intensities through our fitting relation, there is no particular reason to choose this specific aperture. In principle, the properties of the shock shell would be best correlated to the gravitational mass driving the collapse of the structure and, hence, the generation and propagation of the accretion shock. However, in a fully three-dimensional picture, it is not straightforward to agree on the definition of this mass. Therefore, the mass aperture could be varied in a sensibly large interval, as long as it is still a good estimate with the mass driving the collapse of the galaxy cluster or group.

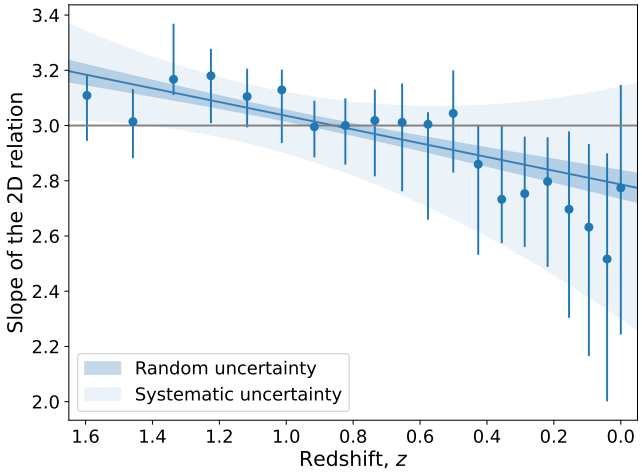


Suppl. Fig. 4. Scaling relation evolution summary, totally analogue to Fig. 2 in the main text, for the fit measuring the masses within $R_{\rm vir}$. Here, all plot elements correspond the fits within $R_{\rm vir}$, except for gray lines, which contain the fits for $2R_{\rm vir}$ that were shown in Fig. 2, for reference. *Top left:* evolution of the coefficient of the shock radius. *Top right:* evolution of the coefficient of Mach number. *Bottom left:* evolution of the independent term. *Bottom right:* evolution of the scatter (in dex) around the three-dimensional scaling relation (blue line), and around the bidimensional scaling relation (orange line).

In Suppl. Fig. 4, we show this by presenting the results for the three-dimensional scaling relation (Eq. 3), taking the masses within $R_{\rm vir}$, instead of $2R_{\rm vir}$. The four panels in the figure are completely analogous to Fig. 2 in the main text, with the addition of the gray lines, which contain the original fits within $2R_{\rm vir}$ for a better visual comparison. Generally, the evolution of the fit coefficients α and β (the coefficients of $\log_{10}R_{\rm sh}$ and $\mathscr{M}$, respectively) with redshift vary minimally, pointing at a consistent behaviour at different radial apertures as stated. Naturally, the normalisation decreases when shifting from $2R_{\rm vir}$ to $R_{\rm vir}$. Regarding the scatter evolution, while at $z \sim 0$ masses can be determined with similar accuracy at both radial volumes, at high redshift choosing the larger aperture reduces very significantly the scatter around our relation.

D Results marginalised over $\mathscr{M}_{\rm sh}$: the 2d relation

The bidimensional relation (i.e., marginalised over the shock intensity, $\log_{10} M(<2R_{\rm vir}) = f(\log_{10} R_{\rm sh})$) is a useful baseline to compare our results with. While in Fig. 2 of the main text we have already compared the magnitude of the scatter around both relations, it is still interesting to study the logarithmic slope of the bidimensional relation. A logarithmic slope of 3 ($M \propto R_{\rm sh}^3$) would indicate a self-similar scaling of the marginalised data (even though, when shock intensity is considered, this additional information is capable of dramatically reducing the scatter). The evolution of this slope of the marginalised relation with cosmic time is shown in Suppl. Fig. 5, which is similar in its presentation to Fig. 2 of the main text. Even though this quantity exhibits a clear evolution, from $\simeq 3.2$ at $z \simeq 1.5$ to $\simeq 2.8$ at $z \simeq 0$, once the uncertainties associated to our limited statistics are considered, the slope of the marginalised relation is consistent with a self-similar behaviour (in the sense that it cannot be ruled out). Far from meaning that $10^{13} M_\odot$ groups and $5 \times 10^{14} M_\odot$ clusters behave likewise (the large scatter around the marginalised relation reflects that the behaviour is far from being that simple), this serves as a sanity check for our shock shell identification procedure.

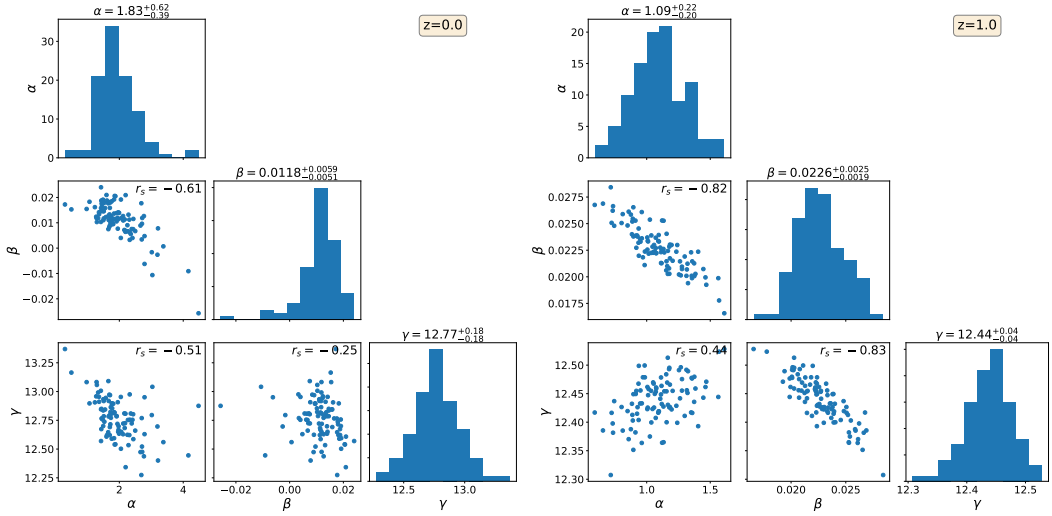


Suppl. Fig. 5. Evolution of the logarithmic slope of the bidimensional relation, $\log_{10} M(<2R_{\rm vir}) = f(\log_{10} R_{\rm sh})$. Here, the dots correspond to the determinations of the slope at each snapshot, with the error bars obtained through bootstrap resampling. The dark contour indicates the 1σ statistical errors associated to the dispersion with respect to the fit, while the light countour contains, in addition to that, the *systematic* uncertainties associated to the magnitude of the error bars.

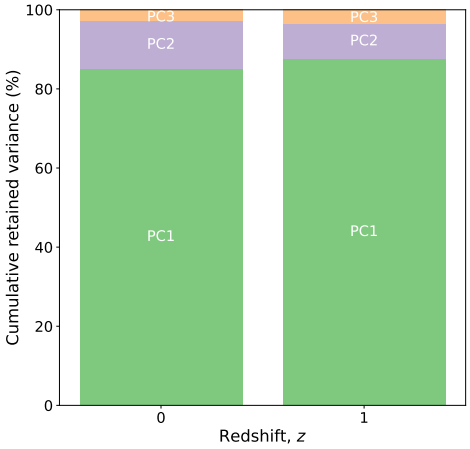
E Further details on the fitted relations

In equations 1 and 2 of the main text we have introduced our best fits for $z = 0$ and $z = 1$. Here, we provide some further information about the results of these fits. In particular, Suppl. Fig. 6 presents graphically the uncertainties and correlations associated to the fit parameters α (the coefficient of $\log_{10} R_{\rm sh}$), β (the cofficient of $\log_{10} M(<2R_{\rm vir})$), and γ (the independent term), for $z = 0$ (left) and $z = 1$ (right). These plots are similar to those of Suppl. Fig. 3, but here each dot corresponds to a different estimation of the fit parameters (which we have produced through bootstrap resampling). Naturally, there is a reasonably high degree of uncertainty on each of the parameters, mostly associated to rather tight correlations amongst themselves. That is to say, upon a particular resampling, the fit might prefer a slightly higher value of the coefficient α, at the expense of a lower value of the coefficient β, hence yielding a noticeable anticorrelation amongst these two variables.

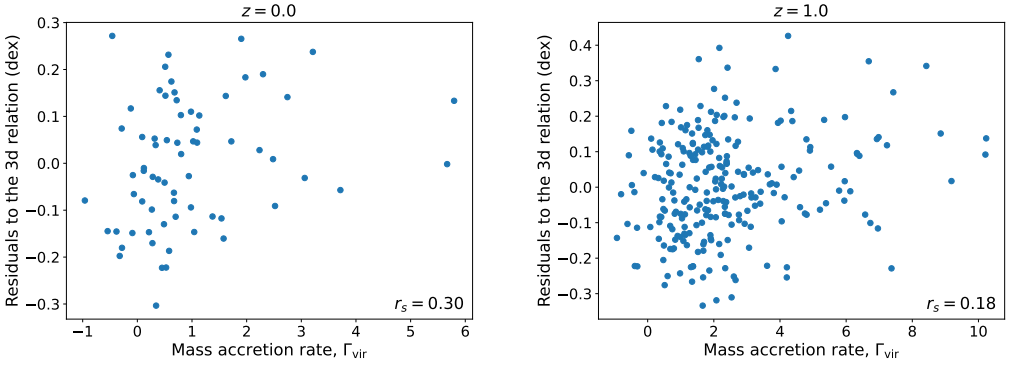
In Suppl. Fig. 7 we show the fraction of variance retained by each of the principal components, again for $z = 0$ (left) and $z = 1$ (right). In both cases, a single principal component is capable of reproducing $\sim 85\%$ of the variance of the standardised variables (implying that an only variable would be enough to describe this data, losing only 15% of the scatter). However, even though the distribution of points in this three-dimensional space is slightly prolate, the second principal component outbalances the third one by a factor of $3-4$ in both cases, implying that, rather than having a component retaining most of the variance



Suppl. Fig. 6. Distribution of the fit parameters in equation (3), at $z = 0$ (left) and at $z = 1$ (right). The panels along the diagonal contain the univariate distributions of each parameter (α being the coefficient of $\log_{10} R_{\rm sh}$, β the cofficient of $\log_{10} M(< 2R_{\rm vir})$, and γ being the independent coefficient), with the value in the title being the median, together with the interquartilic ranges to measure the spread of the distribution. The off-diagonal panels represent the covariances between the fit parameters, with the Spearman rank correlation coefficient (r_s) written at the top of each panel.



Suppl. Fig. 7. Fraction of variance explained by each of the principal components, at $z = 0$ (left) and at $z = 1$ (right).



Suppl. Fig. 8. Correlation between the residuals with respect to the three-dimensional relation (equations 1 and 2; the ones shown in the upper and lower panels of Fig. 1) and the virial, total mass accretion rate, for the best-fits at $z = 0$ and $z = 1$, respectively, for the left-hand side and right-hand side panels. Written at the bottom right of each panel is the Spearman correlation coefficient between this pair of variables.

and two uncorrelated, similarly important variables encapsulating the residuals, the addition of a second variable is capable of breaking most of the degeneracy and hence providing much more precise estimates.

F Relation of the accretion shock location to the virial accretion rate

While simple models for the collapse of an overdensity, assuming for example constant accretion rates, predict that the location of the accretion shock is determined by the accretion rate [24], the reality in a complex, cosmological context, where the growth of an object is the result of the integrated, time-dependent accretion rate, with anisotropic accretion and mergers playing a very significant role, makes these correlations much more insignificant [e.g., 27, their figure 6]. While in the panels of Suppl. Fig. 3 it can already be seen that R_{sh}/R_{vir} is only very loosely correlated to the accretion rate proxy Γ_{vir} (see its definition in the Methods section), in Suppl. Fig. 8 we extend on this topic, by presenting the residuals of our three-dimensional relation (at $z = 0$, left-hand side panel; and at $z = 1$, right-hand side panel) as a function of the accretion rate. These two components are essentially uncorrelated, implying that the addition of information about the mass accretion rate over the last dynamical times brings very limited or no new information.

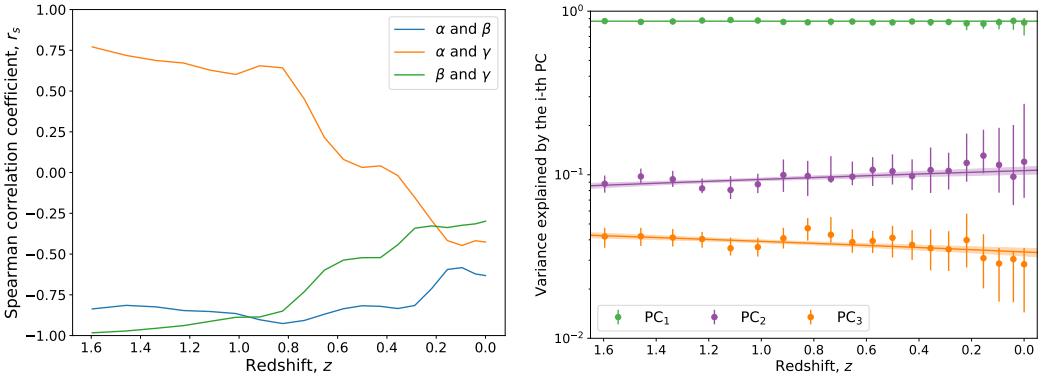
G Evolution of the explained variance fractions, uncertainties and correlations amongst the fit parameters

While the evolution of the intrinsic uncertainties in the determination of each of the coefficients can be seen in Fig. 2 of the main text, similarly to what we have shown in Suppl. Fig. 6 for two specific snapshots of the simulation, in the left-hand side panel of Suppl. Fig. 9 we present the evolution with decreasing redshift of the correlation amongst the fit parameters, obtained in an analogous manner to what we have described in Supplementary Section E. Generally speaking, most of the times the large uncertainties in the parameters seen in Fig. 2 of the main text are not directly relatable to uncertainty in the final solution, but rather to an important covariance or degeneracy amongst the parameters. Calibrating these relations on larger simulation volumes, both because they would host larger structures (therefore, increasing our range of values of radii, Mach numbers and masses) and because they would contain richer statistics, would help to bring down these correlations and therefore decrease the uncertainty in our parameters.

On the other hand, the right-hand side panel of Suppl. Fig. 9 presents the evolution of the variance retained by each of the principal components. While the results do not change dramatically with cosmic time, and therefore the interpretation of the relation can be mantained, there is a clear trend for PC_2 to become more relevant with respect to PC_3 as the redshift decreases, meaning that the relation becomes more and more oblate as cosmic time progresses.

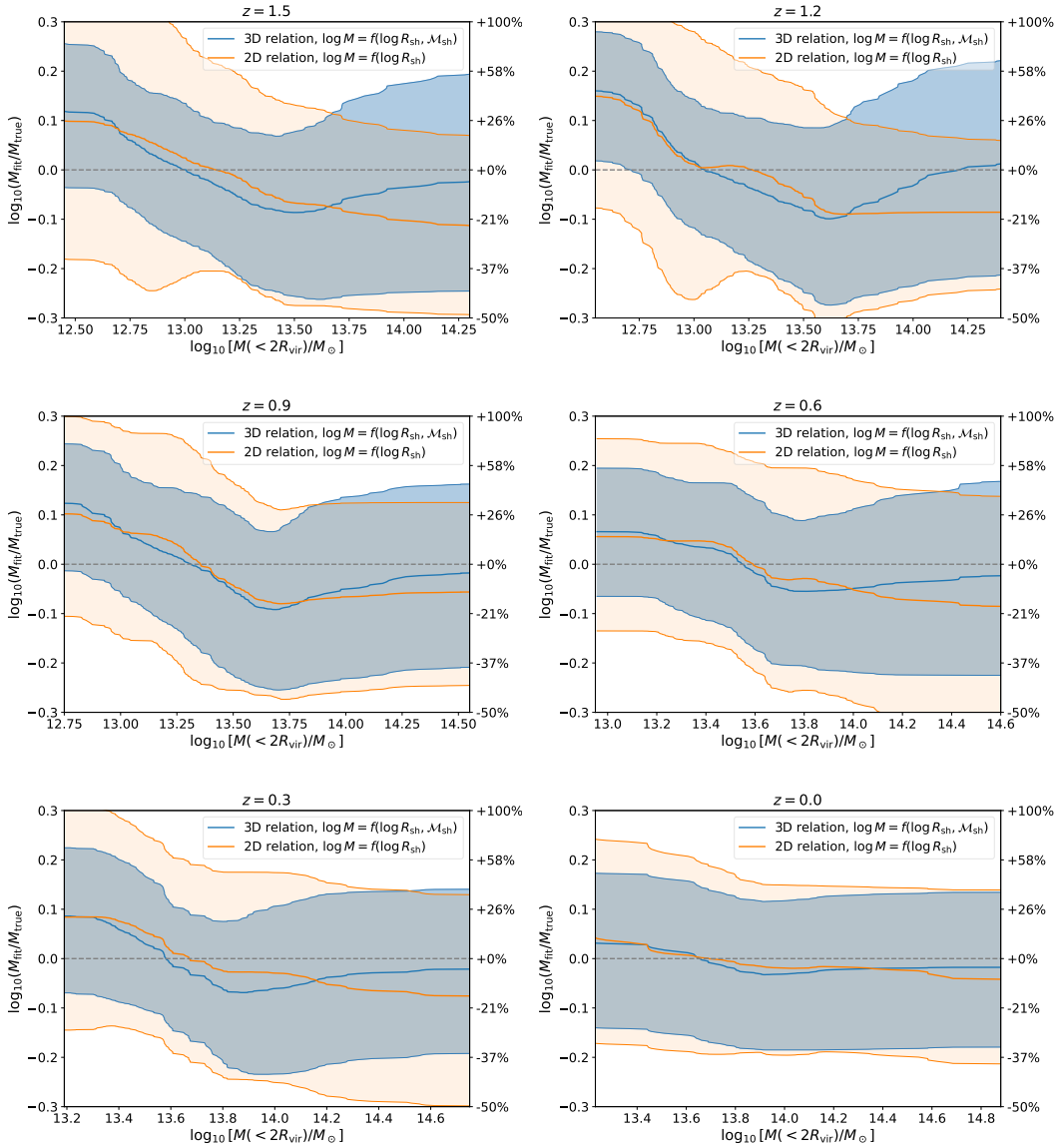
H Intrinsic scatter of the relation: redshift evolution of its mass dependence

In the right-hand side panel of Fig. 3 of the main paper, we had presented the mass-dependence of the residuals with respect to our best-fit relation at $z = 0$. This figure showed that there is a small bias with mass (of a hundredths of a dex, corresponding to a $\lesssim 7\%$ bias), as a consequence that the underlying relation might be non-linear. Interestingly, as we show in the different panels of Suppl. Fig. 10, the behaviour of this small bias with mass is consistent accross our whole redshift range. This hints at



Suppl. Fig. 9. *Left-hand side panel:* evolution of the Spearman rank correlation coefficients amongst the fit parameters α, β and γ introduced in equation (3) The lines represent the data at each snapshot, smoothed with a Gaussian filter with standard deviation set to the Δz between a pair of snapshots, to better visualise the trend. *Right-hand side panel:* evolution of the fraction of variance explained by each of the three principal components. The data points are the values computed at each snapshot, with the $(16-84)\%$ confidence intervals obtained by bootstrap resampling. The lines contain the least-squares fits to this evolution.

a behaviour that could be easily corrected, by calibrating this bias, and therefore would contribute to lower the scatter of our mass determinations. However, we leave this endeavour for future work with enhanced statistics.



Suppl. Fig. 10. Distribution of the residuals ($\log_{10} M_{\rm fit}/M_{\rm true}$) with respect to our fitted three-dimensional scaling relation, at different cosmic times (from $z = 1.5$ to $z = 0$ in increments of $\Delta z = 0.3$). Each panel is similar to the right-hand side panel of Fig. 3 of the main text, but here we show only the contours enclosing the $(16-84)\%$ percentiles around the three-dimensional relation (blue) and the two-dimensional relation (ignoring the information about shock intensity; orange). The solid line corresponds to the mean value of the residuals.

8.3 Future directions

The most immediate continuation of the work presented in this Chapter with respect to turbulence is the extension of the analyses in Paper A3 to the larger-volume simulations also referenced in Sec. 7.3. While our current runs (generally resolving $\sim 60\%$ of the virial mass at $\Delta x \sim 20\,\mathrm{kpc}$ resolution) are insufficient for a proper study of turbulent flows, new runs with higher numerical resolution are already underway. A simulation with a representative number of highly-resolved galaxy clusters and groups will enable the detailed study of several aspects related to the evolution of turbulent flows in the ICM, including:

- Dependence and evolution of the (compressive, solenoidal and total) velocity spectra with dynamical state, cluster-centric distance, and other properties of the cluster. Some effort in this direction has been previously performed, for instance, by Vazza et al. (2011), who find differences in the power spectra and third-order structure functions according to their dynamical state classification. Enabled by the use of a large sample of clusters, we hope to be able to study in more detail the evolution of the turbulent velocity fields and their energetics during the assembly of galaxy clusters, and additionally connect these properties with parameters of cluster mergers (e.g., mass ratio, impact parameter, relative velocity, etc..)

  Linked to this, on a more technical side, it would also be interesting to study the effects of resolution and the refining strategy on the resulting turbulence statistics. To this aim, a series of simulations of a single cluster using either adaptive or static grid refinement have already been run.

- The relation between hydrostatic mass bias and different parameters and observables associated to cluster assembly. Since the evolution of mass bias has been shown to be highly non-trivial (see, for instance, Bennett and Sijacki 2022), it is of great interest to study its relation with observable cluster properties in order to try to reduce the uncertainties associated to its scatter.

  In close connection to this, with the tools developed within this PhD Thesis, we also plan to study the evolution of the non-thermal pressure support in the ICM in the near future.

Regarding the study of shocks in galaxy clusters, in the mid-term, it would be interesting to perform a more detailed study of the effect of different physics to the propagation and the evolution of accretion and merger shocks. For instance,

the relative orientations between shock fronts and magnetic fields have been shown to have a strong impact on the acceleration of cosmic rays (Brüggen et al. 2012, Banfi, Vazza, and Wittor 2020). On the other hand, the physical viscosity of the medium where shocks propagate (Sijacki and Springel 2006) and the presence of additional, non-gravitational physics (Kang et al. 2007) can also determine the evolution, fate and distribution of these features. Magnetic fields and feedback mechanisms are already included in MASCLET, making this study a feasible aim on a short/mid range. However, physical viscosity is not present in the code, and even its effective value in the ICM is still a matter of debate (see, e.g., the discussion in Roediger et al. 2013).

With the MHD version of MASCLET, the amplification of a primordial magnetic field by the small-scale dynamo mechanism will also be able to be studied in detail. The close connection between turbulence and dynamical state, already hinted in Paper A3, suggests that the evolution of magnetic fields in the ICM will itself be connected to the assembly of the cluster, making this topic a natural continuation of the work presented here. Nevertheless, although tentative studies have been performed within this PhD Thesis, properly resolving the small-scale dynamo mechanism in large volume simulations is challenging, especially when AMR and the divergence-cleaning algorithm (which is known to be more dissipative; Balsara and Kim 2004) are involved, demanding resolution to be pushed to even finer scales than for the study of hydrodynamic turbulence.

Finally, a topic that has not been explored during this PhD Thesis corresponds to bridging the gap between these theoretical analyses and observations. In particular, using the full-radiative transfer code SPEV (Mimica et al. 2009, Cuesta-Martínez, Aloy, and Mimica 2015, Mimica et al. 2016) that has already been introduced in Sec. 7.3, it would be possible, for instance, to study the relation between X-ray surface brightness fluctuations and the turbulent velocity field. Regarding non-thermal radio emission, while in previous works by the group using hydrodynamical simulations (Planelles et al. 2018) it was estimated by making assumptions on the magnetic field, this study could now be extended to include physically-realistic magnetic fields.

CHAPTER 9

Cosmic voids

Contents

COSMIC VOIDS represent an opposite end to galaxy clusters in the cosmic web description, by being regions that generally evolve by expanding along the three spatial directions. The study of the properties of voids, together with the structures (mainly, galaxies) that inhabit them, triggers great interest in Astrophysics and Cosmology. This Chapter summarises the results of a work aimed to apply to cosmic voids similar techniques to the ones used in Paper A1 for galaxy clusters.

9.1 Can matter enter voids? Inflows in underdense regions

This Section is based on Vallés-Pérez, Quilis, and Planelles (2021)[1], and the full paper can be found in Appendix A4.

[1] D. Vallés-Pérez, V. Quilis, and S. Planelles. "Void Replenishment: How Voids Accrete Matter Over Cosmic History." In: *Astrophys. J. Lett.* 920.1, L2 (Oct. 2021), p. L2. DOI: 10.3847/2041-8213/ac2816

Overview

Context. Cosmic voids are underdense regions that fill up most of the volume in the Universe and emerge in regions comprising negative initial density fluctuations, which subsequently expand as matter around them collapses to form walls, filaments and haloes.

Aims. Study the nature of the mass flows through the boundaries of voids, to determine whether their velocity field is purely outflowing or more complex patterns arise.

Methods. From the outputs of a cosmological simulation especially set up to describe with enhanced detail the regions that will form voids and their surroundings, a sample of voids defined as the largest possible ellipsoids around expanding, density minima, possibly surrounded by steep density gradients is extracted. The gas mass fluxes through the boundary of voids are estimated in post-processing using a pseudo-Lagrangian approach.

Results. Contrary to the common expectation, around 10% of the gas mass in voids at $z = 0$ has been inflowed from overdense regions, reaching larger fractions for some voids (e.g., $\sim 35\%$ at the 84-percentile). Furthermore, by tracking dark matter inflows, a significant fraction of the mass entering voids lingers on long periods of time reaching inner void-centric radii.

Conclusions. The results reported in this Letter suggest that, if voids are defined from the density field as the largest possible regions around density minima, then it is not possible to affirm that their velocity field is purely outflowing, even for the largest ones. The existence of inflows into voids, a fraction of them coming from gas that has inhabited denser regions and has subsequently been unbound, may have important consequences for the scenario of galaxy formation within voids.

9.2 Future directions

The evolution of galaxies in low-density environments has triggered a lot of attention recently, as it is yet to be seen in how far their properties differ from those of galaxies in denser regions, such as filaments or clusters (Hahn et al. 2007, Kreckel et al. 2011, van de Weygaert and Platen 2011, Ricciardelli et al. 2014, Domínguez-Gómez et al. 2023). On the observational side, a significant effort is being conducted aiming to characterise statistically complete sets of galaxies inhabiting voids, including their dark and baryonic mass assembly histories, gas

properties, etc. This is the case, for instance, of the Calar Alto Void Integral-field Treasury surveY (CAVITY) project.[2]

Producing and studying the properties of void galaxies in cosmological simulations comes with an additional layer of complexity, since both, particle-based SPH and Eulerian AMR simulations are naturally biased towards describing high-density regions. While this is an inherent limitation for finite-mass methods, in AMR simulations it is in principle possible to define a set of refinement criteria to gain resolution precisely in low-density regions. This was attempted, for instance, by Ricciardelli, Quilis, and Planelles (2013) and subsequent works by the group, where a first level of refinement is set based on a low-resolution simulation on the Lagrangian surroundings of underdense regions, and standard, pseudo-Lagrangian AMR is used therein. However, as shown in the aforementioned references, this was insufficient to produce a reasonable population of void galaxies. Therefore, it is imperative that future work is devoted to find a way to sample void regions with sufficient numerical resolution while keeping the computational cost (i.e., the number of resolution elements) at a reasonable level.

Once the new simulations are available, it will be possible to study the evolution of gas within voids and void haloes in great detail, for example by using the tracer particles introduced in MASCLET during the course of this PhD Thesis. By doing this, it will be possible, on the one hand, to shed more light on the nature of gas inflows into voids. Additionally, it will allow exploring to which extent this gas can be accreted or interact with void galaxies, and how this affects their evolution.

[2]https://cavity.caha.es.

Part III

Conclusions

CHAPTER 10

Summary, concluding remarks and outlook

This PhD Thesis presents the results of theoretical research on the nature, the properties, and the evolution of several aspects related to cosmic velocity flows, with especial attention devoted to galaxy clusters, in the first place, and cosmic voids, secondly, as the two main astrophysical objects of study. This research has been primarily conducted by means of the analysis of cosmological simulations and has required the development of new numerical tools for computationally-intensive analyses, which have been released as open-source software packages for the use of the scientific community.

Therefore, the work reported in this dissertation, which is mainly contained in the six publications that constitute Appendix A and can be generally framed within the field of Computational or Numerical Cosmology, is a combination of both theoretical and numerical research. Below, I summarise the main findings, results and outputs of this Thesis. In its numerical facet, these include:

- **The vortex code.** In order to enable analyses on the velocity field of our Eulerian AMR simulations, which do not sample space with uniform resolution, a new algorithm based on the resolution of elliptic PDEs for a scalar and a vector potential which yield the Helmholtz decomposition (compressive+solenoidal) of a vector field has been designed, implemented, tested, and publicly released. The code has been subsequently extended with a multi-scale filter to extract the turbulent part of a velocity field,

following the ideas of the algorithm of Vazza, Roediger, and Brüggen (2012) and Vazza et al. (2017).

  - **The particle-based version of vortex.** In order to allow these algorithms to be applied to a wider range of simulations, a fork of the original VORTEX code, named VORTEX-P has been developed. The velocity field of a particle-based (SPH, MFM, etc.) simulation can be mapped to an ad-hoc AMR grid hierarchy, and the same algorithms as in the original version can be applied. While the user-ready version of this code is still not fully finished, it is already publicly available and being applied to a range of simulations. A publication describing it has been submitted to the journal and recently accepted.

- **ASOHF, an adaptive spherical overdensity halo finder** first introduced by Planelles and Quilis (2010), has been thoroughly redesigned and rewritten in order to improve several of its features, including its parallel performance and memory efficiency, its capabilities of dealing with substructure; and to include new ones, such as the possibility to look for galaxies within haloes, to perform a domain decomposition, or to build merges trees robust to the loss of a halo in some intermediate snapshots. This new version of the code has been publicly released and has been applied to a variety of simulations.

  - **The galaxy-finding capabilities of ASOHF** have been tested against a DIANOGA simulation of a massive cluster, and compared to the results of SUBFIND in order to contrast the properties of the galaxies found by each code. In general terms, ASOHF is capable of finding virtually all galaxies with a significant underlying DM halo, while it misses most of the galaxies that contain a small amount of DM particles within the half-mass radius, which is expected given the nature of the algorithm. However, the properties (e.g., mass distribution) of ASOHF galaxies match to a large extent those of SUBFIND, and around 85% of the galaxies of the former are neatly matched to a galaxy of the latter.

- **Cosmological simulations.** During the course of this Thesis, I have been able to become a user and a contributor of the MASCLET cosmological code, by means of which I have been able to perform several simulations and tests. Besides the ability of gaining a deep understanding of the code, I have also contributed to its development in several areas, including the

generation of initial conditions, improvements to the AMR time-stepping scheme for DM and hydrodynamics, or the addition of tracer particles.

Meanwhile, in a more theoretical aspect, the main results include:

- **An exploratory analysis aiming to examine several measures of the accretion rates** of galaxy clusters and/or groups has been performed, finding that different proxies are loosely correlated amongst themselves. Accretion rates are in clear correspondence with merger events, which are the main contribution to the mass growth of clusters. Furthermore, we have introduced a novel method to describe the spatial (angular) distribution of accretion flows, which are estimated in our Eulerian simulations using a pseudo-Lagrangian post-processing, and reveal a highly anisotropic distribution of infalling gas, from which the main contributions can be extracted and characterised via a multipolar expansion.

- **The assembly/dynamical state of dark matter haloes** has been studied through a broad redshift interval ($5 \gtrsim z \gtrsim 0$), devising a characterisation scheme based on properties at a given time measured from the complete, three-dimensional data yielded by simulation snapshots. This classification turns out to be highly correlated with evolutionary properties of haloes, such as their accretion rates or the timings since the last minor or major merger, hinting that, provided similar insight can be obtained from two-dimensional data from observations, this information could be used to constrain the assembly history of galaxy clusters.

- **The excitation of turbulent motions within the ICM** has been studied in relation with the assembly history of galaxy clusters. We find that, while our simulations show some limitations in resolving the turbulent cascade due to their non-constant resolution, the spectra approximate to a Kolmogorov-like power law. Additionally, some global and local statistics of turbulent motions appear to be in clear correlation with assembly history indicators, e.g. accretion rates. Furthermore, the study of enstrophy as an indicator of the presence of solenoidal turbulent motions tends to confirm a scenario previously proposed by Vazza et al. (2017), where vortical motions are generated by baroclinicity within shocks, and are then amplified by compression and brought downstream of the shocks by vortex stretching.

- **The study of the accretion shocks of simulated galaxy groups and clusters** has yielded a relatively tight correlation between cluster

total mass, shock radius and shock intensity, measured through its Mach number. Using a sample of clusters and groups to calibrate this relation, we have put forward the idea that, provided that future observational facilities are capable of estimating the size and intensity of these features, they could be used as an independent proxy for the total mass of the cluster.

- **The study of the velocity field in and around cosmic voids** has revealed that a significant fraction of them, including some of the largest ones in our simulation, are not purely outflowing, but rather have a complex structure including significant and long-lived inflows. This may have important implications for the formation of galaxies within observed voids, and reflects the complexity in defining these structures.

In the *Future directions* section of each chapter in Part II, the main immediate and more mid- and long-term extensions of the work undertaken in this Thesis and discussed in this dissertation have already been presented. In a brief summary, the numerical tools designed and developed during the duration of the PhD, as well as the cosmological code MASCLET of which I have become a user and contributor, are codes under active, continuous development and optimisation. On the computational side, the main future lines of work include new parallelisation strategies to make the codes more efficient and scalable along with the current trends in high-performance computing architectures, together with an enhanced interoperability with external codes (support for input from other simulation codes, in the case of the analysis codes; and native support for widely-used analysis code, in the case of MASCLET). A more in-depth discussion of the future lines of work in this regard can be found in Sec. 6.3.

Regarding the theoretical and physical goals of this Thesis, the set of simulations that are already in preparation will allow extending and broadening the results presented in this dissertation and the publications that make up Appendix A. Together with the open version of ASOHF and the one of VORTEX for particle-based data, future extensions of the techniques and results on smaller samples obtained during the course of the last four years will likely be able to shed more light on the connection between assembly properties and the presence of non-linear hydrodynamic features (turbulence and shocks) in galaxy clusters, and their relation with observations. A thorough description of these future directions is contained in Secs. 7.3 and 8.3 in what regards galaxy clusters, while the possible extensions of the present work in the field of cosmic voids are discussed in Sec. 9.2.

Part IV

Appendices

APPENDIX A

Publications

The format of this Thesis, as a compilation of articles, requires all the publications to be included in the Appendix. Permission to re-use published material has been obtained from all journals. In the following, the publications are included in chronological order of their acceptance data:

- **Vallés-Pérez, D.**, Planelles, S., & Quilis, V. (2020).
  On the Accretion History of Galaxy Clusters: temporal and spatial distribution.
  Monthly Notices of the Royal Astronomical Society, 499(2), 2303-2318.
  License to publish granted to Oxford University Press.

- **Vallés-Pérez, D.**, Planelles, S., & Quilis, V. (2021).
  Unravelling cosmic velocity flows: a Helmholtz-Hodge decomposition algorithm for cosmological simulations.
  Computer Physics Communications, 263, 107892.
  Published by Elsevier B.V under a CC-BY-NC-ND 4.0 license.

- **Vallés-Pérez, D.**, Planelles, S., & Quilis, V. (2021).
  Troubled cosmic flows: turbulence, enstrophy and helicity from the assembly history of the intracluster medium.
  Monthly Notices of the Royal Astronomical Society, 504(1), 510-527.
  License to publish granted to Oxford University Press.

- **Vallés-Pérez, D.**, Quilis, V., & Planelles, S. (2021).
  Void Replenishment: How Voids Accrete Matter Over Cosmic History.
  The Astrophysical Journal Letters, 920(1), L2.
  © IOP Publishing. Reproduced with permission of the AAS. All rights reserved.

- **Vallés-Pérez, D.**, Planelles, S., & Quilis, V. (2022).
  The halo-finding problem revisited: a deep revision of the ASOHF code.
  Astronomy & Astrophysics, 664, A42.
  Published by EDP Sciences under a CC-BY 4.0 license.

- **Vallés-Pérez, D.**, Planelles, S., Monllor-Berbegal, Ó. & Quilis, V. (2023).
  On the choice of the most suitable indicator for the assembly state of dark matter haloes through cosmic time.
  Monthly Notices of the Royal Astronomical Society, 519(4), 6111–6125.
  License to publish granted to Oxford University Press.

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **499,** 2303–2318 (2020) doi:10.1093/mnras/staa3035
Advance Access publication 2020 October 2

On the accretion history of galaxy clusters: temporal and spatial distribution

David Vallés-Pérez,[1] Susana Planelles[1] and Vicent Quilis[1,2]★

[1]*Departament d'Astronomia i Astrofísica, Universitat de València, E-46100 Burjassot (València), Spain*
[2]*Observatori Astronòmic, Universitat de València, E-46980 Paterna (València), Spain*

Accepted 2020 September 28. Received 2020 September 1; in original form 2020 July 8

ABSTRACT

We analyse the results of an Eulerian adaptive mesh refinement cosmological simulation in order to quantify the mass growth of galaxy clusters, exploring the differences between dark matter and baryons. We have determined the mass assembly histories (MAHs) of each of the mass components and computed several proxies for the instantaneous mass accretion rate (MAR). The mass growth of both components is clearly dominated by the contribution of major mergers, but high MARs can also occur during smooth accretion periods. We explored the correlations between MARs, merger events, and clusters' environments, finding the mean densities in $1 \leq r/R_{200m} \leq 1.5$ to correlate strongly with Γ_{200m} in massive clusters that undergo major mergers through their MAH. From the study of the dark matter velocity profiles, we find a strong anticorrelation between the MAR proxies Γ_{200m} and α_{200m}. Last, we present a novel approach to study the angularly resolved distribution of gas accretion flows in simulations, which allows to extract and interpret the main contributions to the accretion picture and to assess systematic differences between the thermodynamical properties of each of these contributions using multipolar analysis. We have preliminarily applied the method to the best numerically resolved cluster in our simulation. Amongst the most remarkable results, we find that the gas infalling through the cosmic filaments has systematically lower entropy compared to the isotropic component, but we do not find a clear distinction in temperature.

Key words: hydrodynamics – methods: numerical – galaxies: clusters: general – large-scale structure of Universe.

1 INTRODUCTION

Galaxy clusters, as the largest and most massive virialized structures in the Universe, are essential pieces, both, for constraining the cosmological parameters and testing the cosmological model and for improving our understanding of structure formation and evolution on galactic scales (Allen, Evrard & Mantz 2011). A precise understanding of the physics of galaxy clusters, with special focus on the baryonic component, is of utmost importance for these purposes (see, for instance, Kravtsov & Borgani 2012 and Planelles, Schleicher & Bykov 2015 for general reviews on galaxy cluster formation).

In particular, the outskirts of galaxy clusters are dynamically active regions, where the infall of baryons and dark matter (DM) feeds the cluster, giving rise to rather complex physical processes of the gaseous component, such as bulk motions, turbulence, clumping, etc. (see Walker et al. 2019 for a recent review, and references therein). The mass growth of cluster-sized DM haloes and their baryonic counterparts is usually split into two contributions, namely mergers and smooth accretion. Merger events have already been studied as a source of energetic feedback to the intracluster medium (ICM; e.g. Planelles & Quilis 2009), introducing deviations with respect to the X-ray (Markevitch, Vikhlinin & Mazzotta 2001; Nagai, Vikhlinin & Kravtsov 2007) and Sunyaev–Zel'dovich (SZ; Yu, Nelson & Nagai 2015) scaling relations and, thus, potentially biasing mass estimations (see Pratt et al. 2019 for a recent review). As for mass accretion flows, we briefly review the current observational and numerical state of affairs in the following paragraphs.

Gas bulk velocities (along the line of sight) can be directly measured from X-ray line shifts for a small number of nearby clusters (e.g. Tamura et al. 2011, who used Suzaku data to constrain the bulk motions in Abell 2256). Recently, Sanders et al. (2020) have used *XMM–Newton* to map the bulk ICM flows in Perseus and Coma clusters. However, these measurements are still restricted to the central regions of clusters, where enough X-ray photons can be collected. The increased sensitivity of ongoing (e.g. eROSITA[1]) and planned (e.g. ATHENA[2]) facilities will likely extend the X-ray observations to outer regions, thus being able to probe the dynamics of gas in cluster outskirts.

Gas motions can also be inferred from microwave observations through the kinetic SZ (kSZ) effect. Recently, Adam et al. (2017) obtained the first resolved map of kSZ in a galaxy cluster. Future, high-resolution kSZ observations will likely provide strong constraints on the dynamics in the outskirts of the ICM. We refer the interested reader to Simionescu et al. (2019) for an extensive review on the possibilities of ICM velocity measurements with X-ray and kSZ.

On the other hand, accretion on to galaxy clusters has been triggering increasing attention in the numerical cosmology commu-

★ E-mail: vicent.quilis@uv.es

[1]https://www.mpe.mpg.de/eROSITA/
[2]https://www.the-athena-x-ray-observatory.eu/

© 2020 The Author(s)
Published by Oxford University Press on behalf of the Royal Astronomical Society

Table 1. Summary of the main properties of the selected clusters at $z = 0$. x, y, and z refer to the DM centre of mass comoving coordinates. The masses $M_{\rm DM}$ and $M_{\rm gas}$ are measured inside the virial radius, $R_{\rm vir}$, defined according to equations (1) and (2). Temperatures and entropies are computed inside $R_{\rm vir}$, assuming hydrostatic equilibrium (equations 59 and 64 in Voit 2005) and taking a mean molecular weight $\mu = 0.6$.

Cluster	x (Mpc)	y (Mpc)	z (Mpc)	$R_{\rm vir}$ (Mpc)	$M_{\rm DM}$ ($10^{13}\,{\rm M}_\odot$)	$M_{\rm gas}$ ($10^{13}\,{\rm M}_\odot$)	$k_B T_{\rm vir}$ (keV)	$K_{e,\,\rm vir}$ (keV cm^2)
CL01	0.1	0.0	0.1	1.99	42.9	4.56	3.27	1230
CL02	− 3.2	4.9	− 14.9	1.26	10.9	1.33	1.34	520
CL03	17.0	− 3.0	9.4	0.96	4.8	0.66	0.77	290
CL04	10.7	− 2.5	2.0	0.95	4.6	0.44	0.74	279
CL06	− 14.6	− 1.0	− 11.0	0.71	1.9	0.22	0.43	161
CL08	− 15.0	− 4.6	1.9	0.61	1.2	0.12	0.41	117

nity, especially during the last five years. For the dark component, Diemer & Kravtsov (2014) found a sharp drop in DM density profiles, corresponding to the so-called 'splashback' radius, which is generated by recently accreted DM particles in their first apocentric passage, as shown in analytical works by Adhikari, Dalal & Chamberlain (2014) and Shi (2016a). Further works by More, Diemer & Kravtsov (2015), Diemer et al. (2017), and Mansfield, Kravtsov & Diemer (2017) have analysed the relation between the splashback radius and mass accretion rates (MARs), showing that faster accreting DM haloes have generally smaller splashback radii. Chen et al. (2020) have used statistical techniques to reduce the dimensionality of the mass assembly histories (MAHs) of DM haloes, finding correlations with halo concentrations and other parameters.

Similarly, several works have studied the imprint of mass accretion on to the ICM. In particular, Lau et al. (2015) have found that the radial profiles of thermodynamical quantities of the ICM (pressure, entropy, temperature, etc.) depend on the MAR. It has been further shown that faster accreting clusters tend to have smaller accretion shock radii (Lau et al. 2015; Shi 2016b), higher ellipticities (Chen et al. 2019; see also Lau et al. 2020, who explore the connection between DM haloes triaxial shapes and several formation history parameters) and more negative residuals with respect to the T_X–M relation (Chen et al. 2019). Several works have also reported that dynamically disturbed systems display larger hydrostatic mass biases and higher levels of gas clumping in outer cluster regions (see e.g. Biffi et al. 2016 and Planelles et al. 2017, respectively).

It has been seen in N-Body simulations that mergers occur primarily through the filament connecting the cluster with its nearest massive neighbour, which is in turn aligned with the major axis of the cluster (e.g. Lee & Evrard 2007; Lee et al. 2008). However, the accretion pattern of the gaseous component has not yet been extensively covered in the literature.

In this paper, we examine a small sample of clusters from a hydrodynamical, Eulerian adaptive mesh refinement (AMR) cosmological simulation including cooling and heating, star formation, and supernova (SN) feedback. The main aim has been studying and characterizing the accretion processes on these clusters, paying special attention to the gaseous component. In that sense, we have characterized their MAHs and computed several proxies for the instantaneous MAR (namely, Γ_{200m} and α_{200m}; see their definitions in Sections 3.1 and 3.4.1, respectively), their relation to clusters' environments and merging histories, and the imprint of accretion on the radial density profiles for massive clusters and for low-mass clusters or groups, highlighting the differences between these two classes. Besides, the angularly resolved distribution of gas accretion flows has been quantified by means of a simple algorithm proposed in this work. Using multipolar analysis, we have been able to extract the main contributions to the accretion picture and to quantify the differences in their thermodynamical quantities.

The manuscript is organised as follows. In Section 2, we present the numerical details about the simulation and the cluster sample. In Section 3, we study and compare several MAR definitions and relate them to the merging history and surrounding densities, while in Section 4 we present a novel method to evaluate and represent the angular distribution of mass accretion flows. We summarize our main findings and conclusions in Section 5. Appendices A and B discuss in more detail some technical issues regarding the symmetric logarithmic scale, used in some representations in this paper, and the real spherical harmonic basis.

2 THE SIMULATION

In this section, the details of the simulation we analyse and the cluster sample extracted from it are briefly covered. The simulation has already been employed in previous works (e.g. Quilis, Planelles & Ricciardelli 2017; Planelles et al. 2018).

2.1 Simulation set-up

The results presented in this paper correspond to the outputs of a cosmological simulation carried out with the Eulerian, AMR code MASCLET (Quilis 2004). MASCLET combines a multigrid particle mesh N-Body implementation for the description of DM dynamics with *high-resolution shock capturing* techniques for the evolution of the gaseous component.

The background cosmology is set by a spatially flat Λ-cold dark matter model, assuming a Hubble parameter $h \equiv H_0/(100\,{\rm km\,s^{-1}\,Mpc^{-1}}) = 0.678$ and an energetic content given by the density parameters $\Omega_m = 0.31$, $\Omega_\Lambda \equiv \Lambda/3H_0^2 = 0.69$, $\Omega_b = 0.048$. A spectral index $n_s = 0.96$ and a normalization $\sigma_8 = 0.82$ characterize the spectrum of the primordial density fluctuations. These parameters are consistent with the latest results by Planck Collaboration (2018).

The simulation domain corresponds to a cubic box of comoving side length 40 Mpc, which is discretised in a coarse grid with 128^3 cells, thus providing a harsh resolution of ~ 310 kpc at the base level ($\ell = 0$). The initial conditions are set-up at $z = 100$ using a CDM transfer function (Eisenstein & Hu 1998), with a constrained realization aimed to produce a massive cluster in the centre of the computational domain (Hoffman & Ribak 1991). A tentative, low-resolution run from the initial conditions until present time is first performed in order to pick the initially refined regions at the AMR levels $\ell = 1, 2,$ and 3, which will get their DM mass distribution sampled by particles 8, 64, and 512 times lighter than the ones used in the base level.

During the evolution of cosmic inhomogeneities, different regions can get refined under a criterion based on the local gaseous and DM

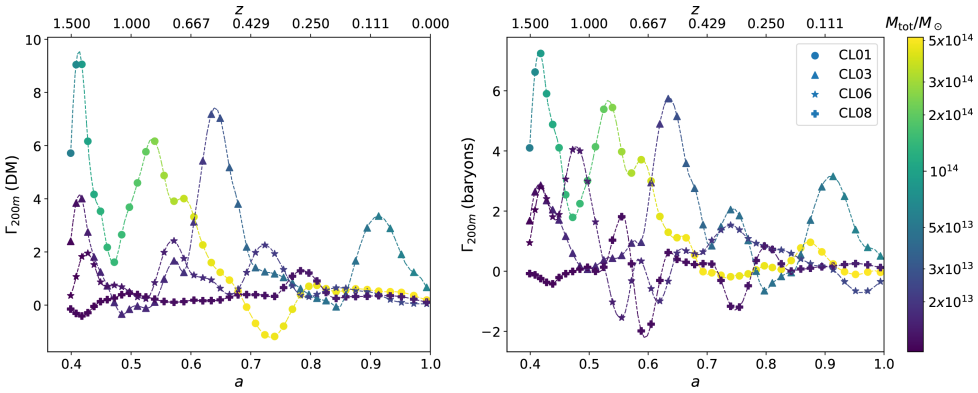


Figure 1. MARs of four of the clusters in the sample, for DM (left-hand panel) and for baryons (right-hand panel). The MARs have been computed using a fourth order Savitzky–Golay filter with window length of 17 points (the MAR curves have been previously resampled with 100 points by linear interpolation, in order to have uniform spacing in the independent variable, $\log a$). Colours encode the total mass (baryons + DM). The shape of the dots refers to each of the clusters according to the legend. The colour bar and the legend in the right-hand panel apply to both plots.

densities. The ratio between cells' side lengths at two consecutive AMR levels is set to $\Delta x_\ell/\Delta x_{\ell+1} = 2$. Up to $n_\ell = 9$ refinement levels are allowed in this run, yielding a best spatial comoving resolution of ~ 610 pc. The lightest DM particles have a mass of $\sim 2 \times 10^6 M_\odot$. This provides a peak mass resolution equivalent to having the domain filled with 1024^3 of such particles.

Our simulation accounts for several cooling processes, such as atomic and molecular cooling for primordial gases, inverse Compton, and free–free cooling. Heating by a UV background of radiation is also included, according to the prescriptions of Haardt & Madau (1996). The abundances are computed under the assumption of the gas being optically thin, in ionization equilibrium, but not in thermal equilibrium (Katz, Weinberg & Hernquist 1996; Theuns et al. 1998). Tabulated, metallicity-dependent cooling rates from Sutherland & Dopita (1993) are employed, with the cooling curves truncated below a temperature threshold of 10^4 K.

Star formation is parametrized following the ideas of Yepes et al. (1997) and Springel & Hernquist (2003). For more details on the particular implementation in this simulation, we refer the interested reader to Quilis et al. (2017). Massive stars produce type-II SN feedback. Feedback mechanisms from type-Ia SN or active galactic nuclei (AGN) are not present in this run. Nevertheless, even though the lack of a central source of energetic feedback (as AGN) could bias the thermal distributions in the inner regions of clusters (see e.g. Planelles et al. 2014; Rasia et al. 2015), it is not expected to have a noticeable impact on their outskirts, which are the focus of this work.

2.2 Structure identification and cluster sample

In order to identify galaxy clusters in our simulation, we have used the DM halo finder ASOHF (Planelles & Quilis 2010; Knebe et al. 2011). ASOHF is a spherical overdensity halo finder especially designed to take advantage of the AMR structure of MASCLET, outputs. Using ASOHF, we identify a total of eight DM haloes with virial masses $M_{\rm DM} > 10^{13}\,{\rm M}_\odot$ by $z = 0$. Two of these haloes have been discarded, as they lie close to the domain boundary and are not faithfully resolved. The remaining six DM haloes and their baryonic counterparts constitute our cluster sample. Two of them have masses above $10^{14}\,{\rm M}_\odot$ and can be fully regarded as galaxy clusters. The rest correspond to low-mass clusters or groups. Their main properties at $z = 0$ are summarized in Table 1.

In addition to a list of objects for each temporal snapshot, ASOHF also provides, given a halo at some code output, a complete list of its *progenitor haloes* (i.e. the haloes at the previous code output which have DM particles in common with it). We refer to this information as the *full merger tree* of a halo. In order to quantify the accretion phenomena, the *reduced merger tree*, containing only the main progenitor of each halo, needs to be built from the previous information.

Amongst all the progenitor haloes, the main progenitor has been picked as the one that contributes the most to the descendant halo mass (i.e. the one which gives the most mass). This strategy is also followed in, e.g. Tormen, Moscardini & Yoshida (2004). Additionally, we have tested other options described in the literature, like tracing the most bound particles back in time (e.g. Planelles & Quilis 2010). These alternative definitions yield remarkably similar results, pointing out the robustness of the reconstructed reduced merger trees.

We follow the objects back in time from $z = 0$ up to $z = 1.5$, for a total of 41 snapshots. No resimulations have been performed.

3 ACCRETION RATES AND THEIR RELATION TO MERGERS AND CLUSTERS' ENVIRONMENTS

Through this section, we cover several topics related to the determination of instantaneous accretion rates (Section 3.1), their relation to clusters' environments and merging histories (Section 3.2) and the impact of accretion on the evolution of the density profiles (Section 3.3). In Section 3.4, we look at the accretion phenomena from a more dynamically motivated perspective, by studying the radial velocity profiles.

3.1 Determination of the mass accretion rates

For each of the DM haloes described in Section 2.2, we determine their boundaries and enclosed masses according to the usual spherical

overdensity definition (Lacey & Cole 1994) with respect to the background matter density:

$$M_{\rm DM}(< R_{\Delta_m}) = \frac{4\pi}{3} R^3_{\Delta_m} \Delta_m \rho_B, \quad (1)$$

with $\rho_B(z) = \Omega_m(z)\rho_c(z)$ and $\rho_c(z) = \frac{3H(z)^2}{8\pi G}$. Along this manuscript we focus on $\Delta_m = 200$, as well as on the virial radius, $R_{\rm vir}$, defined as the radius enclosing an overdensity (Bryan & Norman 1998):

$$\Delta_{\rm vir,m} = \frac{18\pi^2 + 82x - 39x^2}{\Omega_m} \quad (2)$$

being $x \equiv \Omega_m(z) - 1$ and $\Omega_m(z) = \frac{\Omega_m(1+z)^3}{\Omega_m(1+z)^3+\Omega_\Lambda}$. The computation of the stellar, gaseous and DM masses for each snapshot constitutes the MAH of the different material components of the cluster.

Aiming to quantify the strength of accretion on to the objects in our sample, we define the MAR as the logarithmic slope of the $M(a)$ curve, where a is the scale factor of the Universe:

$$\Gamma_\Delta(a) = \frac{{\rm d}\log M_\Delta}{{\rm d}\log a}. \quad (3)$$

This definition is more extended in theoretical works (e.g. Adhikari et al. 2014), while most numerical studies opt for replacing the derivative by a quotient difference over a fixed, wide interval $[a_1, a_0]$, in order to avoid the contaminating effects of the intrinsically noisy nature of the MAH of clusters (Diemer & Kravtsov 2014; More et al. 2015; Mansfield et al. 2017; Chen et al. 2019). In these cases, the resulting MAR oughts to be interpreted as an average over several Gyr in the accretion history of the cluster. We shall denote this definition of the MAR as $\Gamma_\Delta^{[a_1,a_0]}$.

Even though the average MAR accounts for the global impact of accretion on the dynamical state of clusters (as shown in Chen et al. 2019), it is not as suitable for analysing a number of aspects, such as the relation of mass infall and clusters' surroundings or punctual events like mergers, as the accretion rates get overly smoothed. In order to measure the *instantaneous* accretion rate through a cluster's history while getting rid of statistical noise, we implement the computation of the derivatives using Savitzky–Golay filters (Savitzky & Golay 1964; Press & Teukolsky 1990), whose parameters have been tuned to offer a compromise between smoothness and locality of the computed derivatives.

The definition of MAR in equation (3) can be applied independently to each material component (i.e. DM and baryons[3]). It is also worth mentioning, however, that this is not the only proxy for the MAR of a cluster (see Section 3.4 for the comparison with an alternative definition).

3.1.1 Baryonic and DM MARs

We have computed the baryonic and DM MARs of the clusters in our sample from $z = 1.5$ to $z = 0$, with respect to the masses measured within R_{200m}. A selection of them is presented in Fig. 1, where the colour scale keeps track of the total mass. In these plots, the derivatives have been computed using fourth-order Savitzky–Golay filters with window length of 17 points.

[3]Baryons account for both, gas and stars. As the total gas mass is not conserved due to star formation, the MAR of gas would be biased low. This effect may be almost negligible in massive clusters, but definitely noticeable in low-mass clusters and groups, where stellar fractions tend to be higher (e.g. Planelles et al. 2013).

The graphs show similar qualitative behaviour for the DM and baryonic MARs, reflecting the fact that gas traces DM to a first approximation. However, the most prominent peaks are typically more pronounced for the dark component than for their baryonic counterparts, implying that gas is generally accreting at a slower pace when compared to DM. This trend has already been pointed out by other studies (see e.g. Lau et al. 2015, where a similar conclusion is drawn from studying the radial velocity profiles of both components). As opposed to collisionless DM, collisional gas is supported by pressure and experiences ram pressure from the ICM (Tormen et al. 2004; Cen, Roxana Pop & Bahcall 2014; Quilis et al. 2017), shocks, etc. that contribute to slow down the infall.

A clear distinction is displayed between massive (CL01 and CL03) and low-mass (CL06 and CL08) objects. Massive clusters often present pronounced peaks in their MAR curves, typically associated to major mergers, which are still frequent as clusters continue growing and collapsing by $z \sim 0$. We analyse in further depth the relation between mergers and accretion rates in Section 3.2. Less massive clusters show flatter curves, pointing out that either they do not undergo merger events as significant as their massive homologues, or they also experience important mass-losses during these events, as a consequence of their shallower potential wells. The latter idea is supported by the fact that the differences between baryonic and dark components are more remarkable in these systems. Clusters and groups with total mass $\lesssim 5 \times 10^{13}\,{\rm M}_\odot$ do not seem to dominate as efficiently their neighbourhoods, and are therefore harassed by other systems.

3.2 Effects of mergers and surrounding densities on the MARs

3.2.1 Identification and classification of mergers

We define mergers as events where two cluster-sized haloes (and their respective baryonic counterparts) encounter and share a significant amount of mass (e.g. Planelles & Quilis 2009). The merger tree of a massive halo can contain many progenitor haloes, most of which either are low-mass infalling substructures or contribute very little to the mass of the descendant halo. In order to identify halo mergers, we have imposed the following conditions:

(i) The distance between the centres of mass of the two progenitor candidates, i and j, is less than the sum of their virial radii, i.e. their spheres of radius $R_{\rm vir}$ intersect:

$$d_{ij} \leq R_{\rm vir,i} + R_{\rm vir,j} \quad (4)$$

(ii) Each of the progenitor haloes gives, at least, 1 per cent of its (DM) mass to the descendant halo.

(iii) Each of the progenitor halo masses is greater than 1/10 of the descendant mass. Mergers with haloes of smaller mass are regarded as smooth accretion.

These conditions are conceptually similar to those of Chen et al. (2019), who nevertheless use more stringent values (R_{500c} instead of $R_{\rm vir}$ and 10 per cent of shared mass between progenitor and descendant halo) in order to assess the merging times. We use the presence/absence of mergers and the maximum mass ratio between the progenitors to distinguish three accretion regimes in the assembly history of a cluster, according to the following classification:

- Major mergers: involve two haloes of comparable mass and are relatively unfrequent. They typically have an important impact on the structure of haloes. We take a mass ratio of 1: 3 as the threshold for these events (Planelles & Quilis 2009; Chen et al. 2019).

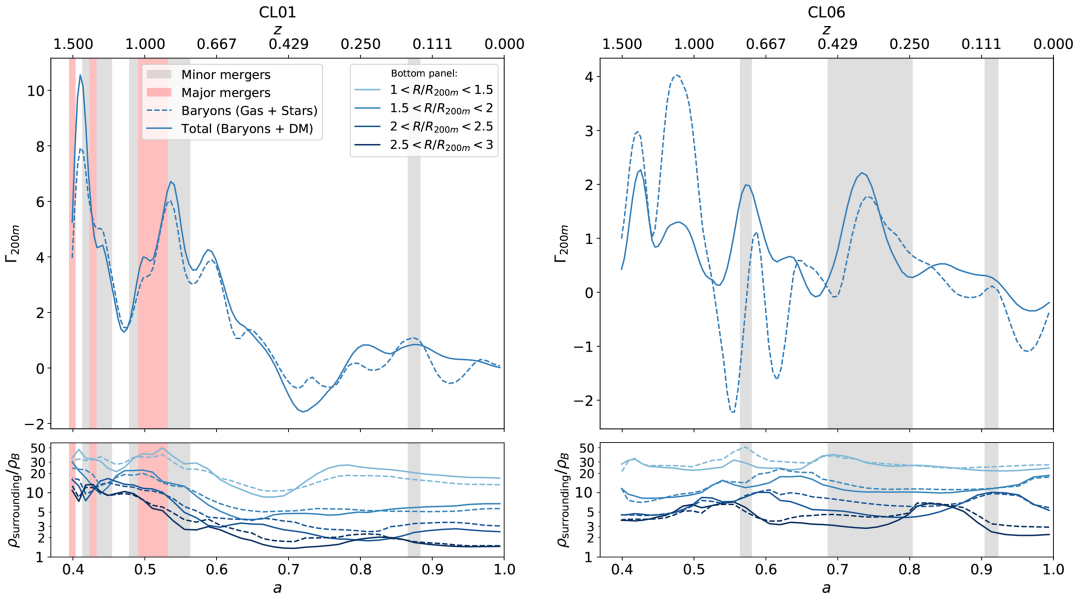


Figure 2. Relation between the MARs, accretion regimes (ongoing merger events) and densities in the surroundings of the cluster, for clusters CL01 (left) and CL06 (right). In each panel, the upper plot shows the MARs for the total mass (the solid line) and baryonic mass (the dashed line). The lower panel displays the surrounding density in several clustercentric radial bins, in units of the background density of the Universe. The dashed line, corresponding to the baryonic density, has been normalized to the cosmic baryon fraction (i.e. it has been multiplied by Ω_m/Ω_b). The legend in the left-hand panel applies to both plots.

• Minor mergers: produce less significant disturbance on the objects, but are generally more frequent. Their mass lower threshold is slightly more arbitrary and varies through the literature. We use a ratio of 1:10 as the threshold for minor mergers, as in Planelles & Quilis (2009).

• Smooth accretion: systems which experience no mergers above the 1:10 mass ratio threshold are considered to undergo smooth accretion.

In our sample, clusters CL01, CL02, and CL03 exhibit periods of major and minor merging activity. CL06 does not experience any major mergers, but only minor ones. Last, no mergers have been identified in CL04 and CL08 and they are therefore smoothly accreting clusters throughout the considered redshift interval, $1.5 \geq z \geq 0$.

3.2.2 *Surrounding densities*

A significant part of the accreted mass in major mergers can end up lying outside the R_Δ boundary of the final halo, and hence, the corresponding spherical overdensity masses are not additive in such events (Kravtsov & Borgani 2012; More et al. 2015). In order to study how matter is deposited in the outskirts of galaxy clusters, and how this effect shapes the MAR curves, we quantify the densities in the surroundings of each cluster and their evolution with cosmic time in four non-overlapping, equally spaced radial bins, covering the region $1 \leq R/R_{200m} \leq 3$.

The results of the joint analyses of MARs, accretion regimes, and surrounding densities are shown in Fig. 2, for clusters CL01 and CL06, as paradigmatic cases of a massive cluster that undergoes numerous major and minor mergers and a low-mass cluster that only suffers minor mergers, respectively. Both panels show the total (the solid lines) and baryonic (the dashed lines) MARs, the accretion regimes (background colour of the plot), and the surrounding densities (lower panels).

In the case of CL01, the differences in the MAR of DM and baryons are small in magnitude. Peaks in the MAR are undoubtedly associated to (major) merger events, as in the case of the displayed peaks around $z \sim 1.4$ and $z \sim 0.8$. It is also interesting to note how high MARs are mantained for a long time after the merger has taken place (particularly salient is the case of the merger at $z \sim 0.8$), as matter deposited outside R_{200m} continues feeding the cluster – in a more quiescent way – for several Gyr. In this respect, the lower panel shows how densities in the $1 \leq R/R_{200m} \leq 1.5$ region keep above $10\rho_B$ until $z \sim 0.5$.

Comparing the surrounding total and baryonic densities, the latter appears to evolve in a much smoother way than the former. At outer radii, $R \gtrsim 2R_{200m}$, baryon surrounding densities (when normalized to the cosmic baryon fraction) tend to be higher than total densities. This reinforces the idea that gas, due to its pressure support, is deposited at larger radii than DM, which can more easily penetrate to inner regions. We find these general trends to be common for all the massive clusters in our sample that suffer major mergers.

On the other hand, CL06 exhibits significant differences in the behaviour of baryonic and total masses. The total MAR experiences peaks in correlation to the minor merger events. The baryonic component roughly follows these peaks, although their magnitude can differ significantly. There are also severe declines in the baryonic mass ($\Gamma_{200m} < 0$), which reflect the inability of low-mass systems and groups to keep their gas inside R_{200m}. This gas is typically expelled to larger cluster-centric radii (as seen in the lower panel) and only slowly reaccreted afterwards.

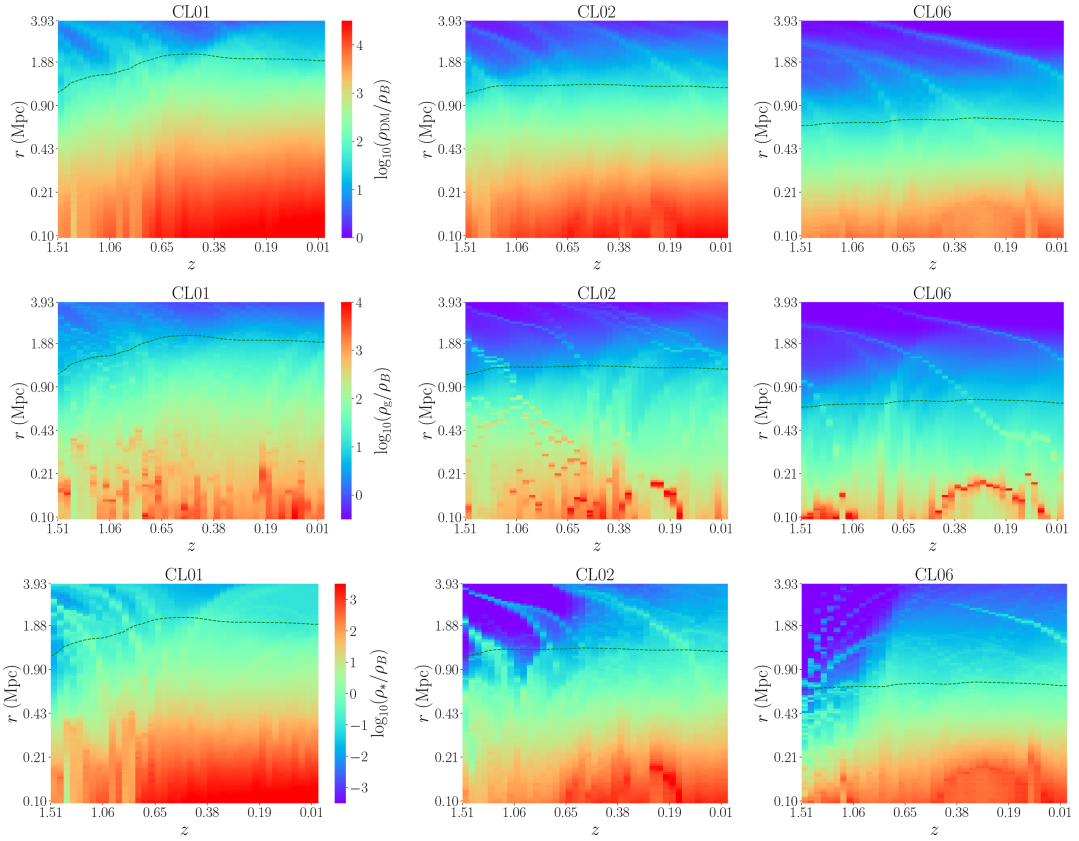


Figure 3. The panels show the evolution of the comoving density profiles of the different material components for three clusters, from $z \simeq 1.5$ to $z = 0$. The profiles have been taken with 100 logarithmically spaced bins, from 100 kpc to 4 Mpc, and taking centre at the potential minima. Densities are always normalized to the background matter density of the Universe and the colour scale is logarithmic. Radial coordinates are comoving. From top to bottom, panels show DM, gas and stellar densities. From left to right, these quantities are presented for CL01, CL02, and CL06. The colour scale in the left plot of each row applies for all plots in the row. The green line in each panel marks the virial radius.

3.2.3 Correlation and time shift between surrounding densities and accretion rates

As seen in the left-hand panel of Fig. 2, the density in the immediate neighbourhood of the cluster shows remarkable resemblance to the behaviour of the MAR, although a time shift between both curves is evident. In order to have an estimation of this shift, we compute the Spearman's rank correlation coefficient,[4] $\rho_{\rm sp}$, of $\rho_{\rm surrounding}(t)$ in the first radial bin with $\Gamma(t+\tau)$, and find the τ that maximizes $\rho_{\rm sp}$. Next, we summarize the results of such analysis:

(i) For cluster CL01, when total masses are considered, an optimal shift of 900 Myr provides a rank correlation of $\rho_{\rm sp} = 0.803$. Restricting to the baryonic component enhances this correlation to $\rho_{\rm sp} = 0.895$, while increasing the time shift to 1.1 Gyr, as pressure effects slow down gas infall. These times are consistent with the shell crossing time of the fastest DM particles. A similar trend is shown by cluster CL03, although in this case the correlations are slightly weaker ($\rho_{\rm sp} = 0.744$ and $\rho_{\rm sp} = 0.691$, respectively).

(ii) CL02, a massive cluster that only experiences a major merger at high redshift and undergoes a quiescent evolution therein, shows weaker correlations: $\rho_{\rm sp} = 0.623$ and $\rho_{\rm sp} = 0.572$, respectively.

(iii) Less massive objects do not display significant correlations between these variables, as a result of their limited ability to capture matter (especially, gas). Indeed, in the lower panel of Fig. 2 a much flatter evolution of the surrounding densities is clearly noticeable.

[4]The Spearman's rank correlation coefficient, $\rho_{\rm sp}$, assesses the monotonicity of the relation between two variables (without the need to assume linearity). $\rho_{\rm sp}$ is valued in $[-1, 1]$, with higher absolute values implying a more monotonic relation and positive (negative) values corresponding to increasing (decreasing) relations.

3.3 Impact of accretion on radial density profiles

In this section, we analyse how mergers and strong accretion rates impact the inner distribution of the different material components in clusters by computing the comoving density profiles of DM, gas and stars. We graphically present these profiles for objects CL01, CL02, and CL06 as a function of redshift in Fig. 3.

Note how the virial radii of CL02 (which only undergoes a major merger at $z \sim 1.4$) and CL06 (which only suffers minor mergers) are roughly constant, while CL01 (undergoing several major mergers) displays an important growth of this boundary.

Relating to this, the inner regions of DM comoving density profiles are mostly constant in time for CL02 and CL06, suggesting that these structures are already collapsed by $z \sim 1$ and the innermost radii ($r \lesssim R_{2500m}$) do not get disturbed by minor mergers and smooth accretion. This result is consistent with More et al. (2015), who find that the mass inside $4r_s$ (being r_s the scale radius of the NFW, Navarro, Frenk & White 1997, profile) evolves relatively slowly for $z \lesssim 1-2$. Conversely, CL01 does experience important disturbances in its DM profile, especially around $z \sim 0.8$. The enhanced MAR during $1.1 \gtrsim z \gtrsim 0.7$, associated to a major merger event (see Fig. 2), appears to substantially increase the central density.

Gas density profiles are clumpier than DM ones and are affected by miscentring issues (specially in the case of CL06). The centres of the gaseous and dark components do not necessarily coincide (see e.g. Forero-Romero, Gottlöber & Yepes 2010; Cui et al. 2016), especially when there is ongoing merging activity (i.e. departures from dynamical equilibrium). Lines of decreasing radii with decreasing redshift reflect the infalling orbits of massive structures being accreted, mainly in galaxy cluster mergers. These streams of matter are better recognized in the stellar component, as the stellar mass is more concentrated towards the centre of the infalling cluster and leaves a sharper imprint on the density profile. The redshifts at which these streams cross the halo $R_{\rm vir}$ boundary appear to be consistent with the periods of mergers according to the classification of Section 3.2.

The profiles of the gaseous component also suggest the presence of gas being deaccreted or expelled outside the virial radius. This is particularly notorious for CL02 and CL06, and hints that dynamical interactions between clusters can extract gaseous matter to outer radii (e.g. through gas sloshing; see e.g. Markevitch et al. 2001; Roediger et al. 2011).

3.4 Velocity profiles. Alternative MAR definitions

The dynamics of accretion on to galaxy clusters and the differential behaviour of DM and gas can be explored from the radially averaged profiles of radial velocity in the cluster outskirts. In Fig. 4, we present such profiles for the massive central cluster CL01, for four evenly spaced redshifts from $z = 1.5$ to $z = 0$. The magnitude represented in the figure corresponds to the physical velocity, $u_r \equiv v_r + H(z)r/(1+z)$, where r and v_r are, respectively, the comoving clustercentric radial coordinate and the peculiar velocity. This quantity has been normalized to the circular velocity at R_{200m}, $V_{\rm circ,200m} = \sqrt{GM_{\rm 200m}/R_{\rm 200m}}$.

The interior of the cluster ($r \lesssim 0.8R_{\rm 200m}$) does not present particularly strong inflows nor outflows, as these regions are already collapsed and relatively stable. Accretion flows are dominant in the cluster outskirts ($r \gtrsim R_{\rm 200m}$), where radial velocities drop sharply and reach a minimum located around $1.5 \lesssim r/R_{200m} \lesssim 2$. The position of the velocity minimum with respect to $R_{\rm 200m}$ does not show a clear redshift evolution in the case of CL01, in consistency with the general behaviour pointed out by Lau et al. (2015). At large radii, radial velocities increase, as the Hubble flow term begins to dominate the dynamics of both material components.

The magnitude of the velocity at the minima shows a remarkable redshift evolution. Both, gas and DM, have larger infall velocities – when compared to the circular velocity – at earlier times and, consequently, their MARs shown in Fig. 2 are larger. Even though

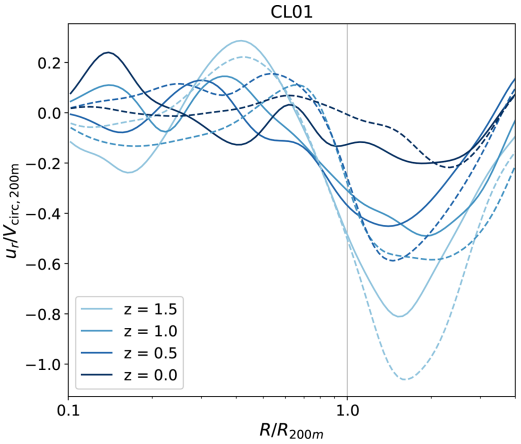


Figure 4. Radially averaged radial physical velocity profiles for the cluster CL01 at 4 different redshifts. The solid lines represent the radial velocities of gas, while the dashed lines correspond to DM. The profiles have been taken with 100 bins logarithmically spaced between $0.1R_{\rm 200m}$ and $4R_{\rm 200m}$ and smoothed with a Gaussian filter with window length of four points.

gas and DM radial velocity curves exhibit a similar pattern in the cluster outskirts, their different magnitudes highlight that DM is being accreted slightly faster, as it is not pressure supported and does not feel hydrodynamical effects, such as ram pressure from the ICM, shock heating, etc. as it falls into the cluster.

3.4.1 *Comparing Γ_Δ to α_Δ*

In Lau et al. (2015), the authors suggest a different proxy for the instantaneous MAR, defined as the radial infall velocity of DM ($v_{\rm r}^{\rm DM}$) at some radius $r = R_\alpha$, in units of the circular velocity at R_Δ:

$$\alpha_\Delta = \frac{v_{\rm r}^{\rm DM}(r = R_\alpha)}{V_{\rm circ,\Delta}} \tag{5}$$

More negative α_Δ indicates more rapid infall of matter, and thus α_Δ should be anticorrelated to Γ_Δ. Lau et al. (2015) report that taking $R_\alpha = 1.25R_{200m}$ maximizes the anticorrelation of α_{200m} and $\Gamma_\Delta^{[a_1,a_0]}$, with $a_1 = 0.67$. Note, however, that $\Gamma_\Delta^{[a_1,a_0]}$ is the averaged MAR over the last ~ 5 Gyr.

In order to compare α_Δ to the instantaneous Γ_Δ MAR proxy defined in equation (3), we have computed α_Δ for all the clusters in our sample and all the simulation outputs available, from $z = 1.5$ to $z = 0$. We find that, taking $R_\alpha = R_{200m}$, the Spearman rank correlation between α_{200m} and Γ_{200m} is $\rho_{\rm sp} = -0.832$. Larger values of $R_\alpha/R_{200m} = 1.25,\ 1.5$ weaken this anticorrelation down to $\rho_{\rm sp} = -0.719$ and $\rho_{\rm sp} = -0.556$, respectively.

In the upper panel of Fig. 5, we present a scatter plot of the two instantaneous MAR proxies, where the dots have been coloured to identify each cluster. All snapshots, from $z = 1.5$ to $z = 0$, are included in this comparison. A non-parametric fit with smoothing splines is shown as the blue line. To rule out any possible dependencies of the scatter in this relation with other variables, the panels below show the distribution of residuals with respect to the fit, $\Delta\alpha_{200m}$, as a function of Γ_{200m}, $M_{\rm 200m}$ and redshift z. No significant residual dependencies on $M_{\rm 200m}$ and z are found. Consistently, the residuals are uncorrelated to these variables, with rank correlation coefficients $\rho_{\rm sp} = -0.038$ and $\rho_{\rm sp} = 0.099$, respectively. Hence,

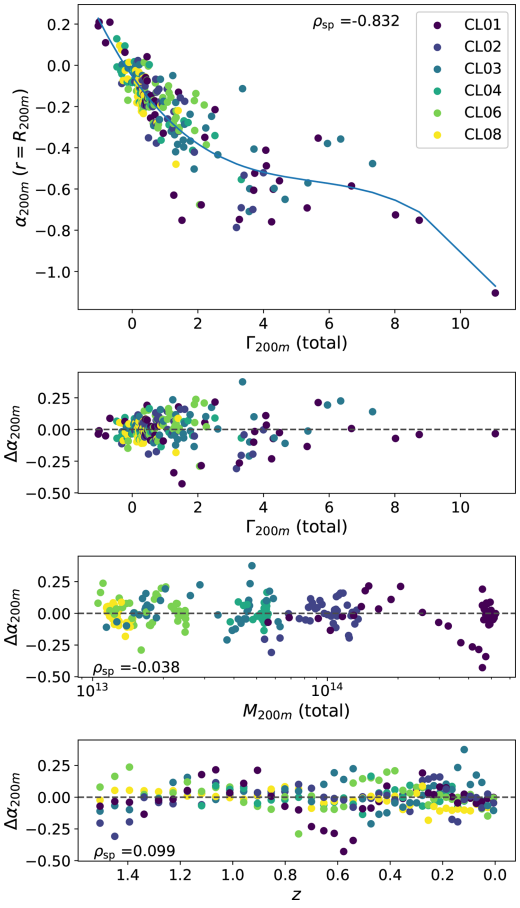


Figure 5. Relation between the instantaneous MAR proxies Γ_{200m} (used in this work) and α_{200m} (Lau et al. 2015), combining all the snapshots in $1.5 \geq z \geq 0$. The upper panel contains the scatter plot of both variables and a non-parametric fit using smoothing splines. The residuals of this fit, $\Delta\alpha_{200m}$ are used in the lower panels to rule out redshift or mass dependencies in this relation, as they appear to be uncorrelated to M_{200m} and z.

α_{200m} and the instantaneous $\Gamma_{200m}(a)$ are consistent probes of the MAR, and their relation is mostly independent of redshift or cluster mass.

4 ANGULAR DISTRIBUTION OF THE MASS FLOWS AND THERMODYNAMICAL PROPERTIES

This section focuses on the characterization of the angular distribution of accreting gas. This topic has not been extensively covered in the literature, but it is of utmost interest in order to assess the complex interplay of clusters with their environments, which shapes the accretion patterns. In Section 4.1, we present a simple method for estimating and presenting the accretion flows. Its results are discussed in Section 4.2.

4.1 Estimation of the mass fluxes through the cluster boundary

Let us consider a cluster at redshift z, delimited by a spherical boundary $R_{\rm bdry}$ (e.g. $R_{\rm vir}$) around it. We quantify the mass flows by computing an estimated flux from the peculiar velocity and density fields as described in the following paragraphs.

We assume each gas cell, with density contrast $\delta_{\rm cell}$ and physical volume $\Delta V_{\rm cell}$, as a particle located at its geometrical centre, with mass $m_{\rm cell} = \rho_B(z)(1+\delta_{\rm cell})\Delta V_{\rm cell}$ and the peculiar velocity $\boldsymbol{v} = a\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}$ given by the corresponding cell-averaged velocity. All cells, regardless of the refinement level they belong to, are considered on equal footing. For each cell, let r be the radial comoving distance to the cluster centre and $v_{\rm r}$ its cluster-centric radial peculiar velocity. With these definitions, we estimate the fluxes accross $r = R_{\rm bdry}$ according to the following rules:

(i) Given a cell within the spherical boundary, $r < R_{\rm bdry}$, we mark it as an escaping cell if $r + \frac{v_{\rm r}}{a}\Delta t > R_{\rm bdry}$.
(ii) A cell outside the radial boundary $r > R_{\rm bdry}$ is marked as an entering cell if $r + \frac{v_{\rm r}}{a}\Delta t < R_{\rm bdry}$.

The time interval, Δt, used for the estimation of the fluxes has to be chosen as a compromise between angular resolution (as higher Δt increases the number of entering and escaping cells) and accuracy (radial flows are not necessarily maintained for arbitrarily large Δt). We have set Δt to the time difference between consecutive snapshots (ranging from ~ 60 to ~ 300 Myr in this particular simulation).

Once the entering and escaping cells have been found, the angular distribution of accreted gas is computed by binning the complete solid angle around the cluster in the cluster-centric spherical angles, ϕ and $\cos\theta$ (so that angular sectors at all latitudes subtend the same solid angle). We define the spherical coordinate system with respect to the major axis of the cluster total mass distribution. This is motivated by N-Body simulations having shown that most of the mergers and accretion of DM occur through the filaments connecting a cluster to its nearest massive neighbour (Lee & Evrard 2007) and, consequently, major axes tend to be aligned with such filaments (Lee et al. 2008). In order to do so, the characterization of the shape of matter distributions is discussed in Section 4.1.2. In our analyses for cluster CL01, we split the solid angle in $n_\phi \times n_\theta = 80 \times 80$ bins. Increasing the number of bins beyond this quantity does not result in any significant enhancement in the description of the mass fluxes, as the resolution gets constrained by the cell sizes at the $r = R_{\rm bdry}$ boundary.

For each entering (escaping) cell, we assign all its mass to the bin corresponding to its angular position, yielding the distribution of accreting (deaccreting) matter. Their subtraction is the net mass flowing across the $r = R_{\rm bdry}$ boundary. Finally, we compute the mass radial flux by normalizing this quantity as follows:

$$j_M = \frac{\Delta M}{R_{\rm bdry}^2 \Delta\Omega \Delta t}, \tag{6}$$

where $\Delta M = \Delta M_{\rm enters} - \Delta M_{\rm escapes}$ and $\Delta\Omega = \frac{4\pi}{n_\phi n_\theta}$. Note that we take, as sign convention, that $j_M > 0$ when matter is infalling (being accreted).

4.1.1 Validity of the method. Spatial and temporal coherence of the radial flows.

The procedure proposed above relies on the implicit assumption that gas velocities measured at one code output are persistent during the time interval, Δt, used to estimate which gas cells cross the cluster boundary. In the following lines, we briefly argue that this

approximation is, indeed, applicable to our system, thus justifying the validity of the method.

The basic scheme of the performed test consists on the comparison of the radial velocity dispersion, σ_r, with the radial velocity, v_r. For each entering cell, we compute its total and radial velocity dispersions as the standard deviation of such quantities in the neighbouring $5 \times 5 \times 5$ cells. Such analysis yields the following conclusions:

(i) The gas flows are eminently radial. The mean radial projection of the velocity of the entering cells has magnitudes $0.8 \lesssim |v_r|/v \lesssim 0.9$ for all code outputs.

(ii) Radial velocity dispersions are consistent, only slightly above the isotropic value, $\sigma_r^{\rm iso} = \frac{\sigma_{3r}}{\sqrt{3}}$. For most of the snapshots, $\frac{\sigma_r}{\sigma_r^{\rm iso}} \sim 1$–$1.1$. Thus, the relative velocity dispersion in the radial direction $\sigma_r/|v_r|$ is much smaller than in the directions tangential to the flow.

(iii) Indeed, $|v_r|/\sigma_r$ takes mean values between 10 and 40, indicating that, in the neighbourhood of a cell, radial gas flows are spatially coherent.

As a result of the radial flows being spatially coherent, one should expect the shear forces between neighbouring cells to be small along the radial direction. Therefore, turbulence is not expected to have a severe impact on the overall distribution of radial flows. However, this analysis does not guarantee the persistence of the flows during an arbitrarily large Δt. Once the spatial coherence of the flows has been confirmed, these can be assumed to be persistent between consecutive snapshots, provided that the angular distributions for consecutive snapshots are temporally coherent, i.e. they show similar structures and temporal changes are gradual.

Note, however, that this analysis has considered the mean values of σ_r and v_r across the whole boundary. It is still possible that turbulence is relevant on the radial flows in small angular regions, thus introducing spurious noise into our flux maps.

4.1.2 *Characterization of clusters' shapes*

The procedure for studying the mass flows presented above requires finding the major axis of the cluster, as we use it to define the spherical coordinate system tied to the cluster.

A widely extended method in the literature to estimate the shape of DM haloes (e.g. Cole & Lacey 1996; Planelles & Quilis 2010; Knebe et al. 2010, among others) relies on finding the eigenvalues and eigenvectors of the shape tensor:

$$\boldsymbol{S} = \frac{1}{M_{\rm tot}} \int_{\mathcal{V}} \rho(\boldsymbol{r}) \boldsymbol{r} \otimes \boldsymbol{r} \mathrm{d}V, \tag{7}$$

where the integration extends to the spherical overdensity boundary definition and $M_{\rm tot} = \int_{\mathcal{V}} \rho(\boldsymbol{r}) \mathrm{d}V$ is the enclosed mass. Even though this approach produces sensible estimates, it might be contradictory that a spherical integration volume is used to characterize the shape of a triaxial object. A more robust method uses an iterative procedure aimed to adapt the integration volume to the – initially unknown – shape of the mass distribution (Zemp et al. 2011). We have implemented a method partially based on the one described in the aforementioned reference, which better suites our purposes. The main steps can be summarized as follows:

(i) The first iteration uses an sphere of radius $R_{\rm bdry}$ as the integration volume. Let us call $\{\hat{\boldsymbol{v}}_1, \hat{\boldsymbol{v}}_2, \hat{\boldsymbol{v}}_3\}$ the orthonormal set of eigenvectors of $\boldsymbol{S}$, where their corresponding eigenvalues are such that $\lambda_1 < \lambda_2 < \lambda_3$. That is to say, the third vector points in the direction of the tentative major axis.

(ii) These values are used to define the new integration volume, which is given by the ellipsoid

$$\frac{(\hat{\boldsymbol{v}}_1 \cdot \boldsymbol{r})^2}{a_1^2} + \frac{(\hat{\boldsymbol{v}}_2 \cdot \boldsymbol{r})^2}{a_2^2} + \frac{(\hat{\boldsymbol{v}}_3 \cdot \boldsymbol{r})^2}{a_3^2} \leq 1, \tag{8}$$

where a_i is the semi-axis length corresponding to $\hat{\boldsymbol{v}}_i$. The semi-axes are rescaled so that their squares are proportional to the corresponding shape tensor eigenvalues. However, this premise does not fix the magnitude of the semi-axes (only their quotients) and, therefore, a choice regarding the normalization has to be made.

In Zemp et al. (2011), the new semi-axes lengths are defined so that the major semi-axis is preserved. This election is motivated by their aim of measuring the shape at different distances, which they label by the major semi-axis length. However, this method shrinks the volume along the direction of the minor and intermediate axis and therefore dramatically increases the enclosed overdensity.

With the aim of describing the overall shape of the cluster without changing significantly the enclosed overdensity, in this work we have chosen to preserve the enclosed volume. Let $r_i \equiv \sqrt{\frac{\lambda_i}{\lambda_3}}$, for $i = 1, 2$. Then, it is easy to check that volume is preserved by rescaling the axes so that

$$a_1 = \sqrt[3]{\frac{r_1^2}{r_2}} R_{\rm bdry}, \quad a_2 = \sqrt[3]{\frac{r_2^2}{r_1}} R_{\rm bdry}, \quad a_3 = \frac{1}{\sqrt[3]{r_1 r_2}} R_{\rm bdry}. \tag{9}$$

(iii) $\boldsymbol{S}$ is estimated in the new volume, yielding a new set of eigenvectors and eigenvalues. Then, step (ii) can be repeated with these new values. This iterating scheme is repeated until convergence.

As in Zemp et al. (2011), convergence is assessed by the change in the semi-axes ratios, r_1 and r_2, in two consecutive iterations. We stop the iterative scheme when the relative change in these magnitudes is smaller than 10^{-3}.

This process is first done for the latest code output, thus characterizing the shape of mass distributions at $z = 0$. In order to trace the principal axes to higher redshifts, the same procedure is repeated on each code output. However, after the principal axes have been determined, it could happen that the eigenvalues have changed their order. In order to prevent sudden rotations of the axes, the principal axes on a given snapshot, $\hat{\boldsymbol{v}}_i$, are matched to the ones in the previous snapshot, $\hat{\boldsymbol{v}}'_{j(i)}$, by pairing up the vectors so that $\sum_i |\hat{\boldsymbol{v}}_i \cdot \hat{\boldsymbol{v}}'_{j(i)}|$ is maximized, i.e. by performing the smallest possible rotation.

4.2 Angular distribution of the mass flows in the principal axes system

As a result of applying the method described in Section 4.1 to cluster CL01, Fig. 6 presents the angular distribution of the mass flows at two characteristic moments in the evolution of this object. The left-hand panel shows the angular distribution of accreting and deaccreting gas at $z \simeq 0.81$, right after a major merger (see Fig. 2), while the baryon accretion rate is still high. The right-hand panel presents the same information for a quiescent stage in the evolution of the cluster, while $\Gamma_{\rm vir}^{\rm baryons} \approx 0$. Besides the flux maps, the adjacent plots show the polar and azimuthal marginal distributions of accreting (the blue line) and deaccreting (the orange line) gas. In order to produce these plots, one needs to represent positive (entering matter) and negative (escaping matter) values that span a broad range of orders of magnitude. To do so, we have implemented a symmetric logarithmic scale (see Appendix A).

The marginal $\cos\theta$ distribution of both panels reveals that gas accretion fluxes are maximum in the polar regions (defined with

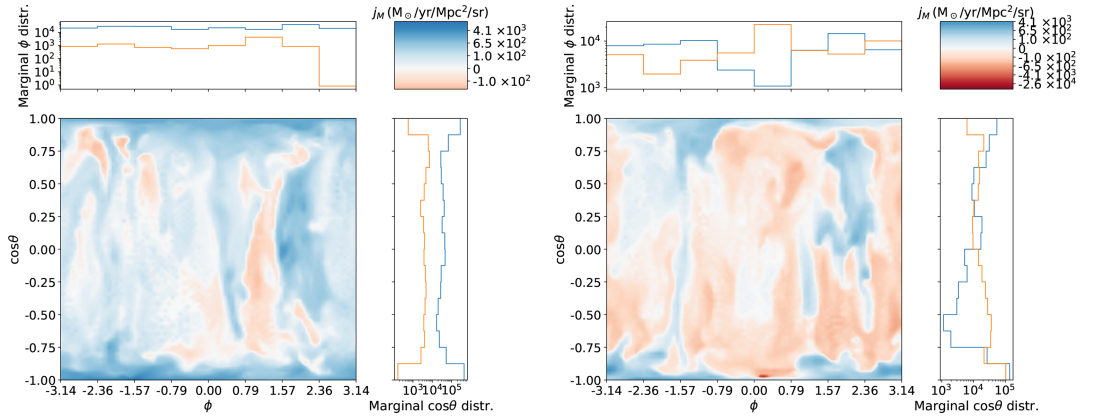


Figure 6. Angular distribution of mass flows through the $r = R_{\rm vir}$ surface of cluster CL01 at $z \simeq 0.81$ (during a major merger; left-hand panel) and at $z \simeq 0.35$ (in the smooth accretion regime; right-hand panel). The colour scale encodes the mass flux density in symmetric logarithmic scale (the orange tones indicate escaping gas, while the blue tones indicate gas being accreted). The upper histogram shows the marginal azimuthal distribution of the entering (blue) and escaping (orange) fluxes, $\int_{-1}^{1} j_M(\theta,\phi)\mathrm{d}(\cos\theta)$. Likewise, the histogram on the right shows the marginal polar distribution of the fluxes, $\int_{-\pi}^{\pi} j_M(\theta,\phi)\mathrm{d}\phi$.

respect to the major axis), corresponding to the gas infalling along the cosmic filaments (Lee & Evrard 2007; Lee et al. 2008). Indeed, the regions limited by $|\cos\theta| > 0.875$ (which correspond to 1/8 of the total solid angle) account for 60 per cent and 54 per cent of the total gas mass inflow, for the snapshots shown in the left-hand and right-hand panels of Fig. 6, respectively.

Outside the polar region, gas accretes in a much smoother manner. However, far from being isotropic, the flows present intricate structures. At $z \simeq 0.81$, during the high-accretion regime, mass inflows dominate almost everywhere, but some structures stretched along the θ direction stand out. At $z \simeq 0.35$, the baryonic component has a net MAR of $\Gamma_{\rm vir}^{\rm baryons} \simeq -0.5$, which could be tempting to interpret as the absence of strong mass flows. However, the right-hand panel in Fig. 6 shows that this interpretation is not a good descriptor of the physical scenario. The low absolute value of Γ emerges as a result of a rather complex counterbalance between strong inflows and outflows. Once again, stretched structures of accretion and deaccretion at nearly constant ϕ are present. Accretion continues dominating in the polar region, while loss of gas is prevalent elsewhere.

4.2.1 *Multipolar expansion of the mass fluxes*

The plots in Fig. 6 depict a complex pattern of gas inflows and outflows. In order to extract quantitative information from these accretion patterns, we develop the mass flux through the $r = R_{\rm bdry}$ boundary in the basis of real spherical harmonics:

$$j_M(\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} c_{lm}\mathcal{Y}_{\rm lm}(\theta,\phi), \tag{10}$$

where $c_{\rm lm} = \oiint \mathrm{d}\Omega\mathcal{Y}_{\rm lm}(\theta,\phi)\, j_M(\theta,\phi) \in \mathbb{R}$ (see Appendix B). However, the discrete sampling of the mass flux in $n_\phi \times n_\theta$ bins constrains the maximum degree, l, that can be faithfully reconstructed. As the spherical harmonics of degree l represent variations on angular scales π/l rad, the sum over l shall be limited to $l_{\rm max} \equiv \min(n_\phi/2, n_\theta)$, i.e. the Nyquist frequency of the grid.[5] In our case, having $n_\phi \times n_\theta = 80 \times 80$, the multipolar expansion ought to be cut at $l_{\rm max} = 40$.

Fig. 7 shows how higher degree components capture the details of the accretion pattern on increasingly smaller angular scales, using the accretion fluxes in the left-hand panel of Fig. 6 as an example. From left to right, columns present the *reconstructed* fluxes in their upper panel, computed by cutting the series at $l_{\rm max} = 4$, 8, 16, and 32. The lower panel corresponds to the residuals with respect to the real flux. Even though low-order degree spherical harmonics represent the dominant contributions, as it will be seen below, higher degrees need to be reached in order for the reconstructed flux to resemble the real one. We quantify the goodness of the fit by computing the rms relative difference between the real and the reconstructed flux. This quantity steadily decreases with increasing $l_{\rm max}$ up to $l_{\rm max} \sim 25$, where it stalls as the high-order multipoles involved vary in angular scales comparable to our sampling of the flux.

4.2.2 *Power spectrum of the angular distribution of accretion fluxes*

In order to extract the contribution of the different angular scales to the accretion flux, we define the *power spectrum of the angular distribution of mass flows* as the average of the square of the c_{lm} coefficients for a given l:

$$P(l) = \frac{1}{2l+1}\sum_{m=-l}^{l} c_{\rm lm}^2 \tag{11}$$

This quantity encodes the magnitude of the contribution of components varying on angular scales $\sim 180^{\rm o}/l$ to the flux. The left-hand panel in Fig. 8 exemplifies this by showing this quantity computed for the same mass flux map in the left-hand panel of Fig. 6.

The overall behaviour of the power spectrum, decreasing with l, reveals that the components corresponding to fluctuations on large angular scales (i.e. low l; especially $l = 0$, 2 and 4) are dominant,

[5] This condition simply corresponds to requiring that the function is sampled twice per period.

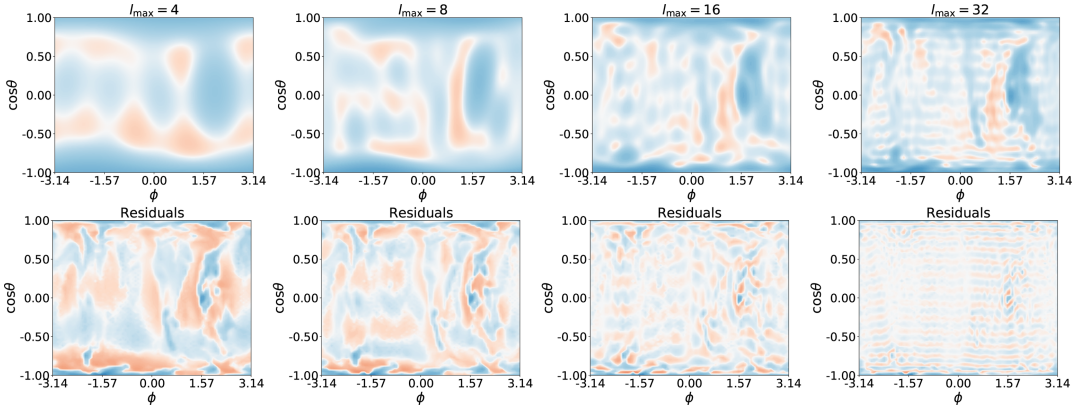


Figure 7. Reconstruction of the mass flux (corresponding to the left-hand panel of Fig. 6) with increasingly higher degree spherical harmonic components: from left to right, $l_{max} = 4,\ 8,\ 16,\ 32$. The upper panels present the flux reconstructed by adding all the terms in the multipolar expansion up to $l = l_{max}$. The lower panels show the residuals of this expansion, i.e. the subtraction of the reconstructed flux from the real flux. The colour scale of all panels has been kept the same as in the left-hand panel of Fig. 6.

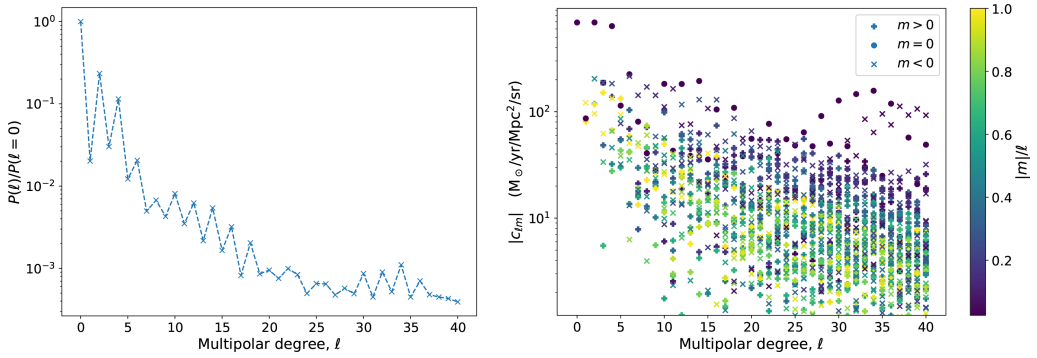


Figure 8. The left-hand panel presents the power spectrum of the mass flows, $P(l)$, normalized to the value corresponding to the monopolar order, $P(l = 0)$. The right-hand panel contains the absolute value of the individual $c_{\rm lm}$ coefficients, as a function of the multipolar degree l. The spherical harmonic order, m, is encoded in the colour scale as $|m|/l$, with the sign given by the shape in the legend. Both representations have been computed for the accretion fluxes of CL01 at $z \simeq 0.81$ through $R_{\rm vir}$, corresponding to the accretion map on the left-hand panel of Fig. 6.

while the small-scale components only provide small corrections. Consequently, the overall accretion and deaccretion patterns can be described, to a great extent, by looking at these simpler, main contributions. Interestingly, the power spectrum shows a clear odd–even effect: even l components are systematically larger than odd l components, implying that parity-even modes are responsible for the bulk of the gas flows through the $R_{\rm vir}$ boundary at this particular code output. Note, however, that although this effect is sustained in time and can be easily understood for the lower degree components ($l = 0$ and 2, which are further discussed below), the same is not true for higher degree components, where this effect is not so clear through all iterations.

On the right-hand panel of Fig. 8, the magnitudes of the $c_{\rm lm}$ coefficients (for the same system and code output) are represented, offering complementary information to the power spectrum (which only gives the average over the orientations of the different multipoles, m, for a given l). In the figure, the horizontal axis encodes the degree, l, while the colour scale is used to indicate the order, m, of the real spherical harmonic. In this particular case, three $m = 0$ components completely dominate the mass flux: namely, $l = 0,\ 2,$ and 4. Their relative weights undergo significant evolution during the redshift interval of this study, $1.5 \leq z \leq 0$. Let us further describe the lowest order components:

The monopolar, isotropic, or smooth component ($l = 0$). This spherically symmetric contribution represents the average isotropic mass flux across the $r = R_{\rm bdry}$ boundary, and hence, can be interpreted as a MAR estimate. In Section 4.2.6, we explore the possibility of using it as a MAR proxy and compare it to Γ_{200m}.

The dipolar components ($l = 1$), or headwind. These multipoles, which account for an excess of accretion flows on a hemisphere and a defect in the antipodal one, can be visualized as the resulting mass flow pattern of an object moving through a dense medium. Although we do not explore it in this manuscript, its usage appears to be promising in the exploration of phenomena like ram pressure stripping (Quilis, Moore & Bower 2000; Quilis et al. 2017).

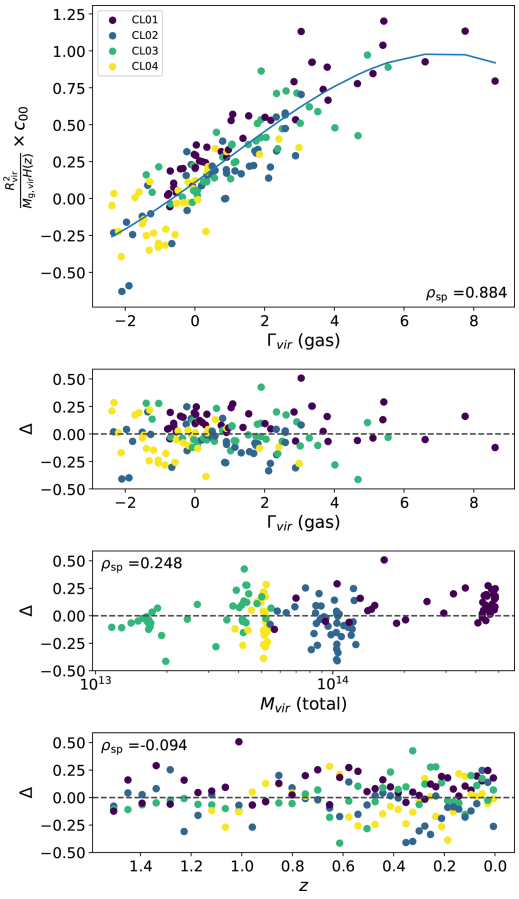


Figure 9. Relation between the instantaneous MAR proxies Γ_{200m} and c_{00}, combining all the snapshots in $1.5 \geq z \geq 0$. The latter has been normalized to cancel the expected dependencies on redshift and mass. The upper panel contains the scatter plot of both variables and a nonparametric fit using smoothing splines (the blue line). The residuals of this fit, Δ, are used in the lower panels to assess whether there is still a redshift or mass dependence in this relation.

The aligned quadrupole or filamentary component ($l = 2$, $m = 0$). This component represents a strong mass inflow through the poles (i.e. through the directions of the major axis) and an outflow near the equatorial regions. Thus, it is the natural candidate to account for the contribution of the mass flux through the cosmic filaments that connect a cluster with a near, massive neighbour.

4.2.3 The monopolar component as an MAR estimate

As already introduced, the monopolar coefficient, c_{00}, is a measurement of the angularly averaged mass flux, and can thus be used in order to quantify the MAR of a cluster. In order to check to which extent does c_{00} agree with other MAR estimates, Fig. 9 presents the analyses of the correlation and the residual dependencies between c_{00} and Γ_{200m}, in the same way we have done for α_{200m} in Section 3.4. Note, however, that c_{00} is a measure of the isotropic mass flux, while Γ_{vir} quantifies the logarithmic increase in the enclosed mass. In order to provide a comparison where no residual dependencies with the mass or redshift are expected, the vertical axes in Fig. 9 show $\frac{R^2_{vir}}{M_{g,vir}H(z)} \times c_{00}$, instead of just c_{00}. This normalization can be understood from the following reasoning:

The MAR, $\dot{M} \equiv \frac{dM}{dt}$, can be got from the mass flux by integrating it over the surface of a sphere of radius R_{bdry}. Since all the non-monopolar spherical harmonics average to zero, introducing the multipolar expansion of $j_M(\theta, \phi)$, equation (10), into the surface integral allows to write the MAR as $\dot{M} = \sqrt{4\pi} R^2 c_{00}$. The same quantity can be obtained from Γ_Δ, as $\dot{M} \equiv \frac{dM}{dt} = \frac{dM}{d\log M}\frac{d\log M}{d\log a}\frac{d\log a}{dt} = M\Gamma H(z)$. Therefore, $R^2 c_{00}$ scales as $M\Gamma H(z)$, justifying our choice to compare Γ^{gas}_{vir} with $\frac{R^2_{vir}}{M_{g,vir}H(z)} \times c_{00}$. Note that, additionally, this normalization cancels out the dimensions of c_{00}.

The upper panel in Fig. 9 shows that, even though there is considerable scatter, these two MAR proxies are tightly correlated ($\rho_{sp} = 0.884$). With this sample, the relation appears to stall at high values of Γ^{gas}_{vir}. However, the reduced number of observations in this high-accretion regime prevents us to draw any robust conclusion to this respect. The lower panels present the distribution of residuals with respect to the independent variable, Γ^{gas}_{vir}, mass and redshift. The residuals appear to be uncorrelated to redshift, suggesting that the normalization applied to c_{00} cancels the redshift evolution of the $\Gamma - c_{00}$ relation. Likewise, the residual dependence with the mass, if any, is weak ($\rho_{sp} = 0.248$).

4.2.4 Relative weight of the smooth and the filamentary contributions

Aiming to assess the relative weight of the filamentary (aligned quadrupolar) and isotropic (monopolar) contributions, we define the following parameter:

$$\beta \equiv \frac{c_{20}}{|c_{00}| + |c_{20}|} \tag{12}$$

Note that β is valued in the interval $[-1,\ 1]$. Positive values imply that the aligned quadrupole contributes to increase the gas mass flux in the regions close to the major semi-axis, corresponding to the gas infalling through the cosmic filaments, while the unlikely scenario in which $\beta < 0$ would correspond to a decreased gas mass flux through these regions. If these two components have comparable weights, $|c_{20}| \sim |c_{00}|$, then $|\beta| \sim 1/2$. Likewise, $|\beta| \sim 1$ ($\beta \sim 0$) indicates that the filamentary (smooth) component is the dominant contribution.

As an example, Fig. 10 presents the evolution of this quantity for the cluster CL01. The value of β fluctuates around the value of $\beta = 1/2$, which separates the filamentary-dominated or smooth-dominated accretion regimes. These values of β do not seem to correlate with the merging regimes studied in Section 3.2. Note, however, that this result has been obtained for a particular cluster, and the method should be applied to a whole sample of massive clusters in order to draw any statistically significant conclusions.

4.3 Thermodynamical properties of the accreted gas

The method described in Section 4.1 can also be applied to measure any property of the accreted and deccreted gas, such as its temperature and entropy. For each of the $n_\phi \times n_\theta$ solid angle bins, its temperature and entropy are computed as the mass-weighted average of such quantities over all the entering cells assigned to the bin. Note,

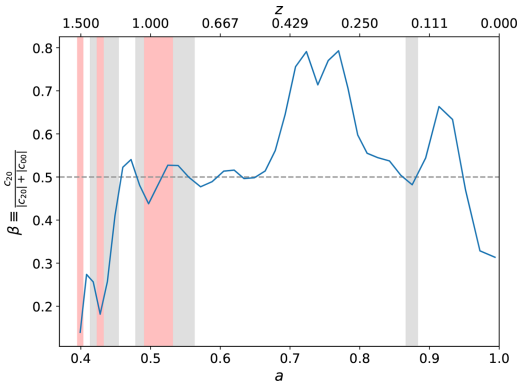


Figure 10. Evolution of the parameter $\beta = \frac{c_{20}}{|c_{00}|+|c_{20}|}$, measuring the relative importance of the filamentary component of the gas mass inflows, for the cluster CL01. The grey-dashed line marks the value $\beta = 1/2$. The grey- and red-shaded regions indicate the accretion regimes, as in Fig. 2.

however, that the angular regions where no gas is inflowing cannot be assigned a temperature nor entropy, resulting in 'gaps' in the corresponding maps.

With the purpose of exploring the thermodynamic differences between the gas accreted from the smooth component and from the filamentary component, we have assigned a temperature and an entropy to each multipolar component as

$$X_{lm} \equiv \frac{\int_{\tilde{\Omega}} d\Omega X(\theta,\phi)\mathcal{Y}_{lm}(\theta,\phi)}{\int_{\tilde{\Omega}} d\Omega \mathcal{Y}_{lm}(\theta,\phi)}, \tag{13}$$

where X represents either the temperature or the entropy, and the integration domain $\tilde{\Omega}$ is the region where $\mathcal{Y}_{lm}(\theta,\phi) > 0$ (and thus the correspondent component represents infall of matter) and $X(\theta,\phi)$ is defined (i.e. at least one gas cell has been marked as entering through the bin and therefore its temperature and entropy can be defined).

The left-hand and central panels in Fig. 11 present, respectively, the evolution of gas temperature and the gas entropy for cluster CL01. In all three panels in the figure, continuous lines represent the evolution of the magnitudes for the smooth component, while the dashed lines correspond to the component infalling through the filaments.

While temperatures of the monopolar and the quadrupolar components do not exhibit a clear distinctive behaviour, their entropies do. Gas being accreted through cosmic filaments tends to have systematically lower entropy than smoothly accreted gas, mainly as a consequence of its higher density. These effects appear to be larger at higher redshifts, while accretion rates are still high for this object (see Fig. 2). The evolution of temperature or entropy of both accreting components does not seem to display a clear relation to merger regimes. However, in order to draw robust conclusions from the analyses of the evolution of thermal properties of the gas accreted smoothly and through filaments, larger samples of massive clusters need to be analysed.

As well as gas mass flows, the method described in Section 4.1 can also be applied to measure the angular distribution of the DM mass flux. The procedure is completely analogous to the one covered for the gas. From the gas and DM mass fluxes, the gas fraction of the accreting material of a given multipolar component, (l, m), is computed as

$$\left(f_g\right)_{lm} \equiv \frac{\int_{\tilde{\Omega}} d\Omega j_M(\theta,\phi)\mathcal{Y}_{lm}(\theta,\phi)}{\int_{\tilde{\Omega}} d\Omega \left[j_M(\theta,\phi) + j_M^{DM}(\theta,\phi)\right]\mathcal{Y}_{lm}(\theta,\phi)}. \tag{14}$$

The right-hand panel of Fig. 11 shows the gas fractions of the matter infalling isotropically and through filaments. These fractions have been normalized to the cosmic baryon fraction (i.e. the grey-dashed line in the plot). The gas fraction of both components fluctuates around the cosmic fraction, being slightly lower during the merger events (due to the collisional nature of gas). The two components show very similar qualitative behaviour, with the gas fractions of matter falling through the cosmic filaments slightly increased with respect to the matter being smoothly accreted. However, these differences are kept small and more statistics are needed in order to draw general conclusions.

5 DISCUSSION AND CONCLUSIONS

In this paper, we have analysed the results of an AMR hydrodynamical coupled to *N*-Body cosmological simulation of a small-volume domain, containing a central, massive galaxy cluster, and several smaller systems. The main focus of this work has been placed on the quantification of the accretion phenomena, with a special focus on the dynamics of the gaseous component. The dynamical scenario depicted by the simulation is essentially dominated by the infall of matter on to the central cluster, which is also the best numerically resolved object. Consequently, most of the focus has been placed on this object, especially in Section 4. Nevertheless, less-massive systems have turned out to be useful in assessing the differences in the MAHs of massive and less-massive clusters.

5.1 Accretion rates

Through these pages, several proxies for the MAR have been compared (see Figs 5 and 9). The most widely adopted MAR proxy in numerical works, $\Gamma_{\Delta}^{[a_1,a_0]}$ (Diemer & Kravtsov 2014), is difficult to relate to an actual observable, since it corresponds to the average accretion rate over several Gyr for typical choices of a_1. The instantaneous MAR proxies α_{Δ} (Lau et al. 2015) and $\Gamma(a)$ (equation 3), however, could in principle be inferred from observations, and can moreover be determined for clusters at any redshift.

- As long as galaxies and gas bulk motions trace DM's, α_{Δ} could be estimated, in the optical band, from the radial velocities of infalling galaxies in clusters' outskirts, or from the radial velocities of gas (from X-ray data, e.g. Sanders et al. 2020; or from kSZ observations, e.g. Adam et al. 2017; see also Simionescu et al. 2019 for a recent review on the topic). In Section 3.4, we have seen that gas systematically infalls with smaller radial velocities than DM, in consistency with Lau et al. (2015), who have computed the averaged radial velocity profiles for a sample of clusters. However, when looking at a specific cluster extracted from our simulation (as we have shown in Fig. 4), we see that the pattern for a single cluster can be fairly more complex than this average behaviour, with gas accreting at larger velocities than DM in some regions. Thus, even though the aforementioned behaviour generally holds, the details of the dynamics in the outskirts of each individual cluster introduces important uncertainties, which make unclear to which extent DM velocities can be inferred from gas or galactic velocities in particular objects, and consequently α_{Δ} can be faithfully measured in observations.

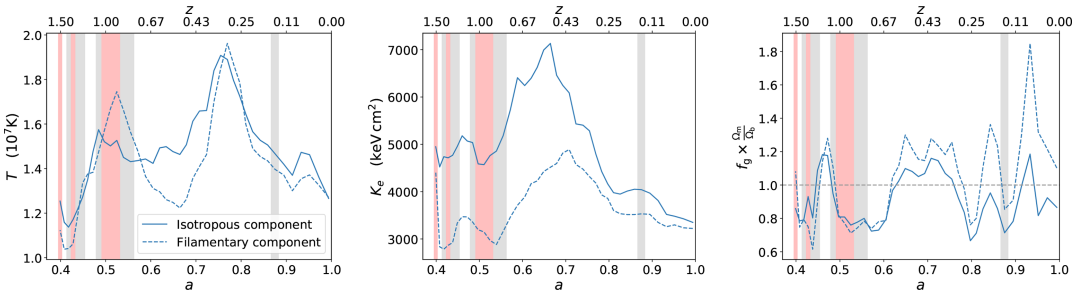


Figure 11. The panels show different thermodynamical properties of the smooth or monopolar component (the solid line) and the filamentary or aligned quadrupolar component (the dashed line) of the mass flows. From left to right, panels show the temperature, the entropy, and the gas fraction of the material being accreted to the cluster. Entropy displays the strongest differential behaviour, with the filamentary component having lower entropy. The shaded regions indicate the merging regimes as in Fig. 2.

• $\Gamma(a)$ cannot be directly measured in observations, since it is defined as a rate of change of the mass. However, in Section 3.2.3, we have obtained a tantalizing correlation between this quantity and the densities in the $1 \leq R/R_{200m} \leq 1.5$ in some massive clusters, suggesting that it might be possible to estimate $\Gamma(a)$ by probing the densities in these regions. However, more statistics are needed to confirm this correlation and to explore the possibility to relate these variables. When computing $\Gamma(a)$ in simulations, it is worth noting that, given that sampling the MAH of clusters introduces statistical noise, the particular differentiation scheme can bias the results: for instance, a Gaussian smoothing of the numerical derivatives leads to overly flattened peaks. This effect can be prevented by using Savitzky–Golay filters.

When comparing α_{200m} to $\Gamma^{\rm total}_{200m}(a)$, we have found that these two quantities display a relatively tight anticorrelation (see Section 3.4.1). If gas velocities are used instead of DM velocities, as suggested above, the Spearman's rank correlation coefficient drops from $\rho_{\rm sp} = -0.832$ to $\rho_{\rm sp} = -0.649$, i.e. gas velocities are not always good indicators of DM velocities.

Last, the coefficient c_{00} presented in Section 4.2.6 is a more direct measurement of the actual mass flux. We have found that once the redshift and mass dependencies are corrected, it correlates strongly to $\Gamma^{\rm gas}_{200m}(a)$, although the scatter continues being large. Even though this parameter is not easily derivable from observational data, it can serve in simulations as a reasonable estimate of the actual gas accretion rate, which can be compared to other proxies.

5.2 Angular distribution of the mass flows

In Section 4, we have presented a novel and general approach to study the accretion phenomena in simulations, through the determination of the angular distribution of the mass flows through the clusters' spherical boundary. The method presented in Section 4.1 allows a simple determination of this quantity by treating cells as pseudo-Lagrangian fluid elements.

The multipolar expansion of the mass flux in the principal axes system of the cluster has been used to extract the elementary contributions to the mass flux. In particular, we have payed special attention to the monopolar and the aligned quadrupolar components, which represent the average isotropic mass flux and the enhancement in the polar regions due to the preferential infall of mass through the cosmic filaments. We have defined a parameter, β, which quantifies the importance of the filamentary component of accretion, with respect to the isotropic component. Applying this analysis to the central, massive cluster in our simulated domain suggests that this geometry of accretion can vary importantly during the accretion history of the cluster, but we have not found any strong relation with merger events in this particular object.

This method can be extended to evaluate properties of the accreted gas, such as its temperature and its entropy. In particular, we have suggested a way to define these properties for each multipolar component. The most direct application of this procedure is assessing the difference between the gas accreted isotropically and the gas accreted through filaments. Our results point out that the gas being accreted through the cosmic filaments has systematically lower entropy than the isotropic component. Recent observational works, Gouin et al. (2020) have studied the multipolar expansion of the projected density in the clusters' outskirts, obtaining a power spectrum which presents similar features to our (3D) spectrum of accretion flows.

The analyses in this paper would importantly benefit from larger samples of clusters in order to enhance the statistical significance of our results and, especially, to pursue further analysis, like relating the accretion phenomena studied in this work to the position of the cluster with respect to the SZ and X-ray mass-observable relations (see e.g. Yu et al. 2015; Chen et al. 2019 for recent studies in this line). Nevertheless, despite this lack of statistics, we have proposed some novel analyses that may offer a new insight into the study of accretion of gas in numerical simulations of cluster formation. With the new generation of X-ray and SZ instruments being able to systematically observe the outskirts of a larger number of galaxy clusters in the next decade, more light will be shed on to the physics happening in these regions.

ACKNOWLEDGEMENTS

We thank the anonymous referee for his/her constructive comments. This work has been supported by the Spanish Ministerio de Ciencia e Innovación (MICINN, grants AYA2016-77237-C3-3-P and PID2019-107427GB-C33) and by the Generalitat Valenciana (grant PROMETEO/2019/071). DVP acknowledges support from the Departament d'Astronomia i Astrofísica de la Universitat de València through an *Iniciació a la Investigació* scholarship. Simulations have been carried out using the supercomputer Lluís Vives at the Servei d'Informàtica of the Universitat de València.. This research has used

the following open-source packages: NUMPY (Oliphant 2006), SCIPY (Virtanen et al. 2020) and MATPLOTLIB (Hunter 2007).

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Adam R. et al., 2017, A&A, 598, A115
Adhikari S., Dalal N., Chamberlain R. T., 2014, J. Cosmol. Astropart. Phys., 2014, 19
Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409
Arfken G., Weber H., Harris F., 2013, Mathematical Methods for Physicists: A Comprehensive Guide. Elsevier Science, Amsterdam
Biffi V. et al., 2016, ApJ, 827, 112
Bryan G. L., Norman M. L., 1998, ApJ, 495, 80
Cen R., Roxana Pop A., Bahcall N. A., 2014, Proc. Natl. Acad. Sci., 111, 7914
Chen H., Avestruz C., Kravtsov A. V., Lau E. T., Nagai D., 2019, MNRAS, 490, 2380
Chen Y., Mo H. J., Li C., Wang H., Yang X., Zhang Y., Wang K., 2020, ApJ, 899, 81
Cole S., Lacey C., 1996, MNRAS, 281, 716
Cui W. et al., 2016, MNRAS, 456, 2566
Diemer B., Kravtsov A. V., 2014, ApJ, 789, 1
Diemer B., Mansfield P., Kravtsov A. V., More S., 2017, ApJ, 843, 140
Eisenstein D. J., Hu W., 1998, ApJ, 496, 605
Forero-Romero J. E., Gottlöber S., Yepes G., 2010, ApJ, 725, 598
Gouin C., Aghanim N., Bonjean V., Douspis M., 2020, A&A, 635, A195
Haardt F., Madau P., 1996, ApJ, 461, 20
Hoffman Y., Ribak E., 1991, ApJ, 380, L5
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Katz N., Weinberg D. H., Hernquist L., 1996, ApJS, 105, 19
Knebe A., Libeskind N. I., Knollmann S. R., Yepes G., Gottlöber S., Hoffman Y., 2010, MNRAS, 405, 1119
Knebe A. et al., 2011, MNRAS, 415, 2293
Kravtsov A. V., Borgani S., 2012, ARA&A, 50, 353
Lacey C., Cole S., 1994, MNRAS, 271, 676
Lau E. T., Nagai D., Avestruz C., Nelson K., Vikhlinin A., 2015, ApJ, 806, 68
Lau E. T., Hearin A. P., Nagai D., Cappelluti N., 2020, preprint (arXiv: 2006.09420)
Lee J., Evrard A. E., 2007, ApJ, 657, 30
Lee J., Springel V., Pen U.-L., Lemson G., 2008, MNRAS, 389, 1266
Mansfield P., Kravtsov A. V., Diemer B., 2017, ApJ, 841, 34
Markevitch M., Vikhlinin A., Mazzotta P., 2001, ApJ, 562, L153
More S., Diemer B., Kravtsov A. V., 2015, ApJ, 810, 36
Nagai D., Vikhlinin A., Kravtsov A. V., 2007, ApJ, 655, 98
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Oliphant T. E., 2006, A guide to NumPy, Vol. 1. Trelgol Publishing, USA
Planck Collaboration, 2018, A&A , 641, A6
Planelles S., Quilis V., 2009, MNRAS, 399, 410
Planelles S., Quilis V., 2010, A&A, 519, A94
Planelles S., Borgani S., Dolag K., Ettori S., Fabjan D., Murante G., Tornatore L., 2013, MNRAS, 431, 1487
Planelles S., Borgani S., Fabjan D., Killedar M., Murante G., Granato G. L., Ragone-Figueroa C., Dolag K., 2014, MNRAS, 438, 195
Planelles S., Schleicher D. R. G., Bykov A. M., 2015, Space Sci. Rev., 188, 93
Planelles S. et al., 2017, MNRAS, 467, 3827
Planelles S., Mimica P., Quilis V., Cuesta-Martínez C., 2018, MNRAS, 476, 4629
Pratt G. W., Arnaud M., Biviano A., Eckert D., Ettori S., Nagai D., Okabe N., Reiprich T. H., 2019, Space Sci. Rev., 215, 25
Press W. H., Teukolsky S. A., 1990, Comput. Phys., 4, 669
Quilis V., 2004, MNRAS, 352, 1426
Quilis V., Moore B., Bower R., 2000, Science, 288, 1617
Quilis V., Planelles S., Ricciardelli E., 2017, MNRAS, 469, 80
Rasia E. et al., 2015, ApJ, 813, L17
Roediger E., Brüggen M., Simionescu A., Böhringer H., Churazov E., Forman W. R., 2011, MNRAS, 413, 2057
Sanders J. S. et al., 2020, A&A, 633, A42
Savitzky A., Golay M. J. E., 1964, Anal. Chem., 36, 1627
Shi X., 2016a, MNRAS, 459, 3711
Shi X., 2016b, MNRAS, 461, 1804
Simionescu A. et al., 2019, Space Sci. Rev., 215, 24
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Sutherland R. S., Dopita M. A., 1993, ApJS, 88, 253
Tamura T., Hayashida K., Ueda S., Nagai M., 2011, PASJ, 63, S1009
Theuns T., Leonard A., Efstathiou G., Pearce F. R., Thomas P. A., 1998, MNRAS, 301, 478
Tormen G., Moscardini L., Yoshida N., 2004, MNRAS, 350, 1397
Virtanen P. et al., 2020, Nat. Methods, 17, 261
Voit G. M., 2005, Rev. Mod. Phys., 77, 207
Walker S. et al., 2019, Space Sci. Rev., 215, 7
Yepes G., Kates R., Khokhlov A., Klypin A., 1997, MNRAS, 284, 235
Yu L., Nelson K., Nagai D., 2015, ApJ, 807, 12
Zemp M., Gnedin O. Y., Gnedin N. Y., Kravtsov A. V., 2011, ApJS, 197, 30

APPENDIX A: THE SYMMETRIC LOGARITHMIC SCALE

When representing the mass fluxes in Section 4.2, or in many other situations where data can be both positive and negative and, at the same time, span a broad range of orders of magnitude, it may be useful to employ a *symmetric logarithmic scale*. Even though some software packages for data visualization (e.g. PYTHON'S MATPLOTLIB, Hunter 2007) include implementations of this scale, its usage is not quite widespread and we have considered covering it here for the interested reader.

The basic underlying idea of our particular implementation of the symmetric logarithmic scale relies on mapping any interval $[-x_{\max}, x_{\max}]$ to the interval $[-1, 1]$ by performing the following continuous transformation:

$$x \mapsto f(x) = \begin{cases} \operatorname{sign}(x)\left[1 + \frac{1-a}{\alpha}\log_{10}\left(\frac{x}{x_{\max}}\right)\right], & \frac{|x|}{x_{\max}} \geq 10^{-\alpha} \\ \frac{a}{10^{-\alpha}x_{\max}}x, & \frac{|x|}{x_{\max}} \leq 10^{-\alpha} \end{cases} \tag{A1}$$

The parameters have been chosen in this way to provide the clearest interpretation. $x_{\max}$ controls the maximum of the scale, i.e. the value that will be mapped to 1. α represents the dynamical range of the representation (the number of orders of magnitude represented in logarithmic scale). Absolute values above $10^{-\alpha}$ are treated logarithmically; below this threshold, linearly. Last, $a = y(10^{-\alpha}x_{\max})$ represents the visual extent of the linear scale. In order for the transformation to be differentiable, a shall be set to $a = (1 + \alpha \ln 10)^{-1}$. However, in many cases differentiability is not necessary, and a can be chosen to better suite the representation purposes. In the plots shown in this work, we have used $\alpha = 3.5$, $a = (1 + \alpha \ln 10)^{-1}$.

APPENDIX B: THE REAL SPHERICAL HARMONICS BASIS

In Section 4.2, we have used the real spherical harmonics basis to study the angular dependence of the mass accretion fluxes. As there is not a general consensus regarding phase and normalization conventions, let us formally define these functions as they have been used in this work.

Let $Y_{\rm lm}(\theta, \phi)$ be the complex spherical harmonic of degree l and order m, with unit square-integral and using the Condon–Shortley phase convention (see e.g. Arfken, Weber & Harris 2013). These functions can be used to expand any square-integrable complex function defined on the unit sphere. However, for real functions, only a half of the coefficients of such expansion are free. Thus, it can be simplified by defining the real spherical harmonics, $\mathcal{Y}_{\rm lm}(\theta, \phi)$, as

$$\mathcal{Y}_{\rm lm}(\theta,\phi) = \begin{cases} (-1)^m \sqrt{2}\ \mathrm{Im}\left[Y_{l,-m}(\theta,\phi)\right], & m < 0 \\ Y_{l0}(\theta,\phi), & m = 0\,. \\ (-1)^m \sqrt{2}\ \mathrm{Re}\left[Y_{l,m}(\theta,\phi)\right], & m > 0 \end{cases} \tag{B1}$$

From the orthonormality properties of the $Y_{\rm lm}$ functions, it is easy to show that any real, square-integrable function defined on the unit sphere, $f(\theta, \phi)$, can be expanded in a series of real spherical harmonics with real coefficients, as

$$f(\theta,\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_{\rm lm} \mathcal{Y}_{\rm lm} \tag{B2}$$

$$\text{with } c_{\rm lm} = \oiint \mathrm{d}\Omega \mathcal{Y}_{\rm lm}(\theta,\phi) f(\theta,\phi).$$

This paper has been typeset from a TEX/LATEX file prepared by the author.

Computer Physics Communications 263 (2021) 107892

Contents lists available at ScienceDirect

Computer Physics Communications

journal homepage: www.elsevier.com/locate/cpc

Unravelling cosmic velocity flows: a Helmholtz–Hodge decomposition algorithm for cosmological simulations☆

David Vallés-Pérez [a,*], Susana Planelles [a,b], Vicent Quilis [a,b]

[a] Departament d'Astronomia i Astrofísica, Universitat de València, E-46100 Burjassot (València), Spain
[b] Observatori Astronòmic, Universitat de València, E-46980 Paterna (València), Spain

ARTICLE INFO

Article history:
Received 17 November 2020
Received in revised form 26 January 2021
Accepted 11 February 2021
Available online 16 February 2021

Keywords:
Turbulence
Large-scale structure of Universe
Galaxies: Clusters: Intracluster medium
Galaxies: Clusters: General
Methods: Numerical
Adaptive Mesh Refinement

ABSTRACT

In the context of intra-cluster medium turbulence, it is essential to be able to split the turbulent velocity field in a compressive and a solenoidal component. We describe and implement a new method for this aim, i.e., performing a Helmholtz–Hodge decomposition, in multi-grid, multi-resolution descriptions, focusing on (but not being restricted to) the outputs of AMR cosmological simulations. The method is based on solving elliptic equations for a scalar and a vector potential, from which the compressive and the solenoidal velocity fields, respectively, are derived through differentiation. These equations are addressed using a combination of Fourier (for the base grid) and iterative (for the refinement grids) methods. We present several idealised tests for our implementation, reporting typical median errors in the order of 1‰-1%, and with 95-percentile errors below a few percents. Additionally, we also apply the code to the outcomes of a cosmological simulation, achieving similar accuracy at all resolutions, even in the case of highly non-linear velocity fields. We finally take a closer look to the decomposition of the velocity field around a massive galaxy cluster.

© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Cosmological structures and, in particular, galaxy clusters, which constitute the most massive structures which have had time to collapse under their own gravity, are dynamically interesting objects from several perspectives. The non-linearity of the evolution of their baryonic component (i.e., ordinary matter, most of which is in the form of a hot, tenuous plasma known as the intra-cluster medium, ICM) couples the different scales in the evolution of cosmic inhomogeneities, producing a plethora of complex hydrodynamical phenomena, such as shock waves and turbulence. Turbulence is an intrinsically multi-scale phenomenon, since bulk motions trigger (magneto-)hydrodynamical instabilities (e.g., [1]) which cascade down to smaller scales until they get dissipated by viscous effects.

Turbulent motions can be [2] and have been recently [3,4] measured on a number of nearby galaxy clusters from X-ray surface brightness fluctuations, which have in turn been connected to signatures of particle acceleration and diffuse radio emission in the ICM [5,6]. Future X-ray facilities, like Athena,[1] will offer unprecedented insight into the dynamics of turbulent motions in the ICM [7]. However, the precise theoretical and numerical description of turbulence in these vast structures is still matter of ongoing research (e.g., [8–11], just to highlight a few recent numerical studies).

Splitting the (turbulent) velocity field in its compressive and solenoidal components, i.e., performing a Helmholtz–Hodge decomposition (HHD), is a crucial step towards exploring the role of turbulence in the ICM, since these components play fundamental and distinct roles in the evolution of cosmic structures. Thus, while the solenoidal component is likely the major responsible for the amplification of cosmic magnetic fields [12,13], compressive turbulence has an important role in generating weak shocks which have consequential effects on the magnetic and thermal evolution of the cluster (e.g., [13,14]). These two components also differ in their spatial distribution, the former tending to be more volume-filling [10,15,16], and even in their spectrum, steeper for the compressive component [8,17–19]. The distinction between the solenoidal and compressive component of the velocity field is also of utmost importance to model the acceleration of cosmic ray particles in the ICM, which will become a vibrant field of observational research with the advent of the new generation of radio telescopes (e.g., SKA [20]). In this line, a lot of effort has been recently put in modelling the acceleration by compressive and solenoidal modes in magnetohydrodynamics (MHD; [21–25])

Cosmological hydrodynamical simulations, as many other applications in computer fluid dynamics, need a huge dynamical

☆ The review of this paper was arranged by Prof. David W. Walker.

* Corresponding author.
E-mail addresses: david.valles-perez@uv.es (D. Vallés-Pérez), susana.planelles@uv.es (S. Planelles), vicent.quilis@uv.es (V. Quilis).

[1] www.the-athena-x-ray-observatory.eu.

https://doi.org/10.1016/j.cpc.2021.107892
0010-4655/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

range to be resolved in order to, for example, form realistic galaxies in a cosmological environment [26,27] or adequately describe turbulence in the ICM (see, e.g., [28] for some graphical examples of the effects of resolution on the ability to capture instabilities; see also [29,30], who present detailed studies of stratified, ICM-like turbulence in numerical grids of varying resolution). While Lagrangian codes are inherently adaptive, Eulerian codes based on high-resolution shock-capturing (HRSC) techniques are especially capable of handling shocks and other types of discontinuities, which are pervasive in the formation of cosmological structures ([31], for a review). That is why, among several other options (see, for example, [32] for a broad review), Adaptive Mesh Refinement (AMR) codes are especially suited for cosmological structure formation.

Previous studies of ICM turbulence have already implemented HHD algorithms. For example, several works using uniform grids (or fixed refinement strategies which, ultimately, allow to resolve the object of interest within a constant resolution) perform the decomposition in Fourier space, where it simply reduces to linear algebra projections (e.g., [10,33]). Additionally, [10] also confront this method with solving a Poisson equation (in Fourier space) to find a scalar potential for the compressive velocity component, reporting more accurate results for the first method. However, any of these two procedures, because of their usage of fast Fourier transform (FFT) algorithms, can only be applied to regular, uniform grids, which are not the common use in cosmological simulations.

In this paper, we propose, implement and test a new algorithm for performing the Helmholtz–Hodge decomposition in a multi-scale AMR grid, which can therefore be applied to the outcomes of a full-cosmological simulation without the need of performing resimulations of specific objects or constrained simulations. Our method decomposes the velocity fields by solving elliptic partial derivative equations (PDEs), which can be addressed iteratively using a wide variety of well-known algorithms (e.g., [34]). Nevertheless, although our primary focus is the application to cosmological structure formation, we emphasise that the approach presented in this work can be directly applied to any application of block-structured AMR, and readily extended to particle-based simulations through a suitable interpolation scheme.

The rest of the manuscript is organised as follows. In Section 2, we describe our method for performing the decomposition and discuss its numerical implementation. In Section 3 we present and describe a set of tests to validate the accuracy of the algorithm, while in Section 4 we apply it to the complex velocity field of a cosmological simulation. Last, in Section 5 we summarise and present our conclusions.

2. Description of the method

The algorithm is based on the Helmholtz–Hodge decomposition (see, e.g., [35–37]), which allows to univocally split any velocity field in three terms,

$$\mathbf{v} = \mathbf{v}_{\rm comp} + \mathbf{v}_{\rm rot} + \mathbf{v}_{\rm harm} \tag{1}$$

where $\mathbf{v}_{\rm comp}$ is the compressive (or irrotational, $\nabla \times \mathbf{v}_{\rm comp} = \mathbf{0}$) velocity field, $\mathbf{v}_{\rm rot}$ is the purely rotational (or solenoidal, $\nabla \cdot \mathbf{v}_{\rm rot} = 0$) velocity field and $\mathbf{v}_{\rm harm}$ is the harmonic velocity field (both irrotational and solenoidal, thus satisfying $\nabla^2 \mathbf{v}_{\rm harm} = \mathbf{0}$ and $\mathbf{v}_{\rm harm} = \nabla\chi$, with $\nabla^2\chi = 0$).

The harmonic component can be shown to be identically null in any domain with periodic boundary conditions, and therefore we will no longer consider it. Because of their defining properties, the compressive component can be written as the gradient of a scalar potential, $\mathbf{v}_{\rm comp} = -\nabla\phi$, while the rotational component can be derived from a vector potential, $\mathbf{v}_{\rm rot} = \nabla \times \mathbf{A}$. From this, it is easy to derive that the scalar and vector potentials can be computed, respectively, as the solutions of the following elliptic PDEs,

$$\nabla^2\phi = -\nabla \cdot \mathbf{v} \tag{2}$$

$$\nabla^2\mathbf{A} = -\nabla \times \mathbf{v} \tag{3}$$

which are formally equivalent to a set of four decoupled Poisson equations (one for ϕ and one for each cartesian component of $\mathbf{A}$) whose sources are the divergence and the components of the curl of the overall velocity field.

Once the potentials have been obtained, the compressive and rotational components of the velocity field can be obtained through differentiation. Note that, in principle, it would only be necessary to find one of the potentials (either ϕ or $\mathbf{A}$), since the other velocity component could be then derived by subtracting to the total velocity (Eq. (1) with $\mathbf{v}_{\rm harm} = \mathbf{0}$). Nevertheless, we have chosen to compute all the potentials in order to keep track of the associated numerical errors.

2.1. Numerical implementation

We have designed a code to perform such decomposition in a multi-resolution, block-structured AMR velocity field. As mentioned before, our code can be easily applied to the outcomes of any AMR simulation (cosmological or not), or even to a particle-based code, from which the continuous velocity field can be defined on an ad-hoc AMR grid structure through a smoothing method (e.g., a particle-mesh, [38]).

In our particular implementation, for the base (coarsest) level, $\ell = 0$, taking advantage of the periodic boundary conditions, Poisson equations are solved in Fourier space,[2] where they just reduce to a multiplication by the Green's function. The basic procedure followed to solve Poisson equations in the base grid can be described in the following steps:

1. The FFT of the source (right-hand side terms in Eqs. (2) or (3)) in the base grid is computed, yielding a set of coefficients F_{lmn}.
2. Poisson's equation is then solved in Fourier space by multiplying by the Green's function, G_{lmn}: $\tilde{\phi}_{lmn} = G_{lmn}F_{lmn}$. The Green's function is given by:

$$G_{\rm lmn} = \frac{(\Delta x/2)^2}{\sin^2\left(\frac{\pi l}{N_x}\right) + \sin^2\left(\frac{\pi m}{N_y}\right) + \sin^2\left(\frac{\pi n}{N_z}\right)} \tag{4}$$

   where Δx is the cell side length and the domain is discretised in $N_x \times N_y \times N_z$ cells [38].
3. The inverse FFT of $\tilde{\phi}_{lmn}$ yields the sought potential at the base level.

[2] Note, however, that this is not a requirement of the method. We solve Poisson's equations in Fourier space as cosmological simulations of sufficiently large volumes typically implement periodic boundary conditions and, in these situations, solving Poisson's equation in Fourier space is much more computationally efficient than using iterative methods. In any case, for non-periodic domains, the base level can be addressed through iterative schemes, just as described below for the refinement levels, if suitable boundary conditions are provided. In the case of non-periodic domains, however, the harmonic term cannot be dropped in Eq. (1). While we cannot compute $\mathbf{v}_{\rm harm}$ by solving elliptic equations, this term can be obtained just by using Eq. (1) to solve for it, once $\mathbf{v}_{\rm comp}$ and $\mathbf{v}_{\rm rot}$ have been found. This is a good approach, as long as our algorithm precisely reconstructs the compressive and rotational velocity fields from an input field, which is tested in Sections 3 and 4.

In subsequent refinement levels, $\ell > 0$, Poisson equations have to be solved taking into account the boundary conditions imposed by the coarser grids the refined patches are embedded into. In order to do so, we use a successive over-relaxation procedure (SOR; see, for example, [34]) on the discretised Poisson equation.

Each AMR patch is first initialised (both in the boundary and in the interior cells) by linear interpolation from the values of the potential at the best-resolved lower level patch available. Then, the interior cells are iteratively updated in a chessboard pattern as

$$\phi^{\text{new}}_{i,j,k} = \omega\phi^{*}_{i,j,k} + (1-\omega)\phi^{\text{old}}_{i,j,k}, \tag{5}$$

where

$$\begin{aligned}\phi^{*}_{i,j,k} = \tfrac{1}{6}\big[&\phi^{\text{old}}_{i+1,j,k} + \phi^{\text{old}}_{i-1,j,k} + \phi^{\text{old}}_{i,j+1,k} + \phi^{\text{old}}_{i,j-1,k} \\ &+\phi^{\text{old}}_{i,j,k+1} + \phi^{\text{old}}_{i,j,k-1} - (\Delta x_\ell)^2 f_{i,j,k}\big],\end{aligned} \tag{6}$$

being Δx_ℓ the cell size at the given refinement level, $f_{i,j,k}$ the source term and $1 < \omega < 2$ the over-relaxation parameter. In order to boost convergence, we set ω according to the Chebyshev acceleration procedure [34]. Aiming to avoid undesirable boundary effects due to the interpolation of the potential boundary conditions, we extend the patches with 3 fictitious cells in all directions, so that these boundary conditions are enforced slightly far away from the region of interest.

Once ϕ and $\mathbf{A}$ are known, the velocity components are found by finite differencing the potentials as defined before. We compute the derivatives (both of the velocity and of the potentials) using an eighth-order scheme with a centred stencil of up to[3] 9 points, which provides robust values of the differential operators of the velocity fields and mitigates the impact of spurious noise.

The code is parallelised according to the OpenMP standard directives. Our implementation is freely available through its GitHub repository.[4]

3. Tests

Aiming to assess the robustness of our HHD method and its implementation, we have designed a battery of tests focused on quantifying to which extent the code is able to accurately identify and disentangle the compressive and rotational velocities. We describe such tests in Section 3.1, and examine their results in Section 3.2.

3.1. Description of the tests

For the four tests described below, we have first established a simple AMR grid structure, i.e., a set of patches at different refinement levels that could reasonably mock the ones generated in an actual simulation.

We have considered a cubic domain of unit length, with origin at the centre of the box, and we have discretised it with 128^3 cells. For each octant, we establish a first refinement patch with twice the spatial resolution covering the central 1/8 of the octant volume. For example, in the first octant ($0 < x, y, z < 1/2$), we set up a patch at $1/8 < x, y, z < 3/8$ with 64^3 cells (and likewise for the remaining 7 octants). Then we add a $\ell = 2$ patch in the central 1/8 volume of each $\ell = 1$ patch, and continue recursively.

In the first three tests we consider up to $\ell_{\max} = 10$ refinement levels, providing a peak resolution of 7.63×10^{-6} (relative to the box size, which is normalised to 1). Fig. 1 presents graphically the grid structure employed in these tests. Below, we describe the velocity fields that we have seeded on these grids.

[3] The stencil length is shortened as cells get closer to the boundary.

[4] https://www.github.com/dvallesp/vortex/.

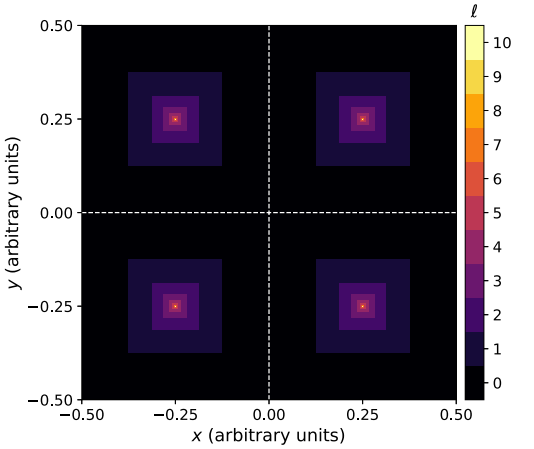


Fig. 1. AMR grid structure for the tests in Section 3. The figure represents a slice at $z = 0.25$, with the colours encoding the highest refinement level at each point in the x–y plane. The white, dashed lines indicate the cartesian x and y axes.

3.1.1. Test 1: constant divergence field

For the first test, we consider a purely compressive velocity field with constant divergence, given in cartesian and spherical coordinates by the analytical expression

$$\mathbf{v} = \omega_0 \left(x\,\hat{\mathbf{u}}_\mathbf{x} + y\,\hat{\mathbf{u}}_\mathbf{y} + z\,\hat{\mathbf{u}}_\mathbf{z}\right) \equiv \omega_0 r\,\hat{\mathbf{u}}_\mathbf{r}. \tag{7}$$

This field has $\nabla\cdot\mathbf{v} = 3\omega_0$ and $\nabla\times\mathbf{v} = \mathbf{0}$, and hence $\mathbf{v} = \mathbf{v}_{\text{comp}}$ and $\mathbf{v}_{\text{rot}} = 0$. It is easy to find an analytic expression for the scalar potential, $\phi = -\frac{\omega_0}{2}\left(x^2+y^2+z^2\right) + \mathcal{C}$, with $\mathcal{C}$ any arbitrary real constant. Likewise, the vector potential ought to be $\mathbf{A} = \nabla\chi$, with χ an arbitrary scalar function or, in particular, $\mathbf{A} = \mathbf{0}$. We have set $\omega_0 = 0.01$.

3.1.2. Test 2: constant curl field

Analogous to the previous one, we have also considered the case of a purely uniformly rotating velocity field, which can be analytically given in cartesian and cylindrical coordinates as:

$$\mathbf{v} = \omega_0 \left(-y\,\hat{\mathbf{u}}_\mathbf{x} + x\,\hat{\mathbf{u}}_\mathbf{y}\right) \equiv \omega_0\rho\,\hat{\mathbf{u}}_{\boldsymbol{\phi}} \tag{8}$$

It is straightforward to show that $\nabla\cdot\mathbf{v} = 0$ and $\nabla\times\mathbf{v} = 2\omega_0\,\hat{\mathbf{u}}_\mathbf{z}$. This velocity field is generated by the potentials $\phi = \mathcal{C}$, being $\mathcal{C} \in \mathbb{R}$ a free constant, and $\mathbf{A} = -2\omega_0\rho\,\hat{\mathbf{u}}_\mathbf{z} + \nabla\chi$, with χ any arbitrary scalar function. As in Test 1, we set $\omega_0 = 0.01$.

These two previous tests (Test 1 and Test 2) are aimed to validate the reconstruction of pure velocity fields (either compressive or rotational, respectively), i.e., to estimate the magnitude of the errors involved in the procedure described in Section 2.1 in idealised situations where there is no cross-talk between rotational and compressive components. Last, note that even though these two velocity fields are not periodic in the mock simulation domain, this will only affect a negligible amount of cells in the domain's boundary.

3.1.3. Test 3: compressive and rotational, periodic field

We have designed a third test, aimed to assess the effects of the cross-talk between the compressive and the rotational components. We have considered the velocity field:

$$\begin{aligned}\mathbf{v} = &[\sin(2\pi x) + \sin(4\pi y) + \sin(6\pi z)]\,\hat{\mathbf{u}}_\mathbf{x} \\ &+[\sin(6\pi x) + \sin(2\pi y) + \sin(4\pi z)]\,\hat{\mathbf{u}}_\mathbf{y} \\ &+[\sin(4\pi x) + \sin(6\pi y) + \sin(2\pi z)]\,\hat{\mathbf{u}}_\mathbf{z}.\end{aligned} \tag{9}$$

The compressive part corresponds to the terms of angular frequency 2π, while the higher frequency ones constitute the rotational component. Also in this case, it is easy to find analytical solutions to ϕ and $\mathbf{A}$:

$$\phi = \frac{1}{2\pi}\left[\cos(2\pi x) + \cos(2\pi y) + \cos(2\pi z)\right] + \mathcal{C} \tag{10}$$

$$\begin{aligned}\mathbf{A} = &\left[-\frac{1}{4\pi}\cos(4\pi z) + \frac{1}{6\pi}\cos(6\pi y)\right]\hat{\mathbf{u}}_{\mathbf{x}} \\ &+ \left[-\frac{1}{4\pi}\cos(4\pi x) + \frac{1}{6\pi}\cos(6\pi z)\right]\hat{\mathbf{u}}_{\mathbf{y}} \\ &+ \left[-\frac{1}{4\pi}\cos(4\pi y) + \frac{1}{6\pi}\cos(6\pi x)\right]\hat{\mathbf{u}}_{\mathbf{z}} + \nabla\chi\,.\end{aligned} \tag{11}$$

3.1.4. Test 4: ICM-like mock velocity field

While the previous tests have checked the ability of the code to reconstruct idealised solenoidal, compressive and mixed velocity fields, we have implemented a last test aimed to assess the ability of the code to capture and reconstruct variations on a broad range of spatial frequencies. The test is, in part, inspired by the one presented by [10, App. A.1.2], but with some differences aimed to mix both velocity fields (instead of generating a purely solenoidal or compressive field), while still having an analytic solution of the HHD to compare with the numerical results.

We have generated our mock, ICM-like velocity field according to the procedure described below:

1. We consider a uniform grid of $(N_x \times 2^{\ell_{\max}})^3$ cells. In that grid, we compute the velocity field $\mathbf{v} = \mathbf{v}_{\rm comp} + \mathbf{v}_{\rm rot}$, with:

$$\mathbf{v}_{\rm comp} = \sum_{i=x,y,z}\sum_{n=N_{\min}}^{N_{\max}} A_n^{\rm comp}\sin\left(\frac{2\pi n}{L}x_i + \psi_{\rm n}^{\rm comp,i}\right)\hat{\mathbf{u}}_{\mathbf{i}} \tag{12}$$

$$\mathbf{v}_{\rm rot} = \sum_{i=x,y,z}\sum_{j\neq i}\sum_{n=N_{\min}}^{N_{\max}} A_n^{\rm rot}\sin\left(\frac{2\pi n}{L}x_j + \psi_{\rm n}^{\rm rot,ij}\right)\hat{\mathbf{u}}_{\mathbf{i}} \tag{13}$$

being L the box side length ($L = 1$ in our case), $A_n^{\rm comp}$ ($A_n^{\rm rot}$) the amplitude of the mode of frequency n of the compressive (solenoidal) component, and $\psi_{\rm n}^{\rm comp,i}$ ($\psi_{\rm n}^{\rm rot,ij}$) the initial phases. For the initial phases, we have generated 9 sets of $N_{\max} - N_{\min} + 1$ random numbers, uniformly sampled from the interval $0 < \psi < 2\pi$. The amplitudes are generated so that the compressive (solenoidal) component follows a Burgers [39] (Kolmogorov [40]) spectrum, $E(k) \propto k^{-2}$ ($E(k) \propto k^{-5/3}$). We set $N_{\min} = 2$ and $N_{\max} = 1024$, so that our mock velocity field presents solenoidal and compressive fluctuations over scales differing almost 3 orders of magnitudes. In order for both components to be relevant, we fix the amplitudes so that $A_n^{\rm rot} = A_n^{\rm comp}$ for $n = 64$. We note that, while the velocity field generated according to this procedure is not the most general one (e.g., one could add oblique plane waves), it is challenging enough in order to show the capability of our code to handle a broad range of spatial frequencies in close-to-realistic conditions.
2. Then, we compute this total velocity field onto the AMR grid structure defined at the beginning of Section 3.1. In this case, we maintain $N_x = 128$ as the resolution of the base grid, and limit the number of refinement patches to[5] $\ell_{\max} = 4$, which is still enough to show the multi-scale capabilities of our code. When computing the value of the velocities on the base level or on AMR levels $\ell < \ell_{\max}$, we average over the uniform grid cells enclosed in the coarser volume element.

This process consistently generates a mixed, solenoidal and compressive, velocity field presenting similar scaling features as the ICM over almost 3 decades in spatial frequency. Therefore, it can robustly show the capability of the code to handle multi-scale (and multi-frequency) velocity signals.

3.2. Results

For each test, we have validated the performance of the code by computing a series of error statistics, which we present below. Let $\mathbf{v}$ be the input velocity field, for which our algorithm returns its compressive and rotational components, $\tilde{\mathbf{v}}_{\rm comp}$ and $\tilde{\mathbf{v}}_{\rm rot}$. Thus, the algorithm recovers a total velocity field $\tilde{\mathbf{v}} = \tilde{\mathbf{v}}_{\rm comp} + \tilde{\mathbf{v}}_{\rm rot}$ which might differ from the original, $\mathbf{v}$, due to the numerical error in the processes involved, namely finite-differencing the velocity field, integrating the elliptical equations and finite-differencing the potentials.

In Tests 1 and 2, where only one velocity component (compressive and rotational, respectively) was present, we have quantified the error in reconstructing these velocity fields by computing the cell-wise relative error as[6]:

$$\varepsilon_r\left(v_{\rm cell}\right) = \sqrt{\sum_{i=1}^{3}\left[\left(\frac{v_{\rm cell}^i}{|\mathbf{v}_{\rm cell}|}\right)^2\left|\frac{\tilde{v}_{\rm cell}^i - v_{\rm cell}^i}{v_{\rm cell}^i + \epsilon}\right|\right]^2} \tag{14}$$

where the subindex i runs over the three cartesian components. Note that we add a small constant, ϵ, in the denominator of the relative error in $v_i^{\rm cell}$ to prevent the overestimation of the error due to the cells with velocities close to 0. We set $\epsilon \equiv 10^{-2}\max(|v_i|)$. For each refinement level, ℓ, we have computed the median error over all the cells, which we use as an estimate of the velocity reconstruction error. We also consider the error percentiles 5, 25, 75 and 95 in order to give a confidence interval (CI) for this error.

The velocity field in Test 1 (Test 2) has null rotational (compressive) component. We have checked this by computing the median value and the 5, 25, 75 and 95 percentiles of $|\mathbf{v}_{\rm rot}|/|\mathbf{v}|$ ($|\mathbf{v}_{\rm comp}|/|\mathbf{v}|$).

The results for Test 1 and Test 2 are summarised in Figs. 2 and 3, respectively. The upper panels show $\varepsilon_r(v)$ as defined above. Both in the constant divergence and the constant rotational test, the median errors are small (typically lower than 1‰), and even the 95-percentile errors do not exceed 1% at any refinement level. The base level, for which FFT is used, has much more precise results, with median relative errors below 10^{-6}. For the AMR levels, we only find a very slight trend to increase the error in more refined levels.

The lower panels in Figs. 2 and 3 show $|\mathbf{v}_{\rm rot}|/|\mathbf{v}|$, for Test 1, and $|\mathbf{v}_{\rm comp}|/|\mathbf{v}|$, for Test 2. For these highly-idealised scenarios, there is virtually no cross-talk between the different components: the rotational velocity in Test 1 accounts for less than 10^{-8} in more than 95% of the cells at any level. Likewise, more than 95% of the cells in Test 2 have compressive velocities less than 10^{-7} relative to the total velocity magnitude.

In Test 3, as both the compressive and the rotational velocity components are present, we have quantified the relative error in disentangling and reconstructing each of these by applying Eq. (14) separately to each component. We present its results in Fig. 4.

[5] The limitation is due to the fact that we need to compute the velocity field in a uniform grid, in the first place. Thus, with $N_x = 128$ and $\ell_{\max} = 4$, this uniform grid consists of 2048^3 cells.

[6] The equation below corresponds to the propagation of the variance in $\{v_i^{\rm cell}\}$ to the function $v^{\rm cell} = \sqrt{\sum_{i=1}^{3}\left(v_i^{\rm cell}\right)^2}$, assuming the velocity components are uncorrelated.

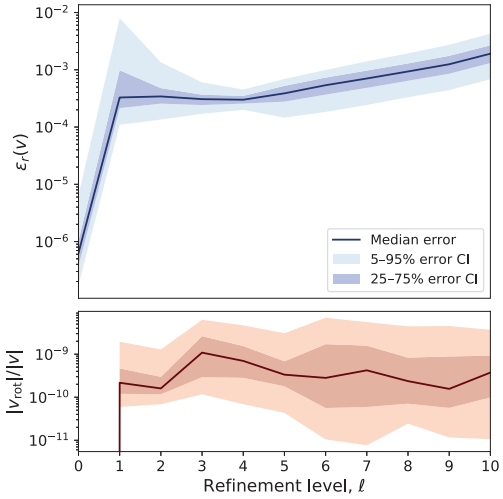


Fig. 2. Results from Test 1. *Upper panel*: median relative error (solid line; defined as in Eq. (14)) and confidence intervals (CIs) in the reconstruction of the velocity field. *Lower panel*: fraction of rotational velocity misreconstructed by the algorithm.

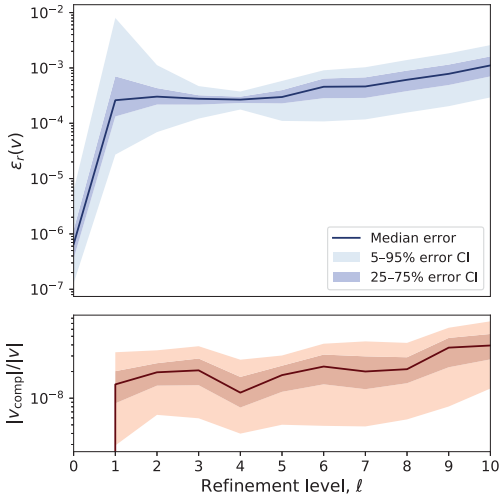


Fig. 3. Results from Test 2. *Upper panel*: median relative error (solid line; defined as in Eq. (14)) and confidence intervals (CIs) in the reconstruction of the velocity field. *Lower panel*: fraction of compressive velocity misreconstructed by the algorithm.

For the compressive velocity, the typical relative errors are below 3×10^{-3}, while the rotational velocity presents median relative errors in the range 10^{-4}–10^{-3}. For the AMR levels, $\ell \geq 1$, the magnitude and behaviour of the errors resembles what has been seen before for Tests 1 and 2. Remarkably, in this example where both components are present, the base level does no longer exhibit a much more precise result, but it shows errors in the same order as the ones in the AMR levels.

Last, for the exceedingly complex Test 4, since the actual decomposition is still known, we follow the same procedure as

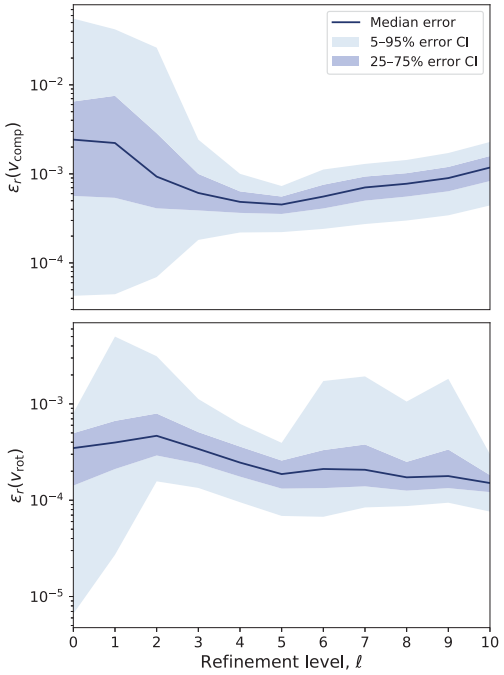


Fig. 4. Results from Test 3. Upper (*lower*) panel presents the relative error in the reconstruction of the compressive (*rotational*) velocity fields.

in Test 3. The detailed results, presented in the same way as for the previous test, can be found in Fig. 5. Compared to Tests 1 – 3, in this case we find errors up to an order of magnitude higher, which is not surprising since in this case we have included a truly multi-scale velocity field, with signal spanning over almost 3 decades in spatial frequency. In any case, our code performs the HHD decomposition with median errors around 1%, and not exceeding 7% even at 95-percentile level, despite the fact that we have introduced oscillations close to the grid nominal frequency.

The results for this last test, together with the ones we present below in Section 4.2, show the ability of our algorithm to perform the HHD in challenging, very non-linear velocity fields.

4. Application to a cosmological simulation

Last, this section focuses on the results of our HHD code when applied to the outcomes of a cosmological simulation, which is described in the paragraphs below. As opposed to the previous highly idealised tests 1 – 3, the velocity field in a full-cosmological simulation exhibits a plethora of complex features due to the non-linear nature of the equations governing its evolution (see, e.g., [26,32,41] for classical and recent reviews).

4.1. Simulation details

The simulation analysed in this paper has been carried out with the cosmological code MASCLET [42], and has been already employed in a series of previous works [43–45]. Here we shall introduce the main details of the simulation, while some topics which are not intimately connected to the analyses in this paper can be found in more detail in the aforementioned references.

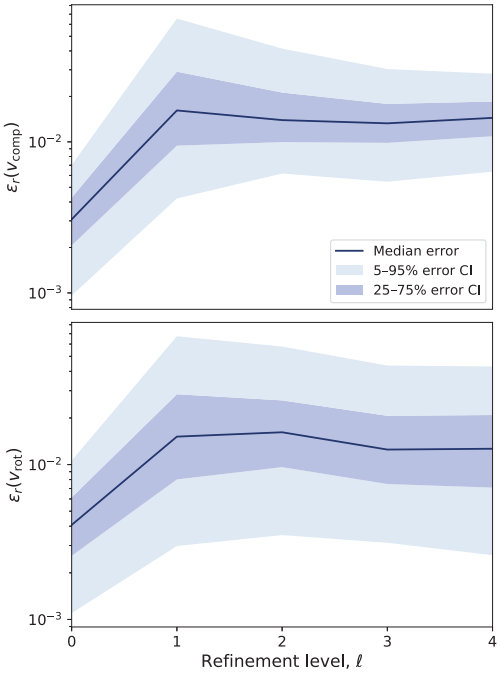


Fig. 5. Results from Test 4. Upper (*lower*) panel presents the relative error in the reconstruction of the compressive (*rotational*) velocity fields.

MASCLET is an Eulerian cosmological code, implementing *high-resolution shock-capturing* techniques for the description of the gaseous component and an N-Body particle-mesh for dark matter (DM). Both components are built into an AMR scheme to gain spatial and temporal resolution in the regions of interest.

We have simulated a cubic domain of comoving side length 40 Mpc, assuming a flat ΛCDM cosmology with the following values of the cosmological parameters: $h \equiv H_0/(100\ \mathrm{km\ s^{-1}\ Mpc^{-1}}) = 0.678$, $\Omega_m = 0.31$, $\Omega_b = 0.048$, $\Omega_\Lambda = 0.69$, $n_s = 0.96$ and $\sigma_8 = 0.82$, which are consistent with the latest values reported by the Planck mission [46]. The domain is discretised in a base grid of 128^3 cells, granting a harsh resolution of ~ 310 kpc at the coarsest level. Regions with large gaseous and/or DM densities can get recursively refined following the AMR scheme. We allow $n_\ell = 9$ refinement levels, each one halving the cell side length with respect to the previous level, providing a peak resolution of ~ 610 pc. The peak DM mass resolution is $\sim 2 \times 10^6 M_\odot$, equivalent to filling the domain with 1024^3 of such particles.

The simulation started at redshift $z = 100$, with the initial conditions set up by a CDM transfer function [47] and generated by a constrained realisation of the gaussian random field aimed to produce a massive cluster in the centre of the computational domain by $z \sim 0$ [48]. By redshift $z \sim 0$, several massive clusters and groups have been formed. Besides gravity, the simulation accounts for a broad variety of feedback mechanisms, which are explained in greater detail in the cited previous works.

4.2. Performance of the code

We have run our HHD algorithm over 80 snapshots of the simulation described above, ranging from $z = 100$ to $z = 0$, and

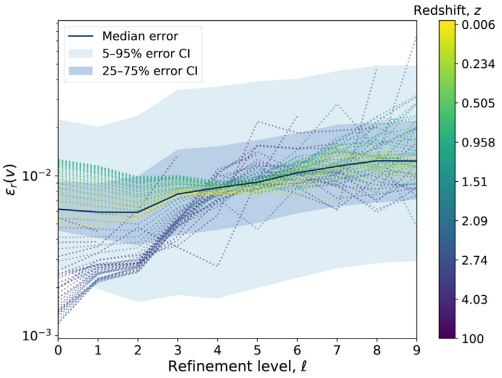


Fig. 6. Results from the tests of the HHD algorithm over MASCLET outputs. The blue solid line and the contours present the median over all the snapshots of the median relative error and the CIs defined as in the figures in Section 3. The dotted lines represent the individual error statistics for each snapshot, with the colourscale encoding the redshift.

computed the cell-wise error[7] as in Eq. (14) and its corresponding percentiles (5, 25, 75 and 95), as done in the tests in Section 3. Fig. 6 presents the overall error statistics, defined as the median of the error statistics over all the code outputs. In order to keep track of the dispersion of the error statistics in different snapshots, we have also plotted the 80 individual error profiles in dotted lines, with the line colours encoding the redshift.

The median relative errors in describing the velocity field as the sum of a compressive and a rotational component are typically in the order of or slightly less than 1% for all refinement levels, with only a small trend to increase the error with the refinement level. Even at the 95% error percentile, the relative errors fall below 5%. At high redshift, the errors at low refinement levels ($\ell \lesssim 3$) tend to be smaller, most likely due to the fact that the velocity field does not present as complex features as it does at more recent redshifts due to its highly non-linear evolution. The behaviour at the most refined levels presents significant scatter and there is not a clear redshift evolution of the error, but its median magnitude keeps below a few percents in all snapshots. Thus, our algorithm is capable of providing a robust multi-scale decomposition of the velocity field in its compressive and rotational velocities, even on highly-complex, non idealised conditions.

4.3. An example: velocity maps and profiles around a massive galaxy cluster

In order to exemplify the ability of the code to split the components of a highly complex velocity field, we focus on a massive galaxy cluster[8], with mass $M_{\mathrm{vir}} \simeq 4.83 \times 10^{14} M_\odot$ and radius $R_{\mathrm{vir}} \simeq 1.99$ Mpc, at $z \simeq 0$. We present in Fig. 7 slices of gas density (top left panel), total velocity (top right), compressive

[7] In this case, as we do not know beforehand the 'true' decomposed velocity fields to compare with the reconstructed ones, we quantify the error by comparing the reconstructed total velocity field, $\tilde{\mathbf{v}} = \tilde{\mathbf{v}}_{\mathrm{comp}} + \tilde{\mathbf{v}}_{\mathrm{rot}}$ to the input one, as defined in Eq. (14). Since, by definition, $\tilde{\mathbf{v}}_{\mathrm{comp}}$ ($\tilde{\mathbf{v}}_{\mathrm{rot}}$) is the gradient of a scalar field (the rotational of a vector field), it is irrotational (solenoidal). We have, indeed, checked that our high-order derivatives verify this, typically much better than 1‰. Therefore, as the decomposition is unique, checking that $\tilde{\mathbf{v}} = \mathbf{v}$ proves the validity of the method.

[8] This same object has been analysed in great detail in [44] (focusing on its observational properties) and [45] (exploring its accretion history).

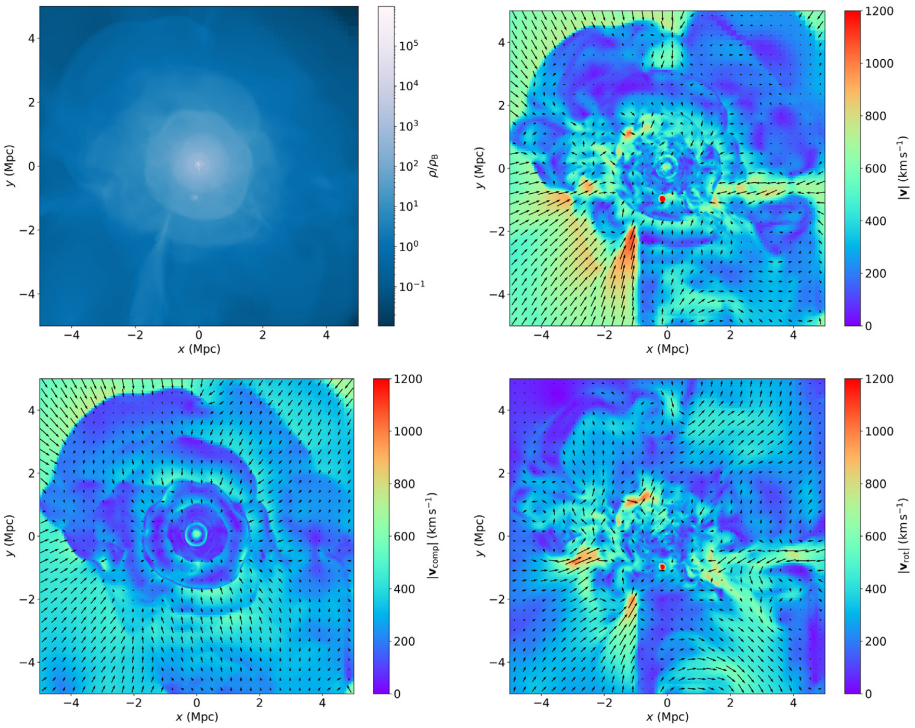


Fig. 7. Maps around a massive cluster. Each map is a mass-weighted projection of the same region around the cluster, 10 Mpc on each side and $\sim$ 150 kpc thick, limiting the image resolution to $\sim$ 10 kpc. *Top left:* Gas density (in units of the background density of the Universe). *Top right:* Total velocity magnitude. *Bottom left:* Compressive velocity magnitude. *Bottom right:* Solenoidal velocity magnitude. The arrows in the velocity maps represent the projection of the corresponding velocity fields in the slice plane.

velocity (bottom left) and solenoidal velocity (bottom right). The velocity maps show both magnitude (encoded in colour) and direction in the slice plane (arrows).

The density map shows that, by $z \sim 0$, the cluster is, indeed, relatively relaxed, sitting in the gravitational potential dominated by the dark matter component, in an approximately spherical shape. Several density discontinuities can be easily discerned, mainly corresponding to (internal) merger shocks and (external) accretion shocks. A filament penetrating to quite inner radii, of around $r \sim 1$ Mpc, is also noticeable in the density plot. The total velocity field displays great complexity, especially in the inner regions of the cluster, where the variations occur on smaller scales (both because the dynamics are more complex and because, correspondingly, the resolution is higher). The filamentary structures appear to present high velocity magnitudes, mainly pointing radially, and velocity discontinuities, hinting the presence of shocks, are ubiquitous.

Compressive velocities show a nearly spherically symmetric pattern, as the cluster smoothly accretes gas from its surroundings. In the outskirts, baryonic matter is accelerated (and compressed) by the gravitational pull of the cluster, until it gets shocked causing the strong accretion shocks. In comparison, the post-shock medium presents very small compressive velocities, as the shock has effectively halted the accretion flows. Part of the energy corresponding to these compressive component gets thermalised, increasing the internal energy (and temperature) of the ICM. However, another important role of shocks is the generation of vorticity (see, e.g., [13]). Indeed, inside the external accretion shocks, the solenoidal component of the velocity fields gets relevant. Eddies develop on a wide range of scales, especially in the cluster central regions. It is also interesting to note how the infalling filament mentioned before presents high values of the rotational velocity, suggesting that matter is infalling following helicoidal trajectories. Being the main aim of this work presenting the algorithm, we may defer a more in-depth analysis of these issues to future work.

Complementarily, in Fig. 8 we present the radial profile of radial total (red line), radial compressive (blue) and radial solenoidal (green) velocities, for 20 kpc $\lesssim r \lesssim$ 4 Mpc $\simeq 2R_{\rm vir}$. In the inner regions ($r \lesssim 0.1$ Mpc) both strong radial compressive and solenoidal flows are present, bounded by an internal shock (clearly visible in the compressive velocity map of Fig. 7). For $r \gtrsim 0.1$ Mpc, the radial compressive velocity clearly dominates. While solenoidal motions are still present, the fact that their radial component is close to zero suggests that these motions tend to occur along the tangential direction.

5. Conclusions

In this paper, we have proposed a novel method to perform a Helmholtz–Hodge decomposition in AMR velocity fields, or virtually in any description which can be smoothed over an ad-hoc hierarchy of grids. Although our primary focus has been cosmological simulations of structure formation, the method is general and could be easily extended to any type of hydrodynamical simulation.

Previous works in the field of numerical cosmology typically use uniform grids and work straightforwardly in Fourier space.

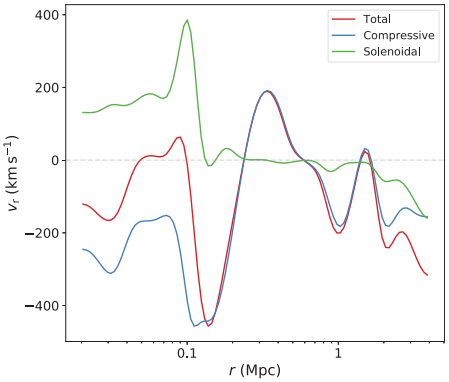


Fig. 8. Radial profiles of radial velocity for the total (*red*), compressive (*blue*) and solenoidal (*green*) velocity components. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

However, this procedure requires to perform constrained simulations (or resimulations) of specific objects of interest (e.g., a galaxy cluster) in order to be able to describe it with a uniform computational grid at a reasonable computational cost. Our algorithm, instead, can be applied to full-cosmological simulations, without the need of performing resimulations and keeping the full description at the maximum resolution at each position.

The performance of the code has been validated in a series of idealised tests, for which the analytical decomposition is known. Our algorithm has shown to succeed in disentangling the compressive and solenoidal velocity components and reconstructing the input velocity field, with typical errors in the order of 1‰ or below (1% in the more complex, ICM-like test). Our errors seem comparable to or even better than the ones displayed by the tests in [10, Appendix A1.2].

For exceedingly complex velocity fields, like the ones generated by actual cosmological simulations at low redshifts, where turbulence is fully developed (e.g., [19]) and velocity fluctuates on many different scales, our tests show that the decomposition can be brought about with median errors below 1%, even resolving scales smaller than the kpc in a domain of several tens of Mpc along each direction.

This procedure, whose implementation has been made publicly available (see Section 2.1), will allow us to further explore the dynamics of the turbulent velocity field in the ICM of large samples of clusters in future works.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We gratefully acknowledge the anonymous referees for their valuable feedback, which has helped us to improve the quality of this manuscript. This work has been supported by the Spanish Ministerio de Ciencia e Innovación (MICINN, grant PID2019-107427GB-C33) and by the Generalitat Valenciana (grant PROMETEO/2019/071). Simulations have been carried out using the supercomputer Lluís Vives at the Servei d'Informàtica of the Universitat de València. This research has made use of the following open-source packages: NumPy [49], SciPy [50], matplotlib [51] and yt [52].

References

[1] J.A. ZuHone, Astrophys. J. 728 (1) (2011) 54, http://dx.doi.org/10.1088/0004-637X/728/1/54, arXiv:1004.3820.
[2] M. Gaspari, E. Churazov, Astron. Astrophys. 559 (2013) A78, http://dx.doi.org/10.1051/0004-6361/201322295, arXiv:1307.4397.
[3] I. Zhuravleva, E. Churazov, A.A. Schekochihin, S.W. Allen, P. Arévalo, A.C. Fabian, W.R. Forman, J.S. Sanders, A. Simionescu, R. Sunyaev, A. Vikhlinin, N. Werner, Nature 515 (7525) (2014) 85–87, http://dx.doi.org/10.1038/nature13830, arXiv:1410.6485.
[4] F. Hofmann, J.S. Sanders, K. Nandra, N. Clerc, M. Gaspari, Astron. Astrophys. 585 (2016) A130, arXiv:1510.08445.
[5] D. Eckert, M. Gaspari, F. Vazza, F. Gastaldello, A. Tramacere, S. Zimmer, S. Ettori, S. Paltani, Astrophys. J. Lett. 843 (2) (2017) L29, http://dx.doi.org/10.3847/2041-8213/aa7c1a, arXiv:1705.02341.
[6] A. Bonafede, M. Brüggen, D. Rafferty, I. Zhuravleva, C.J. Riseley, R.J. van Weeren, J.S. Farnes, F. Vazza, F. Savini, A. Wilber, A. Botteon, G. Brunetti, R. Cassano, C. Ferrari, F. de Gasperin, E. Orrú, R.F. Pizzo, H.J.A. Röttgering, T.W. Shimwell, Mon. Not. R. Astron. Soc. 478 (3) (2018) 2927–2938, http://dx.doi.org/10.1093/mnras/sty1121, arXiv:1805.00473.
[7] M. Roncarelli, M. Gaspari, S. Ettori, V. Biffi, F. Brighenti, E. Bulbul, N. Clerc, E. Cucchetti, E. Pointecouteau, E. Rasia, Astron. Astrophys. 618 (2018) A39, http://dx.doi.org/10.1051/0004-6361/201833371, arXiv:1805.02577.
[8] F. Miniati, The Matryoshka Run. I.I., ApJ 800 (1) (2015) 60, http://dx.doi.org/10.1088/0004-637X/800/1/60, arXiv:1409.3576.
[9] W. Schmidt, J.F. Engels, J.C. Niemeyer, A.S. Almgren, Mon. Not. R. Astron. Soc. 459 (1) (2016) 701–719, http://dx.doi.org/10.1093/mnras/stw632, arXiv:1603.04711.
[10] F. Vazza, T.W. Jones, M. Brüggen, G. Brunetti, C. Gheller, D. Porter, D. Ryu, Mon. Not. R. Astron. Soc. 464 (1) (2017) 210–230, http://dx.doi.org/10.1093/mnras/stw2351, arXiv:1609.03558.
[11] L. Iapichino, C. Federrath, R.S. Klessen, MNRAS 469 (3) (2017) 3641–3655, http://dx.doi.org/10.1093/mnras/stx882, arXiv:1704.02922.
[12] F. Vazza, M. Brüggen, C. Gheller, P. Wang, Mon. Not. R. Astron. Soc. 445 (4) (2014) 3706–3722, http://dx.doi.org/10.1093/mnras/stu1896, arXiv:1409.2640.
[13] D.H. Porter, T.W. Jones, D. Ryu, Astrophys. J. 810 (2) (2015) 93, http://dx.doi.org/10.1088/0004-637X/810/2/93, arXiv:1507.08737.
[14] C. Federrath, G. Chabrier, J. Schober, R. Banerjee, R.S. Klessen, D.R.G. Schleicher, PRL 107 (11) (2011) 114504, http://dx.doi.org/10.1103/PhysRevLett.107.114504, arXiv:1109.1760.
[15] C. Federrath, R.S. Klessen, W. Schmidt, Astrophys. J. 692 (1) (2009) 364–374, http://dx.doi.org/10.1088/0004-637X/692/1/364, arXiv:0710.1359.
[16] L. Iapichino, W. Schmidt, J.C. Niemeyer, J. Merklein, Mon. Not. R. Astron. Soc. 414 (3) (2011) 2297–2308, http://dx.doi.org/10.1111/j.1365-2966.2011.18550.x, arXiv:1102.3352.
[17] C. Federrath, J. Roman-Duval, R.S. Klessen, W. Schmidt, M.M. Mac Low, Astron. Astrophys. 512 (2010) A81, http://dx.doi.org/10.1051/0004-6361/200912437, arXiv:0905.1060.
[18] C. Federrath, Mon. Not. R. Astron. Soc. 436 (2) (2013) 1245–1257, http://dx.doi.org/10.1093/mnras/stt1644, arXiv:1306.3989.
[19] F. Miniati, Astrophys. J. 782 (1) (2014) 21, http://dx.doi.org/10.1088/0004-637X/782/1/21, arXiv:1310.2951.
[20] J.A. Acosta-Pulido, I. Agudo, A. Alberdi, J. Alcolea, E.J. Alfaro, A. Alonso-Herrero, G. Anglada, P. Arnalte-Mur, Y. Ascasibar, B. Ascaso, R. Azulay, R. Bachiller, A. Baez-Rubio, E. Battaner, J. Blasco, C.B. Brook, V. Bujarrabal, G. Busquet, M.D. Caballero-Garcia, C. Carrasco-Gonzalez, J. Casares, A.J. Castro-Tirado, L. Colina, F. Colomer, I. de Gregorio-Monsalvo, A. del Olmo, J.F. Desmurs, J.M. Diego, R. Dominguez-Tenreiro, R. Estalella, A. Fernandez-Soto, E. Florido, J. Font, J.A. Font, A. Fuente, R. Garcia-Benito, S. Garcia-Burillo, B. Garcia-Lorenzo, A. Gil de Paz, J.M. Girart, J.R. Goicoechea, J.F. Gomez, M. Gonzalez-Garcia, O. Gonzalez-Martin, J.I. Gonzalez-Serrano, J. Gorgas, J. Gorosabel, A. Guijarro, J.C. Guirado, L. Hernandez-Garcia, C. Hernandez-Monteagudo, D. Herranz, R. Herrero-Illana, Y.D. Hu, N. Huelamo, M. Huertas-Company, J. Iglesias-Paramo, S. Jeong, I. Jimenez-Serra, J.H. Knapen, R.A. Lineros, U. Lisenfeld, J.M. Marcaide, I. Marquez, J. Marti, J.M. Marti, I. Marti-Vidal, E. Martinez-Gonzalez, J. Martin-Pintado, J. Masegosa, J.M. Mayen-Gijon, M. Mezcua, S. Migliari, P. Mimica, J. Moldon, O. Morata, I. Negueruela, S.R. Oates, M. Osorio, A. Palau, J.M. Paredes, J. Perea, P.G. Perez-Gonzalez, E. Perez-Montero, M.A. Perez-Torres, M. Perucho, S. Planelles, J.A. Pons, A. Prieto, V. Quilis, P. Ramirez-Moreta, C. Ramos Almeida, N. Rea, M. Ribo, M.J. Rioja, J.M. Rodriguez Espinosa, E. Ros, J.A. Rubiño-Martin, B. Ruiz-Granados, J. Sabater, S. Sanchez, C. Sanchez-Contreras, A. Sanchez-Monge, R. Sanchez-Ramirez, A.M. Sintes, J.M. Solanes, C.F. Sopuerta, M. Tafalla, J.C. Tello, B. Tercero, M.C. Toribio, J.M. Torrelles, M.A.P. Torres, A. Usero, L. Verdes-Montenegro, A. Vidal-Garcia, P. Vielva, J. Vilchez, B.B. Zhang, The spanish square kilometre array white book, 2015, arXiv e-prints (Jun. 2015), arXiv:1506.03474.
[21] Y. Fujita, M. Takizawa, C.L. Sarazin, Astrophys. J. 584 (1) (2003) 190–202, http://dx.doi.org/10.1086/345599, arXiv:astro-ph/0210320.

[22] R. Cassano, G. Brunetti, Mon. Not. R. Astron. Soc. 357 (4) (2005) 1313–1329, http://dx.doi.org/10.1111/j.1365-2966.2005.08747.x, arXiv:astro-ph/0412475.
[23] G. Brunetti, A. Lazarian, Mon. Not. R. Astron. Soc. 412 (2) (2011) 817–824, http://dx.doi.org/10.1111/j.1365-2966.2010.17937.x, arXiv:1011.1198.
[24] A. Pinzke, S.P. Oh, C. Pfrommer, Mon. Not. R. Astron. Soc. 465 (4) (2017) 4800–4816, http://dx.doi.org/10.1093/mnras/stw3024, arXiv:1611.07533.
[25] G. Brunetti, F. Vazza, PRL 124 (5) (2020) 051101, http://dx.doi.org/10.1103/PhysRevLett.124.051101, arXiv:2001.07718.
[26] E. Bertschinger, ARA & A 36 (1998) 599–654, http://dx.doi.org/10.1146/annurev.astro.36.1.599.
[27] T. Naab, J.P. Ostriker, ARA & A 55 (1) (2017) 59–109, http://dx.doi.org/10.1146/annurev-astro-081913-040019, arXiv:1612.06891.
[28] F. Vazza, E. Roediger, M. Brüggen, Astron. Astrophys. 544 (2012) A103, http://dx.doi.org/10.1051/0004-6361/201118688, arXiv:1202.5882.
[29] R. Mohapatra, C. Federrath, P. Sharma, Mon. Not. R. Astron. Soc. 493 (4) (2020) 5838–5853, http://dx.doi.org/10.1093/mnras/staa711, arXiv:2001.06494.
[30] R. Mohapatra, C. Federrath, P. Sharma, Mon. Not. R. Astron. Soc. 500 (4) (2021) 5072–5087, http://dx.doi.org/10.1093/mnras/staa3564, arXiv:2010.12602.
[31] A.M. Bykov, K. Dolag, F. Durret, Cosmological.Shock. Waves, Space Sci. Rev, Rev 134 (1–4) (2008) 119–140, http://dx.doi.org/10.1007/s11214-008-9312-9, arXiv:0801.0995.
[32] K. Dolag, S. Borgani, S. Schindler, A. Diaferio, A.M. Bykov, Space Sci. Rev. 134 (1–4) (2008) 229–268, http://dx.doi.org/10.1007/s11214-008-9316-5, arXiv:0801.1023.
[33] A.G. Kritsuk, ÅNordlund, D. Collins, P. Padoan, M.L. Norman, T. Abel, R. Banerjee, C. Federrath, M. Flock, D. Lee, P.S. Li, W.-C. Müller, R. Teyssier, S.D. Ustyugov, C. Vogel, H. Xu, Astrophys. J. 737 (1) (2011) 13, http://dx.doi.org/10.1088/0004-637X/737/1/13, arXiv:1103.5525.
[34] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Art Scient. Comput. (1992).
[35] G. Arfken, H. Weber, F. Harris, Elsevier Sci. (2013).
[36] G. Kowal, A. Lazarian, Astrophys. J. 720 (1) (2010) 742–756, http://dx.doi.org/10.1088/0004-637X/720/1/742, arXiv:1003.3697.
[37] S.K. Harouna, V. Perrier, in: J.-D. Boissonnat, P. Chenin, A. Cohen, C. Gout, T. Lyche, M.-L. Mazure, L. Schumaker (Eds.), Curves and Surfaces, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012, pp. 311–329.
[38] R.W. Hockney, J.W. Eastwood, Computer simulation using particles, 1981.
[39] J. Burgers, verh. k, Akad. Wet. Amst. Addeel. Nat. 17 (1) (1939).
[40] A. Kolmogorov, Akad. Nauk. SSSR Dokl. 30 (1941) 301–305.
[41] S. Planelles, D.R.G. Schleicher, A.M. Bykov, Space Sci. Rev. 188 (1–4) (2015) 93–139, http://dx.doi.org/10.1007/s11214-014-0045-7, arXiv:1404.3956.
[42] V. Quilis, Mon. Not. R. Astron. Soc. 352 (2004) 1426–1438, http://dx.doi.org/10.1111/j.1365-2966.2004.08040.x, arXiv:astro-ph/0405389.
[43] V. Quilis, S. Planelles, E. Ricciardelli, Mon. Not. R. Astron. Soc. 469 (1) (2017) 80–94, http://dx.doi.org/10.1093/mnras/stx770, arXiv:1703.09446.
[44] S. Planelles, P. Mimica, V. Quilis, C. Cuesta-Martínez, Mon. Not. R. Astron. Soc. 476 (4) (2018) 4629–4648, http://dx.doi.org/10.1093/mnras/sty527, arXiv:1802.09458.
[45] D. Vallés-Pérez, S. Planelles, V. Quilis, Mon. Not. R. Astron. Soc. 499 (2) (2020) 2303–2318, http://dx.doi.org/10.1093/mnras/staa3035, arXiv:2009.13882.
[46] Planck Collaboration, Astron. Astrophys. 641 (2020) A6, http://dx.doi.org/10.1051/0004-6361/201833910, arXiv:1807.06209.
[47] D.J. Eisenstein, W. Hu, Astrophys. J. 496 (2) (1998) 605–614, http://dx.doi.org/10.1086/305424, arXiv:astro-ph/9709112.
[48] Y. Hoffman, E. Ribak, Astrophys. J. 380 (1991) L5–L8, http://dx.doi.org/10.1086/186160.
[49] T.E. Oliphant, A Guide to NumPy, Vol. 1, Trelgol Publishing, USA, 2006.
[50] P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, K. Jarrod Millman, N. Mayorov, A.R.J. Nelson, E. Jones, R. Kern, E. Larson, C. Carey, İ. Polat, Y. Feng, E.W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E.A. Quintero, C.R. Harris, A.M. Archibald, A.H. Ribeiro, F. Pedregosa, P. van Mulbregt, Nat. Methods 17 (2020) 261–272, http://dx.doi.org/10.1038/s41592-019-0686-2.
[51] J.D. Hunter, Comput. Sci. Eng. 9 (2007) 90–95, http://dx.doi.org/10.1109/MCSE.2007.55.
[52] M.J. Turk, B.D. Smith, J.S. Oishi, S. Skory, S.W. Skillman, T. Abel, M.L. Norman, Astrophys. J. Suppl. 192 (1) (2011) 9, http://dx.doi.org/10.1088/0067-0049/192/1/9, arXiv:1011.3514.

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **504,** 510–527 (2021) doi:10.1093/mnras/stab880
Advance Access publication 2021 March 26

Troubled cosmic flows: turbulence, enstrophy, and helicity from the assembly history of the intracluster medium

David Vallés-Pérez [1]★ Susana Planelles [1,2] and Vicent Quilis [1,2]

[1] *Departament d'Astronomia i Astrofísica, Universitat de València, E-46100 Burjassot (València), Spain*
[2] *Observatori Astronòmic, Universitat de València, E-46980 Paterna (València), Spain*

Accepted 2021 March 24. Received 2021 March 23; in original form 2021 February 19

ABSTRACT

Both simulations and observations have shown that turbulence is a pervasive phenomenon in cosmic scenarios, yet it is particularly difficult to model numerically due to its intrinsically multiscale character which demands high resolutions. Additionally, turbulence is tightly connected to the dynamical state and the formation history of galaxies and galaxy clusters, producing a diverse phenomenology which requires large samples of such structures to attain robust conclusions. In this work, we use an adaptive mesh refinement (AMR) cosmological simulation to explore the generation and dissipation of turbulence in galaxy clusters, in connection to its assembly history. We find that major mergers, and more generally accretion of gas, is the main process driving turbulence in the intracluster medium. We have especially focused on solenoidal turbulence, which can be quantified through enstrophy. Our results seem to confirm a scenario for its generation which involves baroclinicity and compression at the external (accretion) and internal (merger) shocks, followed by vortex stretching downstream of them. We have also looked at the infall of mass to the cluster beyond its virial boundary, finding that gas follows trajectories with some degree of helicity, as it has already developed some vorticity in the external shocks.

Key words: hydrodynamics – turbulence – methods: numerical – galaxies: clusters: general – galaxies: clusters: intracluster medium.

1 INTRODUCTION

Within the hierarchical paradigm of cosmological structure formation (Press & Schechter 1974; Gott & Rees 1975; see also Kravtsov & Borgani 2012 and Planelles, Schleicher & Bykov 2015 for recent reviews), the assembly history of galaxy clusters is dominated by (major) merger events, which account for most of their mass growth (e.g. Vallés-Pérez, Planelles & Quilis 2020) and have consequential effects on their thermal structure (e.g. Planelles & Quilis 2009; ZuHone 2011). Galaxy cluster mergers and, more generally, gas accretion are the main energy source fuelling different complex hydrodynamical processes in the intracluster medium (ICM), such as turbulence and a rich phenomenology of shock waves (see e.g. Quilis, Ibáñez & Sáez 1998; Ryu et al. 2003; Zhang et al. 2020). Thus, a precise understanding of these phenomena is necessary for a proper description of the bulk of clusters' baryonic mass, which, furthermore, is essential for the usage of galaxy clusters as cosmological probes (for a review, see Allen, Evrard & Mantz 2011) and to correctly model galaxy formation processes in a cosmological context (Naab & Ostriker 2017).

Constraining the amplitude and spectrum of turbulence, and elucidating their evolution with cosmic time and possible dependencies on clusters' mass, dynamical state, formation history, etc. is a fundamental step towards a correct description of many phenomena: non-thermal pressure (Nelson, Lau & Nagai 2014b; Shi & Komatsu 2014; Vazza et al. 2018b; Angelinelli et al. 2020) leading to hydrostatic mass bias (Nelson et al. 2014a; Biffi et al. 2016; Shi et al. 2016), amplification of cosmic magnetic fields (Subramanian, Shukurov & Haugen 2006; Iapichino et al. 2008; Cho 2014; Beresnyak & Miniati 2016; Vazza et al. 2018a; Brzycki & ZuHone 2019; see also Donnert et al. 2018 for a recent review), chemical and thermal mixing (Ruszkowski & Oh 2010; ZuHone 2011; Shi et al. 2020), star formation (Kretschmer & Teyssier 2020), cosmic ray acceleration and radio emission (Fujita, Takizawa & Sarazin 2003; Cassano & Brunetti 2005; Brunetti & Lazarian 2011; Brunetti & Vazza 2020), and heating due to viscous dissipation, which could help to alleviate the cooling flow problem (Zhuravleva et al. 2014a; Valdarnini 2019; Shi et al. 2020).

Turbulence can be sourced, not only from gas accretion and mergers, but also due to galaxy motions (Faltenbacher et al. 2005; Kim 2007; Ruszkowski & Oh 2011), feedback from active galactic nuclei (AGNs; Gaspari, Ruszkowski & Sharma 2012; Gaspari et al. 2018; cf. Sayers et al. 2021), interaction of a compact, cool core with the surrounding ICM (Valdarnini 2011, 2019), etc. According to the classical Kolmogorov (1941) model for fully developed, homogeneous, and isotropic turbulence, bulk motions induced by these large-scale structure (LSS) formation processes break down to smaller scales due to different instabilities, transferring energy from the injection scale (the characteristic scale of bulk motions) down to the dissipation scale, where this energy is converted to heat, magnetic field amplification, and cosmic ray acceleration, among others. While this model could be rather idealized (the ICM is far from homogeneous; see Shi, Nagai & Lau 2018; Shi & Zhang 2019; Mohapatra, Federrath & Sharma 2020, 2021 for several recent exploratory studies on stratified, ICM-like turbulence), it constitutes

★ E-mail: david.valles-perez@uv.es

© 2021 The Author(s)
Published by Oxford University Press on behalf of Royal Astronomical Society

a simple baseline for exploring the turbulent phenomena in galaxy clusters.

Direct observations of turbulent flows in galaxy clusters have yet been elusive, in part due to the unfortunate fate of the Hitomi mission (Kitayama et al. 2014). Indirect detections of turbulence have been reported by measurements of X-ray surface brightness fluctuations (e.g. Churazov et al. 2012; Zhuravleva et al. 2014a; see also Gaspari & Churazov 2013; Zhuravleva et al. 2014b for theoretical investigations) or Sunyaev–Zel'dovich (SZ) signal fluctuations (Khatri & Gaspari 2016). We refer the interested reader to Simionescu et al. (2019) for a recent review on the current constraints and detectability prospects of gas motions in the ICM. In the next one to two decades, future X-ray missions (XRISM,[1] ATHENA[2]) will offer an unprecedented level of insight on to the ICM turbulent motions. Potentially, future mm/sub-mm facilities (e.g. ATLAST, see Klaassen et al. 2020; or SKA,[3] Acosta-Pulido et al. 2015) will also be able to map the (line-of-sight) velocity of clusters through the kinetic SZ effect (e.g. Adam et al. 2017; Sayers et al. 2019), complementing and establishing promising synergies in constraining the magnitude and distribution of turbulent flows in individual clusters.

In the meanwhile, numerical simulations of cosmological structure formation are amongst our best tools for exploring the physics of the ICM in order to interpret and lead future observations, as shown in the references above. A particularly relevant and fundamental issue is that of extracting a bulk and a turbulent part from an input velocity field (i.e. performing a Reynolds decomposition; Adrian, Christensen & Liu 2000). This is especially challenging in the case of the ICM because of its strong multiscale character: while the outer accretion shocks which bound clusters can have curvature radii of a few Mpc, galactic processes, which have consequential effects for the whole cluster, occur on kpc scales or below. Vazza, Roediger & Brüggen (2012) proposed an iterative algorithm for uniform grid data, further employed by Vazza et al. (2017) and extended to SPH simulation outputs by Valdarnini (2019), to constrain the outer scale of turbulence (roughly equivalent to the injection scale), which in turn allows to obtain a mean, bulk velocity field and a small-scale, turbulent velocity field. Recently, Shi et al. (2018) presented a new approach, based on the wavelet decomposition of the velocity field, which allows not only to disentangle these two components, but also to construct a local spectrum of the velocity field.

Since the viscous scales are many orders of magnitude smaller than the resolution achievable by current, state-of-the-art simulations, small-scale fluctuations at the end of the energy cascade are dissipated by numerical viscosity. Several techniques have been applied to overcome this artefact and to extend the turbulent cascade, at least, to consistently model the relevant features of ICM turbulence: either constrained adaptive mesh refinement (AMR) simulations with an additional refinement scheme based on local vorticity ($\boldsymbol{\omega} = \nabla \times \mathbf{v}$) or velocity jumps (Iapichino & Niemeyer 2008; Vazza et al. 2009b, 2011), or explicit subgrid models for the unresolved part of the turbulent cascade (e.g. Schmidt et al. 2016; see also Schmidt 2015 for a review). Other authors have used uniform grids (or static refinement techniques which, ultimately, lead to a virial volume resolved within a fine, fixed grid; e.g. Miniati 2014, 2015; Vazza et al. 2017). While these techniques have been successful in reproducing the expected properties of ICM turbulence, there are also evident downsides: for both, the refinement schemes based on vorticity or velocity jumps and subgrid models, one has to perform dedicated simulations, whose primary intent is to explore turbulent phenomena and which involve a much higher computational demand. As for static refinement simulations, these techniques are typically applied to single objects, rather than to a whole, cosmological population of clusters.

Large samples of clusters need to be analysed in order to achieve statistically significant conclusions. For this aim, given the extremely high computational demand of performing high-resolution simulations including full physics (taking into account that phenomena associated with feedback processes, such as galaxy motions or AGN feedback, are important for correctly modelling ICM turbulence), it can be challenging to perform dedicated simulations of large cosmological volumes including subgrid models or ad hoc refinement criteria, or either performing resimulations of a large number of objects.

In this paper, we perform an exploratory work over an AMR simulation without subgrid modelling or ad hoc refinement strategies, but with a base grid resolution which allows to resolve the bulk of gas mass within the virial volume of clusters with resolutions of $\Delta x \sim 20\,\mathrm{kpc}$ or better. For this aim, we have developed several algorithms that take full advantage of the multiresolution description of AMR simulations. Then, following our study of the accretion history of galaxy clusters in Vallés-Pérez et al. (2020), in this work we connect the evolution of several global and local indicators of turbulence to the assembly history of the objects, mainly focusing on solenoidal turbulence, and we look into the mechanisms for its generation and dissipation.

The rest of the manuscript is organized as follows. In Section 2, we describe our main numerical tools for analysing the velocity field in high-resolution simulation outputs. We detail the main features of our simulation, and the objects within it that will constitute the primary focus of this work, in Section 3. Our main results are then divided into Section 4 and Section 5, where we present, respectively, a global and a local description of turbulence in the ICM. Last, we further discuss several aspects of our results in Section 6 and present our conclusions in Section 7. Appendix A discusses the convergence of volume-averaged quantities in pseudo-Lagrangian AMR sampled outputs.

2 METHODS

In order to provide a multiscale characterization of the properties of the three-dimensional velocity field, we have designed and implemented several algorithms that take full advantage of the AMR description. In Section 2.1, we briefly describe our algorithm for splitting the velocity field in its rotational and compressive components. In Section 2.2, we present our filtering strategy to disentangle the bulk motions from the purely turbulent velocity field.

2.1 Helmholtz–Hodge decomposition in an AMR multigrid

In order to split the velocity field in its compressive and solenoidal components, we perform a Helmholtz–Hodge decomposition (HHD). Our implementation, presented in Vallés-Pérez, Planelles & Quilis (2021), does so by solving an elliptic equation for the scalar potential, ϕ, which generates the compressive velocity, $\mathbf{v}_{\mathrm{comp}} = -\nabla\phi$; and one elliptic equation for each Cartesian component of the vector potential, $\mathbf{A}$, which is responsible for the solenoidal (or rotational) velocity component, $\mathbf{v}_{\mathrm{rot}} = \nabla \times \mathbf{A}$.

The elliptic equations are addressed using standard methods for solving Poisson's equation in AMR cosmological simulations: namely, fast Fourier transform (FFT; see e.g. Press et al. 1992)

[1] http://xrism.isas.jaxa.jp/en/
[2] https://www.the-athena-x-ray-observatory.eu/
[3] https://www.skatelescope.org/

techniques for the base level,[4] and successive overrelaxation (SOR; Young 1954) for the refinement patches. All the derivatives are computed using high-order stencils to damp their contamination due to high-frequency noise. A more complete description and a set of tests can be found in Vallés-Pérez et al. (2021).

2.2 Multiscale filtering

Splitting the velocity field into a bulk (or coherent) and an inherently turbulent (uncorrelated, or *small-scale*) component, i.e. performing a Reynolds decomposition, can be fairly subtle due to the absence of a univoque definition of turbulence (Adrian et al. 2000). An early approach for this aim, used in cosmological simulations (e.g. Dolag et al. 2005; Vazza et al. 2009b), consisted of filtering out the turbulent motions by subtracting the local mean velocity, defined over a fixed spatial scale. Another widely used strategy (e.g. Lau, Kravtsov & Nagai 2009) uses the velocity dispersion in radial shells as a proxy to quantify the level of turbulence around a particular cluster. Although these algorithms are conceptually simple and allow to effectively split the velocity fields in a bulk and a turbulent component, their fixed lengths do not fully capture the essentially multiscale nature of turbulent phenomena.

More recently, Vazza et al. (2012, 2017) proposed an algorithm to iteratively constrain the outer scale[5] for turbulence, $L(\mathbf{x})$, which then allows to extract the purely turbulent velocity field, $\delta\mathbf{v}(\mathbf{x})$, without explicitly fixing a filter length. While in Vazza et al. (2017) the authors use a fixed refinement technique, so that the virial volume of their cluster is simulated within a uniform grid of constant, $\Delta x \sim 20\,\mathrm{kpc}$ resolution, here we aim to extend the algorithm to a purely AMR velocity field, which allows us to straightforwardly apply the filter to clusters extracted from full-cosmological simulations without the need of performing resimulations. While the 'physical' steps are parallel to those of Vazza et al. (2012), the multigrid, multiresolution description of an AMR simulation prevents us from using several handy tools (e.g. boxcar averages, convolutions, etc.) in the original formulation and, consequently, also increases substantially the computational cost. The basic steps of our implementation are as follows:

(i) For each volume element in the computational domain, at a given refinement level ℓ, the outer scale length, $L(\mathbf{x})$, is initialized to $L_{n=0} = 3\Delta x_\ell$, where Δx_ℓ is the cell size at level ℓ and n represents the algorithm iteration.

(ii) For each cell, at position $\mathbf{x}$, $L(\mathbf{x})$ grows until convergence is reached according to the following iterative scheme:

(a) The bulk velocity at $\mathbf{x}$, at the iteration n, is computed as

$$\langle\mathbf{v}\rangle_n(\mathbf{x}) = \frac{\iiint_{V'_n}\rho(\mathbf{x}')\mathbf{v}(\mathbf{x}')\mathrm{d}V'}{\iiint_{V'_n}\rho(\mathbf{x}')\mathrm{d}V'} = \frac{\sum_{V'_n} m_i\mathbf{v}_i}{\sum_{V'_n} m_i}, \tag{1}$$

where $\rho(\mathbf{x}')$ is the gas density at $\mathbf{x}'$, m_i is the mass of the volume element i, and V'_n is the spherical volume consisting of all the cells at positions $\mathbf{x}'$ such that $|\mathbf{x}' - \mathbf{x}| \leq L_n$ (at the maximum available resolution at each position).

(b) The turbulent velocity field at $\mathbf{x}$, at the iteration n, is computed as $\delta\mathbf{v}_n(\mathbf{x}) = \mathbf{v}(\mathbf{x}) - \langle\mathbf{v}\rangle_n(\mathbf{x})$.

(c) Unless a stopping condition is triggered (see below), $L(\mathbf{x})$ is increased according to

$$L_{n+1} = \max\left(L_n + \Delta x_\ell,\, [1+\chi]\, L_n\right), \tag{2}$$

where we fix $\chi = 0.05$. This condition, which has been tested experimentally, prevents a slow convergence when a high-resolution region has $L(\mathbf{x}) \gg \Delta x_\ell$.

(iii) For each cell, the iterative scheme is stopped whenever at least one of the following conditions is met:

(a) The fractional change in $\delta\mathbf{v}$ between two consecutive iterations, which we define as

$$\Delta = \max_{i=1,2,3}\left|\frac{\delta v^i_n}{\delta v^i_{n-1}} - 1\right|, \tag{3}$$

falls beyond a fixed tolerance parameter, $\Delta \leq \Delta_{\mathrm{tol}}$. When this condition is achieved, L_n represents the maximum correlation scale of the velocity field at $\mathbf{x}$. We have set $\Delta_{\mathrm{tol}} = 0.05$, although in our tests the turbulent velocity field does not depend strongly on this threshold, as it was also reported by Vazza et al. (2012) and Valdarnini (2019).

(b) A shocked cell, which we define as a cell with Mach number $\mathcal{M} \geq 1.3$ (e.g. Quilis et al. 1998; Ryu et al. 2003; Vazza, Brunetti & Gheller 2009a; Planelles & Quilis 2013; Martin-Alvarez, Planelles & Quilis 2017), enters the volume V'_n. We detect shocks with the shock finder presented in Planelles & Quilis (2013), which uses the one-dimensional discontinuities in temperature to solve for $\mathcal{M}$. The shock exclusion is well motivated by the fact that shocks introduce velocity discontinuities, which importantly bias the bulk velocity determination.

(c) The volume V'_n intersects the computational domain's boundary. Even though we could explicitly enforce the periodic boundary conditions to get rid of this limitation, this condition only biases the results in a negligible amount of volume close to the boundary, where we do not expect to find objects of interest.

Once this procedure is repeated for all the volume elements, we have a complete, multiscale description of the coherence scale, $L(\mathbf{x})$, and the bulk and turbulent[6] velocity fields. Naturally, the turbulent velocity field can also be decomposed into its rotational and compressive components using the algorithm described in Section 2.1, yielding the solenoidal and compressive turbulent velocity fields, respectively. The routines for applying this filtering scheme have been included in the publicly available code VORTEX[7] (Vallés-Pérez et al. 2021).

3 THE SIMULATION

The remaining pages of this manuscript are based on the results of the codes described in Section 2 applied to a high-resolution cosmological simulation, which is described below. In Section 3.1,

[4] While we use FFT for the base level, taking advantage of the periodic boundary of cosmological simulations, this is not an imposition of the method. If non-periodic boundary conditions are given, the base level could also be addressed with iterative methods, as we do for the refinement levels.

[5] In turbulence theory, the outer scale, $L(\mathbf{x})$, is the spatial scale at which turbulent motions are injected, and it corresponds to the largest scale inside the inertial range. In this text, we shall refer to the *outer scale* or *injection scale* interchangeably. We refer the reader to Landau & Lifshitz (1987, sections 26–38) for an introductory text on hydrodynamic turbulence.

[6] Through this manuscript, we shall also refer to $\delta\mathbf{v}(\mathbf{x})$ as the *filtered* or the *small-scale* velocity field.

[7] https://github.com/dvallesp/vortex

we outline the structures which constitute the primary focus of this work.

The simulation analysed in this paper has been carried out with MASCLET, an Eulerian AMR, *high-resolution shock-capturing* hydrodynamics, coupled to particle-mesh *N*-body, cosmological code (Quilis 2004). This same simulation has been employed in several previous works (Quilis, Planelles & Ricciardelli 2017; Planelles et al. 2018; Vallés-Pérez et al. 2020). Here, we describe the main features of the simulation. For more information about some details not directly connected to our analyses, we refer the reader to the aforementioned references.

The simulation corresponds to a cubic, periodic domain of comoving side length 40 Mpc, with a flat Λ cold dark matter (CDM) cosmology set up by a Hubble parameter $h \equiv H_0/(100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}) = 0.678$, composition given by the matter, baryon and dark energy parameters $\Omega_\mathrm{m} = 0.31$, $\Omega_b = 0.048$, $\Omega_\Lambda = 0.69$, and primordial fluctuation spectrum set according to a spectral index $n_s = 0.96$ and amplitude $\sigma_8 = 0.82$. Thus, the cosmology is consistent with the latest values reported by the Planck Collaboration (2020). The domain has been discretized in a base grid of 128^3 cells, which yields a harsh resolution of ~ 310 kpc. Regions with large gaseous and/or DM densities get recursively refined following a pseudo-Lagrangian AMR scheme with up to $n_\ell = 9$ refinement levels, providing a peak resolution of nearly ~ 610 pc. With four species of DM particles, the best DM mass resolution is $\sim 2 \times 10^6\,\mathrm{M}_\odot$, equivalent to filling the domain with 1024^3 of such particles.

The initial conditions were set up by a CDM transfer function (Eisenstein & Hu 1998) at redshift $z = 100$. A constrained realization of the Gaussian random field, according to the procedure of Hoffman & Ribak (1991), was used to produce a massive cluster in the centre of the box. Besides gravity, the simulation accounts for several cooling mechanisms (free–free, inverse Compton, and atomic and molecular cooling for a primordial gas), as well as heating by a UV background of radiation (Haardt & Madau 1996). Star formation and type-II supernova feedback are parametrized according to Yepes et al. (1997) and Springel & Hernquist (2003). Even though this run does not include AGN feedback, which could constitute an important source of turbulent motions in the innermost regions of galaxy clusters (see e.g. Vazza et al. 2012), this drawback is not especially relevant for this work, where our primary intent is performing an exploratory analysis showing the capabilities of our algorithms. We may postpone further analyses, including AGN feedback and magnetic fields, to a future work.

3.1 Structure finding

We have identified the structures in this computational domain by means of the spherical overdensity DM halo finder ASOHF (Planelles & Quilis 2010; Knebe et al. 2011). At $z \simeq 0$, there are two massive galaxy clusters (with virial masses $M_\mathrm{vir,DM} > 10^{14}\,\mathrm{M}_\odot$), which we shall hereon refer to as CL01 and CL02.

In Vallés-Pérez et al. (2020), we have analysed the accretion histories of these two objects. CL01 is a massive cluster which suffers several major and minor merging periods through its recent history, the most recent major merger having occurred at around $z \sim 0.9$ (see fig. 2 in Vallés-Pérez et al. 2020), and only experiencing quiescent accretion therein, with low accretion rates from $z \sim 0.4$ on. Likewise, CL02 is a smaller, $\sim 10^{14}\,\mathrm{M}_\odot$ cluster which has experienced a major merger at $z \sim 1.4$ and a minor merger at around $z \sim 0.2$, the latter not having any significant impact on the cluster's structure and, thus, probably having had reduced dynamical relevance, as far as turbulence is concerned. A summary of the properties of these two objects, at $z \simeq 0$, is shown in Table 1.

Table 1. Main properties of clusters CL01 and CL02 at $z = 0$. Virial radii, R_vir, are defined with respect to the DM distribution, according to the standard spherical overdensity definition (Lacey & Cole 1994) with the virial overdensity given by Bryan & Norman (1998). M_DM and M_gas are measured inside R_vir. Temperatures and entropies are computed inside R_vir, assuming hydrostatic equilibrium (equations 59 and 64, Voit 2005) and with a mean molecular weight $\mu = 0.6$.

Cluster	R_vir (Mpc)	M_DM ($10^{13}\,\mathrm{M}_\odot$)	M_gas ($10^{13}\,\mathrm{M}_\odot$)	$k_\mathrm{B}T_\mathrm{vir}$ (keV)	$K_{e,\,\mathrm{vir}}$ (keV cm^2)
CL01	1.99	42.9	4.56	3.27	1230
CL02	1.26	10.9	1.33	1.34	520

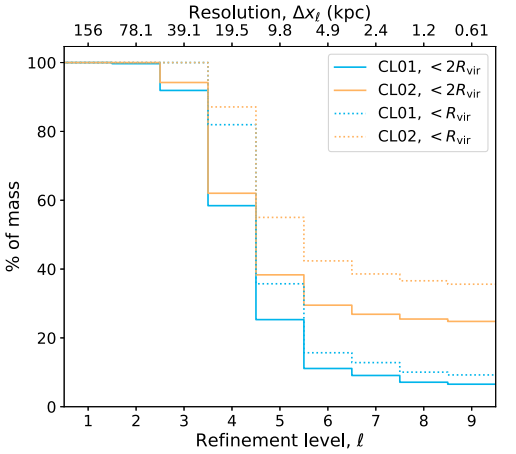


Figure 1. Description of the resolution coverage of each cluster (within $2R_\mathrm{vir}$, solid lines; and within R_vir, dotted lines) at $z \simeq 0$. For each refinement level, ℓ, or resolution, Δx_ℓ, the vertical axis indicates the fraction of mass inside the considered volume in cells with resolution Δx_ℓ or better.

As the description of the turbulent flows is strongly dependent on the resolution of the numerical grid, we present the (cumulative) fraction of gas mass (within $2R_\mathrm{vir}$, solid lines; and R_vir, dashed lines) resolved, by $z \sim 0$, at each refinement level in Fig. 1. For both clusters, $\gtrsim 60$ per cent of the mass inside $2R_\mathrm{vir}$ is resolved in patches with resolution at least $\Delta x_4 \sim 20$ kpc. This resolution can be directly compared to previous works which employ ad hoc resimulations and fixed refinement techniques (e.g. Vazza et al. 2017, who resolve the virial volume of their clusters within a uniform grid with an equivalent resolution; see also Miniati 2014, 2015, who achieve a ~ 10 kpc resolution throughout the virial volume of their cluster).

If we restrict to the virial volume, virtually all the mass is resolved at resolutions equal or better than ~ 40 kpc, and $\gtrsim 80$ per cent at resolutions at least ~ 20 kpc. Additionally, almost 10 per cent (CL01) and 40 per cent (CL02) of the virial mass is resolved with resolutions below the kpc.

4 GLOBAL STATISTICS OF TURBULENCE

In this section, we analyse several global indicators of turbulence. In particular, we explore the second-order velocity structure functions (Section 4.1), their evolution (at fixed scales) in relation to accretion

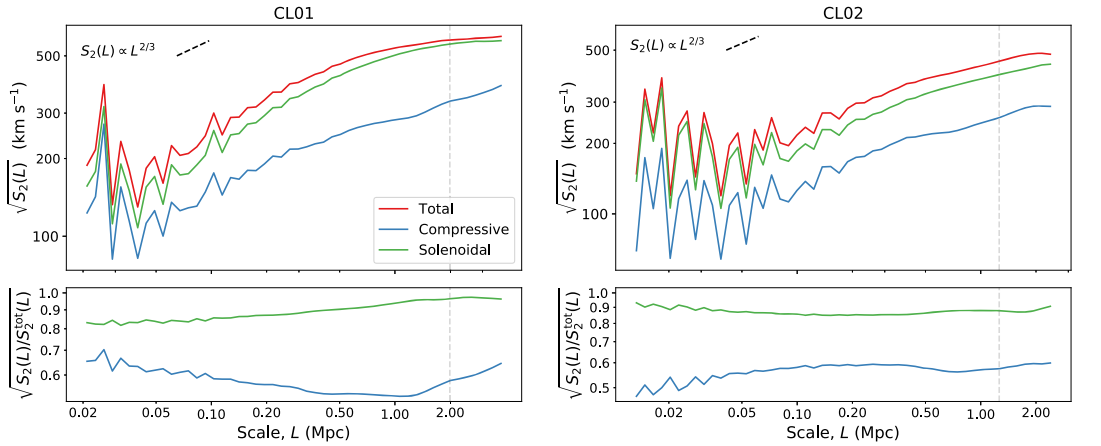


Figure 2. Second-order structure functions for clusters CL01 (left) and CL02 (right) at $z \simeq 0$. The red, green, and blue lines correspond to the total, solenoidal, and compressive velocity fields, respectively. The grey, dashed, vertical lines mark the scale of the virial radii. The black, dashed lines indicate the slope expected by the Kolmogorov scaling (Kolmogorov 1941). The square root of the second-order structure functions is represented to show the results in units of velocity.

rates and merger events (Section 4.2) and the scaling relations and mechanisms for the generation of solenoidal turbulence (Section 4.3).

4.1 Turbulence spectra: structure functions

Structure functions quantify the magnitude of the velocity fluctuations on different scales over the cluster volume (or mass), and therefore offer a straightforward way to define a global statistic of turbulence for a given object. The most direct definition of the structure function of order p, $S_p(L)$, is given by the expression

$$S_p(L) = \langle |\mathbf{v}(\mathbf{x} + L\hat{\mathbf{n}}) - \mathbf{v}(\mathbf{x})|^p \rangle_{\mathbf{x},\hat{\mathbf{n}}}\,, \tag{4}$$

where $\mathbf{v}$ will hereon denote the (*unfiltered*, i.e. total) peculiar velocity field, $\hat{\mathbf{n}}$ is a unit vector and the average is carried both in positions, $\mathbf{x}$, inside the volume of interest and in directions, $\hat{\mathbf{n}}$. Using the HHD algorithm presented in Vallés-Pérez et al. (2021) (see Section 2.1), we will also compute these structure functions separately for the compressive and solenoidal velocity components. While some authors (e.g. Valdarnini 2011; Miniati 2014) further decompose the structure functions in a longitudinal and a transverse component, we shall not pursue such decomposition here, as we are only interested in the magnitudes of the velocity fluctuations at each scale.

In order to correctly ponder the dynamically relevant regions, we adopt a mass weight when performing the average over $\mathbf{x}$. In particular, our procedure for computing $S_p(L)$ can be summarized in the following steps:

(i) We choose N_s mass-weighted random points (hereon, the *sampling points*) inside the considered volume (in our case, inside a sphere twice the virial radius around each cluster). The weighted sampling is performed by applying Smirnov's inverse transformation method (see e.g. Devroye 1986), i.e. we compute the cell-wise cumulative normalized mass distribution inside $2R_{\rm vir}$, draw uniformly sampled random numbers, $s \in [0, 1]$, and apply the inverse cumulative function to these numbers.

It is worth noting that gas clumps and substructures can importantly bias the random selection of points, especially in this simulation as the lack of AGN feedback could produce overcooling (see e.g. Kravtsov, Nagai & Vikhlinin 2005; Eckert et al. 2012; Zhuravleva et al. 2013; Planelles et al. 2014). In order to avoid our sampling points to concentrate in highly overdense gas clumps, whose dynamics may differ from those of the bulk ICM, we first apply the inhomogeneity identification technique described by Zhuravleva et al. (2013). In particular, we radially split the cluster in $N_{\rm bins} = 20$ logarithmic bins,[8] and tag as clumps all the cells whose logarithmic density is further than $f_{\rm cut} = 3.5$ standard deviations from the volume-weighted median log-density.

(ii) Around each of the sampling points, we pick N_f volume-weighted random points (hereon, the *field points*). In order to effectively sample all the scales, we extract the field points around each sampling point, $\mathbf{x}_s$, using the same method as above but weighting the probability of each cell by $V_{\rm cell}/|\mathbf{x}_s - \mathbf{x}_{\rm cell}|^3$. It is easy to show that this weight ensures equal number of points per logarithmic radial bin.

(iii) Finally, we bin logarithmically the distances among points, L, and average over all the possible pairs at each bin according to expression (4).

In our tests, taking $N_s = 5000$ mass-weighted points is enough to adequately sample the mass distribution within a precision of 5 per cent or better. We take as well up to $N_f = 5000$ volume-weighted field points around each sampling point, yielding a maximum of $N_s \times N_f = 2.5 \times 10^7$ pairs of points to compute the structure function. We then bin L in 50 logarithmic intervals, so that the final value of $S_p(L)$ at each L has been computed as the average over $\lesssim 5 \times 10^5$ pairs. In order to check whether these statistics are robust enough, we have performed 20 bootstrap iterations of these procedures, verifying a small variance in our results (typically in the order of ~ 1 per cent, and $\lesssim 10$ per cent for any bin).

We present in Fig. 2 the second-order ($p = 2$) structure functions for the total velocity, as well as for the compressive and the rotational

[8] We note that the substructure-excised density profiles are not sensitive to the particular choice of $N_{\rm bins}$ in the range $N_{\rm bins} \in [15, 100]$.

velocity components, for clusters CL01 (left-hand panel) and CL02 (right-hand panel) at $z \simeq 0$. For comparison, in both panels we show the slope predicted by the Kolmogorov (1941) model for fully developed isotropic, incompressible turbulence, $S_2(L) \propto L^{2/3}$. Both components (as well as the total velocity) present a well-defined inertial range, whose slope roughly agrees with the Kolmogorov prediction. The structure functions for the solenoidal component are consistently higher than their compressive counterparts, mainly as a consequence that solenoidal flows are much more ubiquitous and volume filling. However, at variance with previous results with static refinements, we do not find a clear difference in the slopes of these components. Miniati (2014, 2015) finds the compressive component to have a steeper spectrum, the opposite trend being reported by Vazza et al. (2017). The lower panels in Fig. 2 show the quotient of the solenoidal and compressive structure functions to the total one. While CL01 presents a higher slope in the solenoidal component, the opposite happens for CL02. This highlights that these slopes could be particularly sensitive to the dynamical state of the cluster, as well as to the numerical scheme and the sampling of $S_p(L)$.

The structure functions flatten at $L \gtrsim 1$ Mpc, placing the characteristic scale for turbulence injection at around $(0.5–1)R_{\rm vir}$, in consistence with Miniati (2014) and Vazza et al. (2017).[9] The energy that bulk motions inject into the turbulent velocity field then cascades down to smaller scales through different instabilities (e.g. Kelvin–Helmholtz; see for example ZuHone 2011), until they get (numerically) dissipated when the eddy sizes are in the order of several times the cell side length. The pseudo-Lagrangian refinement approach of our AMR implementation implies that different mass elements in the cluster get different dissipation scales, originating oscillations in the small-scale end of the structure functions. These oscillations are also present in the 'Lagrangian AMR' runs of Miniati (2014), and limit the extent to which structure functions are a useful statistic to study turbulence in full-cosmological simulations using pseudo-Lagrangian approaches.

4.1.1 *Characteristic time-scales of turbulent eddies*

The time-scale in which eddies of size L cascade down to the dissipation scale can be estimated by the so-called *eddy turn-over time*, $\tau_{\rm eddy} \sim L/\mathcal{V}$, where $\mathcal{V}$ is a characteristic velocity at the scale L. In the framework of the Kolmogorov (1941) model, this quantity is expected to scale as $\tau_{\rm eddy} \propto L^{2/3}$. It is customary through the bibliography (e.g. Miniati 2014) to just choose a length ($L \sim R_{\rm vir}$) and a characteristic velocity representing the overall cluster dynamics ($v_{\rm vir} \equiv \sqrt{GM_{\rm vir}/R_{\rm vir}}$) to have a time-scale estimate. Even though these can only be taken as order-of-magnitude estimates, we argue that a more self-consistent estimation can be got by using the $S_2(L)$ spectrum, as $\tau_{\rm eddy}(L) \sim L/\sqrt{S_2(L)}$, since it is a property intrinsic to the gas velocity field (and not of the gravitational potential well, as the one derived from the circular velocity inside a clustercentric radius L, $v_{\rm circ}(L)$).

Fig. 3 presents the eddy turn-over time-scales for clusters CL01 and CL02 (red and green dots, respectively) at redshift $z \simeq 0$. The time-scales for both clusters are remarkably similar through all the scale range, despite the difference in masses. Across the inertial

[9] Other works (e.g. Vazza et al. 2012) report slightly smaller injection scales, in the range $(0.1–0.3)R_{\rm vir}$. In any case, turbulence is not expected to be injected by bulk motions from any specific length, but from a wide range of scales instead.

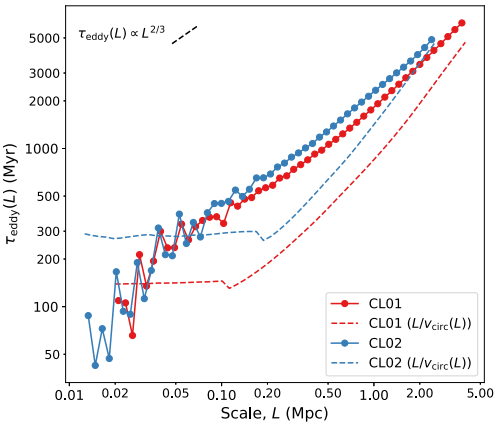


Figure 3. The dots joined by the solid lines show the eddy turn-over times for different scales, $\tau_{\rm eddy}(L)$, estimated using the structure functions shown in Fig. 2 for cluster CL01 (red lines) and CL02 (blue lines). For comparison, the dashed lines show the time-scales estimated from the circular velocities inside spheres $r < L$ around the cluster centre.

range, $\tau_{\rm eddy}(L)$ scales in good agreement with the Kolmogorov prediction. For comparison, we also show, in dashed lines, the time-scale derived from the circular velocity inside a radius L. Assuming this magnitude to be dominated by the DM mass distribution, and taking an NFW (Navarro, Frenk & White 1997) profile for DM density, it is easy to show that the time-scale computed that way should scale as $L^{3/2}/\log(L) \sim L^{3/2}$ for sufficiently large L ($L \gg r_s$, being r_s the scale radius of the NFW profile). Indeed, this is the behaviour observed in our clusters. Interestingly, these two time-scales seem to approach for $L \gtrsim R_{\rm vir}$, making $R_{\rm vir}/v_{\rm vir}$ an incidentally good guess for the characteristic time-scale of the turbulent energy cascade for injection scales comparable to the virial radius. However, we emphasize that the time-scale $\tau_{\rm eddy} = L/\sqrt{S_2(L)}$ better captures the multiscale nature of turbulent motions.

By fitting the inertial ranges of the time-scale curves to a power law, for each of the clusters, we get the following fitting formula for the cascade time-scales of eddies of size L:

$$\tau_{\rm eddy}(L) \approxeq (1755 \pm 78)\,{\rm Myr} \left(\frac{L}{1\,{\rm Mpc}}\right)^{0.647 \pm 0.026} \quad (5)$$

Naturally, these fitting formulae have to be taken with caution, since a proper statistical analysis, which we may defer for a future work, must be performed in order to evaluate its universality. Nevertheless, within the context of this work and focusing at recent redshifts, we shall use this expression to interpret the results in the next sections.

4.1.2 *Behaviour of the different phases*

Although the average statistics of the ICM velocity field show general properties which roughly agree with the Kolmogorov scaling, the behaviour can vary significantly for different phases. While some authors perform a phase-space splitting and study the properties of the velocity fields of the different phases (e.g. Schmidt et al. 2016), here we find more clear and reproducible to take a geometric criterion to study the dependence of the velocity structure functions. Thus, we consider the following regions: *core* ($0 < r/R_{\rm vir} < 0.1$), *off-core*

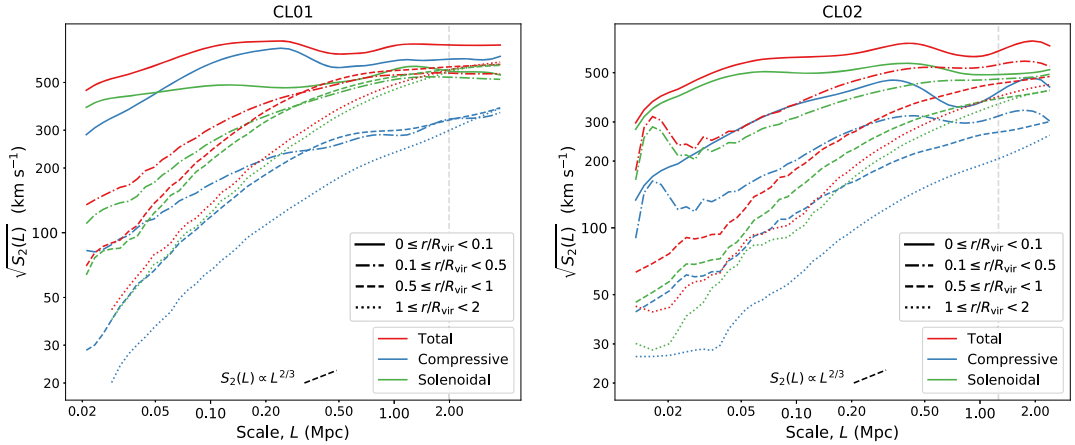


Figure 4. Second-order structure functions for different regions of clusters CL01 (left) and CL02 (right) at $z \simeq 0$. The solid, dot–dashed, dashed, and dotted lines correspond to the core, off-core, virial, and outskirts regions, as defined in Section 4.1.2. The rest of notations are the same as in Fig. 2.

($0.1 < r/R_{\rm vir} < 0.5$), *virial* ($0.5 < r/R_{\rm vir} < 1$), and *outskirts* ($1 < r/R_{\rm vir} < 2$). Similar choices are made in previous works (e.g. Miniati 2014). For each phase, we only count the sampling points inside the considered region, while, naturally, field points can be in any position within the $2R_{\rm vir}$ volume. The results for CL01 and CL02 are presented in Fig. 4. In this case, we have smoothed the functions using a Savitzky & Golay (1964) filter to alleviate the oscillations, whose origin has already been commented, and focus on the general scaling.

For both clusters, at large clustercentric distances, the small-scale velocity fluctuations are strongly suppressed, as these regions are typically resolved within refinement levels $\ell = 2$ or $\ell = 3$ (thus, with an effective resolution of ~ 40 kpc or worse, and placing the numerical dissipation scale at several hundreds of kpc), and, consequently, the structure functions are much steeper than Kolmogorov's. Although these trends do vary significantly from cluster to cluster (and with cosmic time), this lack of small-scale power, especially for the solenoidal component, is also reported by Miniati (2014) in their 'Lagrangian AMR' run.

On the other hand, cluster cores are extremely sensitive to the feedback mechanisms which are accounted for in the simulation (see, for example, Rasia et al. 2015 and Planelles et al. 2017). In this case, the lack of AGN feedback prevents the injection of random motions in the innermost regions of clusters. Despite this effect, fluctuations on these inner regions are considerably higher than in the rest of regions, both because of numerical (these regions are likely to be mostly resolved with cell sizes of $\lesssim$kpc, placing the dissipation scale below $\sim$10kpc) and physical (SNe feedback and inner, merger shocks can be important sources of turbulent motions) reasons. Valdarnini (2011, 2019) suggests a scenario where small-scale turbulence is generated in the cluster central regions due to the interaction of a dense, compact core with the surrounding ICM in runs with cooling but not central sources of energy. This could, indeed, be the case of our clusters. Both CL01 and CL02 have flat density profiles in the central ~ 100 kpc and lower temperatures in their cores than the surrounding ICM, thus corresponding to cool, compact cores (Burns et al. 2008). Therefore, this scenario is capable of providing a plausible explanation for the flattening of the structure functions in inner radii.

4.2 Evolution of the velocity fluctuations on different scales

As seen through Section 4.1, the global statistics of the velocity field in pseudo-Lagrangian AMR simulated galaxy clusters present systematic deviations from those obtained in other works using fixed grids. However, we shall argue that these differences emerge noticeably wherever the resolution is low (i.e. in low-density regions such as cluster outskirts). Even though our clusters simulated in a full-cosmological environment can lack small-scale velocity power in a fraction of their volume, the *ansatz* behind the AMR strategy based on local density implies that this has indeed a small contribution to the total energetics and dynamical evolution of the cluster, since the mass fraction corresponding to this large volume is generally reduced.

As mentioned in Section 4.1.2, the actual behaviour of the structure functions is highly dependent on the dynamical state, and thus it is interesting to investigate its time evolution. In order to avoid the systematic effects seen in the velocity structure functions due to the AMR grid structure, we will focus on the values of the structure function for certain pre-defined scales, large enough to neglect the effect of the aforementioned oscillations. Hereon we will refer to these quantities as the *velocity fluctuations at the scale L*. Thus, even though the slope of our structure functions could be biased with respect to uniform grid runs, a large part of these systematics is cancelled out by focusing on constant scales, and the evolution of the velocity fluctuations can be properly assessed.

In particular, we compute $S_2(L)$ through each of the 41 code outputs from $z \simeq 1.5$ to $z \simeq 0$ at the scales $L/R_{\rm vir} = 0.1$, 0.5, 1, and 1.5. The procedure is identical to the one described in Section 4.1, but in this case we only generate uniformly distributed field points at radial distances corresponding to the four considered scales around each sampling point (with a given tolerance, which we have fixed to 10 per cent).

4.2.1 Relation to accretion and mergers

Accretion and mergers are important sources of energetic feedback to the ICM, and are directly connected to the dynamical state of the cluster (see e.g. Quilis et al. 1998; Planelles & Quilis 2009; Lau

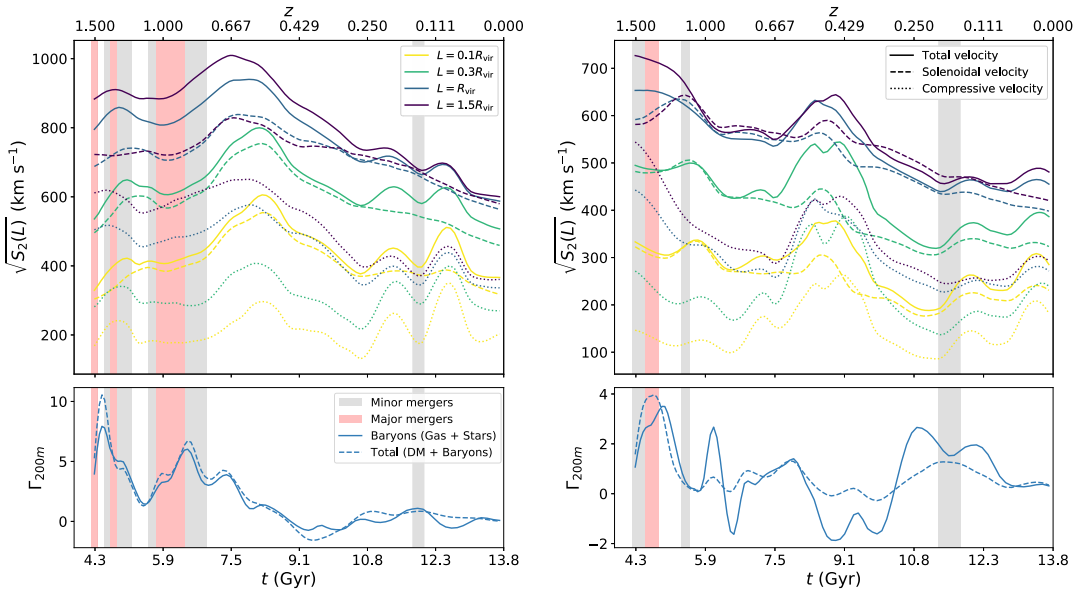


Figure 5. Top panels show the evolution, from $z \simeq 1.5$ to $z \simeq 0$, of the velocity fluctuations on scales $L = 0.1R_{\rm vir}$ (yellow), $0.3R_{\rm vir}$ (green), $R_{\rm vir}$ (blue), and $1.5R_{\rm vir}$ (purple), for clusters CL01 (left) and CL02 (right). The lower panels show the baryonic (solid lines) and total (dashed lines) instantaneous MARs. The velocity fluctuation curves have been smoothed using a Savitzky–Golay filter of the same order and window length as the MARs, for both curves to have a similar level of (temporal) locality. The legends apply to both panels.

et al. 2015; Chen et al. 2019; Vallés-Pérez et al. 2020). In order to investigate their role in generating turbulence, we will correlate the velocity fluctuations on the scales defined above with the merger periods and the instantaneous accretion rates.

Following Diemer & Kravtsov (2014), we define the mass-accretion rate (MAR) as the logarithmic rate of change of the enclosed mass with respect to the scale factor. Operationally, and as opposed to other works which use an averaged MAR over a wide time interval, in Vallés-Pérez et al. (2020) we defined an instantaneous MAR,

$$\Gamma_\Delta(a) = \frac{{\rm d}\log M_\Delta}{{\rm d}\log a} \tag{6}$$

computed from the sparse snapshot sampling using Savitzky & Golay (1964) filters for performing the differentiation, in order to mitigate the contamination of sampling noise on the numerical derivatives. For the analyses in this work, we will be interested in the baryonic (thus, M_Δ being the combined mass of gas and stars) and total (baryonic and dark matter) MARs.

We split the recent ($1.5 \gtrsim z \geq 0$) history of the clusters in three merging regimes, namely 'ongoing major merger', 'ongoing minor merger', and 'smoothly accreting'. For classifying the mergers, we use the mass ratio of the two most massive progenitor clusters (Planelles & Quilis 2009). Mergers above a 1:3 mass ratio are regarded as *major mergers*. Those between 1:3 and 1:10 are tagged as *minor mergers*. Mergers below the 1:10 threshold are not considered, and a cluster not experiencing any merger above this threshold at a given redshift is classified as smoothly accreting.

Fig. 5 presents the joint analysis of velocity fluctuations on different scales, accretion rates, and merging regimes. For the case of CL01 (left-hand panel), an important correlation between the evolution of velocity fluctuations on different scales and MARs is noticed. At high redshifts, while the cluster is fastly accreting gas from the mergers, velocity fluctuations are mantained at high values. In fact, the major merger at around $z \sim 0.9$ and the corresponding peak of the baryonic MAR at $z \sim 0.8$ injects a large amount of kinetic power on large scales, resulting in the enhancement of the corresponding $S_2(L)$ in the following ~Gyr ($S_2(1.5R_{\rm vir})$ peaks at $z \sim 0.67$, 1.1 Gyr later than the MAR peak). The delay in the peaks at smaller scales directly reflects how the energy is cascading to smaller scales through fluid instabilities. For example, at $0.1R_{\rm vir}$ scales, the total velocity fluctuation peaks at $z \sim 0.55$ (800 Myr after the peak at the largest scales). On the other hand, after the minor merger at $z \sim 0.15$, which does not have a severe impact on the MARs, the compressive velocity fluctuations are slightly enhanced at all scales (but not so the solenoidal ones).

As for cluster CL02, the most remarkable feature is perhaps the increase of the velocity fluctuations at all scales during $0.7 \gtrsim z \gtrsim 0.4$. Incidentally, during this period the cluster presents negative baryonic MARs, while the DM halo mass remains fairly constant (and so the total MAR is close to zero). This effect could be due to e.g. gas sloshing in the DM potential well, or due to a merger-accelerated shock scenario (Zhang et al. 2020). Although we do not pursue a detailed explanation of the dynamical origin of this phenomenon, visual inspection of density and Mach number slices around the object, coupled with the fact that the increase in velocity fluctuations is driven by the compressive component of the velocity field, tend to suggest the latter as the most plausible.

The examples above depict a complex and varied phenomenology where, nevertheless, galaxy cluster mergers and accretion phenomena seem to dominate the evolution of the velocity fluctuations through cosmic time, acting thus as primary sources of turbulence

in the ICM. In any case, these mechanisms could be punctuated, specially at the smallest scales and in the innermost regions of galaxy clusters, by sources of feedback such as SNe and/or AGN energy injection (the latter not been accounted for in this simulation).

4.3 Enstrophy and helicity

We now consider the evolution of two quantities intimately related to solenoidal turbulence, which as we have seen is the dominant ICM turbulence component. These two quantities are computed from the pseudo-vector *vorticity*, i.e. the curl of the velocity field, $\boldsymbol{\omega} = \nabla \times \mathbf{v}$. From this quantity, we consider the scalar *enstrophy*,

$$\epsilon = \frac{1}{2}\boldsymbol{\omega}^2, \tag{7}$$

which has been employed in many previous studies (Porter, Jones & Ryu 2015; Iapichino, Federrath & Klessen 2017; Vazza et al. 2017; Wittor et al. 2017; Valdarnini 2019) as a proxy of solenoidal turbulence; and the pseudo-scalar[10] *helicity* (e.g. Moffatt 2014),

$$\mathcal{H} = \mathbf{v} \cdot \boldsymbol{\omega}, \tag{8}$$

which is an interesting quantity from several points of view, as it will be discussed below, which however has not yet been applied, to our knowledge, to previous studies of cosmic flows. When focusing on a particular cluster, we shall compute $\mathcal{H}$ in its rest frame. The evolution of these quantities can be obtained, in the cosmological case, from the curl of the equation for the evolution of peculiar velocity in comoving coordinates. After some manipulation, this yields the following equation for the evolution of the (peculiar) vorticity pseudo-vector:

$$\frac{\partial \boldsymbol{\omega}}{\partial t} + \frac{1}{a}(\mathbf{v}\cdot\nabla)\boldsymbol{\omega} = -H\boldsymbol{\omega} - \frac{1}{a}\boldsymbol{\omega}(\nabla\cdot\mathbf{v}) + \frac{1}{a}(\boldsymbol{\omega}\cdot\nabla)\mathbf{v} + \frac{1}{\rho^2 a}\nabla\rho\times\nabla P, \tag{9}$$

P being the thermal pressure of the gas. That is to say, local vorticity changes due to advection, cosmic expansion, local fluid expansion, vortex stretching, and baroclinicity (reading, from left to right, the terms in equation 9). From this equation, one can straightforwardly get the equation for the evolution of enstrophy in the comoving frame, since $\frac{\partial\epsilon}{\partial t} = \boldsymbol{\omega}\frac{\partial\boldsymbol{\omega}}{\partial t}$ and therefore

$$\frac{\partial\epsilon}{\partial t} + \frac{1}{a}\nabla\cdot(\epsilon\mathbf{v}) = -2H\epsilon - \frac{1}{a}\epsilon(\nabla\cdot\mathbf{v}) + \frac{1}{a}\boldsymbol{\omega}(\boldsymbol{\omega}\cdot\nabla)\mathbf{v} + \frac{\boldsymbol{\omega}}{\rho^2 a}(\nabla\rho\times\nabla P). \tag{10}$$

Note this equation is equivalent to equation 3 in Porter et al. (2015), but in this case in terms of the peculiar magnitudes of the fluid (removing the cosmological background), and without the magnetic and dissipative terms (since we do not include magnetic field nor we can explicitly evaluate the viscosity in our numerical scheme). The terms in equation (10) can be interpreted in a one-to-one correspondence to the ones in equation (9). Writing $\frac{\partial\epsilon}{\partial t} = \sum_i F_i$, we have:

- The advection term, $F_{\rm adv} = -\frac{1}{a}\nabla\cdot(\epsilon\mathbf{v})$, whose volume integral equates to the net enstrophy inflow.
- The cosmic expansion term, $F_{\rm cosm} = -2H\epsilon$, which dilutes the local enstrophy proportionally to the cosmic expansion rate, $H(t) \equiv \dot{a}/a$.
- The peculiar expansion/compression term, $F_{\rm pec} = -\frac{1}{a}\epsilon(\nabla\cdot\mathbf{v})$. Although we shall generally refer to this term as the 'expansion' term, note it is positive (negative) wherever the fluid contracts, $\nabla\cdot\mathbf{v} < 0$ (expands, $\nabla\cdot\mathbf{v} > 0$).
- The vortex stretching term, $F_{\rm vs} = \frac{1}{a}\boldsymbol{\omega}(\boldsymbol{\omega}\cdot\nabla)\mathbf{v}$, which appears when a gas element is accelerated in the direction of its vorticity (and therefore the vortex tube is stretched).
- The baroclinic term, $F_{\rm baroc} = \frac{\boldsymbol{\omega}}{\rho^2 a}(\nabla\rho\times\nabla P)$. Looking at the corresponding term in equation (9), this is the only term which can generate vorticity even if $\boldsymbol{\omega} = \mathbf{0}$.

Similarly, an equation for the evolution of helicity can be derived from equation (9), given that $\frac{\partial\mathcal{H}}{\partial t} = \frac{\partial\mathbf{v}}{\partial t}\boldsymbol{\omega} + \mathbf{v}\frac{\partial\boldsymbol{\omega}}{\partial t}$:

$$\frac{\partial\mathcal{H}}{\partial t} + \frac{1}{a}\nabla\cdot(\mathcal{H}\mathbf{v}) = -2H\mathcal{H} + \frac{1}{2a}(\boldsymbol{\omega}\cdot\nabla)v^2 + \frac{\mathbf{v}}{\rho^2 a}\cdot(\nabla\rho\times\nabla P) - \frac{1}{a}\nabla\cdot(\Phi\boldsymbol{\omega}) - \frac{1}{\rho a}\nabla\cdot(P\boldsymbol{\omega}). \tag{11}$$

Thus, the local helicity changes due to advection, cosmic expansion, vortex stretching, and baroclinicity, as for the previous magnitudes (reading the terms from left to right; note, however, that there is no peculiar expansion/compression term). The last two terms represent the contribution of the acceleration of pre-existing vortices due to gravity and thermal pressure gradients. In the case of a barotropic fluid, i.e. $P = P(\rho)$, and in the limit where the term due to cosmology can be neglected (i.e. $H\mathcal{H}$ is much smaller than the rest of terms in equation 11), the volume-integrated helicity is a conserved quantity (see e.g. Webb 2018). Thus, the changes in helicity inside a cluster volume can be directly traced to either accretion (or, generally, mass flows) or baroclinicity.

4.3.1 *Evolution of the volume-averaged enstrophy and helicity*

In order to study the driving mechanisms for the evolution of the volume-averaged helicity and enstrophy through our considered redshift interval, $1.5 \gtrsim z > 0$, we have computed the volume average of each of the source terms present in equations (10) and (11), respectively. In Appendix A, we discuss the convergence of the volume-averaged quantities. For each snapshot, we perform the integrals out to the virial radius (at the given snapshot; thus, time-dependent), since a pre-defined, constant comoving radius would imply accounting for a changing composition (in terms of phases), e.g. very non-virialized regions at earlier redshifts, when $R_{\rm vir}$ is smaller. In order to keep track of the magnitude of the *pseudo-evolution* due to the time-dependent integration volume, we add an ad hoc pseudo-evolution source term, which is computed at each snapshot as the difference between the magnitudes average over the new and the previous integration domains. For conciseness, we only present the results for the most massive object, CL01. The results for CL02, although clearly different from the ones for the object we study here due to their different assembly histories, display similar trends. As for the integrated source terms, we have checked that our results do not depend strongly on resolution.

Fig. 6 presents the evolution of enstrophy (top panel), its source terms as defined in equation 10 (second panel), the evolution of helicity (third panel), its source terms as in equation 11 (fourth panel), and, for comparison, the MARs (bottom panel). The solid line in the upper panel represents the total (i.e. unfiltered) enstrophy averaged

[10] In this paper, we refer to *pseudo-scalar* quantities are those which change sign under parity, i.e. under the change of coordinates $\mathbf{x} \mapsto -\mathbf{x}$.

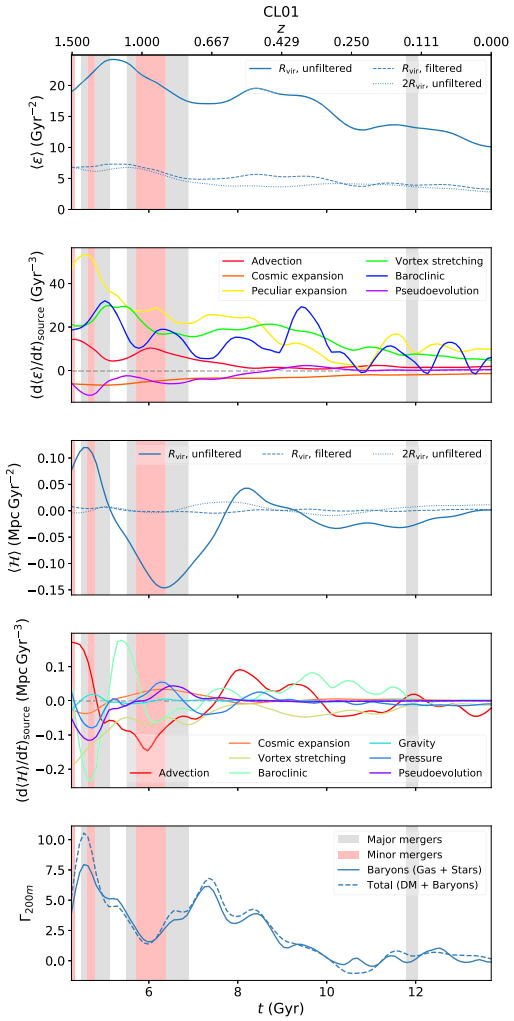


Figure 6. The first and third panels present the evolution of the volume-averaged enstrophy and helicity with cosmic time (or redshift, see legend in the upper panel), respectively. The solid and dashed lines present the unfiltered (total) and filtered (turbulent) quantities averaged over $R_{\rm vir}$, while the dotted lines correspond to the unfiltered quantity inside the larger, $2R_{\rm vir}$, volume. The second and fourth panels quantify the importance of each of the source terms in equations 10 (enstrophy) and 11 (helicity), respectively. The lower panel presents the baryonic (solid) and total (dashed) MARs, while background colours correspond to the merging regimes.

over the virial volume. This quantity undergoes strong evolution across the history of the cluster, intimately tied to the evolution of the accretion rates. Generally, the cluster-averaged enstrophy peaks after major mergers roughly 1 Gyr after the MARs do, thus displaying a delay effect similar to what we have seen in the evolution of the velocity fluctuations (Section 4.1). At recent times, as the net accretion rate halts, the enstrophy falls to values around 10 Gyr^{-2}, which are consistent with the values Vazza et al. (2017) obtain for a similar, slightly less massive cluster. Thus, stirring of the ICM due to mergers, and more generally accretion of gas, appears to be the primary source of solenoidal turbulence in CL01 (Shi & Komatsu 2014; Nelson et al. 2014b; Miniati 2015). Focusing on a larger comoving volume ($2R_{\rm vir}$ around the cluster, dotted lines in the same panel), the evolution of enstrophy is much milder, and enstrophy values themselves are lower. This is again due to, both, numerical and physical reasons. On the one hand, the larger cell sizes in clusters' outskirts suppress the energy cascade at larger scales, as vortices smaller than several times the local resolution length cannot be resolved. Although a handful of methods to overcome this inconvenience exist (either through subgrid modelling, e.g. Schmidt et al. 2016; Kretschmer & Teyssier 2020; or through ad hoc refinement based on local vorticity, e.g. Iapichino & Niemeyer 2008), the dynamical relevance of the possible vortices that we could miss in such low-density regions is reduced and, indeed, Vazza et al. (2017) also find a reduced value of the enstrophy in their large volume, as lower density regions typically experience less intense dynamics. The dashed lines present the evolution of the turbulent enstrophy (i.e. the one computed from the small-scale [turbulent] velocity field, $\delta\boldsymbol{\omega} \equiv \nabla \times \delta\mathbf{v}$). While displaying a similar evolution to the corresponding quantity from the total velocity field, we find that the turbulent enstrophy contributes around $\sim$1/3 to the total cluster-integrated enstrophy budget.

The source terms shown in the second panel correspond to the integrated, unfiltered enstrophy inside the virial volume (solid line in the upper panel). There are several prominent reasons for the sum of the source terms not to add up to the gradient of the enstrophy curve. First, even though we have not quantified it, numerical dissipation causes small vortices at the end of the turbulent cascade to fade, effectively acting as a sink term. Second, the source terms are computed on the sparsely saved simulation outputs (roughly, each $\sim$ 300 Myr), which does not allow to properly integrate them for the smallest scales (say, below $\sim$100 kpc, c.f. Fig. 3). This same effect is, indeed, seen in e.g. Porter et al. (2015), Vazza et al. (2017), and also in Wittor et al. (2017), who investigate the Lagrangian evolution of enstrophy using passively advected tracer particles in post-processing.[11] Nevertheless, the values at the sparsely saved snapshots can still be used to interpret the relative importance of the different mechanisms.

The pseudo-evolution term (purple line) is generally small, compared to the rest of sources, implying that our definition of a Lagrangian region for quantifying enstrophy is not a bad choice and the changes in its value are dominated by cosmic flows and inner dynamics. Cosmic expansion, as well, is barely important when compared to the rest of mechanisms. Its effect is only marginally relevant ($\lesssim$ 10 per cent) at higher redshifts, when $H(t)$ is larger. The primary mechanism of enhancement of cluster's enstrophy appears to be compression (yellow line; positive and thus $\langle\nabla\cdot\mathbf{v}\rangle < 0$), which especially dominates during (major) merger epochs. The integral of the advection term (red line) corresponds to the net flux of enstrophy. This quantity is positive during the mergers, as the cluster is feeding from the already stirred ICM of the infalling clusters. As the accretion rates decrease, this term becomes irrelevant, slightly negative (as the gas is advected from lower density regions, where the enstrophy is lower). These terms, however, imply redistribution, but not properly speaking *sources* of vorticity. Vortex stretching and baroclinicity are persistent and non-negligible during the whole cluster's

[11] When comparing our results to Wittor et al. (2017) note, however, that their results present much narrower and taller peaks, as they focus on different families of tracer particles mostly associated with gas clumps, while we focus on cluster-wide quantities.

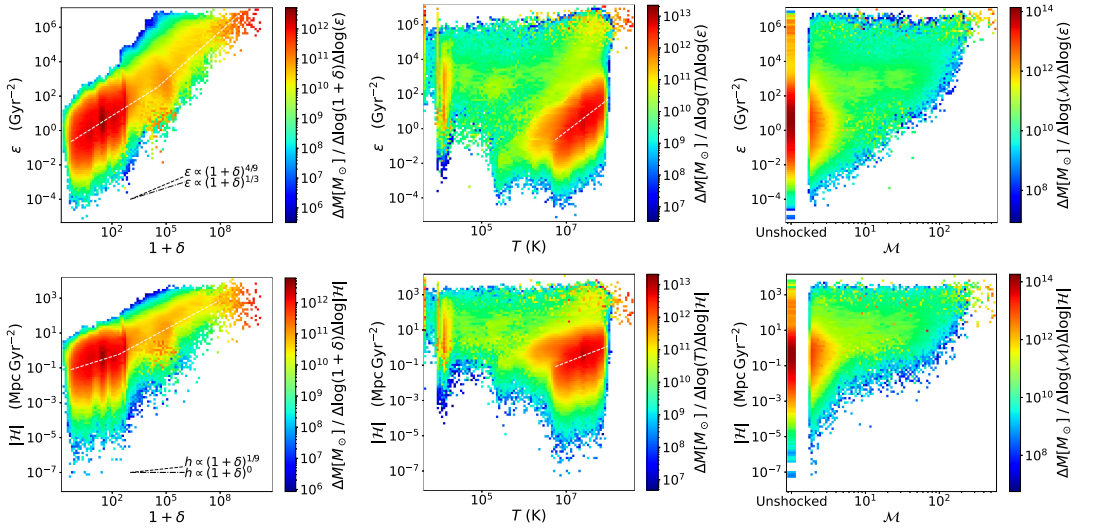


Figure 7. Phase-space density maps, for CL01 at $z \simeq 0$, of enstrophy (top row, ϵ) and helicity magnitude (bottom row, $|\mathcal{H}|$) versus overdensity (left-hand column, $1+\delta \equiv \rho/\rho_B$), temperature (central column, T) and Mach number (right-hand column, $\mathcal{M}$). All fluid elements inside $2R_{\rm vir}$ are considered for producing the phase map. The colour encodes the phase-space density (the hotter the colour is, the larger the phase-space density). The black lines in the left column represent the expected slope if there is no physical correlation between the variables, and the only effect is the resolution one, assuming a Kolmogorov (dashed) and a Burgers (dash–dotted) spectrum. The white lines correspond to broken power-law fits (left-hand column) and simple power-law fits (middle column). The Mach number axis scale is cut at $\mathcal{M} \simeq 1.3$ and cells below this threshold are regarded as 'unshocked'.

history. While the former seems to be enhanced after the mergers, we do not see a clear trend for the later. We will come back to these sources in Section 5.2, when examining their spatial distribution.

As for helicity, because of not being a positive-definite quantity, contrary to enstrophy, and being a pseudo-scalar, it should average to zero as long as the cluster is spherically symmetric (and, thus, invariant under reflections). Therefore, interpreting this quantity can be slightly more subtle. As we will see in Section 5.1, the role of helicity can be better disentangled locally. The global evolution of its volume-integrated value (third panel in Fig. 6) reveals how mergers, again, induce strong variations in helicity. The fourth panel in Fig. 6 shows the helicity source terms evolution. Advection dominates the helicity variations during the mergers period, as the cluster accretes the ICM of the merged clusters which may already have developed some degree of helicity. Once discussed this more 'geometric' term, baroclinicity and vortex stretching remain as the primary sources of helicity generation, while the acceleration of pre-existing vortices due to gravity and pressure gradients is slightly smaller in magnitude. Nevertheless, the evolution is fairly more complex than what we have seen for enstrophy. As covered for the enstrophy evolution, cosmic expansion and pseudo-evolution have a generally small contribution. Note how the helicity corresponding to the turbulent velocity field averages to zero over the cluster's volume. That behaviour is expected since, once the bulk flows (e.g. almost laminar infall of gas in the outskirts) have been removed, the residual velocity field is much more isotropic.

4.3.2 Phase maps

Finally, we explore the phase-space distribution of enstrophy and helicity against several thermodynamical quantities of the gas (namely, density and temperature; from which the scaling with other quantities, e.g. pressure or entropy, can be directly derived), as well as Mach number to elucidate their relation to shocks. We shall as well restrict the presentation of our results to the cluster CL01 at the most recent snapshot, at redshift $z \simeq 0$. For this aim, we present in Fig. 7 the phase-space density maps for the aforementioned variables. In all cases, we have considered all the volume elements lying inside $2R_{\rm vir}$, at the best resolution available at each point.

The left-hand panels in Fig. 7 represent the enstrophy-overdensity (upper panel) and helicity-overdensity (lower panel) phase-space densities. The general trend of both maps shows a tendency for denser gas to be more vortical (higher enstrophy) and helical (higher helicity). A relevant question in this scope, however, is whether this trend is physical or only due to a resolution effect (i.e. denser regions get resolved by smaller cell sizes and smaller vortices can develop, thus enhancing the vorticity magnitude). This effect can be ruled out by a simple scaling argument. If the velocity field scales according to a Kolmogorov (1941) spectrum, then one expects the vorticity to depend on the scale L as $\omega_L \sim L^{-2/3}$. On the other hand, the pseudo-Lagrangian refinement scheme provides $(1+\delta) \sim L^{-3}$. Therefore, the dependence of vorticity on density due to pure resolution effects ought to be $\omega \sim (1+\delta)^{2/9}$. Bringing this result to the enstrophy and helicity definitions, the expected scaling is $\epsilon \sim (1+\delta)^{4/9}$ and $\mathcal{H} \sim (1+\delta)^{1/9}$. If we assumed a steeper, Burgers (1939) spectrum, in which $\omega_L \sim L^{-1/2}$, then we shall expect shallower dependencies, $\epsilon \sim (1+\delta)^{1/3}$ and $\mathcal{H} \sim (1+\delta)^0$. As ICM turbulence typically lies in between these regimes (e.g. Miniati 2014; Vazza et al. 2017), if the dependency on density could be explained by resolution alone, then ϵ and $\mathcal{H}$ are expected to scale with logarithmic slopes between these two predictions.

The bulk of ICM's mass sits in the overdensity range $1+\delta \lesssim 10^3$–10^4. In this range, the logarithmic slope of the $\epsilon-(1+\delta)$ phase-space

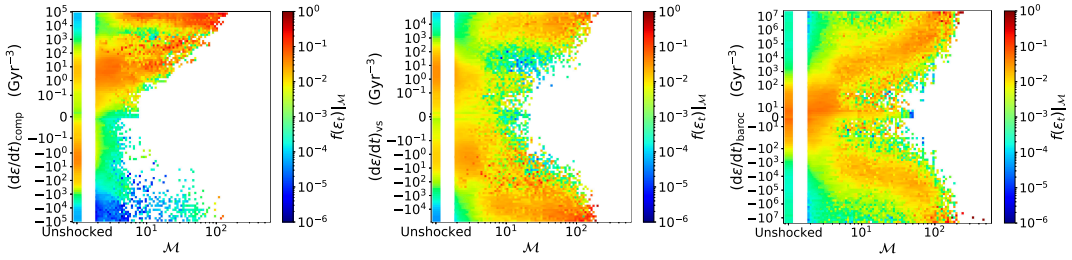


Figure 8. Conditional distributions, for CL01 at $z \simeq 0$, of the compressive (left-hand panel), vortex stretching (centre panel), and baroclinic (right-hand panel) enstrophy source terms strength as a function of the shock Mach number. That is, each column expresses the probability distribution of the values of the source terms, for all cells within $2R_{\rm vir}$ within a given $\mathcal{M}$ range. Mach number axis scale is cut at $\mathcal{M} \simeq 1.3$ and cells below this threshold are regarded as 'unshocked'.

density map is much shallower than through the whole overdensity range, but still above the expectation by the scaling reasoning due to resolution. We have fitted the mass-weighted mean $\epsilon - (1 + \delta)$ relation to a broken power law, which yields a logarithmic slope 0.658 ± 0.038 for $1 + \delta \lesssim 10^4$ (see white dashed lines in the figure), thus pointing at a genuinely physical effect beyond resolution. A similar pattern is observed for helicity magnitude, whose observed slope for the bulk ICM (0.329 ± 0.059) is inconsistent with the resolution effect alone. This result is not surprising, since denser gas (usually located towards smaller clustercentric radii, c.f. Fig. 4) varies on smaller spatial and temporal scales.

When relating our two measures of solenoidal turbulence, ϵ and $|\mathcal{H}|$, to temperature (central column in Fig. 7), the resolution effect becomes less relevant. This is mainly due to the fact that the radial dependence of temperatures on the ICM is much weaker than that of densities (see e.g. the profiles in Planelles & Quilis 2009). There is a strong tendency for most of the cluster gas mass to be increasingly vortical for higher temperatures. In the region where most of the cluster's mass is concentrated, enstrophy follows a steep power law, with slope 1.884 ± 0.065. A similar behaviour is seen in the $|\mathcal{H}| - T$ map, with roughly half the slope (as $|\mathcal{H}| \propto \omega$ and $\epsilon \propto \omega^2$). However, if the whole phase-density map is considered, the low-helicity end presents a drop at $T \gtrsim 10^6$ K, which is not clearly reflected in enstrophy. This shows that helical motions are partially suppressed for hotter gas. This effect can be much more easily interpreted in terms of the spatial distribution (Section 5.1), since hotter gas tends to reside, preferentially, in more central cluster regions (i.e. inside $R_{\rm vir}$).

The right column of Fig. 7 corresponds to the phase-space density of enstrophy and helicity magnitude with respect to the Mach number of shocked cells. Unshocked (and weakly shocked, $\mathcal{M} \lesssim 5$) cells have a mass-weighted enstrophy distribution which peaks around $\epsilon \sim 1$–$10\,{\rm Gyr}^{-2}$, in agreement with the values displayed in Fig. 6 at recent redshifts. However, while most of the cluster mass resides in unshocked and weakly shocked regions, as strong shocks are only dominant in low-density environments (as the accretion shocks in the outskirts), we find these regions to have high levels of enstrophy, in spite of being resolved with lower spatial resolution. This would, indeed, suggest that we are capturing turbulent motions in the cluster outskirts, despite not being implementing any ad hoc refinement scheme (cf. Vazza et al. 2009b; Iapichino et al. 2017). The fundamental difference, which allows us to recover a high level of vorticity even in the outskirts, is the base grid resolution (which, in turn, determines the 'mass resolution' of the pseudo-Lagrangian AMR approach). While the base grid cells in Iapichino et al. (2017) are $2\,{\rm h}^{-1}$Mpc on each side, ours are ~ 310 kpc. Therefore, this scheme is able to resolve with considerably higher resolution these outer regions without the need of additional refinement criteria.

In order to elucidate and disentangle the role of the main enstrophy source terms, we present in Fig. 8 the conditional distributions of the source terms as a function of the Mach number, which we have denoted $f(\epsilon_t)|_{\mathcal{M}}$ in the figures. We have focused on peculiar expansion/compression (left-hand panel), vortex stretching (middle panel), and baroclinicity (right-hand panel), which, as it has been seen in Section 4.2.1, are the dominant sources of cluster-wide enstrophy variation. We show conditional distributions, instead of phase maps as done above, since it better serves our purpose of analysing the trends of the enstrophy production/dissipation rates as a function of the shock's Mach number.

The compressive component is perhaps the simplest one to analyse. While unshocked cells have a relatively symmetric distribution of $\left(\frac{\partial \epsilon}{\partial t}\right)_{\rm comp}$, with a subtle preference towards negative values (expansion), shocked cells are, unsurprisingly, dominated by positive values of the source term, since these cells have $\nabla \cdot \mathbf{v} < 0$. The extremely high values of $\left(\frac{\partial \epsilon}{\partial t}\right)_{\rm comp}$ for strong shocks ($\mathcal{M} \gtrsim 10$), which are mostly located in the outskirts of the cluster, highlight the thermodynamically irreversible compression happening at the external, accretion shocks as a primary mechanism for enstrophy generation.

The distribution of values of the vortex stretching enstrophy source mechanism presents a clear distinction between unshocked and weakly shocked gas ($\mathcal{M} \lesssim 5$) and strong shocks ($\mathcal{M} \gtrsim 5$), and is fairly symmetrical under a change of sign, only with a subtle tendency towards more negative values in strong shocks. As it will be seen in Section 5.2, this term is typically negative inside of the external shocks.

The baroclinic term also presents a more or less symmetric distribution with respect to its sign, being slightly more positive (for all $\mathcal{M}$). Its magnitude clearly increases with the shock Mach number, although in this case the trend is much more continuous than for the vortex stretching term. Especially in the strongest shocks, baroclinicity is very efficient in inducing changes in vorticity, with a net effect of increasing the enstrophy. This is expected since, for strong enough shocks, the Rankine–Hugoniot jump conditions (see e.g. Landau & Lifshitz 1987) imply a mild density gradient but a steep pressure gradient. Therefore, the pressure gradient will be strongly aligned with the shock normal direction, while the density gradient can more easily have a non-negligible component tangential to the shock.

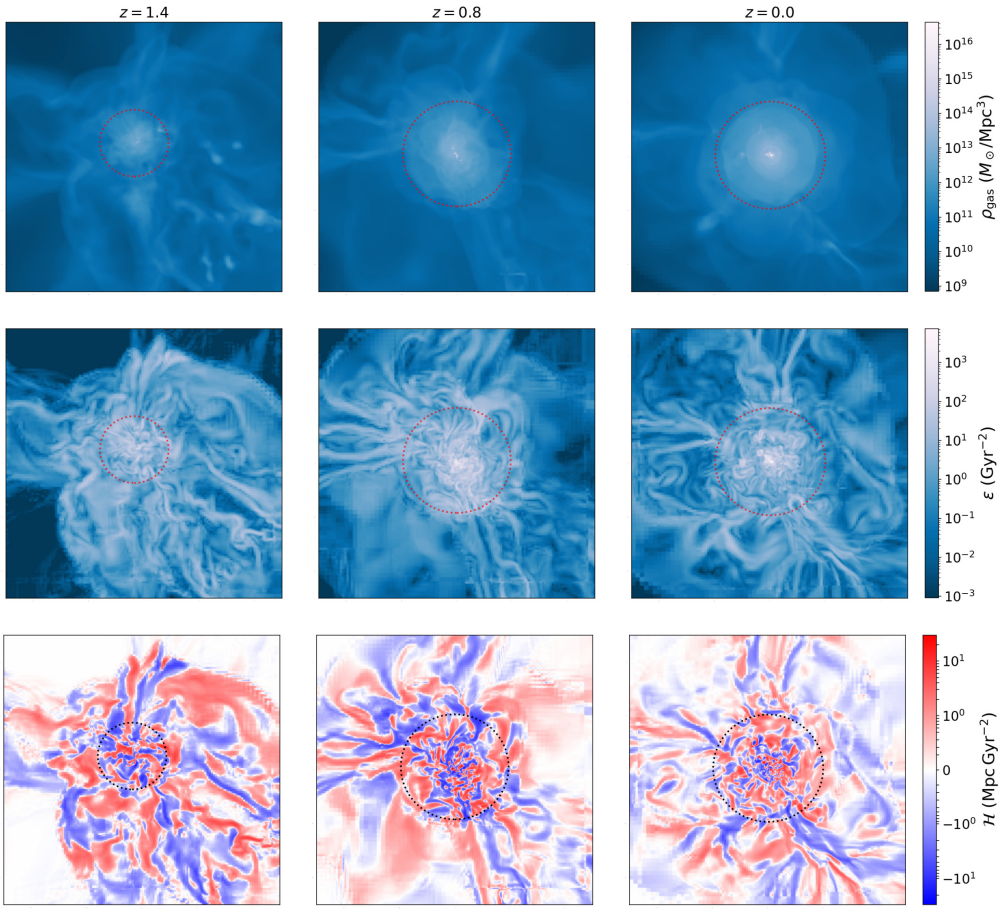


Figure 9. Maps of gas density (top row), enstrophy (middle row), and helicity (bottom row) of a thin ($\sim$ 150 kpc comoving) slice through cluster CL01 at three different redshifts: $z \simeq 1.45$ (left-hand panels) and $z \simeq 0.80$ (middle panels), while the cluster is undergoing strong accretion associated with mergers, and $z \simeq 0$, while the cluster is relaxed with negligible accretion rates. Each panel represents the same comoving region, 10 Mpc (comoving) on each side.

5 LOCAL DESCRIPTION OF TURBULENCE

Even though global statistics are informative of the general state of the cluster and are certainly useful when extending these analysis to large samples of clusters in order to attain statistically meaningful conclusions, some features associated with turbulence are much better captured by looking at their local distribution. In this section, we especially focus on solenoidal turbulence, which as we have seen is the dominant component. In particular, we analyse the distribution and relations amongst (filtered and unfiltered) enstrophy and helicity (Section 5.1), and the enstrophy source terms (Section 5.2).

5.1 Distribution of enstrophy and helicity

In Section 4.3, it has been seen that denser and hotter gas tends to be more vortical and helical. To complement this, we now study the spatial distribution of these two magnitudes inside and around the cluster. In Fig. 9, we show thin slices through the centre of the cluster of gas density (upper panels), enstrophy (middle panels), and helicity (lower panels); and for three redshifts: at $z \simeq 1.45$, 0.80 (both just after major mergers), and $z \simeq 0$ (during a smooth accretion phase), respectively, in the columns left to right. In all plots, we indicate the virial volume with a dotted circle.

At high redshift, the cluster, which has a noticeably smaller virial radius (1.36 Mpc) and around a quarter of its total mass at $z = 0$, is undergoing multiple mergers, resulting in an aspherical shape. The strongest external shocks, located at more than $2R_{\rm vir}$ in most directions, are clearly visible as density discontinuities, and so are the filaments, which are a dominant contribution to the accretion flows. The main merger has been produced by the infall of a structure from the right of the represented slice, where the cluster atmosphere is more extended. Complex gas structures appear in this region due to the stripping of the merged cluster's gas. A large part of this gas, which has ended up lying beyond $R_{\rm vir}$, will be reaccreted in the subsequent ~Gyr, mantaining the MARs high (Vallés-Pérez et al. 2020).

At this stage, enstrophy is clearly bounded by the external shocks, especially in those regions where there have been no mergers by (left and top borders of the shown slices). This suggests that

the baroclinic mechanism in the accretion shocks drives the first generation of vorticity to the hitherto pristine gas (since it is the only term in equation 9 whose magnitude is not proportional to the vorticity magnitude), closely followed by strong compression which further enhances its magnitude. Filaments also present high levels of enstrophy, which most likely is the main contribution to the advection term. However, correctly accounting for and interpreting the contribution of filamentary accretion of gas may be complex and we leave it for future work.

Helicity naturally follows a similar pattern. Being defined as a pseudoscalar quantity, it presents ubiquitous changes of sign (especially around velocity discontinuities), but it is however an interesting quantity to explore the infall of matter to the cluster. Note that the filaments (better discernible in the density map, towards the left and top borders of the slice) are considerably helical: while enstrophy spans several orders of magnitude from the core to the filaments, filaments have only ∼1/10 helicity magnitudes, compared to these of the core. We will explore this effect in more detail below, when showing the normalized radial profiles.

After the second relevant major merger sequence has taken place, at $z \simeq 0.8$ (central column), the cluster has considerably increased its virial radius after reaccreting the dark matter particles and baryons which ended up beyond the spherical overdensity boundary, and the shape has become slightly more spherical. The main merged substructure has infallen from the bottom-right direction in the represented slice. In that direction, there is a clear gradient of enstrophy magnitude: while the gas near to the resulting cluster's virial boundary has considerably high levels of enstrophy with a complex spatial distribution, enstrophy is greatly reduced behind the shock. This can be interpreted as an evidence of the predominance of the compressive source term during mergers: enstrophy is being generated in the head-on collision of the two ICMs, primarily by the weak shocks arising during the merger. Besides the merger region, gas is also being intensely accreted through most of the cluster boundary. Note, for example, in the top-left of the slice, how the magnitude of helicity is enhanced towards the outskirts, where gas is infalling more laminarly. The complex patterns in helicity inside the virial volume are, at least partially, due to the presence of internal shocks and shear motions, which produce the frequent sign changes.

At $z \simeq 0$, no important mergers have occurred recently and the ICM has relaxed to a nearly spherical shape, with extended accretion shock radii as well. A series of inner shocks, propagating outwards, are reflexed as density discontinuities (see also the compressive and solenoidal velocity magnitude maps in Vallés-Pérez et al. 2021). Solenoidal motions, showing eddies on a wide range of scales, dominate in the volume bounded by the outermost shocks, not only in the central regions of the cluster, but also near the shocks and around filaments. Helicity is especially enhanced in these filaments, while its magnitude is less relevant elsewhere (due to the absence of bulk flows in the collapsed regions).

Complementarily, in Fig. 10 we present radial profiles of enstrophy (red lines) and helicity magnitude (blue lines), in solid (dashed) lines for the unfiltered (filtered) velocity fields. The small-scale (or turbulent) enstrophy accounts for ∼ 50 per cent of the total enstrophy at all radii, only with a subtle decrease towards larger radii (from ∼ 60 per cent in the core to ∼ 40 per cent in the outskirts), which may well be explained by the decreased resolution in the outskirts: as the flows in these regions are typically resolved with larger cells, the multiscale filtering algorithm (Section 2.2) may converge to an outer-scale length, $L(\mathbf{x})$, a few times the cell side length, reducing the magnitude of the turbulent velocity field. However, the magnitudes of the small-scale helicity are only 10–30 per cent those

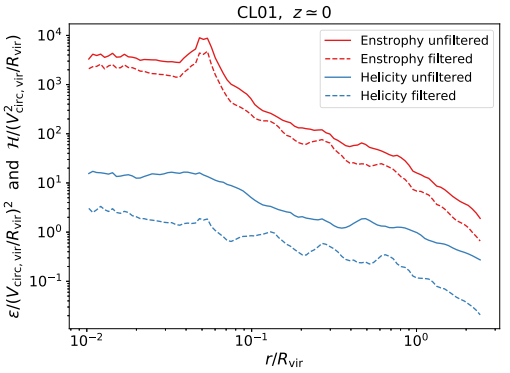


Figure 10. Radial profiles of enstrophy (red) and helicity magnitude (blue) for the cluster CL01 at $z \simeq 0$. The solid (dashed) lines correspond to the unfiltered (filtered) velocity field. The quantities are normalized in units of $R_{\rm vir}$ and $V_{\rm circ,vir} \equiv \sqrt{GM_{\rm tot,vir}/R_{\rm vir}}$.

of the unfiltered magnitude. Therefore, as suggested in Section 4.3, helicity is primarily contributed by the eddies being developed within (initially) laminarly infalling flows.

The comparison of the enstrophy and helicity magnitudes, when normalized to the cluster-wide units (i.e. we take $R_{\rm vir}$ and $V_{\rm circ,\,vir}$ as the length and velocity units), clearly reflects that helical motions are much more pervasive towards the cluster outskirts: the ratio between $\frac{\mathcal{H}}{V^2_{\rm circ,vir}/R_{\rm vir}}$ and $\frac{\epsilon}{(V_{\rm circ,vir}/R_{\rm vir})^2}$ grows with radius, from below 1 per cent in the cluster core to ∼ 30 per cent in the off-virial region, where laminar motions due to the accretion of gas are dominant.

5.2 Distribution of the enstrophy source terms

Complementarily to the study of the cluster-wide trends of the enstrophy source terms (Section 4.3.1), the examination of their spatial distribution can help in distinguishing their roles in the evolution of enstrophy. Focusing on CL01, Fig. 11 presents the spatial distribution of the compressive (top panels), vortex stretching (middle panels), and baroclinic (bottom panels) enstrophy source terms for the same comoving region as in Fig. 9. We have focused on the snapshots at $z \simeq 0.8$ (left-hand column) and $z \simeq 0$ (right-hand column). In the maps, we have overplotted Mach number $\mathcal{M} \geq 2$ contours.

The compressive mechanism is responsible for the most part of enstrophy generation in the outskirts (beyond $R_{\rm vir}$), where the other source terms are smaller. This is due to the presence of strong accretion shocks (with large, negative $\nabla \cdot \mathbf{v}$). At more recent redshifts, the relaxed dynamical state of the cluster and its nearly spherical shape evidence an onion-like structure of enstrophy sources and sinks around the cluster. A (Lagrangian) gas element radially infalling to the cluster would experience a succession of compressions (at the accretion and merger shocks) and rarefactions, producing enstrophy increases and decreases, respectively. Since shocks are a dissipative phenomenon, this mechanism causes a net, thermodynamically irreversible generation of enstrophy (see also the Lagrangian study of Wittor et al. 2017).

The main source of enstrophy in the unshocked regions, however, is the vortex stretching mechanism; while its magnitude is most typically negative inside the shock contours. This, together with the fact that this term is mostly positive for unshocked cells (see

Fig. 8 in Section 4.3.2), portrays how the stretching of vortices, primarily by the accretion flows, channels the enstrophy generated in shocks to inner regions in the cluster. This fact is also supported by the correlation between the magnitude of the compressive term in shocks and that of the vortex stretching term in the immediate post-shock region. For instance, at $z \simeq 0$ the strongest compression in the external shock seems to occur in the bottom-left direction of the plot; correspondingly, the strongest vortex stretching downstream of that shock also occurs in the same direction.

Last, the baroclinic term presents complex features and variations on short spatial scales, especially towards the cluster core. Nevertheless, quite generally its value is positive in (and downstream of) the strongest shocks, in line with the behaviour observed in Section 4.3.2. Its magnitude inside the virial volume of the cluster tends to be smaller than the vortex stretching and the compressive terms, especially at $z \simeq 0$.

6 DISCUSSION

Finally, we further discuss a number of additional issues, connected to the results shown in Sections 4 and 5 in relation to several models suggested in the recent literature.

In Section 4.2, we have focused on the velocity fluctuations over particular comoving lengths and their evolution with cosmic time. We have found the velocity fluctuations on fixed scales to be primarily determined by the presence/absence of merger events, especially major mergers, with the accretion rates determining to a fairly good extent the behaviour of these velocity fluctuations. Recently, Shi et al. (2018) proposed a two-phase mechanism for turbulence decay after a major merger, consisting of a first phase of fast-decay, which lasts for ~ 1 Gyr, and a subsequent phase of secular decay. A similar effect is seen in our velocity fluctuation evolution graph (Fig. 5). Their conclusions, as well as ours, are based on the study of a single galaxy cluster, so investigating it in large samples of clusters, with different merger configurations and accretion histories would definitely help to elucidate its physical origin. Therefore, although we indeed detect a similar behaviour, we leave this research direction for future work.

By studying the evolution of enstrophy sources through cosmic time (Section 4.3.1), we have found the primary, cluster-wide mechanisms for enstrophy increase to be peculiar compression, vortex stretching, and baroclinicity, while the rest of mechanisms tend to have more modest contributions, in general terms. However, the cluster-averaged values of the source terms do not represent the full physical picture. In order to better discriminate the conditions under which each of the mechanisms operates, we have followed two paths.

First, we have explored the phase-space distribution of the main enstrophy source terms, mentioned above, with the shock Mach number. These three source terms have clearly distinctive behaviour with $\mathcal{M}$, which already gives information about where they do take place. Naturally, the larger the Mach number is, the more efficient the compressive mechanism is in enhancing the local enstrophy. On the other hand, vortex stretching most usually removes enstrophy from strong shocks, and increases it in unshocked regions. Last, we have seen that baroclinicity most often is a positive source of enstrophy, and its magnitude increases continuously with $\mathcal{M}$. Complementarily, we have also looked at the spatial distribution of these enstrophy source terms (Section 5.2).

Our results seem to add evidence supporting the mechanism for generation of enstrophy (thus, solenoidal turbulence) proposed by Vazza et al. (2017), and further backed by the studies of Wittor et al. (2017) using Lagrangian passive tracer particles. This mechanism,

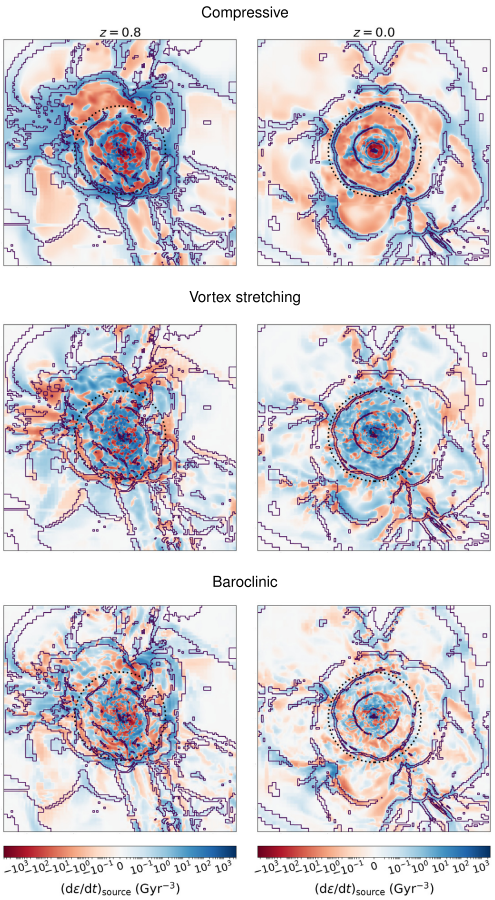


Figure 11. Maps of the compressive (top row), vortex stretching (middle row), and baroclinic (bottom row) enstrophy source terms, for $z \simeq 0.8$ (during a major merger; left-hand panels) and for $z \simeq 0$ (smoothly accreting; right-hand panels). The slices represent the same region as those in Fig. 9. The contours indicate the shocked regions ($\mathcal{M} > 2$).

which especially applies to smooth accretion, can be summarized in two stages:

(i) As low-density gas, with negligible amounts of enstrophy, infalls for the first time into a cluster, enstrophy is mainly generated at the outermost, accretion shocks by the compressive and the baroclinic mechanism. This stage can repeat at inner radii, if merger shocks are present.

(ii) Enstrophy is then accreted with the infalling gas. The dominant mechanism in this stage is vortex stretching, which transports the enstrophy downstream of shocks, where it had been generated.

We note that some level of baroclinicity is always necessary in the first place, since it is the only term in equation (9) capable of generating vorticity from gas with $\boldsymbol{\omega} = \mathbf{0}$. Note that at high redshift, in the linear regime and even in the mildly non-linear regime (as long as the Zel'Dovich 1970 approximation is applicable), the velocity field can be written as the gradient of a potential flow,

thus being identically irrotational. As a consequence, pristine gas from low-density regions suffers a first generation of enstrophy due to baroclinicity, most significantly in the outermost shocks, closely followed by an intense amplification due to compression.

While this two-stage mechanism describes well the generation of enstrophy by smooth accretion, the picture for mergers can be slightly different. First, the gas accreted from an infalling cluster has already some degree of enstrophy. Therefore, the advection mechanism is also important during these events (indeed, the red line in the second panel of Fig. 6 gets enhanced during mergers). Also, in major mergers the collision of the two ICMs generates an important amount of compression.

7 CONCLUSIONS

The intrinsically multiscale nature of turbulent phenomena poses challenging scientific and technical (numerical) problems. While a proper description of the turbulent cascade requires high-resolution simulations, a statistically meaningful analysis demands samples of (ideally) hundreds of clusters. Performing resimulations of large samples (e.g. Cui et al. 2018), or dedicated simulations including specific refinement criteria (e.g. Vazza et al. 2017) or phenomenological closures to model physical dissipation (e.g. Schmidt et al. 2016; Iapichino et al. 2017), is computationally expensive. In this work, we have instead explored the possibility (and the associated limitations) of studying turbulence in galaxy clusters directly extracted from cosmological simulations, neither performing resimulations nor adding phenomenological subgrid modelling of the turbulent cascade down to the dissipation scale (which introduces uncertainties due to the involved physics and the numerical scheme). With this aim, we have focused on two massive clusters from a moderate-sized cosmological simulation. Our main conclusions and findings are summarized below:

(i) Turbulent motions in, and around, galaxy clusters can be adequately captured by pseudo-Lagrangian AMR descriptions, relying upon the ansatz that less dense regions experience milder dynamics, provided that the mass resolution is fine enough. This assertion is key to allow the exploration of large samples of clusters at a reasonable computational cost.

(ii) The non-constancy of resolution causes several systematic effects to show up in some turbulence indicators, such as structure functions, which thereby limit the extent to which computing these global statistics can be useful in our case. These effects are due to the fact that (i) the smallest scales are only resolved at certain regions (thus introducing an 'environment' bias) and (ii) different AMR levels have different dissipation scales. To overcome these limitations, we have taken a complementary look, focusing on the velocity fluctuations on some particular, fixed comoving lengths (so as to avoid the effect of the previously mentioned systematics) and studying their evolution with cosmic time.

(iii) Velocity fluctuations at different scales, $\sqrt{S_2(L)}$, are primarily determined by the accretion rates, showing prominent peaks in their evolution after mergers (especially, major mergers). The injection scale, although difficult to constrain (due to its non-uniqueness) can be placed around $\lesssim R_{\rm vir}$, in consistence with previous findings. The peaks in $\sqrt{S_2(L)}$ at smaller scales occur with a delay roughly consistent with the cascade time-scale, $L/\sqrt{S_2(L)}$. The evolution of the velocity fluctuations after merger events provides suggestive evidence in favour of the two-phase mechanism for turbulence decay proposed by Shi et al. (2018), although larger statistics are required to properly investigate its origin.

(iv) Solenoidal (incompressible) turbulence is the dominant component in the ICM. Using enstrophy as a proxy for solenoidal turbulence, we have studied its generation mechanisms and their evolution with cosmic time, which is in turn intimately connected to the assembly history of the ICM. The evolution of the sources and their spatial distribution provide strong evidence supporting the two-stage mechanism for enstrophy generation suggested by Vazza et al. (2017), which combines baroclinicity and compression at shocks, with vortex stretching downstream of them. Nevertheless, we highlight that: (i) baroclinicity is responsible for the first generation of enstrophy, and (ii) while this scenario naturally accounts for enstrophy generation during smooth accretion, the situation might be notably different during mergers.

(v) Complementarily, we have introduced helicity, $\mathcal{H} \equiv \mathbf{v} \cdot \boldsymbol{\omega}$, as a quantity which highlights eddies being developed within bulk motions of the gas infalling into the cluster. While this magnitude is more difficult to interpret, due to not being positive defined (contrary to enstrophy), we have shown that:

> (a) Helical motions are irrelevant in cluster cores, in the sense that they contribute little to the total enstrophy budget, but much more pervasive towards the outskirts, associated with the fact that gas inside $\sim R_{\rm vir}$ presents milder bulk flows than gas in the outskirts by recent redshifts.
>
> (b) Helicity maps highlight the presence of cosmic filaments, suggesting that matter does not fall in straight, radial trajectories to cluster, but rather with complex geometries.
>
> (c) When computed from the small-scale, turbulent velocity field, helicity averages to zero through the cluster volume (as opposed to enstrophy). This reflects how helicity captures the interaction of bulk and turbulent velocity fields, and is suitable to investigate accretion flows.

We plan to thoroughly examine the role of helicity in cluster outskirts and filaments in future works.

As we have mentioned through the text, the particular details about the generation and evolution of ICM turbulence are extremely sensitive to the assembly history of the cluster, as well as to its environment. Thus, one of the main shortcomings of this work is the lack of statistics. In a forthcoming work, we will extend these analyses to a large sample of massive clusters, in order to extract statistically robust conclusions about the scaling relations of turbulence-related quantities with cluster mass, the differences between varying dynamical states, etc. A thorough and deep understanding of the description of turbulent phenomena in galaxy clusters is key to explain some observable effects, such as surface brightness fluctuations, acceleration of cosmic ray particles, amplification of magnetic fields, or non-thermal pressure leading to mass bias, etc., some of which can in turn be crucial for the usage of clusters for precision cosmology.

ACKNOWLEDGEMENTS

We thank the anonymous referee for his/her constructive comments, which have helped to improve the quality of this manuscript. This work has been supported by the Spanish Ministerio de Ciencia e Innovación (MICINN, grant PID2019-107427GB-C33) and by the Generalitat Valenciana (grant PROMETEO/2019/071). Simulations have been carried out using the supercomputer Lluís Vives at the Servei d'Informàtica of the Universitat de València. This research has made use of the following open-source packages: NUMPY (Oliphant

2006), SCIPY (Virtanen et al. 2020), MATPLOTLIB (Hunter 2007), and YT (Turk et al. 2011).

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Acosta-Pulido J. A. et al., 2015, preprint (arXiv:1506.03474)
Adam R. et al., 2017, A&A, 598, A115
Adrian R. J., Christensen K. T., Liu Z. C., 2000, Exp. Fluids, 29, 275
Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409
Angelinelli M., Vazza F., Giocoli C., Ettori S., Jones T. W., Brunetti G., Brüggen M., Eckert D., 2020, MNRAS, 495, 864
Beresnyak A., Miniati F., 2016, ApJ, 817, 127
Biffi V. et al., 2016, ApJ, 827, 112
Brunetti G., Lazarian A., 2011, MNRAS, 412, 817
Brunetti G., Vazza F., 2020, Phys. Rev. Lett., 124, 051101
Bryan G. L., Norman M. L., 1998, ApJ, 495, 80
Brzycki B., ZuHone J., 2019, ApJ, 883, 118
Burgers J., 1939, Akad. Wet. Amst. Addeel. Nat., 17, 1
Burns J. O., Hallman E. J., Gantner B., Motl P. M., Norman M. L., 2008, ApJ, 675, 1125
Cassano R., Brunetti G., 2005, MNRAS, 357, 1313
Chen H., Avestruz C., Kravtsov A. V., Lau E. T., Nagai D., 2019, MNRAS, 490, 2380
Cho J., 2014, ApJ, 797, 133
Churazov E. et al., 2012, MNRAS, 421, 1123
Cui W. et al., 2018, MNRAS, 480, 2898
Devroye L., 1986, Non-Uniform Random Variate Generation. Springer-Verlag, New York, NY, USA
Diemer B., Kravtsov A. V., 2014, ApJ, 789, 1
Dolag K., Vazza F., Brunetti G., Tormen G., 2005, MNRAS, 364, 753
Donnert J., Vazza F., Brüggen M., ZuHone J., 2018, Space Sci. Rev., 214, 122
Eckert D. et al., 2012, A&A, 541, A57
Eisenstein D. J., Hu W., 1998, ApJ, 496, 605
Faltenbacher A., Kravtsov A. V., Nagai D., Gottlöber S., 2005, MNRAS, 358, 139
Fujita Y., Takizawa M., Sarazin C. L., 2003, ApJ, 584, 190
Gaspari M., Churazov E., 2013, A&A, 559, A78
Gaspari M., Ruszkowski M., Sharma P., 2012, ApJ, 746, 94
Gaspari M. et al., 2018, ApJ, 854, 167
Gott J. R. I., Rees M. J., 1975, A&A, 45, 365
Haardt F., Madau P., 1996, ApJ, 461, 20
Hoffman Y., Ribak E., 1991, ApJ, 380, L5
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Iapichino L., Niemeyer J. C., 2008, MNRAS, 388, 1089
Iapichino L., Adamek J., Schmidt W., Niemeyer J. C., 2008, MNRAS, 388, 1079
Iapichino L., Federrath C., Klessen R. S., 2017, MNRAS, 469, 3641
Khatri R., Gaspari M., 2016, MNRAS, 463, 655
Kim W.-T., 2007, ApJ, 667, L5
Kitayama T. et al., 2014, preprint (arXiv:1412.1176)
Klaassen P. D. et al., 2020, in Marshall H. K., Spyromilio J., Usuda T., eds, Proc. SPIE Conf. Ser. Vol. 11445, Ground-based and Airborne Telescopes VIII. SPIE, Bellingham, p. 114452F
Knebe A. et al., 2011, MNRAS, 415, 2293
Kolmogorov A., 1941, Akademiia Nauk SSSR Doklady, 30, 301
Kravtsov A. V., Borgani S., 2012, ARA&A, 50, 353
Kravtsov A. V., Nagai D., Vikhlinin A. A., 2005, ApJ, 625, 588
Kretschmer M., Teyssier R., 2020, MNRAS, 492, 1385
Lacey C., Cole S., 1994, MNRAS, 271, 676
Landau L., Lifshitz E., 1987, Fluid Mechanics. Vol. 6. 2nd edn. Butterworth-Heinemann, Oxford
Lau E. T., Kravtsov A. V., Nagai D., 2009, ApJ, 705, 1129
Lau E. T., Nagai D., Avestruz C., Nelson K., Vikhlinin A., 2015, ApJ, 806, 68
Martin-Alvarez S., Planelles S., Quilis V., 2017, Ap&SS, 362, 91
Miniati F., 2014, ApJ, 782, 21
Miniati F., 2015, ApJ, 800, 60
Moffatt H. K., 2014, Proc. Natl. Acad. Sci., 111, 3663
Mohapatra R., Federrath C., Sharma P., 2020, MNRAS, 493, 5838
Mohapatra R., Federrath C., Sharma P., 2021, MNRAS, 500, 5072
Naab T., Ostriker J. P., 2017, ARA&A, 55, 59
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Nelson K., Lau E. T., Nagai D., Rudd D. H., Yu L., 2014a, ApJ, 782, 107
Nelson K., Lau E. T., Nagai D., 2014b, ApJ, 792, 25
Oliphant T. E., 2006, A Guide to NumPy. Vol. 1. Trelgol Publishing, USA
Planck Collaboration, 2020, A&A, 641, A6
Planelles S., Quilis V., 2009, MNRAS, 399, 410
Planelles S., Quilis V., 2010, A&A, 519, A94
Planelles S., Quilis V., 2013, MNRAS, 428, 1643
Planelles S., Borgani S., Fabjan D., Killedar M., Murante G., Granato G. L., Ragone-Figueroa C., Dolag K., 2014, MNRAS, 438, 195
Planelles S., Schleicher D. R. G., Bykov A. M., 2015, Space Sci. Rev., 188, 93
Planelles S. et al., 2017, MNRAS, 467, 3827
Planelles S., Mimica P., Quilis V., Cuesta-Martínez C., 2018, MNRAS, 476, 4629
Porter D. H., Jones T. W., Ryu D., 2015, ApJ, 810, 93
Press W. H., Schechter P., 1974, ApJ, 187, 425
Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical Recipes in FORTRAN. The Art of Scientific Computing, Cambridge University Press, Cambridge (UK)
Quilis V., 2004, MNRAS, 352, 1426
Quilis V., Ibáñez J. M., Sáez D., 1998, ApJ, 502, 518
Quilis V., Planelles S., Ricciardelli E., 2017, MNRAS, 469, 80
Rasia E. et al., 2015, ApJ, 813, L17
Ruszkowski M., Oh S. P., 2010, ApJ, 713, 1332
Ruszkowski M., Oh S. P., 2011, MNRAS, 414, 1493
Ryu D., Kang H., Hallman E., Jones T. W., 2003, ApJ, 593, 599
Savitzky A., Golay M. J. E., 1964, Anal. Chem., 36, 1627
Sayers J. et al., 2019, ApJ, 880, 45
Sayers J., Sereno M., Ettori S., Rasia E., Cui W., Golwala S., Umetsu K., Yepes G., 2021, preprint (arXiv:2102.06324)
Schmidt W., 2015, Living Rev. Comput. Astrophys., 1, 2
Schmidt W., Engels J. F., Niemeyer J. C., Almgren A. S., 2016, MNRAS, 459, 701
Shi X., Komatsu E., 2014, MNRAS, 442, 521
Shi X., Zhang C., 2019, MNRAS, 487, 1072
Shi X., Komatsu E., Nagai D., Lau E. T., 2016, MNRAS, 455, 2936
Shi X., Nagai D., Lau E. T., 2018, MNRAS, 481, 1075
Shi X., Nagai D., Aung H., Wetzel A., 2020, MNRAS, 495, 784
Simionescu A. et al., 2019, Space Sci. Rev., 215, 24
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Subramanian K., Shukurov A., Haugen N. E. L., 2006, MNRAS, 366, 1437
Turk M. J., Smith B. D., Oishi J. S., Skory S., Skillman S. W., Abel T., Norman M. L., 2011, ApJS, 192, 9
Valdarnini R., 2011, A&A, 526, A158
Valdarnini R., 2019, ApJ, 874, 42
Vallés-Pérez D., Planelles S., Quilis V., 2020, MNRAS, 499, 2303
Vallés-Pérez D., Planelles S., Quilis V., 2021, Comput. Phys. Commun., 263, 107892
Vazza F., Brunetti G., Gheller C., 2009a, MNRAS, 395, 1333
Vazza F., Brunetti G., Kritsuk A., Wagner R., Gheller C., Norman M., 2009b, A&A, 504, 33
Vazza F., Brunetti G., Gheller C., Brunino R., Brüggen M., 2011, A&A, 529, A17
Vazza F., Roediger E., Brüggen M., 2012, A&A, 544, A103
Vazza F., Jones T. W., Brüggen M., Brunetti G., Gheller C., Porter D., Ryu D., 2017, MNRAS, 464, 210

Vazza F., Brunetti G., Brüggen M., Bonafede A., 2018a, MNRAS, 474, 1672
Vazza F., Angelinelli M., Jones T. W., Eckert D., Brüggen M., Brunetti G., Gheller C., 2018b, MNRAS, 481, L120
Virtanen P. et al., 2020, Nat. Methods, 17, 261
Voit G. M., 2005, Rev. Mod. Phys., 77, 207
Webb G., 2018, Helicity in Fluids and MHD. Springer International Publishing, New York City, p. 21
Wittor D., Jones T., Vazza F., Brüggen M., 2017, MNRAS, 471, 3212
Yepes G., Kates R., Khokhlov A., Klypin A., 1997, MNRAS, 284, 235
Young D., 1954, Trans. Am. Math. Soc., 76, 92
Zel'Dovich Y. B., 1970, A&A, 500, 13
Zhang C., Churazov E., Dolag K., Forman W. R., Zhuravleva I., 2020, MNRAS, 494, 4539
Zhuravleva I., Churazov E., Kravtsov A., Lau E. T., Nagai D., Sunyaev R., 2013, MNRAS, 428, 3274
Zhuravleva I. et al., 2014a, Nature, 515, 85
Zhuravleva I. et al., 2014b, ApJ, 788, L13
ZuHone J. A., 2011, ApJ, 728, 54

APPENDIX A: VOLUME AVERAGES, RESOLUTION DEPENDENCE, AND CONVERGENCE TEST

Vorticity is not a scale-invariant magnitude, neither in the classical Kolmogorov theory ($\omega_L \propto L^{-2/3}$) nor in actual simulations of ICM turbulence (see e.g. Vazza et al. 2017). Thus, in order for any results obtained in either simulations or observations to be physically meaningful, it is of utmost importance to make sure that cluster-integrated vorticity (or any similar magnitude, such as enstrophy or helicity) is defined in such a way that it does not depend strongly on numerical resolution.

In particular, in this study we have considered the volume-averaged enstrophies and helicities. Note, however, that our refinement scheme consistently samples with increased resolution the overdense regions (where dynamics vary on shorter spatial scales). Thus, in this case,[12] the volume weight of these quantities naturally privileges the denser regions, since only in these regions smaller eddies will have been able to develop. Note therefore that performing a mass-weight on this 'mass-sampled' data would effectively correspond to performing a double-weight, and should not be a converging quantity as the resolution gets increased.

In order to check this is the case, we present in Fig. A1 the resolution dependence of several quantities (namely vorticity, enstrophy, helicity, and helicity magnitude) integrated in volume and in mass, i.e. $\frac{1}{V}\int_{\mathcal{V}} X \mathrm{d}V$ and $\frac{1}{M}\int_{\mathcal{V}} X \mathrm{d}M$, respectively. The solid lines

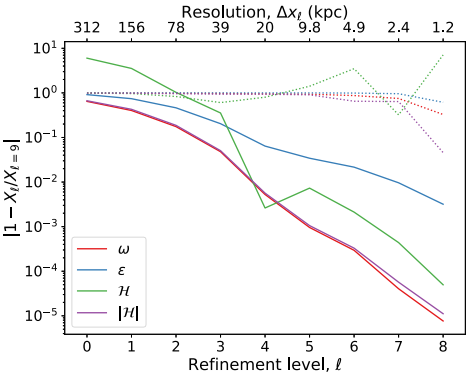


Figure A1. Resolution dependence of several cluster-averaged quantities, defined as a volume-weighted average (solid lines) or a mass-weighted average (dotted lines). The vertical axis represents the fractional variation of the quantity computed using data up to the ℓ-th refinement level with respect to the quantity computed using the maximum resolution data at each point.

show the fractional difference in each quantity, when performing the volume integral up to a given refinement level, with respect to its value at the maximum available resolution ($\ell = 9$). The dotted lines present the same quantity, for the mass-weighted averages. While the mass-averages do not converge when increasing the resolution (i.e. the 'errors', $|1 - X_\ell/X_{\ell=9}|$ are always of order unity), all volume-averaged quantities do converge rapidly. Therefore, their values do not depend strongly on the resolution of the simulation and are the physically meaningful quantities in this scope.

We find necessary to emphasize that this consideration is valid as long as we work with mass-sampled simulation data. Naturally, different refinement strategies (e.g. Vazza et al. 2012, who add a refinement criterion based on local vorticity) or fixed grids could rescue small-scale eddies in low-density regions, and thus enhance the volume-averaged vorticity. The only point of the reasoning above is thus to guarantee that, while we work with mass-sampled data (such as standard AMR or SPH), our results do not depend strongly on the resolution of a specific object and, thus, it is possible to compare different objects from the same or from different simulations.

[12]This discussion could in principle be applied to SPH data, as well.

This paper has been typeset from a TEX/LATEX file prepared by the author.

The Astrophysical Journal Letters, 920:L2 (8pp), 2021 October 10 https://doi.org/10.3847/2041-8213/ac2816
© 2021. The American Astronomical Society. All rights reserved.

Void Replenishment: How Voids Accrete Matter Over Cosmic History

David Vallés-Pérez[1], Vicent Quilis[1,2], and Susana Planelles[1,2]
[1] Departament d'Astronomia i Astrofísica, Universitat de València. Burjassot (València), E-46100, Spain; david.valles-perez@uv.es
[2] Observatori Astronòmic, Universitat de València. Paterna (València), E-46980, Spain
Received 2021 July 28; revised 2021 September 14; accepted 2021 September 16; published 2021 October 5

Abstract

Cosmic voids are underdense regions filling up most of the volume in the universe. They are expected to emerge in regions comprising negative initial density fluctuations, and subsequently expand as the matter around them collapses and forms walls, filaments, and clusters. We report results from the analysis of a cosmological simulation specially designed to accurately describe low-density regions, such as cosmic voids. Contrary to the common expectation, we find that voids also experience significant mass inflows over cosmic history. On average, 10% of the mass of voids in the sample at $z\sim 0$ is accreted from overdense regions, reaching values beyond 35% for a significant fraction of voids. More than half of the mass entering the voids lingers on periods of time $\sim$10 Gyr well inside them, reaching inner radii. This would imply that part of the gas lying inside voids at a given time proceeds from overdense regions (e.g., clusters or filaments), where it could have been preprocessed, thus challenging the scenario of galaxy formation in voids, and dissenting from the idea of them being pristine environments.

Unified Astronomy Thesaurus concepts: Large-scale structure of the universe (902); Voids (1779); Accretion (14); Galaxy environments (2029); Computational astronomy (293)

Supporting material: animation

1. Introduction

Cosmic voids are underdense regions filling up most of the volume in the universe (Zeldovich et al. 1982). According to the accepted paradigm of cosmological structure formation, they emerge in regions comprising negative initial density fluctuations (Sheth & van de Weygaert 2004), and subsequently expand as the matter around them collapses and forms walls, filaments, and clusters (see van de Weygaert & Platen 2011 and van de Weygaert 2016 for recent, general reviews). This leads to coherent outflows (van de Weygaert & van Kampen 1993; Padilla et al. 2005; Ceccarelli et al. 2006; Patiri et al. 2012), making them a pristine environment with notable applications for cosmology (Dekel & Rees 1994; Park & Lee 2007; Lavaux & Wandelt 2010, 2012; Bos et al. 2012; Pisani et al. 2019) and galaxy formation (Hahn et al. 2007; Kreckel et al. 2011; van de Weygaert & Platen 2011; Ricciardelli et al. 2014a).

The dynamics of cosmic voids are dominated by their expansion and consequent depletion of gas and dark matter (DM), as revealed by the coherent outflows found both in simulations (van de Weygaert & van Kampen 1993; Padilla et al. 2005; Ceccarelli et al. 2006) and observations (Bothun et al. 1992; Patiri et al. 2012; Paz et al. 2013), and also expected from analytical models of isolated voids (Bertschinger 1985; Sheth & van de Weygaert 2004; Baushev 2021). However, in a fully cosmological environment, it should be in principle possible to expect coherent streams of matter —gas and DM—to unbind from dense structures and end up penetrating inside low-density regions. As a matter of fact, a handful of scenarios for unbinding mass do exist, such as galaxy cluster mergers (Behroozi et al. 2013) or strong shocks that can extend up to a few virial radii (e.g., Zhang et al. 2020).

In this Letter, we explore this scenario with a Λ cold dark matter (ΛCDM) cosmological simulation of a large volume domain, especially designed to describe matter in and around voids. The rest of the Letter is organized as follows. In Section 2, we describe the simulation and the void finding algorithm. In Section 3, we present our results regarding the existence of mass inflows through voids' boundaries. Finally, we summarize the implications of these results in Section 4.

2. Methods

2.1. The Simulation

The results reported in this paper proceed from a cosmological simulation of a periodic domain, $100\,h^{-1}$ Mpc along each direction, produced with MASCLET (Quilis 2004), an Eulerian, adaptive mesh refinement hydrodynamics coupled to a particle-mesh N-body code. The Eulerian hydrodynamic scheme in MASCLET, based on high-resolution shock-capturing techniques, is capable of providing a faithful description of the gaseous component in low-density regions, such as cosmic voids.

Structures evolve on top of a flat ΛCDM cosmology consistent with the latest Planck Collaboration (2020) results. Dark energy, matter, and baryon densities are specified by $\Omega_\Lambda = 0.69$, $\Omega_m = 0.31$, and $\Omega_b = 0.048$, relative to the critical density $\rho_c = \frac{3H_0^2}{8\pi G}$. The Hubble parameter, $H_0 = 100\,h\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, is set by $h = 0.678$. The initial conditions were set up at $z = 100$, by evolving a power spectrum realization with spectral index $n_s = 0.96$ and normalization $\sigma_8 = 0.82$ using Zeldovich's (1970) approximation.

A low-resolution run on a grid of 128^3 cells was performed in order to identify the regions that would evolve into cosmic voids by $z \simeq 0$. Back to the initial conditions, the seeds of voids and their surroundings were sampled with higher numerical resolution according to the procedure introduced in Ricciardelli et al. (2013, hereafter RQP13). The regions at $z = 100$ comprising the DM particles that end up in zones with[3] $\rho/\rho_B < 10$ by $z = 0$ are thus mapped with a first level of mesh refinement ($\ell = 1$), with half the cell size and DM particles eight times lighter than those of the base grid; therefore, with

[3] $\rho_B \equiv \Omega_m \rho_c$ is the background matter density of the universe.

DM mass resolution $6.4 \times 10^9 M_\odot$ and spatial resolution 390 h^{-1} kpc. In order to capture the structures forming in cosmic voids, subsequent levels of refinement ($\ell \geqslant 2$, up to $n_\ell = 10$) are created following a pseudo-Lagrangian approach that refines cells where density has increased by a factor of 8 with respect to the previous, lower-resolution level. Besides gravity and hydrodynamics, the simulation includes standard cooling and heating mechanisms, and a phenomenological parameterization of star formation (Quilis et al. 2017).

2.2. The Void Finder

We have identified the sample of cosmic voids in our simulation with a void finder based on the one presented by RQP13, which looks for ellipsoidal voids using the total density field ($\rho_{\rm tot}$) and the gas velocity field ($\boldsymbol{v}$), as underdense ($\rho_{\rm tot} < \rho_{\rm B}$), peculiarly expanding ($\nabla \cdot \boldsymbol{v} > 0$) regions surrounded by steep density gradients. While the original void finder in RQP13 did not assume any prior on the void's shape, which could therefore develop highly complex, nonconvex, and non-simply connected shapes, such a precise definition of a void boundary is counterproductive for assessing mass fluxes in postprocessing, since it limits the validity of the pseudo-Lagrangian approach (see Section 2.3).

In order to have voids with smooth surfaces, our void finder looks for voids as ellipsoidal volumes around density minima, using the same thresholds on total density, total density gradient, and gas velocity divergence as RQP13. While voids are not generally ellipsoidal, by using the same threshold values as RQP13 we ensure that our algorithm looks for the largest possible ellipsoid inside actual, complex-shaped voids, thus providing a robust, stable, and conservative definition of these structures, which is readily comparable with their identification in observational data (Foster & Nelson 2009; Patiri et al. 2012). The voids are found and characterized one at a time, using the 128^3 base grid, in the steps summarized below.

Protovoid finding. A tentative center is chosen, as the most underdense, positive velocity divergence cell not yet inside an already found void. The initial protovoid is a 5^3 cells cube around this cell. The protovoid is then grown iteratively in the directions of its six faces, one extra cell along each direction at a time, if the following conditions are met by all the "new" cells:

$$\delta_{\rm tot} < 0, \quad \nabla\delta_{\rm tot} < (\nabla\delta)^{\rm max}, \quad \nabla \cdot \boldsymbol{v} > 0, \tag{1}$$

where $\delta_{\rm tot} \equiv \frac{\rho_{\rm tot}}{\rho_B} - 1$ is the total density contrast. The threshold on the density gradient is set to $(\nabla\delta)^{\rm max} = 0.25\ {\rm Mpc}^{-1}$, in consistency with RQP13. Once the protovoid has been determined, the center of the void is adjusted to the center of mass defect of the protovoid. A first approximation to the shape of the ellipsoid is determined by computing and diagonalizing the inertia tensor.

Growth of the ellipsoidal void. The initial ellipsoid is subsequently grown iteratively to find the maximal ellipsoid that fits inside the actual void. This is performed by repeatedly applying the following two substeps:

1. The shape of the ellipsoid is adjusted iteratively, in a similar manner to what is done in the galaxy cluster's literature (Zemp et al. 2011; Vallés-Pérez et al. 2020). In particular, the new eigenvalues of the inertia tensor yield the orientation of the void, while the new eigenvectors are used to compute the new semiaxes. These semiaxes are rescaled proportionally, so as to preserve the volume of the ellipsoid. The process is iterated, adjusting the integration volume, until convergence, which is assessed by the change in the semiaxes lengths.
2. Once the shape has been found, the ellipsoidal void is grown at constant shape, by multiplying each of its semiaxes by a factor of $1 + \chi$. We have fixed $\chi = 0.05$, although this parameter does not have a severe impact on the resulting void population while kept small.

This two-step iteration is repeated until either one of the stopping conditions in RQP13 (maximum density, density gradient, or negative velocity divergence) is met by a cell, or the mean slope of the total density field at the boundary exceeds the prediction of the universal density profile (Ricciardelli et al. 2013, 2014b). It can be easily shown that, assuming that the spherical profiles in RQP13 can be extended to ellipsoidal shells, this condition can be applied by requiring:

$$3\left(\frac{\rho(R)}{\rho(<R)} - 1\right) > 1.37 - 0.25z, \tag{2}$$

with $\rho(<R)$ the mean density inside the ellipsoid and $\rho(R)$ computed from the newly added cells in the growing step. The numerical coefficients, which are derived from the fit in RQP13, are valid for $2.5 \geqslant z \geqslant 0$.

Void sample and merger tree. In order to produce the final sample of voids, we start from the latest code output (at $z = 0$) and select an initial sample of voids taking care of the overlaps. To do so, we iterate through the voids found by the algorithm described above, from the largest and emptiest to the smallest and densest, and accept those that do not overlap more than 50% with the volume occupied by previously accepted voids.

Then, we trace this initial sample back in time by building their merger tree. To assess which is the best progenitor candidate for a void, we find the one that maximizes the *volume retention* defined as ${\rm VR} \equiv \frac{N_{A\cap B}}{\sqrt{N_A N_B}}$, with N referring to the number of cells, and A and B being some void and one of its parent candidates, respectively. This approach is equivalent to the *particle retention* defined by Minoguchi et al. (2021), but here applied to cells (Sutter et al. 2014). We did not find strong overall variations when using other figures of merit defined on Minoguchi et al. (2021), although there can be variations in a small number of individual voids.

With this procedure, we are able to obtain a sample of 207 voids, with equivalent radius at $z = 0$ larger than 5 Mpc (the largest of them reaching ~18 Mpc), which can be traced back, at least, to $z = 1.5$; and 179 of them are traced back down to $z = 2.5$. The overall statistics (radii, ellipticities, and mean overdensities) of the void sample are displayed in Figure A1.

2.3. The Pseudo-Lagrangian Approach and Its Validity

To assess mass fluxes in postprocessing, we take a pseudo-Lagrangian approach, by interpreting each volume element in the simulation (up to a certain refinement level) as a tracer particle, and advecting these tracer particles using the gas velocity field between each pair of code snapshots. This technique is analog to the one applied by Vallés-Pérez et al. (2020) for galaxy clusters.

In practical terms, at each code snapshot we take all (nonrefined and nonoverlapping) gas cells, each one at a

position $\boldsymbol{x}_\mathbf{n}$ and compute their updated position with an explicit, first-order step. For consistency, we have performed the same analysis with the dark matter particle distribution. Since voids experience mild dynamics with large dynamical times (associated to their low densities), this is a sensible approach. Then, in order to compute the accretion mass flux around a void between a pair of code outputs, we consider the mass of all the dark matter particles (or gas pseudo-particles) that were outside the void in the previous iteration, and inside the same volume at the latter (and vice versa for the decretion mass flux).

The previously discussed procedure will be valid as long as the timestep involved (the timespan between two consecutive outputs of the simulation) fulfills $\Delta t \lesssim \mathcal{L}/\mathcal{V}$, with $\mathcal{L}$ a characteristic scale of the surface of the void, and $\mathcal{V}$ a characteristic velocity for these cosmic flows. By choosing an ellipsoidal, instead of a complex, irregular shape, $\mathcal{L}$ can be taken of the order of the smallest semiaxis (several Mpc). On the other hand, by using the same volume for both consecutive iterations, we ensure that we are detecting the mass elements that are being dynamically accreted onto the void, and not accounting for the elements that may appear inside or outside the ellipsoid due to its change between iterations.

We have also taken a conservative approach for the sake of showing the robustness of our results. Since DM particles can be traced, we can compare the total mass accretion in a certain redshift interval, when computed using our pseudo-Lagrangian method, and when computed by tracing the actual evolution of the particles in the simulation, i.e., checking which particles were outside the volume at the first iteration of the interval, and are inside that same volume at the last iteration. The results, shown in Figure A2, confirm that (i) the pseudo-Lagrangian approach works well, on a statistical level, on the DM particle distribution, with reasonable scatter (0.3–0.5 dex) between the estimated and the actual accretion inflow; and (ii) on these low-density environments, gas and DM dynamics exhibit remarkably similar results, producing a scatter between DM and gas accretion rates typically below 0.2 dex. This serves as a confirmation of the applicability of our pseudo-Lagrangian approach for estimating gas accretion.

3. Results

A summary of the results of this analysis over the whole cluster sample is shown in Figure 1, where we also present the decretion rates for comparison. The raw fluxes (gas or DM mass, entering or leaving the void per unit time) are displayed in the top panel, where it can be seen that accretion[4] flows onto voids, while smaller in magnitude than the decretion flows typically by a factor of $1/6-1/3$, are present in the void sample in a statistical sense.

In the middle panel we have normalized these fluxes to the mass of the given material component (gas or DM) in the void at each time, to be read as the percentage of gas or DM void's mass that leaves or enters the void per gigayear. By performing this normalization, the robust mean values of the fluxes of DM and gas match each other, reflecting that both components undergo remarkably similar dynamics. This is expected, since gas in these regions has low temperature and pressure and thus behaves closer to a collisionless fluid (further validating the applicability of the pseudo-Lagrangian approach; see Section 2.3). While normalized decretion flows show little

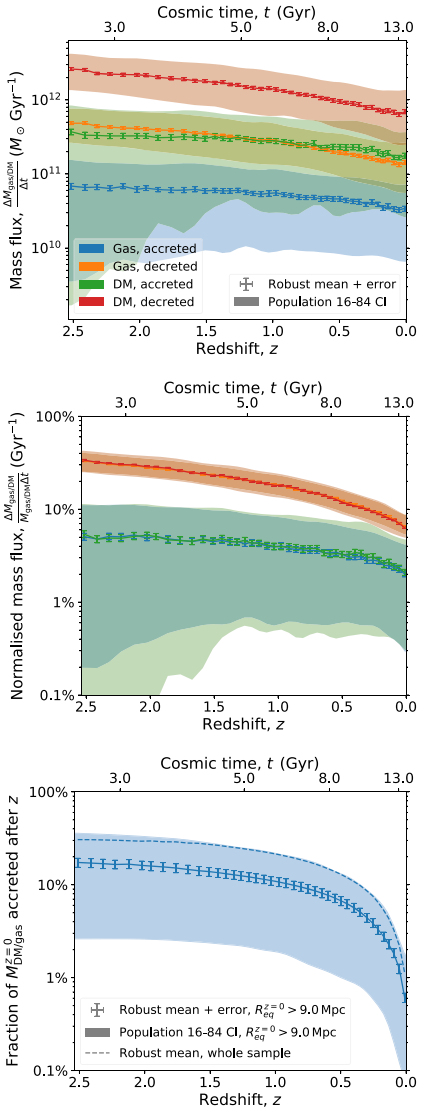


Figure 1. Evolution of the accretion and decretion rates in the void sample. Top panel: evolution of the mass fluxes (gas and DM mass, entering or leaving the void per unit time, according to the color legend) as a function of cosmic time. The shaded regions delineate the 16–84 percentiles (referred to as "confidence intervals" [CI] in the legend), as an indication of the scatter in these quantities, while the error bars represent the error of the mean value. Middle panel: mass fluxes normalized to the void's gas or DM mass, averaged over the cluster sample. Same legend as above. Bottom panel: anticummulative accreted gas mass fraction, i.e., the mass that has been accreted after a given redshift z as a fraction of the gas mass of the void at $z = 0$. Errors are given only for the $R_{\rm eq}^{z=0} > 9$ Mpc for clarity, being similar in magnitude for the whole sample.

[4] Although these mass inflows would not be triggered by the peculiar gravitational fields, but by the external, large-scale structure bulk and shear velocity flows, since they move inwards in the voids and remain within them for long times, we refer to them as accretion flows in analogy with mass flows in massive objects.

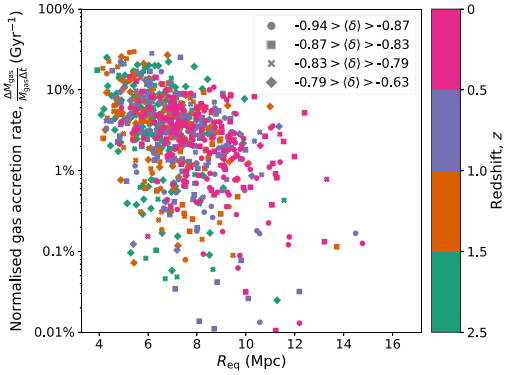


Figure 2. Normalized gas accretion rate, i.e., gas inflow per gigayear in units of the void's gas mass (vertical axis) as a function of the size of the void (equivalent radius; horizontal axis). The gas accretion rates are computed on four redshift intervals, encoded in the figure according to the color scale on the right. The shapes of the data points refer to the mean overdensity of the void, according to the binning specified in the legend.

scatter, and evolve from nearly 30% Gyr^{-1} at $z = 2.5$ to $\sim$6% Gyr^{-1} at $z \simeq 0$, accretion fluxes are smaller by a factor of $\sim 1/6$ at high redshift, but do not decrease as sharply and reach $\sim 1/3$ of the mean decretion values at $z = 0$. It might be argued that part of these accretion flows could be due to void-in-cloud processes (Sheth & van de Weygaert 2004; Sutter et al. 2014, i.e., voids collapsing in a larger-scale overdense environment), even though our sample building strategy, from $z = 0$ backwards in time, should exclude most of them since they would not have survived until $z = 0$. Nevertheless, we have checked that the same results hold when restricting the sample to large voids ($R_{\rm eq}^{z=0} > 9$ Mpc) only with a slight decrease in the accretion rates, less than a factor of 2, at low redshifts with respect to the whole sample. At high redshifts, there are no differences between large voids and the whole sample. Therefore, the accretion signal detected here cannot be ascribed to cloud-in-void processes alone. Smaller voids may experience stronger inflows, since they are more sensitive to external influences by a larger-scale velocity field.

Last, the bottom panel presents the accumulated accreted mass, as a function of the present-day void's gas or DM mass, from a redshift z up to $z = 0$. When focusing on the large-void subsample, on average, up to 17% of the void's current mass has been accreted after $z = 2.5$ (reaching beyond 35% at percentile 84), and the average void has suffered a mass inflow 10% of its current mass after $z = 1$. Interestingly, in their general analysis of the cosmic web, Cautun et al. (2014) found that $\sim$20% of the mass in voids at $z = 0$ belonged to walls and filaments at $z = 2$. Despite the similarity of the result, note that their interpretation is subtly different: while Cautun et al. (2014) ascribe this result to an artifact due to the difficulty of identifying tenuous structures within voids as they become emptier, our result corresponds to an actual inflow (matter initially outside the void, which crosses its boundary at a given time).

As the central result of this Letter, in Figure 2 we present the gas accretion rates (gas mass accreted, normalized by the voids' mass and per unit time) as a function of the void size (equivalent radius), computed on four redshift intervals (from $z = 2.5$ to $z = 1.5$, and three subsequent intervals with $\Delta z = 0.5$ thereon) that are encoded by the color scale in the figure. A significant fraction of voids, at any redshift, presents relevant accretion rates (above a few percent per gigayear, which are sufficient to impact their composition and dynamics). The highest accretion rates are seen in the smallest voids. As mentioned above, smaller voids are more prone to externally induced flows due to larger-scale influences: while $R_{\rm eq} \sim 5$ Mpc voids show mean values $\sim$8% Gyr^{-1}, this rate lowers to $\sim$(1–2)% Gyr^{-1} in the case of the largest voids. Nevertheless, large voids with exceptionally large accretion rates also exist, even at low redshifts. Naturally, the abundance of small voids makes it possible for some of them to show extreme accretion rate values. The point markers encode the mean overdensity, $\langle\delta\rangle \equiv \langle\rho\rangle/\rho_B - 1$ of the voids, according to the legend, to check whether there is any trend between this property and the accretion rates. All voids in the sample, and even the large and rapidly accreting ones, have very small overdensities, thus ruling out the fact that our results could be contaminated by a bad delineation of the void's wall. Indeed, as discussed in Section 2.2, our void identification technique has aimed to be conservative enough to exclude these possible effects.

To better visualize the effect, we show in Figure 3 the gas density field around a large void ($R_{\rm eq} \simeq 12$ Mpc) with high accretion rates. Blue and green contours correspond to isodensity surfaces of $\rho/\rho_{\rm B} \simeq 3$ and 0.5, roughly enclosing mean total densities $15\rho_{\rm B}$ and $4\rho_{\rm B}$, respectively, in order to give context of the distribution of matter around the void. The orange shadow highlights the void, with mean total density $\rho/\rho_{\rm B} \simeq 0.1$. On top, we overplot the accretion velocity field, that is, the velocity vectors in the region around the void where they point toward it. This representation clearly exemplifies the presence of coherent, large-scale streams of matter flowing toward cosmic voids from higher-density regions. Complementarily, in Figure A3 we show an animation of a gas density slice through a large ($R_{\rm eq} \simeq 14$ Mpc) void that undergoes significant accretion, with the velocity field overplotted with arrows. The animation exemplifies the nature of these inflows: the velocity field around the void consists of its own induced outflow (dominant through most of its boundary), plus the bulk and shear flows originated by the surrounding structures (see also Aragon-Calvo & Szalay 2013 for a detailed study of the hierarchical nature of velocity fields in and around voids).

Tracing the newly accreted DM particles in time, we find that nearly 50% remain inside the original volume of the void for up to $\sim$10 Gyr, and a significant fraction of them reach inner radii. The same behavior is expected for the gas, therefore granting the accreted gas, which may have been preprocessed outside the void, a long enough timespan to play a crucial role in the formation and evolution of void galaxies.

4. Conclusions

The findings reported in this Letter challenge the common accepted picture on the evolution of cosmic voids and could consequently have a direct potential impact on the understanding of galaxy formation and evolution in low-density environments. Hence, the uncontaminated and pristine void domains could be altered by the entrance of chemically and thermodynamically processed gas. Future effort should be devoted to confirm these results with other simulation codes and void identification strategies (e.g., those based on the

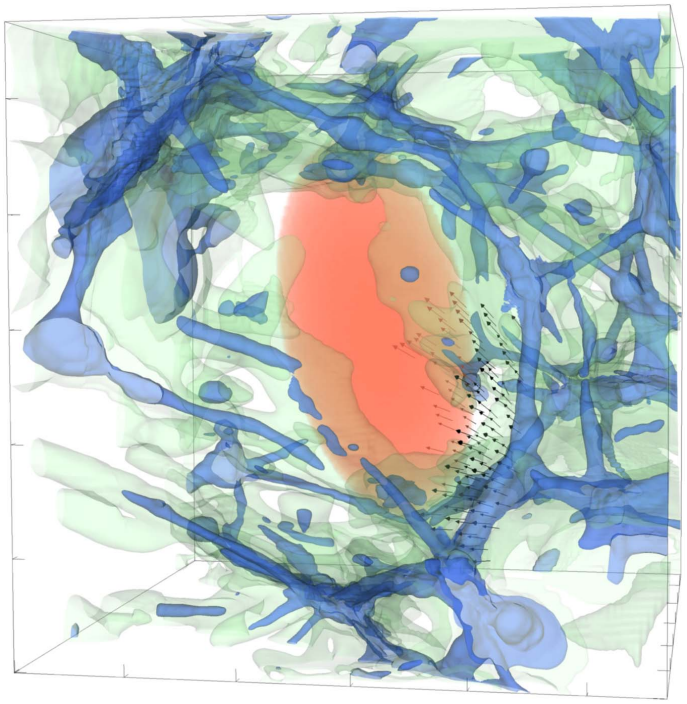


Figure 3. Close look at a large ($R_{\rm eq} \simeq 12$ Mpc) void that undergoes significant accretion by redshift $z \simeq 0$. The box corresponds to a cubic domain $\sim$56 Mpc (comoving) along each direction. The ticks on the axes are spaced 10 Mpc for visual reference. Blue and pale green contours represent the gas density, with blue and green regions approximately corresponding to collapsed (cluster and filaments) matter and diffuse gas around it, respectively. The orange, shaded area shows the location of the ellipsoidal void, which is well delineated by the surrounding matter. The arrows represent the gas velocity field around the surface of the void in those regions where gas is being accreted toward it.

watershed transform Platen et al. 2007; Neyrinck 2008; see, for example, the comparison project of Colberg et al. 2008).

The authors thank the referee for constructive criticism. This work has been supported by the Spanish Agencia Estatal de Investigación (AEI, grant PID2019-107427GB-C33) and by the Generalitat Valenciana (grant PROMETEO/2019/071). D.V. acknowledges partial support from Universitat de València through an *Atracció de Talent* fellowship. Simulations have been carried out using the supercomputer Lluís Vives at the Servei d'Informàtica of the Universitat de València.

Appendix
Additional Material

This appendix contains three additional figures. Figure A1 presents a summary of the statistical properties of our void sample. Figure A2 presents the results on the checks of the robustness of the pseudo-Lagrangian approach (see Section 2.3). Finally, Figure A3 contains an animation displaying the evolution of a density slice with the velocity field overplotted, showing an example of mass inflows through void boundaries.

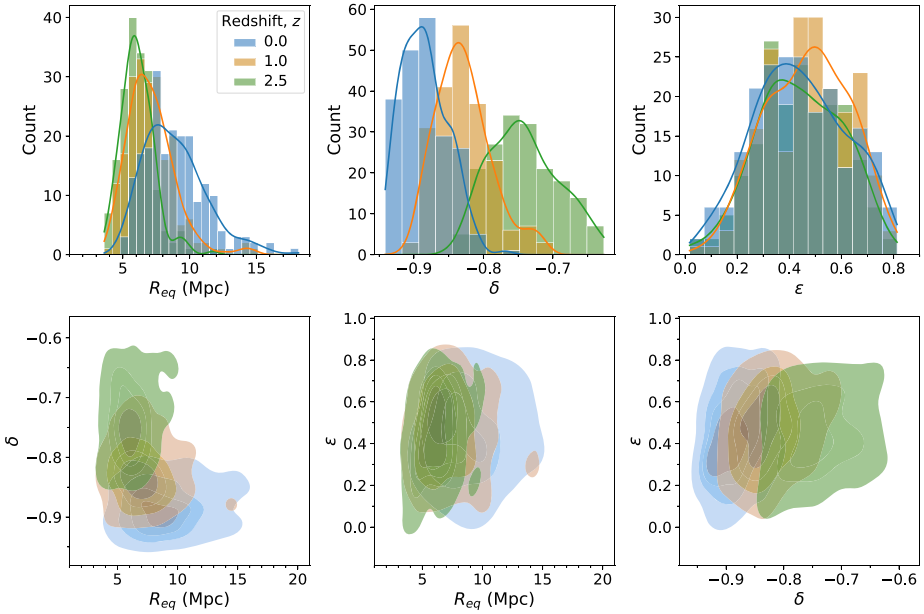


Figure A1. Summary of general properties of the void sample. Top row: distribution of the equivalent radius ($R_{\rm eq}$, left panel), mean overdensity (δ, middle panel), and ellipticity (ε, right panel) for the voids sample at redshifts $z \simeq 0$ (blue), 1 (orange), and 2.5 (green). Bottom row: joint distribution of each pair of variables, according to the same color palette. The continuous lines in the histograms, and the color contours in the joint distributions have been obtained by means of a Gaussian kernel density estimation procedure. Darker colors imply higher density of voids in the corresponding parametric space.

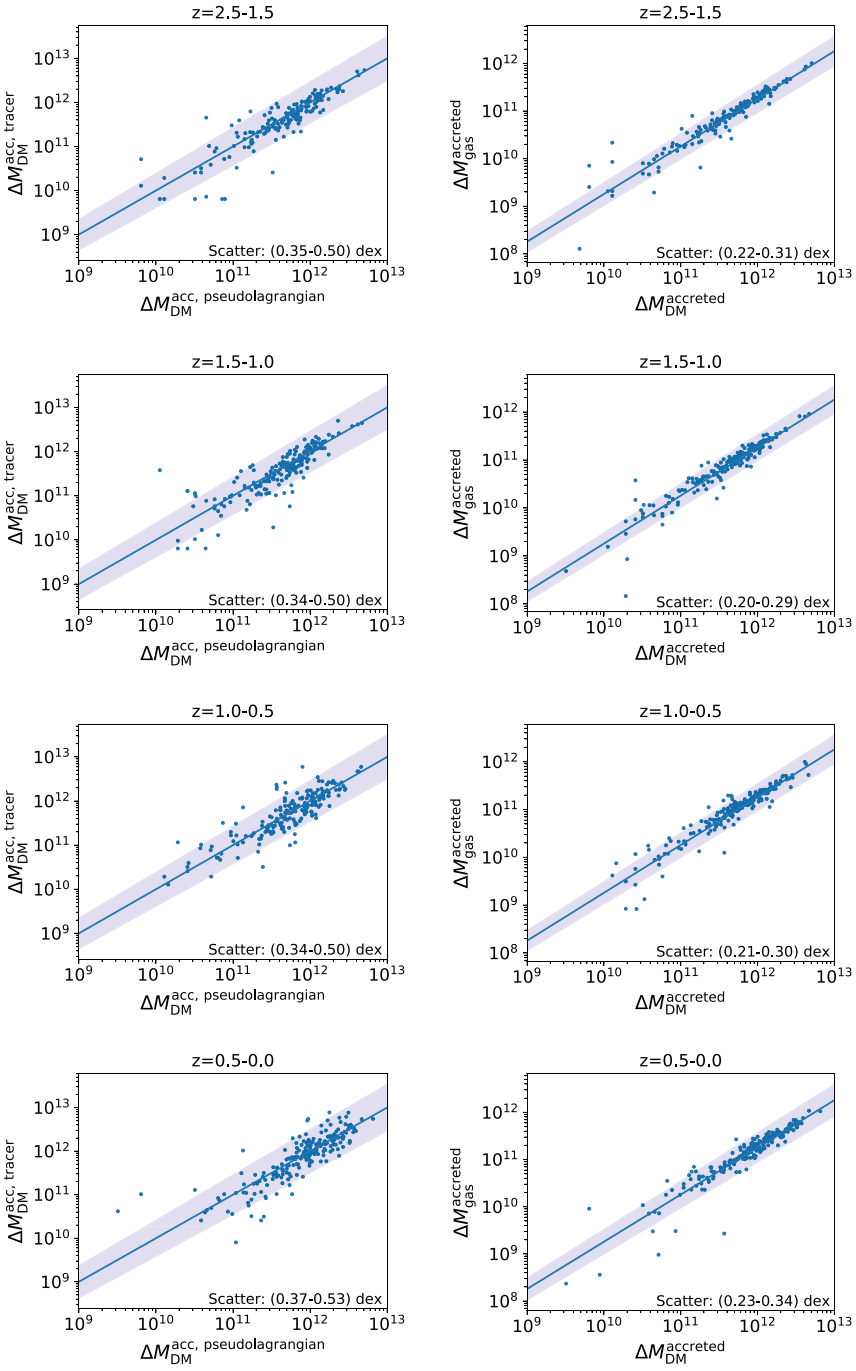


Figure A2. Left-hand column: correlation between the accreted DM mass computed according to our pseudo-Lagrangian method, and when computed by explicitly tracing DM particles in the simulation, in the redshift interval specified at the top of each panel. Right-hand column: tight correlation between the gas and the DM accreted mass, both computed according to our pseudo-Lagrangian algorithm.

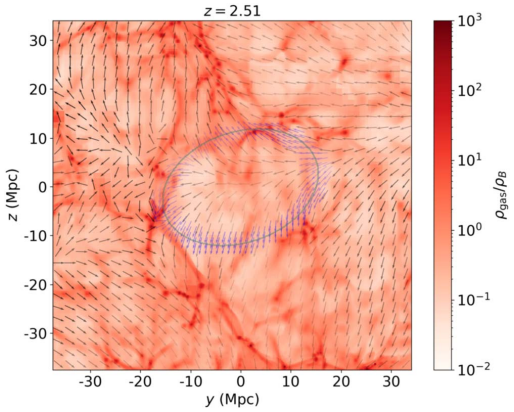


Figure A3. Animation showing the evolution of a density slice (∼9 Mpc thick) through a large ($R_{eq} \simeq 14$ Mpc) void and its environment, from $z = 2.5$ to $z = 0$. The colors encode the gas density (in units of the background density) according to the adjacent color scale. The green ellipse represents the slice through the void's ellipsoid. Note that some part of the void's wall may appear to be inside the void due to projection effects. Nevertheless, the void overdensity is low, below $\langle\delta\rangle \sim -0.8$, thus ensuring that overdense regions are well excluded. Arrows represent the gas velocity field in the slice. Note that the size of the arrows is constant for better visualization, while the magnitude of the projected velocity vector is encoded in the opacity of the arrow (more opaque arrows imply higher velocity magnitudes). The velocity field around the void's boundary is sampled with higher density of arrows and colored blue for better visualization. Matter entering the void through its rightmost corner can be clearly visualized in the animation. This figure is available as an animation in the HTML version of the article. *Still frame*: first frame of the animation. *Duration*: 8 s.

(An animation of this figure is available.)

ORCID iDs

David Vallés-Pérez https://orcid.org/0000-0003-2656-5985
Vicent Quilis https://orcid.org/0000-0002-2852-5031
Susana Planelles https://orcid.org/0000-0002-0105-4815

References

Aragon-Calvo, M. A., & Szalay, A. S. 2013, MNRAS, 428, 3409
Baushev, A. N. 2021, MNRAS, 504, L56
Behroozi, P. S., Loeb, A., & Wechsler, R. H. 2013, JCAP, 2013, 019
Bertschinger, E. 1985, ApJS, 58, 1
Bos, E. G. P., van de Weygaert, R., Dolag, K., et al. 2012, MNRAS, 426, 440
Bothun, G. D., Geller, M. J., Kurtz, M. J., Huchra, J. P., & Schild, R. E. 1992, ApJ, 395, 347
Cautun, M., van de Weygaert, R., Jones, B. J. T., et al. 2014, MNRAS, 441, 2923
Ceccarelli, L., Padilla, N. D., Valotto, C., & Lambas, D. G. 2006, MNRAS, 373, 1440
Colberg, J. M., Pearce, F., Foster, C., et al. 2008, MNRAS, 387, 933
Dekel, A., & Rees, M. J. 1994, ApJL, 422, L1
Foster, C., & Nelson, L. A. 2009, ApJ, 699, 1252
Hahn, O., Carollo, C. M., Porciani, C., & Dekel, A. 2007, MNRAS, 381, 41
Kreckel, K., Platen, E., Aragón-Calvo, M. A., et al. 2011, AJ, 141, 4
Lavaux, G., & Wandelt, B. D. 2010, MNRAS, 403, 1392
Lavaux, G., & Wandelt, B. D. 2012, ApJ, 754, 109
Minoguchi, M., Nishizawa, A. J., Takeuchi, T. T., & Sugiyama, N. 2021, MNRAS, 503, 2804
Neyrinck, M. C. 2008, MNRAS, 386, 2101
Padilla, N. D., Ceccarelli, L., & Lambas, D. G. 2005, MNRAS, 363, 977
Park, D., & Lee, J. 2007, PhRvL, 98, 081301
Patiri, S. G., Betancort-Rijo, J., & Prada, F. 2012, A&A, 541, L4
Paz, D., Lares, M., Ceccarelli, L., Padilla, N., & Lambas, D. G. 2013, MNRAS, 436, 3480
Pisani, A., Massara, E., Spergel, D. N., et al. 2019, BAAS, 51, 40
Planck Collaboration 2020, A&A, 641, A6
Platen, E., van de Weygaert, R., & Jones, B. J. T. 2007, MNRAS, 380, 551
Quilis, V. 2004, MNRAS, 352, 1426
Quilis, V., Planelles, S., & Ricciardelli, E. 2017, MNRAS, 469, 80
Ricciardelli, E., Cava, A., Varela, J., & Quilis, V. 2014a, MNRAS, 445, 4045
Ricciardelli, E., Quilis, V., & Planelles, S. 2013, MNRAS, 434, 1192
Ricciardelli, E., Quilis, V., & Varela, J. 2014b, MNRAS, 440, 601
Sheth, R. K., & van de Weygaert, R. 2004, MNRAS, 350, 517
Sutter, P. M., Elahi, P., Falck, B., et al. 2014, MNRAS, 445, 1235
Vallés-Pérez, D., Planelles, S., & Quilis, V. 2020, MNRAS, 499, 2303
van de Weygaert, R. 2016, The Zeldovich Universe: Genesis and Growth of the Cosmic Web, 308, 493
van de Weygaert, R., & Platen, E. 2011, IJMPS, 1, 41
van de Weygaert, R., & van Kampen, E. 1993, MNRAS, 263, 481
Zeldovich, I. B., Einasto, J., & Shandarin, S. F. 1982, Natur, 300, 407
Zeldovich, Y. B. 1970, A&A, 500, 13
Zemp, M., Gnedin, O. Y., Gnedin, N. Y., & Kravtsov, A. V. 2011, ApJS, 197, 30
Zhang, C., Churazov, E., Dolag, K., Forman, W. R., & Zhuravleva, I. 2020, MNRAS, 494, 4539

A&A 664, A42 (2022)
https://doi.org/10.1051/0004-6361/202243712
© D. Vallés-Pérez et al. 2022

Astronomy & Astrophysics

The halo-finding problem revisited: a deep revision of the ASOHF code

David Vallés-Pérez[1], Susana Planelles[1,2], and Vicent Quilis[1,2]

[1] Departament d'Astronomia i Astrofísica, Universitat de València, 46100 Burjassot (València), Spain
e-mail: david.valles-perez@uv.es
[2] Observatori Astronòmic, Universitat de València, 46980 Paterna (València), Spain

Received 5 April 2022 / Accepted 4 May 2022

ABSTRACT

Context. New-generation cosmological simulations are providing huge amounts of data, whose analysis becomes itself a pressing computational problem. In particular, the identification of gravitationally bound structures, known as halo finding, is one of the main analyses. Several codes that were developed for this task have been presented during the past years.
Aims. We present a deep revision of the code ASOHF. The algorithm was thoroughly redesigned in order to improve its capabilities of finding bound structures and substructures using both dark matter particles and stars, its parallel performance, and its abilities of handling simulation outputs with vast amounts of particles. This upgraded version of ASOHF is conceived to be a publicly available tool.
Methods. A battery of idealised and realistic tests are presented in order to assess the performance of the new version of the halo finder.
Results. In the idealised tests, ASOHF produces excellent results. It is able to find virtually all the structures and substructures that we placed within the computational domain. When the code is applied to realistic data from simulations, the performance of our finder is fully consistent with the results from other commonly used halo finders. The performance in substructure detection is remarkable. In addition, ASOHF is extremely efficient in terms of computational cost.
Conclusions. We present a publicly available deeply revised version of the ASOHF halo finder. The new version of the code produces remarkable results in terms of halo and subhalo finding capabilities, parallel performance, and low computational cost.

Key words. large-scale structure of Universe – dark matter – galaxies: clusters: general – galaxies: halos – methods: numerical

1. Introduction

Over the past four decades, numerical simulations of cosmic structure formation have grown significantly in size, dynamical range, and accuracy of the physical model ingredients (see e.g. Vogelsberger et al. 2020; Angulo & Hahn 2022, for recent reviews). In addition to a precise description of the evolution of the dark matter (DM) component of the Universe, current cosmological simulations have also improved their modelling of the complex baryonic physical processes that shape the properties of the gaseous and the stellar components (see e.g. Planelles et al. 2015, for a review). On the other hand, the outstanding development of computing facilities has led to an increasing computational power and to important advances in algorithms and techniques. This progress has allowed current cosmological simulations to reach a significant level of mass and force resolution, complexity, and realism.

The excellent predictive power of these simulations makes them essential tools in cosmology and astrophysics: they are crucial not only for testing the accepted cosmological paradigm, but for interpreting and analysing how different physical processes inherent to the cosmic evolution affect the observational properties of the Universe we inhabit. To properly exploit the unprecedented capabilities of current advanced cosmological simulations, equivalently complex and sophisticated structure finding algorithms are also required. A proper identification and characterisation of the population of DM haloes and subhaloes, including their abundances, shapes and structure, physical properties, and merging histories, is decisive for understanding the formation and evolution of cosmic structures.

Even when a DM halo is simply a locally overdense gravitationally bound structure embedded within the global background density field, the definition of its boundary and, hence, its mass is unavoidably arbitrary. The situation is even more accentuated in the case of subhaloes that are located within larger-scale overdensities, called hosts. In an attempt to overcome these issues, a significant number of halo-finding methods and techniques have been developed in the past decades (e.g. FoF[1], Davis et al. 1985; SO[2], Cole & Lacey 1996; HOP, Eisenstein & Hut 1998; BDM[3], Klypin et al. 1999; Subfind, Springel et al. 2001; Dolag et al. 2009; AHF[4], Gill et al. 2004; Knollmann & Knebe 2009; ASOHF[5], Planelles & Quilis 2010; Velociraptor, Elahi et al. 2011, 2019; HBT[6], Han et al. 2012, 2018; Rockstar, Behroozi et al. 2013; to cite a few).

In general, however, all these algorithms can be broadly divided into three main families of codes: those based on the SO method (Press & Schechter 1974; Cole & Lacey 1996), those relying on the FoF algorithm in 3D configuration space

[1] Friends of friends.
[2] Spherical overdensity.
[3] Bound density maxima.
[4] Amiga halo finder.
[5] Adaptive spherical overdensity halo finder.
[6] Hierarchical bound tracing.

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
This article is published in open access under the Subscribe-to-Open model. Subscribe to A&A to support open access publication.

(Davis et al. 1985), and those based on the 6D phase-space FoF (e.g. Diemand et al. 2006). Algorithms in the first class involve locating peaks in the density field and finding the extent of the halo, either by growing spheres until either the enclosed density falls below a given threshold or other properties are met. On the other hand, codes in the second and third categories link particles that are close to each other, either in configuration or in phase-space. Knebe et al. (2011, 2013) reported an exhaustive review and comparison of some of these halo finders. These studies showed that while all codes are able to identify the location of mock isolated haloes and some of their properties (e.g. the maximum circular velocity, $v_{\rm max}$), important differences arose for small-scale haloes, especially substructures (Onions et al. 2012, 2013; Hoffmann et al. 2014), or when some properties such as spin and shape were determined.

In addition to DM structures, modern simulations also track the formation of stellar particles from cold gas. These particles cluster and form stellar haloes, or galaxies. While the formation of galaxies is tightly linked to their underlying DM haloes, their different evolutionary histories, properties, and dynamics imply that they deserve to be studied as objects in their own right. Consistently, specific algorithms for these tasks have recently been developed (e.g. Navarro-González et al. 2013; Cañas et al. 2019).

Planelles & Quilis (2010) presented ASOHF, a halo finder based on the SO approach. Although ASOHF was especially designed to be applied to the outcomes of grid-based cosmological simulations, it was adapted to work as a stand-alone halo finder on particle-based simulations. The performance of ASOHF in different scenarios was demonstrated in Knebe et al. (2011). In addition, the code has been employed in a number of works (Planelles & Quilis 2013; Quilis et al. 2017; Martin-Alvarez et al. 2017; Planelles et al. 2018; Vallés-Pérez et al. 2020, 2021). The incessant improvements of cosmological simulations, aided by the ever-growing available computing power, have enormously increased the number of particles and, consequently, the richness of small-scale structures. This demanded a revisit of the process of halo-finding in ASOHF to ensure that small structures are well captured and their properties can be recovered in an unbiased way, and also to guarantee that the code is able to tackle these amounts of data within reasonable computational times.

In this paper we present an upgraded, faster, and more memory-efficient version of ASOHF that is capable of efficiently dealing with the new generation of cosmological simulations that include huge numbers of DM particles, haloes, and substructures. Amongst the main improvements, we present a smoother density interpolation that lowers the computational cost by decreasing the number of spurious density peaks, the addition of complementary unbinding procedures, a new scheme for searching for substructures, the ability of identifying and characterising stellar haloes, and a domain decomposition approach that can lower the computational cost and computing time. On the performance side, the code has been profoundly overhauled in terms of parallelisation and memory requirements. It is now able to analyse simulations with hundreds of millions of particles on desktop workstations within a few minutes at most. This version of ASOHF is publicly available.

The paper is organised as follows. In Sect. 2 we describe the main procedure on which our halo finder relies to identify the samples of haloes, subhaloes, and stellar haloes, as well as additional features such as the domain decomposition scheme and the merger tree. The performance and scalability of ASOHF in some idealised but rather complex tests is shown in Sect. 3. Additionally, in Sect. 4 we test the performance of the code against actual simulation data and compare its performance to other well-known halo finders, and we show the capabilities of ASOHF as a stellar halo finder. Finally, in Sect. 5 we discuss and summarise our results. Appendix A further describes one of the unbinding schemes, while in Appendix B we discuss our estimation of the gravitational binding energy and most-bound particle by sampling.

2. Algorithm

While the original algorithm was introduced by Planelles & Quilis (2010), a large number of modifications and upgrades have been undertaken in order to provide a fast, memory-efficient, and flexible code that is able to tackle a new generation of cosmological simulations, with an increase of several orders of magnitude in the number of DM particles, haloes, and rich substructure. Here we describe the main steps of our halo-finding procedure. The implementation of ASOHF for shared-memory platforms (OpenMP), written in Fortran, is publicly available through the GitHub repository of the code[7]. The following subsections describe the input data (Sect. 2.1), the process of identifying density peaks (Sect. 2.2), the characterisation of haloes using particles (Sect. 2.3), the substructure identification scheme (Sect. 2.4), the characterisation of stellar haloes (Sect. 2.5), and several additional features and tools of the ASOHF package (Sect. 2.6). Table 1 contains a summary of the parameters of the code that can be configured.

2.1. Input data

Generally, the input data for ASOHF consist of a list of DM particles, containing the three-dimensional positions and velocities, masses, and a unique integer identifier of each of the $N_{\rm part}$ particles. While ASOHF was originally envisioned to be coupled to the outputs of the cosmological code MASCLET (Quilis 2004; Quilis et al. 2020), it can work as a fully stand-alone halo finder. The reading routine is fully modular and can easily be adapted to suite the input format of the user[8].

2.2. Grid halo identification

The identification of haloes relies on the analysis of the underlying continuous density field, which is obtained by means of a grid interpolation from the particle distribution. Originally inherited from MASCLET, since its original version, ASOHF uses an Adaptive Mesh Refinement (AMR) hierarchy of grids to compute the density field on different scales. This enables capture of a large dynamical range in masses and radii. In this section we describe the mesh creation procedure (Sect. 2.2.1), which has been optimised to allow the code to handle even hundreds of thousands of refinement patches in large simulations. Next, we present the new density interpolation scheme (Sect. 2.2.2) that mitigates the effect of sampling noise, and the halo-finding process over the grid (Sect. 2.2.3). It is worth noting that to avoid arbitrariness in defining a given structure, the new version of ASOHF does not deal with peaks within haloes at this stage. This is considered in a later step in Sect. 2.4.

2.2.1. Mesh creation

A coarse grid of size $N_x \times N_y \times N_z$ covers the whole domain of the input simulation. On top of this base grid, an arbitrary number of

[7] https://github.com/dvallesp/ASOHF

[8] For more information, check the code documentation in https://asohf.github.io

Table 1. Summary of the main parameters that can be tuned to run ASOHF.

Parameter	(Symbol)	Description and remarks
Parameters for halo finding: mesh creation and halo identification		
Base grid size	N_x	Typically set to $N_x = \sqrt[3]{N_{\rm part}}$ or $2\sqrt[3]{N_{\rm part}}$
Number of refinement levels	n_ℓ	Peak resolution will be $L/(N_x \times 2^{n_\ell})$, typically set to match the force resolution of the simulation
Number of particles to flag a cell as 'refinable'	$n_{\rm part}^{\rm refine}$	
Fraction of refinable cells to extend the patch	$f_{\rm refinable}^{\rm extend}$	A patch is grown along a direction only if more than this fraction of the newly added cells is refinable
Minimum size of the patch to be accepted	$N_{\rm min}^{\rm patch}$	Only patches with minimum dimension above this are accepted
Base grid refinement border		Exclude these many cells close to the domain boundary (≥1)
AMR grids refinement border		Idem., in each AMR grid (≥0)
Kernel order for interpolating density from particles		Either 1 (linear kernel) or 2 (quadratic kernel; recommended)
Particle species		Assign kernel size by particle mass, local density, or none
Minimum number of particles per halo	$n_{\rm min}^{\rm halo}$	Discard haloes below this number of particles (e.g. 25)
Stellar haloes		
Component used for mesh halo finding		Use only DM or DM+stars for identifying density peaks
Kernel width for stars	$\ell_{\rm stars}$	Interpolate stars in a cloud of radius $L/(N_x \times 2^{\ell_{\rm stars}})$
Minimum number of stellar particles per stellar halo		
Density increase (from inner minimum) to cut the halo	$f_{\rm min}$	See Sect. 2.5
Maximum radial distance without stars to cut the halo	$\ell_{\rm gap}$	See Sect. 2.5
Minimum density (in units of $\rho_B(z)$) to cut the halo	f_B	See Sect. 2.5

Notes. The first block contains the parameters that have effect on the identification of haloes, and the second block refers to the stellar halo-finding procedure.

mesh refinement levels is generated by placing patches covering the regions with highest particle number density, each level halving the cell size with respect to the previous one. In particular, we flag as refinable any cell hosting more than a minimum number of particles ($n_{\rm part}^{\rm refine}$). The mesh creation routine then examines all the refinable cells, from densest to least dense, and tries to extend the patch in each direction if the fraction of refinable cells amongst the added cells exceeds $f_{\rm refinable}^{\rm extend}$. Only patches with a given minimum size ($N_{\rm min}^{\rm patch}$) are accepted; otherwise, the region is not refined. These quantities ($n_{\rm part}^{\rm refine}$, $f_{\rm refinable}^{\rm extend}$, and $N_{\rm min}^{\rm patch}$), as well as the number of refinement levels (n_ℓ), are free parameters that can be tuned to find a balance between memory usage and resolution. Generally, the latter can be fixed so that the peak resolution matches the force resolution of the simulation. We present in Sect. 3.4.1 a test showing how these parameters work with varying particle resolutions.

2.2.2. Density interpolation

In our experiments, we find that interpolating particles into cells using the standard cloud-in-cell (CIC) or triangular-shaped cloud (TSC) schemes at the resolution of each AMR patch has a detrimental effect for the purpose of density peak finding because it produces a very large number of peaks due to shot noise. While these false peaks would be removed in subsequent steps when halo properties are refined with particles (Sect. 2.3), they produce an overwhelming computational burden and load imbalance between different threads.

To avoid this, we compute the local density field by spreading each particle, using linear or quadratic kernels, in a cubic cloud with the same volume as the particle sampled in the initial conditions. That is to say, if different particles species (in terms of mass) are present, each one is spread into the AMR grids according to their corresponding kernel sizes, that is, the kernel radius $\Delta x_{{\rm kern},i}$ is set by

$$\Delta x_{{\rm kern},i} = L/2^{\lfloor \log_8 \frac{M_{\rm box}}{m_{p,i}} \rfloor}, \tag{1}$$

with L the box size, $M_{\rm box}$ the total mass in the box, and $m_{p,i}$ the mass of the particle i. Alternatively, for simulations with equal-mass particles, the kernel size can be chosen to be determined by the local density in the base grid cell occupied by each particle. This procedure naturally produces a smooth density field that is free of spurious noise while still capturing local features such as the density peaks associated with a hierarchy of substructures.

2.2.3. Halo finding

Haloes are pre-identified as peaks, that is, local maxima, in the density field, from the coarsest to the finest AMR levels. To do this, ASOHF lists all the (non-overlapping) cells at each AMR level that simultaneously fulfil the following two criteria. First, the cell density is higher than the virial density contrast at a given redshift, $\Delta_{\rm m,vir}(z)$, computed according to the prescription of Bryan & Norman (1998). Incidentally, we also consider cells whose overdensity exceeds $\Delta_{\rm m,vir}(z)/6$ because the density interpolation could smooth the peaks in some situations. We verify with the tests in Sect. 3.1 that further lowering this value does not allow to recover any more haloes. Additionally, the cell must correspond to a local maximum of the density field, that is, its density is higher than that of the 3^3 neighbouring cells.

The algorithm then iterates over them with decreasing overdensity. For each cell, the procedure can be summarised as follows. First, check whether the cell is inside a previously identified halo. If it is, skip it; if not, continue (note that we do not search for substructure at this stage). Then, refine the location of the density peak using the information in higher

(i.e. finer) AMR levels, if available, and check that the refined position does not overlap with previously identified haloes. Finally, grow spheres of increasing radii until the enclosed overdensity, measured using the density field, falls below $\Delta_{\rm m,vir}$.

This produces a first list of isolated haloes (i.e. that are not substructure), with an estimation of their positions, radii, and masses, which will be refined further by using the whole particle list information in the following step. At this step, haloes can overlap, although no halo centre can be placed inside another halo. These peaks are processed within the substructure-finding step (Sect. 2.4).

2.3. Halo refinement using particles

After they are identified, the halo properties are refined using the particle distribution in several steps as described below. We refer to Sect. 3.1 for a test that shows the capabilities of ASOHF to identify and recover the properties of haloes in an unbiased way. While the basic process (collecting particles, unbinding, and determining the virial radius), common to most SO halo finders, was present in the original version, the new version of ASOHF includes many refinements to improve the quality of the recovered catalogues (e.g. recentring of the density peak or additional unbinding procedures), as well as other features to significantly decrease the computational cost and compute certain halo properties.

Recentring of the density peak. Within the scheme described in Sect. 2.2, the density peak location, as identified within the grid, has an uncertainty of the order of the cell size of the finest grid covering the peak. In order to improve this determination, the first step consists of refining the position of the density peak using particles.

To do this, we consider a cube centred on the estimated density peak, with a side length of $2\Delta x_\ell$, with Δx_ℓ the cell size of the finest patch used to locate the density peak within the grid. We compute the density field in this volume using an ad hoc 4^3 cells grid, and select as the corrected centre the position of the largest local maximum of the density field. The process is iterated while more than 32 particles are found in the cube. Each time, the cube side lenth is halved, and it is centred on the peak of the previous step. The position found by this procedure is not further changed in the process of halo finding.

Selection of particles. In order to alleviate part of the computational burden associated with traversing the whole particle list, only the particles in a sphere with mean density $\langle\rho\rangle = \min(\Delta_{\rm vir,m}(z),\, 200)\,\rho_B(z)$, with $\rho_B(z)$ the background matter density at redshift z, are kept for each halo. Subsequently, this list is sorted by increasing distance to the halo centre.

Additionally, to boost performance by decreasing the number of array accesses and distance calculations, which are an important performance penalty as the number of particles in the simulation increases, particles are sorted according to their x-component before the whole process of halo finding starts. Then, only a small fraction of the particle list is traversed in order to select the halo particle candidates, in a way similar to a tree search algorithm in one dimension. This effect becomes increasingly stronger as the volume being analysed increases.

Unbinding. A common feature of configuration-space halo finders is the necessity of pruning the particles that are dynamically unrelated to the haloes because particles are initially collected using only spatial information. We perform two complementary, consecutive unbinding procedures that we describe below.

Local escape velocity unbinding. First, the Poisson equation is solved in spherical symmetry for the particle distribution (see Appendix A), yielding the gravitational potential $\phi(r)$. The escape velocity is then estimated as $v_{\rm esc}(r) = \sqrt{-2\phi(r)}$. Finally, any particle with velocity (relative to the halo centre of mass reference frame) exceeding its local escape velocity should be flagged as unbound and pruned from the halo.

However, because the bulk velocity of the halo can be severely contaminated by the unbound component, we unbind particles in an iterative way by pruning particles with a speed relative to the centre of mass higher than $\beta v_{\rm esc}(r)$. In successive steps, we take $\beta = 8$, 4, and finally $\beta = \beta_{\rm final} \equiv 2$. Each time, the gravitational potential generated by bound particles alone is recomputed. The process is iterated with the last value of β until no new particles are removed in an iteration.

We set $\beta_{\rm final} = 2$ instead of removing all particles whose speed exceeds the local escape value because these *marginally unbound* particles may become bound at a later step and will not drift away from the cluster immediately (see e.g. the discussion in Knebe et al. 2013).

Velocity space unbinding. The standard deviation of the particle velocities within the halo, σ_v, with respect to the centre-of-mass velocity, is computed. Particles whose three-dimensional velocity differs by more than $\beta\sigma_v$ from the centre-of-mass velocity are pruned. Like in the previous scheme, β is lowered progressively, from $\beta = 6$ to $\beta = \beta_{\rm final} = 3$. In each step, the centre-of-mass velocity is recomputed for bound particles.

We refer to Sect. 3.3 for two specific tests that show the complementarity of these two unbinding schemes.

Determining spherical overdensity boundaries. The spherical overdensity boundaries $\Delta_m \equiv \frac{\langle\rho\rangle}{\rho_{\rm B}(z)} = 200$, 500, 2500, $\Delta_c \equiv \frac{\langle\rho\rangle}{\rho_{\rm crit}(z)} = 200$, 500, 2500 (where $\rho_{\rm crit}(z)$ is the critical density for a flat universe) and $\Delta_{\rm vir}$ (Bryan & Norman 1998) are precisely determined. The centre-of-mass velocity inside the refined $R_{\rm vir}$ boundary, as well as the maximum circular velocity, its radial position, and the corresponding enclosed mass are also computed. Haloes with fewer than a user-specified minimum number of particles ($n_{\rm min}^{\rm halo}$) are regarded as poor haloes and are removed from the list.

Determining halo properties. Finally, several properties of the halo are computed inside $R_{\rm vir}$. They include centre of mass position and velocity, velocity dispersion, kinetic energy, gravitational energy (either by direct sum or using a sampling estimate, see Appendix B), and most-bound particle (which serves as a proxy for the location of the potential minimum), specific angular momentum, inertia tensor and principal axes of the best-fitting ellipsoid, mass-weighted radial speed, enclosed mass profile, and list of bound particles, from which any other possible information can be directly computed.

2.4. Substructure finding

Once all the non-substructure haloes are identified, we proceed to search for substructures, which we define as systems centred on density peaks within previously identified haloes (either isolated haloes or substructures detected at a coarser level of refinement). The process of substructure finding is nearly parallel to that of finding non-substructure haloes. We therefore outline

the differences here with the procedure described above. The most remarkable distinction lies in the choice of halo boundary. While non-substructure (i.e. isolated) haloes can be well characterised by an enclosed density threshold, such a threshold may not exist if the halo is embedded in a larger-scale overdensity. While some finders use a change in the slope of the density profile to set the boundary of substructure (e.g. AHF, Knollmann & Knebe 2009), we find that this rise in the slope is not always found and leads to arbitrarily large substructure radii. Instead, we use the Jacobi radius, $R_{\rm J}$, defined in Binney & Tremaine (1987) as the saddle point of the effective potential generated by the host-satellite system, as an estimate of the substructure extent (the region of space in which the attraction towards the satellite is stronger than that of the host). We note that this new scheme for substructure characterisation is integrally new to the revised version of the finder.

Search on the grid. Cells above the virial density contrast are considered as candidate substructure centres if they are a strictly defined local maximum (their density is higher than in any other of the 3^3-1 neighbouring cells), they do not belong to a previously identified substructure at the same grid level, and they are not within $2\Delta x_\ell$ of their host halo centre, where Δx_ℓ is the grid cell size at the given refinement level. This last step is enforced to avoid arbitrarily detecting the same peak as a subhalo of itself, while allowing to recover increasingly more central substructures as long as refinement levels that cover the central region of the host are available.

After recentring, we choose as host for each centre candidate the halo (at the highest hierarchy level[9]) that minimises the distance between host and substructure centres, D. We then compute M, the mass of the host within a sphere of radius D, by means of a cubic interpolation from the previously saved enclosed mass profiles, and obtain a rough estimate of the substructure extent by numerically solving Eq. (2),

$$\left(\frac{R_{\rm J}}{D}\right)^3 - \frac{m}{3M+m} = 0, \tag{2}$$

where m is the mass of the substructure candidate within a sphere of radius $R_{\rm J}$. This equation is a simplification of the exact definition of the Jacobi radius (see below, Eq. (3); see also Binney & Tremaine 1987) under the assumption $m \ll M$ (and $r_{\rm J} \ll D$). We note that an approximate version of this definition was already implemented by `MHF` (Gill et al. 2004).

Refinement with particles. The procedure is analogous to the one described above. The sole difference is that the boundary of the halo (for measuring all halo properties) is taken as the Jacobi radius. In this respect, we use particles to refine this boundary with the non-approximate expression yielding $R_{\rm J}$, as given by Binney & Tremaine (1987),

$$f(x) \equiv \frac{1}{(1-x)^2} - \frac{g(x)}{x^2} + \left[1+g(x)\right]x - 1 = 0, \tag{3}$$

with $x \equiv R_{\rm J}/D$ and $g(x) \equiv m/M$. After $R_{\rm J}$ has been identified, all halo properties can be computed from the bound particles as discussed in Sect. 2.3. We refer to Sect. 3.2 for a test of the substructure identification capabilities of `ASOHF`.

[9] That is, if a peak, candidate for corresponding to a substructure, is inside a halo and inside a subhalo, the corresponding structure is regarded as a subsubhalo if it is finally accepted.

2.5. Stellar haloes

In addition to identifying DM haloes, the new version of `ASOHF` is also able to characterise and produce catalogues of stellar haloes (i.e. galaxies) if such particles are provided. The identification of stellar haloes in the first place relies on the identification of the underlying DM haloes and subhaloes, which are typically more massive and less concentrated than their stellar counterpart (e.g. Pillepich et al. 2014). However, it is worth emphasising that stellar haloes are then characterised by `ASOHF` as independent objects. The procedure iterates over all previously found haloes and subhaloes, and performs the following steps.

Selection of particles. All stellar particles inside the virial volume of the DM halo are collected using the same tree-like search from the list of particles sorted along the x-coordinate as described in Sect. 2.3. All the bound DM particles identified in the halo-finding step are recovered, and the whole list of DM+stellar particles is sorted by increasing distance to the centre of the DM halo.

Determining a preliminary boundary of the stellar halo. In order to compute half-mass radii, we first need to place an outer boundary on the stellar halo. This is especially important in the case of central galaxies, where we wish to avoid that the masses, radii, and other properties of the resulting galaxy are affected by the presence of satellites or intracluster light. For each stellar halo candidate, we compute its spherically averaged stellar density profile, and place a radial cut at the smallest radius that fulfils at least one of the following conditions.

First, the stellar halo is cut if stellar density increases by more than a factor $f_{\rm min}$ from the previous (inner) density minimum. This indicates the presence of a massive satellite. As a second condition, the radial cut may also be triggered by the stellar density falling below a given threshold, which we parametrise in terms of the background matter density as $f_B\rho_B(z)$. The last condition is the presence of a gap in radial space, that is, a comoving distance larger than $\ell_{\rm gap}$ without any stellar particle.

These three conditions are complementary, present a small dependence on the free parameters, and are conservative enough to avoid splitting a real stellar halo into several pieces. In our test, we find that the resulting galaxy catalogue is fairly independent of the density increase parameter, which can be varied in the range $f_{\rm min} \in [2, 20]$. The results do not depend strongly on $f_{\rm B} \sim 1$ because stellar density profiles usually present a sharp boundary. However, we note that too low values ($f_{\rm B} \ll 1$) should be avoided because they may add strong contamination by intracluster light stellar particles. Finally, $\ell_{\rm gap}$ can be set to a conservative value, $\sim(5-10)$ kpc.

Unbinding. The unbinding steps are conceptually similar to those discussed in Sect. 2.3 for DM haloes. We stress a few subtleties here, however.

Local escape velocity unbinding. The procedure is similar to what we discussed for DM haloes. In this case, we consider all particles inside the previously found fiducial radius to solve Poisson's equation. However, we only unbind stellar particles because all DM particles were already bound to the underlying DM halo by construction.

Velocity space unbinding. From this point on, we remove DM particles and only consider the stellar ones. While this procedure is analogous to that performed for the DM halo, the fact that we only consider stellar particles at this stage implies that we account for the fact that DM and stellar components might correspond to distinct kinematic populations.

Determining half-mass radius and recentring. From the list of bound particles, the half-mass stellar radius is determined as the distance to the first particle whose enclosed bound mass exceeds half the total mass of bound particles inside the fiducial radius. Up to this point, the DM halo centre had been used as provisional centre. Henceforward, we perform an iterative recentring (analogous to the one described in Sect. 2.3) to the peak of stellar density by iteratively considering the largest stellar density peak inside the half-mass radius sphere. Finally, the half-mass radius is recomputed from this centre, which by definition yields a tighter radius than the first estimate.

Characterising halo properties. Finally, `ASOHF` computes a series of properties of the galaxy. All these properties, which include the stellar-mass inertia tensor, stellar angular momentum, and the velocity dispersion of stellar particles, are given with reference to the half-mass radius. The code can also output the whole list of stellar particles in the halo for further analyses. We refer to Sect. 4.2 for results of the stellar halo finding capabilities of `ASOHF`.

2.6. Additional features and tools

In addition to the main code, the `ASOHF` package includes a series of complementary tools (mainly implemented in `python3` with the usage of standard libraries) to set up a domain decomposition for running `ASOHF` (Sect. 2.6.1) and to compute merger trees from `ASOHF` catalogues (Sect. 2.6.2) as well as a library to load all `ASOHF` outputs into `Python`.

2.6.1. Domain decomposition

As we show in Sect. 3.4.1, the wall-time of our halo-finding code scales proportionally to the number of DM particles and haloes. Especially when large spatial volumes are analysed, the performance can therefore be greatly boosted by decomposing the domain. Each domain can be run by an independent `ASOHF` process, and the resulting catalogues (together with any other output files) can be merged to create a single catalogue representing the whole input domain. The absence of communication between the different domains allows the user to run the different domains either sequentially in the same machine or concurrently in different machines, thus effectively allowing the user to run `ASOHF` in distributed memory platforms.

To enable this procedure, `ASOHF` allows the user to specify a rectangular subdomain as an input parameter. All particles outside this domain are discarded by the reader routine, so as to reduce memory usage. The `python` script `setup_domdecomp.py`, included within the `ASOHF` package, automates this task by creating the necessary folders, executables, and parameter files for each domain. The number of divisions in each direction determines a fiducial domain for each task. Each of these fiducial domains, which cover the whole domain without overlapping with each other, is enlarged in all directions by an overlap length, which is also a free parameter. To avoid losing objects close to these boundaries of the domains, the overlap length can be safely set to the largest expected size of a halo in the simulation (∼3 Mpc for standard cosmologies). Once the domain decomposition is set up, we provide example shell scripts for running all the domains, either sequentially or concurrently (we provide an example `slurm` script).

When the halo-finding procedure has concluded for a given snapshot, the output files of each domain can be merged using the `merge_domdecomp_catalogues.py` script, which keeps all haloes whose centre lies on the fiducial domain. When substructure is present, the position of the progenitor halo (or the first non-substructure halo up the hierarchy) is considered to decide whether the substructure is kept in the merged catalogue. While the overlap between adjacent domains, if sufficiently large as discussed above, ensures that no structure is lost by the decomposition, the strategy of only keeping haloes whose density peak lies within the fiducial domain, or whose host fulfils these conditions, guarantees that no halo is identified twice.

2.6.2. Merger trees

Also included in the `ASOHF` code package, the `mtree.py` `python3` script allows building merger trees from the catalogues and particle list files. For each pair of consecutive snapshots (the prev and the post iterations) of the simulations, for which the halo catalogues have already been produced, the script first identifies for each post halo all the prev haloes in a sphere with radius equivalent to the maximum comoving distance travelled by the fastest particle in either the prev or the post iteration. These are referred to as the progenitor candidates.

For each of the progenitor candidates, the intersection of its member particles with the post halo is computed using the unique IDs of the particles. This process is performed in $O(N_{\rm prev}+N_{\rm post})$ for each intersection, instead of $O(N_{\rm prev}N_{\rm post})$, by sorting the particle lists by ID at the moment of reading the catalogues, thus boosting the performance of the `mtree.py` code. For instance, for a simulation with $\sim 10^8$ particles and ∼20 000 haloes, it takes 1−2 min on 16 threads to connect the haloes between each pair of snapshots.

We count as progenitors all haloes that contributed more than a fraction $f_{\rm given}$ to the descendent mass, which we arbitrarily set to a sufficiently low value, such as $f_{\rm given} = 10^{-3}$, for the merger tree to be complete. For each progenitor above this threshold, we report the following quantities.

Contribution to the descendent halo, $M_{\rm int}/M_{\rm post}$, where $M_{\rm int}$ is the mass of the particles in the post and the prev halo. This quantity needs to be interpreted carefully for substructures because we count the particles of a substructure as also belonging to the host. Therefore, a substructure most typically receives contributions of close to 100% from its host halo in the previous snapshot.

Retained mass, $M_{\rm int}/M_{\rm prev}$, that is, the fraction of prev mass given to the descendent halo. Again, in the presence of substructures, a host halo will be quoted as retaining 100% of a substructure in the previous iteration in most situations.

A third figure of merit is the normalised shared mass, $M_{\rm int}/\sqrt{M_{\rm prev}M_{\rm post}}$, which is the geometric mean of the two above quantities. This quantity has the virtue of suppressing the links of very massive haloes with very small haloes and vice versa because they are either suppressed by $M_{\rm prev}$ or $M_{\rm post}$.

Additionally, to track the main branch of the merger tree, we use the (approximate) determination of the most-bound particle introduced in Sect. 2.3 and described in Appendix B. Thus, we verify whether the most strongly bound particle of the post halo is in the prev candidate, and vice versa. The two-fold check is necessary in the presence of substructure because the most strongly bound particle of a substructure most usually lies within the host halo.

It may happen at some frequency, especially in the case of small haloes and substructures in very dense environments, that a halo is lost in one iteration and recovered afterwards. In order to avoid losing the main branch of a DM halo due to this spurious effect, the `mtree.py` script allows linking these

lost haloes to their progenitors, skipping an arbitrary number of iterations.

2.6.3. Python readers

All the information contained in ASOHF outputs can easily be loaded into python for analysis purposes, making use of the included readers.py library for DM and stellar haloes files, which allows the user to read and structure the data in several useful formats. Particle lists can also be loaded from python. These readers take care of the different indexing conventions between Fortran and python.

3. Mock tests and scalability

Before testing ASOHF on actual simulation data and comparing its results to other halo finders, we have quantified the performance of the key procedures of the code in some idealised but rather complex tests. In particular, we focused on identifying isolated haloes (Sect. 3.1) and substructures (Sect. 3.2), the unbinding procedures (Sect. 3.3), and the performance of the code (Sect. 3.4).

3.1. Test 1. Isolated haloes

The first idealised test aims to prove the ability of the code to identify haloes in a broad range of masses, without overlaps nor substructure. The setup of the test is as follows.

We considered a flat ΛCDM cosmology, with $\Omega_{\rm m} = 0.31$, $\Omega_\Lambda = 0.69$, $h \equiv H/(100\,{\rm km\,s^{-1}\,Mpc^{-1}}) = 0.678$ and $\sigma_8 = 0.82$, at redshift $z = 0$, and a cubic domain of side length $L = 40$ Mpc. This domain was populated with $N_{\rm part}$ particles of equal mass, $m_{\rm p} = \rho_{\rm B}(z=0)L^3/N_{\rm part}$. Halo masses were drawn from a Tinker et al. (2008) mass function by inverse transform sampling, setting a lower mass limit of $M_{\rm min} = 50\,m_{\rm p}$. The sampling was constrained so as to produce one halo with mass higher than $8\times10^{14}M_\odot$. The number of haloes was set by integrating the mass function from $M_{\rm min}$. Their corresponding virial radii were computed and haloes were placed at random positions, avoiding overlaps and crossing the box boundaries. Each halo was then realised with particles by sampling a Navarro-Frenk-White profile (NFW; Navarro et al. 1997) for the radial coordinate, using the concentration-mass ($c_{\rm vir} - M_{\rm vir}$) relation modelled by Ishiyama et al. (2021), and assuming spherical symmetry for the angular coordinates. The NFW profiles are extended up to 1.5 $R_{\rm vir}$ to avoid a sharp cut in the density profile, and the remaining particles up to $N_{\rm part}$, after sampling all haloes, were placed at random positions outside them. For this test, we used $N_{\rm part} = 128^3$, so that 1528 haloes with a mass higher than $M_{\rm min} \approx 6\times10^{10}\,M_\odot$ were generated and realised with particles of mass $m_p = 1.2\times10^9\,M_\odot$. These parameters are varied in Sect. 3.4 when we consider the scalability of the code.

While it is complex due to the high number of haloes, this test is idealised in the sense that it lacks any large-scale structure (LSS), such as filaments connecting haloes. This limitation of the test design implies that it is more challenging to detect small isolated haloes because they may only occupy one base grid cell and may not be refined enough to be detected as a density peak[10]. Therefore, the key parameter in this test is the base grid resolution, N_x. The remaining parameters were fixed to $n_\ell = 4$, $n_{\rm part}^{\rm refine} = 3$, and $N_{\rm min}^{\rm patch} = 14$ and have very limited impact on the outcomes of the test.

[10] The situation is different for a realistic simulation output, where a web of filaments surrounding a low-mass halo may more easily trigger the creation of a refinement patch covering it.

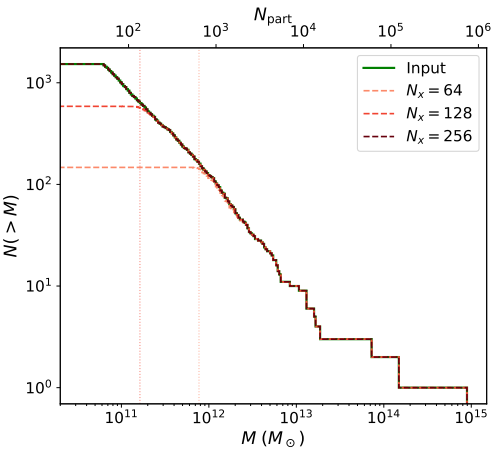


Fig. 1. Results from Test 1. Input cumulative mass function (green line) compared to ASOHF results with $N_{\rm x} = 64$ (light red), $N_{\rm x} = 128$ (red), and $N_{\rm x} = 256$ (dark red). The vertical lines mark the completeness limit, at 90%, of the catalogue produced by ASOHF. This value is not reported with $N_{\rm x} = 256$ because the code is able to detect all haloes.

Table 2. Completeness limits of ASOHF halo finding for Test 1 (in mass and number of particles, $M_{\rm lim}$ and $N_{\rm part}^{\rm lim}$), at 90%.

$N_{\rm x}$	$M_{\rm lim}$ ($M_\odot$)	$N_{\rm part}^{\rm lim}$
64	7.71×10^{11}	638
128	1.63×10^{11}	134
256	All	All

Notes. With $N_{\rm x} = 256$, all haloes are detected, and so we do not report these limits.

Figure 1 presents the mass functions (unnormalised; number of haloes with a mass higher than M, $N(>M)$) of the catalogues generated by ASOHF for $N_{\rm x} = 64$, 128, and 256 (in light red, red, and dark red lines, respectively), compared to the input (thick green line). Due to the overabundance of small haloes, it is difficult to interpret the number of detected haloes as a measure of the performance of the algorithm. Instead, we quote the completeness limit of the sample detected by ASOHF, defined as the highest mass $M_{\rm lim}$ so that the recovered mass functions differs by more than some fraction $1-\alpha$ from the input mass function. These results are listed at a completeness $\alpha = 0.9$ in Table 2, and they are represented as vertical dotted lines in Fig. 1. With increasing resolution, the algorithm is capable of systematically detecting lower-mass haloes. It is able to detect all 1528 haloes when $N_x = 256$ is used.

The precision of this identification is shown in Fig. 2, where we have matched the input and output catalogues (for the $N_x = 256$ case) and test the ability of ASOHF to obtain a precise estimate of radii (left panel), masses (middle panel), and halo centres (right panel). The left and central panels show that we obtain an unbiased estimate of radii and masses on average, although the scatter becomes larger when there are fewer than ∼1000 particles. An amplitude of ∼1.5% for the radius and

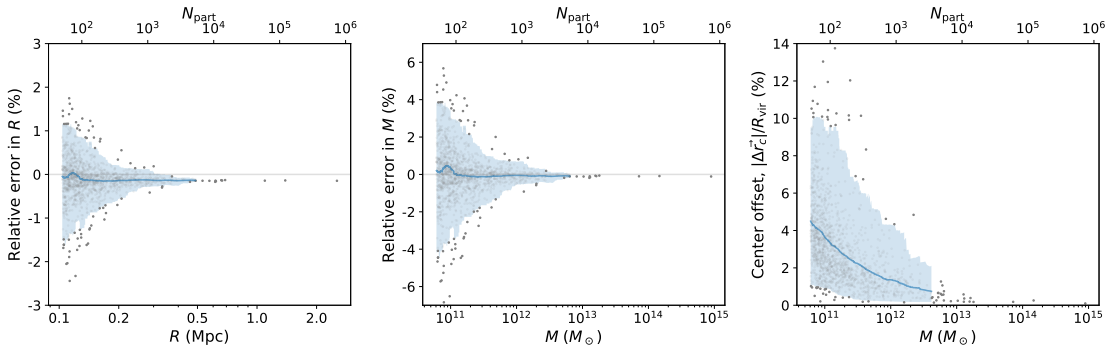


Fig. 2. Precision of ASOHF in recovering basic halo properties in Test 1, with $N_x = 256$. In each panel, dots represent individual haloes, the blue line presents the smoothed trend by using a moving median, and the shaded region encloses the 2σ confidence interval around it. *Left panel*: relative error in the determination of the virial radius. *Middle panel*: relative error in the determination of the virial mass. *Right panel*: centre offset, in units of the virial radius.

~4.5% for the mass, at the 95% confidence level, is reached for haloes of 50–100 particles. This is mostly associated with the uncertainty in determining the halo centre, as shown in the right panel of Fig. 2. For haloes with fewer than ~1000 particles, the median offset between the input and the recovered density peak differs by 2% of the virial radius, which increases up to $\lesssim$10% at the 97.5 percentile error for haloes with ~50 particles. This effect is expected because the density in a sphere with a given comoving radius decreases with decreasing mass for NFW haloes. Related to this, there is an unavoidable source of uncertainty in the test set-up, especially for low-mass haloes, since we are sampling the particles from a probability distribution when creating the test (thus, the nominal centre may differ from the centre of the realisation with particles).

3.2. Test 2. Substructure

To assess the ability of the code to detect substructure, we designed a second test consisting of a large, massive halo rich in substructure, with a similar procedure as in the previous test. In particular, we considered the same domain and cosmology, and placed a halo with a mass of $10^{15} M_\odot$ in its centre. In order to being able to span a wide range in substructure masses, we used $N_{\rm part} = 512^3$ DM particles, each with a mass of $m_{\rm p} = 1.9 \times 10^7 M_\odot$. The particles in this host halo were generated in the same way as in Test 1, up to 1.5 $R_{\rm vir}$.

Subsequently, we placed $N_{\rm subs} = 2000$ substructures, with masses drawn from the same Tinker et al. (2008) mass function, from $M_{\rm min} = 50\, m_p = 9.4 \times 10^8\, M_\odot$ and constraining the sample to have at least one large subhalo, with mass above $10^{13}\, M_\odot$. Subhaloes were placed uniformly inside the host volume, avoiding overlaps, and were populated with particles following the same procedure as for isolated haloes. These particles were superimposed on the particle distribution of the host. The remaining particles up to $N_{\rm part}$ were placed outside the host halo by sampling a uniform random distribution to constitute a homogeneous background. It is worth stressing that even though the particles of each halo were sampled from a spherically symmetric NFW profile, the resulting realisation of the halo can depart strongly from spherical symmetry due to sample variance (especially relevant in smaller haloes, which are the most abundant). Therefore, the test contains many non-spherical haloes, including elongated systems that might resemble tidally stripped haloes.

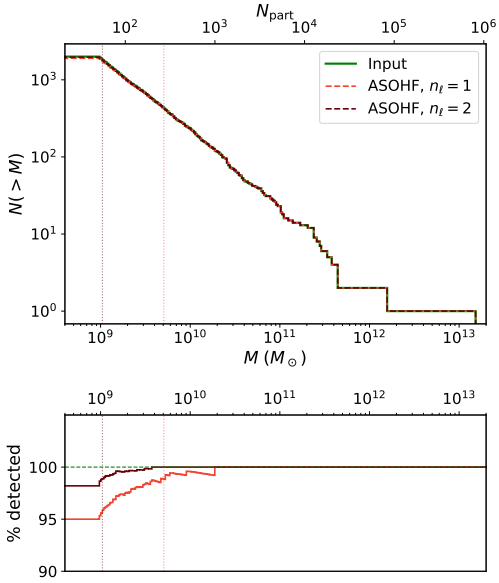


Fig. 3. Results from Test 2. *Upper panel*: input substructure cumulative mass function (green line) compared to ASOHF results with $n_\ell = 1$ (red) and $n_\ell = 2$ (dark red). The vertical lines mark the completeness limit, this time at 99%, of the catalogue produced by ASOHF. *Bottom panel*: fraction of substructures with input mass higher than M detected by ASOHF, with the same colour codes as above.

We ran ASOHF on this particle distribution using a base grid of $N_{\rm x} = 512$ cells in each direction and $n_\ell = 1$ and 2 refinement levels, with a threshold of $n_{\rm min}^{\rm halo} = 25$ particles per cell to flag it as refinable and accepting all patches with at least $N_{\rm min}^{\rm patch} = 14$ cells in each direction. The results for the detection capabilities of ASOHF are shown in Fig. 3. The input and recovered mass functions visually overlap, therefore the results are better assessed in the bottom panel, which shows their quotient. In general terms, just one refinement level allows detecting 95% of the substructures, while using two levels for the mesh increases this fraction

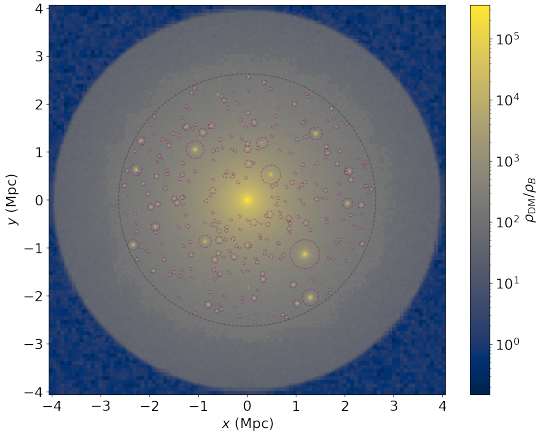


Fig. 4. Projection, 460 kpc thick, of the DM density field interpolated by ASOHF in Test 2, from which the pre-identification of haloes over the AMR grids is performed. The plot shows the maximum value in the projection direction. The black circle marks the virial radius of the host halo, and each purple circle corresponds to a substructure identified by ASOHF. The radius of each circle matches the Jacobi radius of the subhalo.

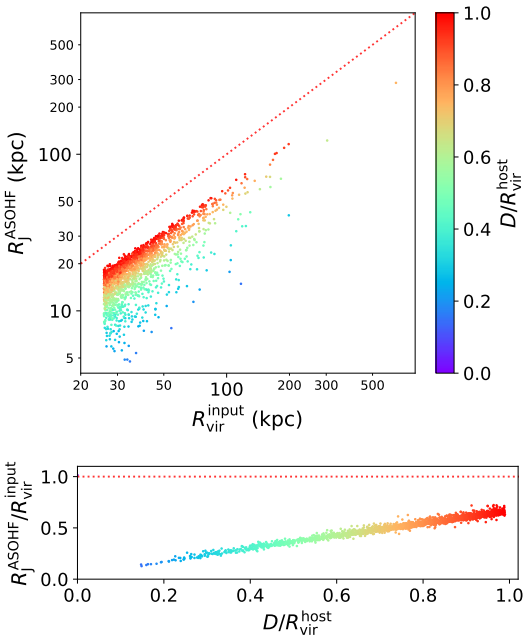


Fig. 5. Relation between virial and Jacobi radii in Test 2. *Top panel*: comparison between the virial radius of the input NFW halo (R_{vir}^{input}) and the Jacobi radius recovered by ASOHF (R_J^{ASOHF}) for all detected substructures. Colours encode the radial position of the substructure centre in the host halo (D) in units of the host virial radius (R_{vir}^{host}). The dotted red line corresponds to the identity relation, $R_J^{ASOHF} = R_{vir}^{input}$. *Bottom panel*: tight linear relation between the quotient $R_J^{ASOHF}/R_{vir}^{input}$ and the radial position of the substructure.

to over 98%. At a more restrictive value of $\alpha = 0.99$, the completeness limits correspond to ∼270 and 55 particles for $n_\ell = 1$ and 2, respectively. Even at a more restrictive threshold, these values are better than those in Test 1 (Table 2) because of the dense environment into which the substructures are embedded. This triggers the refinement of these regions more easily.

Figure 4 presents a thin (∼460 kpc) slice of the density field as computed by ASOHF in order to pre-identify haloes. The adaptiveness allows simultaneously capturing small substructures within the dense host halo while removing sampling noise in underdense regions, which would increase and unbalance the computational cost. Purple circles represent the extent of the substructure, that is, a sphere of radius R_J. We note that in contrast to the virial radius of isolated haloes, the Jacobi radius naturally depends on the location of the substructure within its host: the more central its position, the smaller R_J in relation to the input virial radius of the NFW halo. This is quantitatively shown in Fig. 5, whose upper panel presents the relation between the input virial radius and the Jacobi radius determined by ASOHF. The trend implies that a satellite moving through a host would present a pseudo-evolution of R_J (and its enclosed mass) even if it moved rigidly through the medium (the smaller the distance to the host, the smaller R_J). While approximate due to the spherical symmetry and Keplerian rotation assumptions, this definition for the boundary of a substructure is reasonable because during the dynamical evolution of the infall of a satellite, it is expected that the particles in the outer layers become unbound through dynamical friction with the host particles. For more detailed studies, the whole list of particles of the satellite before its infall can be tracked (similarly to e.g. Tormen et al. 2004). In the lower panel, we show the tight linear correlation between the ratio of radii, R_J/R_{vir}, and the distance to the host centre, D/R_{vir}^{host}. This can be parametrised roughly as

$$\frac{R_J}{R_{vir}} = (0.073 \pm 0.002) + (0.592 \pm 0.003)\frac{D}{R_{vir}^{host}}, \quad (4)$$

although the parameters naturally depend on the particular density profile of the host, and the scatter would be increased by the asphericity of real haloes.

3.3. Test 3. Unbinding

The unbinding procedure is a critical step in all halo finders based on configuration space because the initial assignment of particles has neglected any dynamical information. Here we present two tests to prove the capabilities of the two complementary unbinding procedures implemented in ASOHF: in Sect. 3.3.1 we study the case of a small halo moving in a dense medium, and in Sect. 3.3.2 we consider a fast stream traversing a halo.

3.3.1. Test 3a. Halo moving in a dense medium at rest

We considered a set-up similar to Test 1 (Sect. 3.1) in terms of box size and number of particles. We placed a halo of virial mass $M_h = 5 \times 10^{13}\, M_\odot$ in the centre of the box, and realised it with particles up to a radial distance of 3 $R_{vir} \approx 2.9$ Mpc. This involved 73 976 particles, which are referred to as halo particles; 41 428 of them lie inside the virial radius.

The remaining $\sim 2 \times 10^6$ particles, corresponding to most of the mass in the box, were then placed in uniformly random positions in a sphere of radius 6 R_{vir} and are referred to as background particles. This amounts to a background of constant density

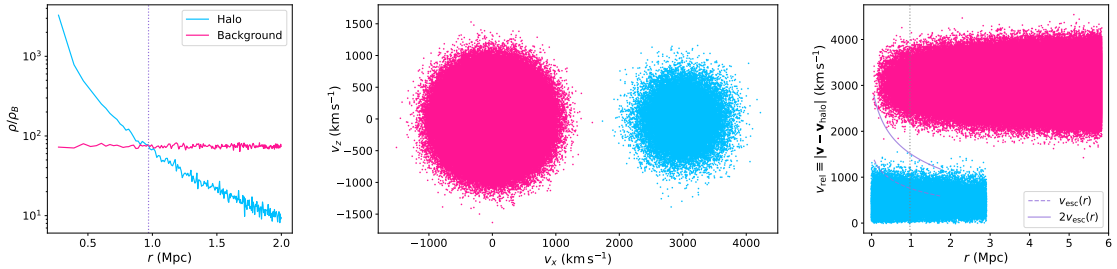


Fig. 6. Results from *Test 3a* (unbinding). *Left panel*: density profile of halo particles (blue) and background particles (pink). The dashed vertical line marks the input virial radius of the halo. *Middle panel*: $v_z - v_x$ phase space, using the same colour coding as in the previous panel. The particle distributions are disjoint in velocity space. *Right panel*: $v_{rel} - r$ phase plot in which the unbinding is performed. The same colour coding as in the previous panels is used, the vertical line marks the input virial radius, and the dashed and solid purple lines are $v_{esc}(r)$ and $2v_{esc}(r)$ at the last iteration of the local velocity speed unbinding, the latter being the threshold velocity for unbinding.

$\sim 75\rho_B$, so that the halo and background densities are approximately equal at the virial radius of the cluster. Therefore, roughly one-fifth of the particles inside the virial volume are background particles. This may bias many of the halo properties (centre of mass position, bulk velocity, angular momentum, etc.).

We gave the halo particles a bulk velocity of 3000 km s^{-1} along the x-axis, so that background particles should be clearly unbound to the halo, and all particles (halo and background) were given a normal velocity dispersion with standard deviation of 300 km s^{-1} to add some noise. We note that it is not the aim of this test to use physically realistic values, but just to show the robustness and performance of the unbinding procedure in a fairly reasonable situation.

Figure 6 summarises the set-up and the results of the test. Its left panel shows the density profiles of the halo particles (blue) and of background particles (pink), which roughly agree with each other at the virial radius of the halo. The central panel in Fig. 6 presents the $v_z - v_x$ phase space with the same colour code and shows that these two components are entirely disentangled in velocity space, but are mixed up in configuration space. Finally, the right panel shows the $v_{rel} - r$ phase plot, where $v_{rel} \equiv |\boldsymbol{v} - \boldsymbol{v}_{halo}|$, which corresponds to the space in which the escape velocity unbinding is performed. In this plot, the dashed purple line represents the local escape velocity at the last iterative unbinding step (when no more new particles are unbound), and the solid purple line is twice this value (which is the last threshold speed used for the escape velocity unbinding, according to the procedure described in Sect. 2.3). The particle-wise results in particular are described below.

The offset between the input centre and the centre detected by ASOHF is $|\Delta \boldsymbol{r}| = 2.3$ kpc, which is 2.3‰ of the virial radius (or 1.5% of the scale radius of the NFW profile). This small miscentrering causes a decrease of 1.3‰ in the virial radius of the host. Two halo particles that were nominally outside the halo are listed as inside, while 29 inside particles appear to be outside.

All background particles were pruned by one of the unbinding methods. Almost all background particles inside the virial volume (all but six) were pruned by the local escape velocity method, as shown by the fact that they lie above the threshold value $2v_{esc}(r)$ (solid purple line) in the right panel of Fig. 6. It is worth noting that the iterative procedure of lowering the threshold, as described in Sect. 2, is crucial to prevent a biased centre-of-mass velocity in the first unbinding steps. The remaining six particles were pruned by the non-local unbinding in velocity space because they lie at more than $3\sigma_v$ of the centre-of-mass velocity. This method is especially useful for unbound components at small halo-centric distances because escape velocities are high near the halo core. Only one halo particle was incorrectly pruned, for being slightly over $3\sigma_v$ of the mean velocity.

This shows the ability of ASOHF of pruning the unbound component of haloes, even when it represents a significant amount of the mass within the halo volume. It also illustrates the situation of a satellite moving through its host. We note that a crucial step to enable unbinding host particles using the iterative procedure described in Sect. 2.3 is that the mass density of halo particles within the radial extent of the halo or satellite is greater than that of background or host particles. Otherwise, the centre-of-mass velocity would converge to that of the background. However, in the case of substructures, this condition is automatically guaranteed by the definition of the Jacobi radius (note that Eq. (2) implies that the substructure is at least three times denser than the mean density of the host within a sphere of radius D).

3.3.2. Test 3b. Fast stream traversing a halo

To illustrate the importance of the complementary velocity space unbinding, we present here a second test, in which a fast stream traverses a halo at rest. The set-up is as follows. A halo of virial mass $M_{vir} = 10^{15}\,M_\odot$ is placed in the centre of a box, using the same particle mass ($m_p \approx 1.2 \times 10^9\,M_\odot$) as in the previous test. The halo is realised up to 3 R_{vir} with $\sim 1.5 \times 10^6$ particles (the halo particles). As for the stream of particles, we considered a curved cylindrical stream, with impact parameter $b = 0.5\,R_{vir}$, curvature radius $r = 2\,R_{vir}$, radius $\Delta b = 250$ kpc, and length $L \approx 8.3$ Mpc (so that it crosses the whole virial volume of the halo). We assigned a density of $200\rho_B$ to the stream, so that it amounts to a mass of $1.29 \times 10^{13}\,M_\odot$, or slightly over 10 000 particles (the stream particles). The situation in configuration space is depicted in the left panel of Fig. 7, where blue (pink) dots represent the $x - z$ positions of halo particles (stream particles) lying within a 20 kpc slice passing through the centre of the halo.

Velocities for the halo particles were drawn from a normal distribution with zero mean and an isotropic $\sigma_{1D} = 300$ km s^{-1}. Stream particles were given a speed of 3000 km s^{-1} along the axis of the stream, plus an isotropic dispersion component as in the halo particles. The middle panel of Fig. 7 shows that while most of the stream particles are nominally unbound, ($v > v_{esc}(r)$),

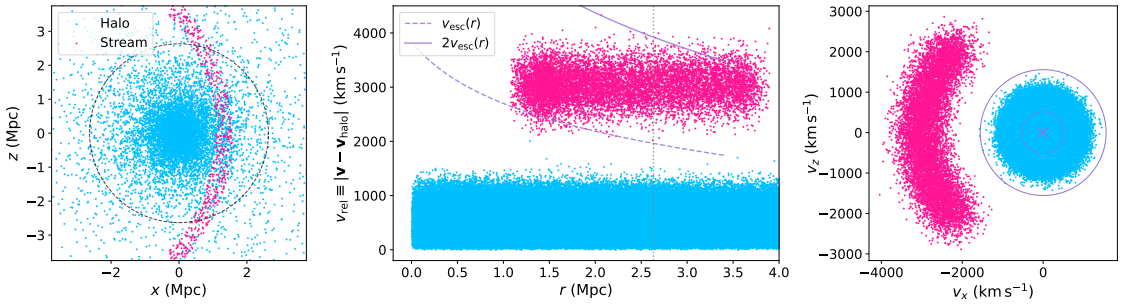


Fig. 7. Results from Test 3b (unbinding). *Left panel*: particles in a 20 kpc slice through the centre of the halo. Blue (pink) dots refer to halo (stream) particles. The dashed grey circle corresponds to the location of the input virial radius. *Middle panel*: $v_{rel} - r$ phase plot, with the same colour coding as in the previous panel, showing that the local escape velocity unbinding does not succeed in pruning the stream. *Right panel*: $v_z - v_x$ phase space, keeping the colour coding as in the previous plots. The purple cross, dashed circle, and solid circle indicate the converged centre-of-mass velocity and the $1\sigma_v$ and $3\sigma_v$ regions, respectively. Particles outside the latter are pruned by the velocity space unbinding.

they do not reach the threshold for unbinding, $2v_{esc}$. While lowering this threshold would remove these particles, loosely unbound particles (i.e. those with $v_{esc}(r) < v < 2v_{esc}(r)$) may still remain within the halo for a dynamical time. This means that lowering the threshold is somewhat aggressive (see also the discussion in Knebe et al. 2013 and Elahi et al. 2019).

However, the situation in velocity space clearly presents two disjoint components, as represented in the right panel of Fig. 7. In this case, the velocity space unbinding is able to remove all stream particles because they lie beyond $3\sigma_v$ of the centre-of-mass velocity after the iterative procedure.

3.4. Scalability

While the previous set of tests demonstrates the ability of ASOHF to provide complete samples of haloes with unbiased properties, here we consider the performance of the code in terms of execution time and memory requirements in more detail. We used a set-up similar to that of Test 1 to assess how the code scales with the number of particles and base grid size (Sect. 3.4.1), and the performance increase with the number of OMP threads (Sect. 3.4.2). All the results given here correspond to the performance of ASOHF on a Ryzen Threadripper 3960X processor with 24 physical cores.

3.4.1. Scaling with N_{part} and N_x

To assess the scaling capabilities of the code with increasing number of particles and size of the base grid, we replicated the set-up in Test 1 (Sect. 3.1) with varying number of particles. To do this, we first considered the case with the largest number of particles, $N_{part} = 1024^3$ (corresponding to a particle mass of $m_p^{1024^3} = 2.4 \times 10^6 M_\odot$) and generated and randomly placed 366 652 haloes with masses higher than $M_{min} = 50\, m_p^{1024^3}$. Then, we realised this halo catalogue with $N_{part} = 32^3$, 64^3, etc. up to 1024^3 DM particles. For each realisation, we only kept the haloes with masses higher than or equal to that corresponding to 50 particles. For each value of N_{part}, we tested $N_x = \sqrt[3]{N_{part}}$ and $N_x = 2\sqrt[3]{N_{part}}$[11] because Test 1 (Sect. 3.1) showed that this latter value was able to recover all haloes. For the AMR grid parameters, we used $n_{part}^{refine} = 8$, and the remaining parameters were kept as in Test 1. The detailed results are presented in Table 3 and are graphically summarised in Fig. 8.

The left panel in Fig. 8 shows the scaling of the 90% mass completeness limit with the number of particles for the two base grid sizes ($N_x = \sqrt[3]{N_{part}}$ in dark red, labelled '*a*' in Table 3; and $N_x = 2\sqrt[3]{N_{part}}$ in light red, labelled '*b*'; the same colour code is kept in the remaining panels). The dashed grey lines are lines of constant number of particles (at fixed total mass in the box), so that it is explicitly shown that using $N_x = \sqrt[3]{N_{part}}$ (runs *a*) the code is able to correctly identify barely all haloes comprising more than ∼140 particles. For the runs *b*, with $N_x = 2\sqrt[3]{N_{part}}$, all haloes were detected and the completeness limit is arbitrarily placed at the minimum mass of the input mass function (50 particles) for representation purposes. Nevertheless, it is worth stressing that in real situations, where an LSS component surrounds the haloes, it is not generally necessary to increase the base grid resolution beyond $\sqrt[3]{N_{part}}$ to detect more smaller haloes, although this may be useful for some particular applications (e.g. identifying haloes within voids).

The wall time taken by each of the runs is presented in the central panel of Fig. 8. Runs 4.1, 4.2, and 4.3, with fewer than one million particles, last for less than a few seconds. These results are therefore biased with respect to the general scaling because the measured times are likely to be dominated by input-output, system tasks, and by the coarse time resolution of the profiling utility used. For large enough task sizes (i.e. for $N_{part} \gtrsim 256^3$), the wall time increases proportionally to the product $N_{part} \times N_{haloes}$, as shown by the dashed lines, which correspond to constant time per halo and particle. When the base grid resolution is increased from $N_x = \sqrt[3]{N_{part}}$ to $N_x = 2\sqrt[3]{N_{part}}$, the scaling is kept, but the normalisation of the relation increases by a factor of 2–3, mostly associated with the fact that more haloes are detected in this case (both real haloes and spurious peaks throughout the interpolated density field, which are later discarded when particles are considered).

The right panel in Fig. 8 depicts the peak memory requirements of the code. When the grid size is fixed at $N_x = \sqrt[3]{N_{part}}$, ASOHF requires ∼180−200 bytes/particle, which allows running the code on desktop-sized workstations for simulations with up to hundreds of millions of particles. When the base grid resolution is doubled, these memory requirements increase up to ∼1 kbyte/particle.

[11] This is equivalent to setting the base grid cell size to the mean particle separation and to half this value, respectively.

Table 3. Results of the scalability tests.

Test		$N_{\rm part}$	$M_{\rm min}^{\rm input}$ ($M_\odot$)	$N_{\rm haloes}^{\rm input}$	N_x	$N_{\rm haloes}^{\rm ASOHF}$	$M_{\rm lim}^{90\%}$ ($M_\odot$)	$n_{\rm part,lim}^{90\%}$	Wall time	Peak RAM	(bytes/part.)
4.1	a	32^3	3.86×10^{12}	31	32	10	9.52×10^{12}	123	10 ms	12.6 MB	385
	b				64	31	All	All	40 ms	36.0 MB	1098
4.2	a	64^3	4.83×10^{11}	213	64	79	1.23×10^{12}	127	280 ms	49.0 MB	187
	b				128	214	All	All	290 ms	247 MB	940
4.3	a	128^3	6.03×10^{10}	1263	128	495	1.73×10^{11}	143	1.64 s	385 MB	184
	b				256	1263	All	All	2.1 s	1.98 GB	946
4.4	a	256^3	7.54×10^{9}	8206	256	3099	2.06×10^{10}	136	19.3 s	3.18 GB	190
	b				512	8211	All	All	37.9 s	15.1 GB	903
4.5	a	512^3	9.43×10^{8}	54 476	512	20 617	2.58×10^{9}	137	13 min 38 s	25.5 GB	190
	b				1024	54 495	All	All	40 min 39 s	155 GB	1155
4.6	a	1024^3	1.18×10^{8}	366 653	1024	136 411	3.25×10^{8}	137	13 h 41 min	223 GB	208

Notes. Each row corresponds to a particular test, characterised by the number of particles ($N_{\rm part}$) and the base grid size (N_x). For each $N_{\rm part}$, the mass function is truncated at $M_{\rm min}^{\rm input}$ and $N_{\rm haloes}^{\rm input}$ are thus generated. For each run, we report the number of haloes detected by `ASOHF` ($N_{\rm haloes}^{\rm ASOHF}$), the 90% completeness limit (in mass and number of particles; $M_{\rm lim}^{90\%}$ and $n_{\rm part,lim}^{90\%}$, respectively), the execution (wall) time and the peak RAM usage. The last column gives the peak RAM per particle.

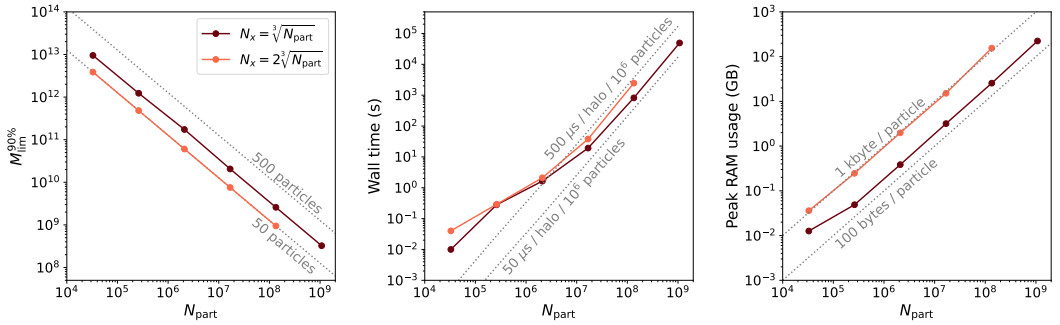


Fig. 8. Summary of the results from the scalability test (Sect. 3.4.1). *Left panel*: scaling of the 90% completeness limit (in terms of mass) with the number of particles in the domain. Dashed grey lines correspond to a constant number of particles. Dark red (light red) lines represent the results for the two base grid sizes according to the legend. *Middle panel*: scaling of the wall time taken by ASOHF with number of particles. Dashed grey lines correspond to a scaling $\propto N_{\rm part} N_{\rm haloes}^{\rm input}$. *Right panel*: scaling of the maximum RAM used ASOHF with number of particles. Dashed grey lines correspond to a constant amount of RAM per particle.

Last, we note that as the number of particles increases, the scaling $\propto N_{\rm part} N_{\rm haloes}$ can greatly benefit from performing a domain decomposition for running `ASOHF`, as described in Sect. 2.6.1. This is especially useful in large domains, where the required overlaps amongst domains (which can be as low as the size of the largest expected halo) correspond to a negligible fraction of the volume. Therefore, when the volume is decomposed in d domains, the CPU time is reduced by a factor d, since the number of particles and haloes in each domain are themselves reduced by a factor d, on average. If all domains can be run concurrently, rather than sequentially, this amounts to an improvement in wall time of a factor of d^2.

3.4.2. Scaling with the number of threads

To explore the performance gain in terms of wall time and memory usage of the OMP parallelisation scheme, we repeated Test 4.5a (see Table 3) with $n_{\rm cores} = 1$, 2, 4, 8, 16, 24, and 32 and 36 OMP threads using nodes equipped with two 18-core CPU Intel® Xeon® Gold 6154. The results are presented in Fig. 9.

The upper panel exemplifies the performance improvement when the number of cores is increased. Red crosses correspond to the actual performance of `ASOHF` with a varying number of threads. We fitted these data to the functional form

$$\Delta t_{\rm Wall} = (\Delta t_{\rm CPU})_{\rm seq} + \frac{(\Delta t_{\rm CPU})_{\rm par}}{n_{\rm cores}^{\alpha}} \tag{5}$$

using a least-squares method, finding that the sequential part amounts for $(\Delta t_{\rm CPU})_{\rm seq} \approx 3.4$ min, while the parallel part would correspond to a CPU time of $(\Delta t_{\rm CPU})_{\rm par} \approx 5.4$ h. Being $(\Delta t_{\rm CPU})_{\rm seq} \ll (\Delta t_{\rm CPU})_{\rm par}$, the scaling of the code is close to optimal (null sequential part, which is represented by the green line in the upper panel of Fig. 9) for $n_{\rm cores} \lesssim 100$, which is a reasonably high number of threads available in typical shared memory nodes. The exponent α resulting from the fit is $\alpha = 1.007 \pm 0.017 \cong 1$, which is consistent with the expected behaviour of the parallel part. It is interesting to note the absence of a significant performance gain when increasing from 16 to 24 cores because the system is comprised of two non-uniform memory access (NUMA) nodes, with 18 physical cores each. When the treads are increased from 16 to 24, the task can no longer be allocated to a single node, so that memory access outside the node is penalised. This is a well-known issue in shared-memory

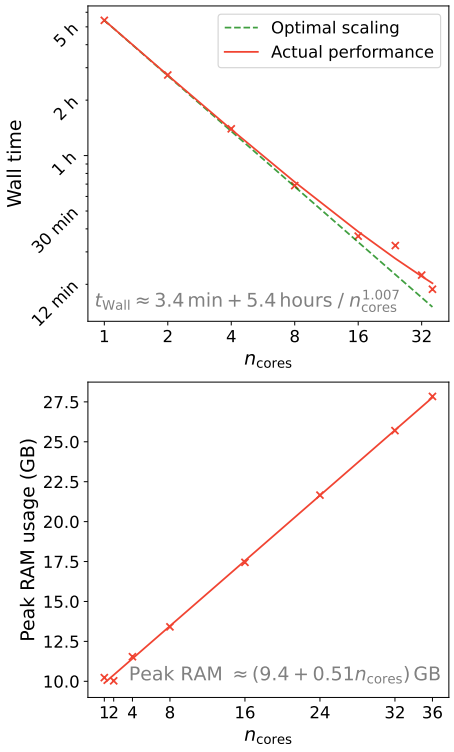


Fig. 9. Parallel performance of ASOHF. *Upper panel*: scaling of the wall time of ASOHF in Test 4.5a (see Table 3) when the number of OMP threads is varied. The solid red line is a fit of the benchmark data (red crosses), and the dashed green line corresponds to an optimal scaling i.e. $t_{\rm Wall} \propto 1/n_{\rm cores}$. *Lower panel*: peak memory usage of ASOHF scaling with the number of OMP threads.

systems, and we advise the interested users to take the architecture of their system into account for optimal performance.

However, increasing the number of OMP threads implies an increased memory usage because it is necessary to replicate part of the data. The lower panel in Fig. 9 displays the (linear) relation between the number of threads and the peak RAM used by the job. To conclude this section, we recall that these results are dependent on the application and the configuration. For example, it is reasonable to expect a smaller memory penalty by increasing the number of threads in simulations of large volumes, where each halo contains only a small fraction of the particles in the domain, as opposed to these test cases, in which a halo contains nearly 25% of the particles.

4. Tests on real simulation data and comparison with other halo finders

Last, in order to evaluate the performance of ASOHF on a cosmological simulation, we have tested our code on outputs from the public suite CAMELS (Villaescusa-Navarro et al. 2021, 2022). The Cosmology and Astrophysics with MachinE Learning Simulations project consists of over 4000 simulations of $(25h^{-1}\,{\rm Mpc})^3 \approx (37.25\,{\rm Mpc})^3$ cubic domains, including DM-only and (magneto)hydrodynamic (MHD) simulations with star formation and feedback mechanisms, and covering a broad space of several cosmological and astrophysical parameters. Alongside with the simulation data, the public release also includes halo catalogues produced with three public halo finders, namely SUBFIND (Springel et al. 2001; Dolag et al. 2009), AHF (Gill et al. 2004; Knollmann & Knebe 2009), and ROCKSTAR (Behroozi et al. 2013), which we use here to examine how does ASOHF compare to other well-known halo finders.

In particular, we used the simulation LH-1 from the IllustrisTNG subset of CAMELS. These simulations were carried out using Arepo (Springel 2010; Weinberger et al. 2020), which is a publicly available code using a tree+particle-mesh (TreePM) scheme coupled to Voronoi moving-mesh MHD for hydrodynamical cosmological simulations, including galaxy formation physics. The TreePM (Bagla 2002) method implemented in Arepo, which combines a particle-mesh method for computing the large-range force with a more accurate tree code (Barnes & Hut 1986) at short distances, allows these simulations to host very many small haloes and substructures even though the number of DM particles, $N_{\rm part}^{\rm DM} = 256^3$, is modest. The comoving gravitational softening length is as low as 2 kpc. While all CAMELS simulations correspond to flat universes with a baryon density parameter $\Omega_b = 0.049$, Hubble constant $h \equiv H_0/(100\,{\rm km\,s^{-1}\,Mpc^{-1}}) = 0.6711$, and spectral index $n_s = 0.9624$, the simulation we chose, name-coded LH-1, assumes a matter density parameter $\Omega_m = 0.3026$ and $\sigma_8 = 0.9394$ as the amplitude of the primordial fluctuations spectrum. These parameters imply a DM particle mass of $m_{\rm part} = 9.8 \times 10^7\,M_\odot$.

4.1. Results analysing only DM particles

We ran ASOHF using only DM particles on the most recent snapshot of this simulation (at redshift $z = 0$), using a base grid with as many cells as particles ($N_x = 256$) and $n_\ell = 6$ refinement levels, so that the peak resolution of the grid is $\Delta x_6 \approx 2.3$ kpc, similar to the gravitational softening length of the simulation. The threshold particle number to mark a cell as refinable was set to 3, and the minimum patch size was set to 14 cells. Density was interpolated with a kernel size determined by the local density, using four kernel levels. We discarded all haloes resolved with fewer than 15 particles. In this configuration, ASOHF detects 11 794 non-substructure haloes, the most massive of them with a DM mass of $7.44 \times 10^{13} M_\odot$, and 1263 substructures (including sub-substructures). The wall-time duration of the test was below 4 min, using eight cores in the same architecture as described in Sect. 3.4.1.

The halo mass function recovered by ASOHF is shown as the solid thick red line in the upper panel of Fig. 10, together with the same quantity obtained from the catalogues of AHF (solid blue), Rockstar (solid green), and Subfind (solid orange). For reference, the thick pale purple line represents a reference Tinker et al. (2008) mass function with the cosmological parameters of the simulation. At first glance, all four halo finders yield comparable mass functions (both amongst them and with the reference one) for $N_{\rm part} \gtrsim 1000$, with some variations at the high-mass end due to the combination of small number counts and subtleties in the mass definitions.

To allow a more precise comparison, the lower panel presents each of the mass functions, normalised to the geometric mean of all four finders, which we take as the baseline for comparison. In this plot, the red line (corresponding to ASOHF) displays the Poisson ($\sqrt{N}$) confidence intervals as the shadowed red area. The magnitude of these uncertainties is similar for the other lines. ASOHF, Subfind, and Rockstar match reasonably well

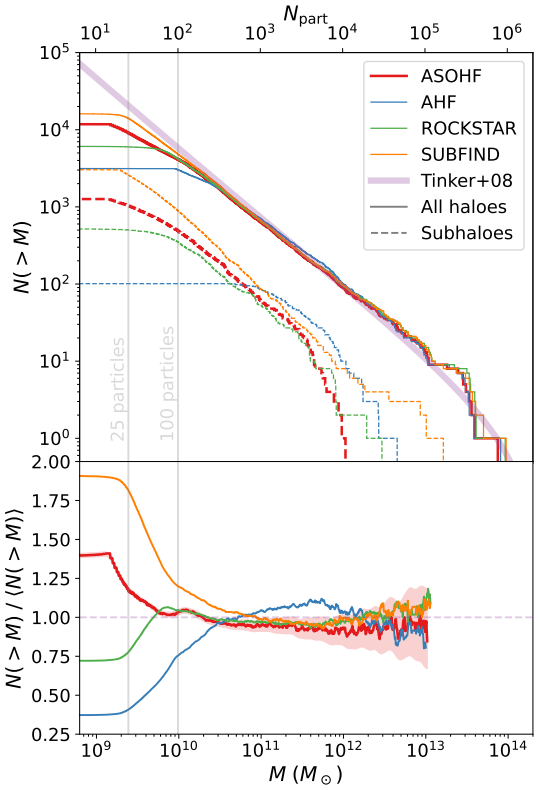


Fig. 10. Comparison of the mass functions of the halo catalogues obtained by ASOHF (red), AHF (blue), Rockstar (green), and Subfind (orange) in the CAMELS Illustris-TNG LH-1 simulation at $z = 0$. *Upper panel*: mass functions obtained by each code, as solid lines. The thick purple line corresponds to a Tinker et al. (2008) mass function at $z = 0$ and with the cosmological parameters of the simulation for reference. Dashed lines present the mass function of subhaloes in the same colour scale. *Lower panel*: mass function of non-substructure haloes, normalised by the geometric mean of this statistic for the four halo finders. The shadow on the red line (ASOHF) corresponds to $\sqrt{N}$ errors in halo counts. The errors are similar for the other finders.

inside their respective confidence intervals, while AHF shows a small $\gtrsim$15% excess with respect to them at $N_{\rm part} \sim 10^4$. Interestingly, Rockstar and ASOHF present the largest similarities in their mass function. They match each other within a few percents all the way down to 50−100 particles, when ASOHF starts to identify a larger number of structures. Compared to them, Subfind starts to present a larger abundance of haloes below $N_{\rm part} \lesssim 1000$, mostly driven by the larger number of substructures, while the AHF halo counts stall below a few hundred particles.

Dashed lines in the upper panel of Fig. 10 present the substructure mass functions for the four halo finders, using the same colour code. In this case, substructure mass functions are not so easily comparable, and thus the differences amongst the finders are exacerbated (see e.g. Onions et al. 2013). In particular, Subfind finds the largest number of substructure, nearly 2.5 times as many as ASOHF, 6 times more than Rockstar and 30 times more than AHF. Even the high-mass end of the subhalo mass function shows important differences, with Subfind having the highest and ASOHF the lowest mass of their respective most massive substructure. This is mainly due to the fundamentally different definitions of substructure. The ASOHF choice of using the Jacobi radius, which approximately delimits the region in which the gravitational attraction of the substructure is stronger than that of the host, is more restrictive than other definitions, for example, using density contours traced by the AMR grid (AHF), reducing the 6D linking-length (Rockstar), or finding the saddle point of the density field (Subfind). This more stringent definition of the substructure extent may indeed also be the reason that ASOHF recovers less substructure than Subfind because many of the small substructures identified by the latter may contain fewer than 15 particles using our more stringent definition based on the Jacobi radius.

Focusing on the main properties of haloes, in Fig. 11 we present the comparison of ASOHF radii and masses to the corresponding magnitudes obtained by AHF (left column), Rockstar (middle column), and Subfind (right column). For each halo finder, we first matched its catalogue to the catalogue produced by ASOHF. Here, we only focus on haloes and exclude substructure because, as mentioned above, the substructure definitions of each algorithm are fundamentally different. The results are also summarised in Table 4.

Surprisingly, and despite their agreement regarding the mass functions, we are only able to match ∼20% of Rockstar haloes to ASOHF's[12], in contrast to ∼80% of matches when comparing to AHF and Subfind. For each property, X, that is to be compared, we fitted the matched data to a power law of the form

$$\log \frac{X^{\rm finder}}{X^*} = \log N + \alpha \log \frac{X^{\rm ASOHF}}{X^*}, \tag{6}$$

where N is the normalisation and α is the index (both are unity if the finders yielded identical results). X^* is a normalisation for the quantities, which we selected as the median over the ASOHF sample: $R^* = 43.2$ kpc and $M^* = 5.18 \times 10^9\, M_\odot$. For each fit, we computed the scatter as the RMS of the relative residuals between the data points and the fit,

$$s = \sqrt{\frac{1}{N_{\rm matched}} \sum_{i=1}^{N_{\rm matched}} \left(1 - \frac{\hat{X}^{\rm finder}(X^{\rm ASOHF})}{X^{\rm finder}}\right)^2}, \tag{7}$$

where $\hat{X}^{\rm finder}(X^{\rm ASOHF})$ is the fitting function from Eq. (6) with the best parameters obtained by least-squares estimation.

The largest discrepancies with ASOHF are seen when it is compared to AHF, which overall finds ∼6% larger radii than ASOHF. This may be well due to the fact that AHF uses all particles (including gas and stars) to determine the spherical overdensity radius, while we only use DM particles. This normalisation offset in radii propagates to a ∼22% offset in masses. There is a significant amount of scatter in radii and masses at the low-mass end, with a population of matched haloes with AHF radii a few times that of ASOHF. This is not seen when comparing with the rest of halo finders.

Although only ∼1200 haloes are matched between Rockstar and ASOHF catalogues, their virial radii and masses match tightly, with ∼4 and 5% offsets, respectively. A few outliers increase the overall scatter figure, which reaches

[12] This is due to some systematic off-centring between the halo centres reported by Rockstar and those reported by the other finders (as well as ASOHF) in the CAMELS suite. Discussing the origin of this off-centring is beyond the scope of this paper.

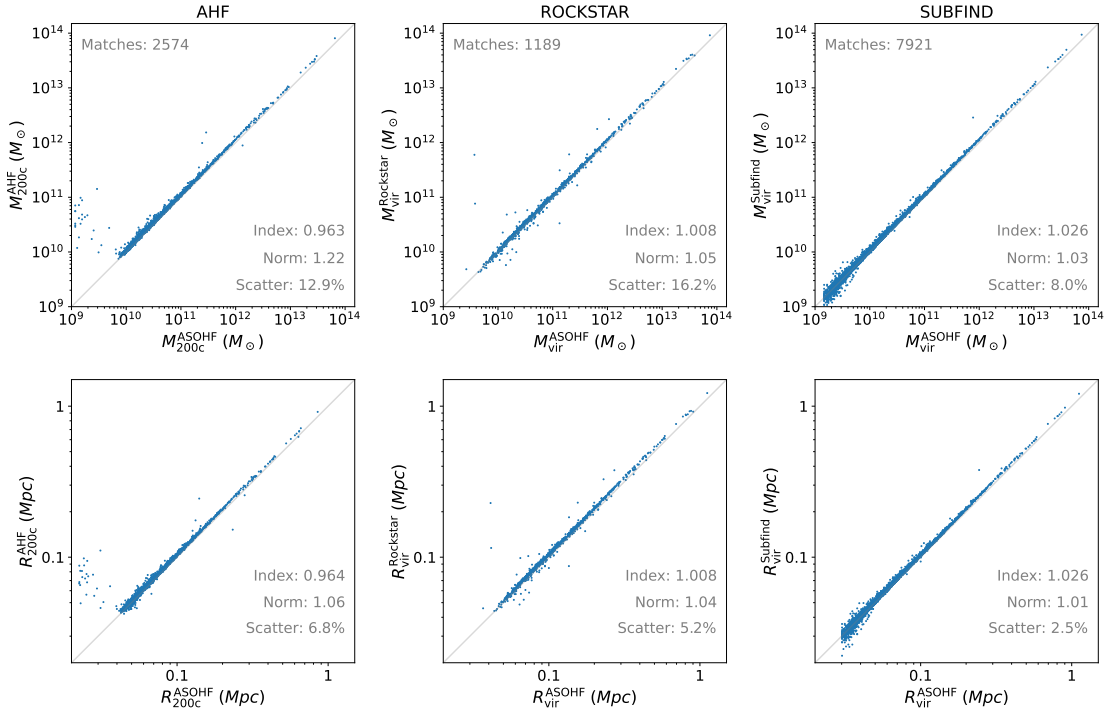


Fig. 11. Comparison of the main properties (masses, *top row*; and radii, *bottom row*) of the matched haloes between ASOHF and AHF (*left column*), Rockstar (*middle column*), and Subfind (*right column*). For the two latter we compare virial radii and masses, and for the former, we use R_{200c} and M_{200c} because the AHF catalogues given in the public release of CAMELS have used this overdensity for marking the boundary of the halo.

Table 4. Summary of the main features of the comparison between ASOHF, AHF, Rockstar, and Subfind.

Halo finder	Halo statistics				Radii fit			Masses fit		
	$N_{\rm haloes}$	$N_{\rm subs}$	$M_{\rm max}$ ($10^{13}M_\odot$)	max($n_{\rm subs}$)	N_R	α_R	s_R	N_M	α_M	s_M
ASOHF	11 794	1263	7.44	122	–	–	–	–	–	–
AHF	3141	101	8.07	7	1.0595(33)	0.9639(45)	6.8%	1.219(13)	0.9627(48)	12.9%
Rockstar	6089	519	9.14	50	1.0369(43)	1.0077(38)	5.2%	1.051(12)	1.0077(38)	16.2%
Subfind	16 129	3037	9.38	343	1.00968(33)	1.02632(61)	2.5%	1.03081(96)	1.02632(61)	8.0%

Notes. From left to right, the columns contain the total number of detected haloes ($N_{\rm haloes}$), the total number of substructures ($N_{\rm subs}$), the mass of the largest halo ($M_{\rm max}$), the maximum number of substructures of a single halo (max($n_{\rm subs}$)), and the normalisation, index, and scatter (N_X, α_X, and s_X as defined in Eqs. (6) and (7)) of the radii and mass comparison between each finder and ASOHF.

~5 and ~16% for radii and masses, respectively. When they are excluded, the relation is quite tight.

Strikingly, despite their very different natures, the Subfind results are the most compatible with ASOHF of the halo finders we considered. With almost 8000 matched haloes, the normalisations as defined in Eq. (6) only show a 1% offset in radii and a 3% offset in mass. The scatter, dominated by the low-mass haloes, is 2.5 and 8%, respectively.

Overall, and despite the difficulties of comparing results of different halo finders on a halo-to-halo basis, which has been thoroughly explored in the literature (see Knebe et al. 2011, 2013; Onions et al. 2012), these analyses show that ASOHF is capable of providing results that generally agree well with other widely used halo finders. The largest discrepancies are seen at the smallest scales.

4.2. Results including stellar particles

By the last snapshot, at $z = 0$, the simulation contains 227 454 stellar and black hole (BH) particles, accounting for ~2‰ of the DM mass in the computational domain. While the mass in stars is too low to significantly impact the process of halo finding (by enhancing DM density peaks that would otherwise be too small to be detected), we can still demonstrate the ability of ASOHF of characterising stellar haloes.

We ran ASOHF on the DM+stellar particles of the same simulation and snapshot as considered above and kept the same values for the mesh creation and DM halo-finding parameters. The stellar halo-finding parameters, which as discussed in Sect. 2.5 do not affect the resulting galaxy catalogues strongly, were set such that all stellar haloes with more than $N_{\rm min}^{\rm stellar} = 15$ stellar

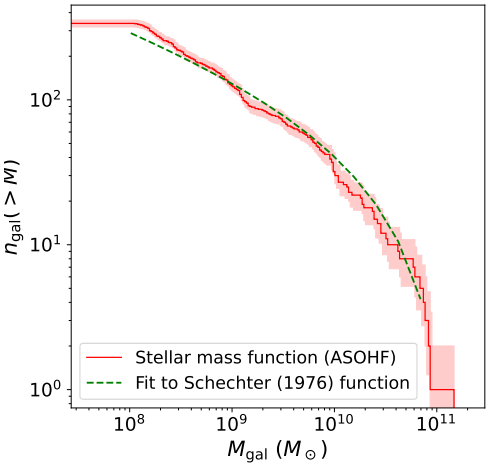


Fig. 12. Mass function of stellar haloes found in the simulation from CAMELS at $z = 0$ (red line). The mass of the galaxy on the horizontal axis, $M_{\rm gal}$, refers to the stellar mass inside the half-mass radius of the stellar halo. The red shading represents the Poissonian, $\sqrt{N}$ uncertainties. The dashed green lines corresponds to the best least-squares fit of the differential mass function to a Schechter (1976) parametrisation.

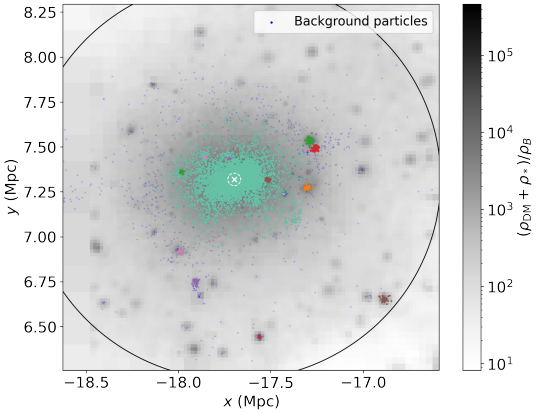


Fig. 13. Example of the stellar halo finding capabilities of ASOHF. The background greyscale colour map is a projection of DM+stellar density around the most massive halo in the simulation. The solid black line indicates its virial radius. The coloured dots correspond to stellar particles, and each colour corresponds to a different stellar halo. The dark blue dots are the particles inside the virial volume of the DM halo that do not belong to any identified stellar structure. For the central galaxy, the white cross and circle represent the stellar density peak (which we regard as centre) and the half-mass radius of the stellar halo, respectively.

particles inside the half-mass radius were kept. We preliminarily cut the stellar haloes whenever the density increased by more than a factor $f_{\rm min} = 5$ from the profile minimum, there was a radial gap of more than $\ell_{\rm gap} = 10$ kpc without any stellar particle or stellar density fell below the background density, $\rho_B(z)$ ($f_B = 1$).

In this configuration, ASOHF identifies 337 stellar haloes, with masses ranging between $1.1 \times 10^8 M_\odot$ and $1.5 \times 10^{11} M_\odot$, and whose cumulative mass distribution is shown in Fig. 12 by a red line (here, $M_{\rm gal}$ makes reference to the stellar mass inside the stellar half-mass radius, defined in Sect. 2.5). The dashed green line presents a fit to a Schechter (1976) function. The fit was performed over the differential mass function, and obtained a chi-squared per degree of freedom $\chi^2_\nu \approx 1.8$ (computed assuming Poissonian errors for number counts). This implies that the fit is not inconsistent with a Schechter (1976) mass function.

An example of such an identification is represented graphically in Fig. 13. In this figure, the background colour encodes the DM+stellar projected density in grey scale, as computed by ASOHF, around the most massive DM halo in the simulation (which is a $\sim 8 \times 10^{13} M_\odot$ group). Each colour represents all the particles of a given stellar halo. A central halo, represented in turquoise, dominates the stellar mass budget. This halo has a half-mass stellar radius of ∼36 kpc, whose extent is represented as the dashed white circle in the figure, around a white cross marking the stellar density peak. Around it, ASOHF detects ten satellite haloes, each represented in a different colour. In dark blue, we plot all the particles inside the virial volume of the main DM halo that do not belong to any stellar halo. Although a few groups of stars on top of some DM haloes that are not identified as galaxies are visible, these objects correspond to poor stellar haloes, with fewer than 15 stellar particles. They were therefore discarded by ASOHF.

We summarise some of the properties of the stellar haloes catalogue in Fig. 14. The left panel contains a histogram of the half-mass stellar radii. The distribution of galaxy sizes peaks at around (6−8) kpc, while a handful of haloes as large as $R_{1/2} \gtrsim 20$ kpc are found, most typically associated with massive DM haloes. Although more rare, some of them correspond to small ($\sim 10^{11} M_\odot$) DM haloes with an extended stellar component.

In the middle panel of Fig. 14, we present the relation of stellar mass inside the half-mass radius (vertical axis) to DM host halo mass (horizontal axis). For reference, we indicate some constant stellar fractions as grey lines. At low stellar mass ($M_{*,1/2} \lesssim 2 \times 10^9$), we find a large scatter between stellar and DM halo masses. The relation becomes tighter at high stellar halo masses, approaching but not reaching 1%. The particularities of these results may have to do with the particular IllustrisTNG feedback scheme implemented in the simulation and numerical resolution. Discussing them is beyond the scope of this work.

Last, the right panel of Fig. 14 contains the stellar 3D velocity dispersion correlated to the stellar mass of the halo. This relation shows a clear increasing trend because more massive haloes (if virialised) have larger kinetic energy budgets or are dynamically hotter. This result provides a dynamical confirmation of the fact that our galaxy identification scheme targets bona fide stellar haloes and not spurious groups of particles. In sum, with these examples, we show the capabilities of ASOHF, not only as DM halo finder, but also as a galaxy finder for cosmological simulations.

5. Conclusions

We have introduced and discussed a revised version of the DM halo finder ASOHF, which constitutes a profound revision of most of the features of the original version that was presented over a decade ago by Planelles & Quilis (2010). The ever-increasing trend in cosmological simulation sizes and resolution, and the accordingly greater abundance and richness of structures at a wide range of scales, required revisiting several aspects related

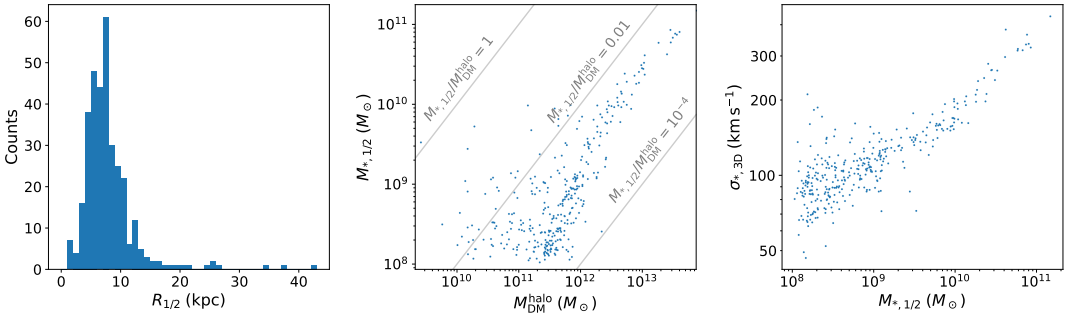


Fig. 14. Summary of some of the properties of the stellar haloes identified by ASOHF in the CAMELS, IllustrisTNG LH-1 simulation at $z = 0$. *Left panel*: distribution of half-mass radii. *Middle panel*: relation of stellar mass to DM halo mass. Diagonal lines correspond to constant stellar fractions. *Right panel*: relation of 3D velocity dispersion to stellar mass.

to the halo-finding process. Because they are the basic building blocks of the LSS of the Universe, an exceptionally precise characterisation of DM haloes is a mandatory condition for bridging the gap between simulations and observations and bringing cosmological predictions from simulations to accuracies below the percent level (e.g. Borgani 2008; Clerc & Finoguenov 2022, for several reviews about cluster cosmology), especially when the smallest scales are involved (e.g. Vogelsberger et al. 2020, and references therein). Related to this, given that the baryonic processes leading to galaxy formation are being exhaustively analysed in numerical works (e.g. Murante et al. 2010; Iannuzzi & Dolag 2012; Vogelsberger et al. 2014; Schaye et al. 2015; Kaviraj et al. 2017; Pillepich et al. 2018; Villaescusa-Navarro et al. 2021, for some large galaxy formation simulation projects), it is of the utmost importance to devise numerical techniques to describe them accurately.

Amongst the main improvements to the halo-finding procedure that were assessed in a battery of complex, idealised tests and in a cosmological simulation, we summarise the main points below. They can be useful not only for ASOHF, but for halo finders in general (especially for those that are based on configuration space).

1. The density interpolation procedure is, naturally, critical to the procedure of halo finding because an overly smoothed interpolation would accidentally smear true density peaks, thus losing small-scale structures, while a density interpolation that is too sharp (e.g. applying a CIC or TSC at the resolution of each grid) increases the computational cost of the process dramatically by producing a large number of spurious peaks due to sampling noise. A compromise between these extremes is achieved by spreading particles in a cloud, whose size can be tuned either by local density or by particle mass (i.e. by the volume sampled by the particle in the initial conditions), but is independent of the grid level. Nevertheless, since any interpolation will tend to be more diffuse than the particle distribution itself, it is generally harmless to also consider density peaks above a certain fraction of $\Delta_{\rm vir}(z)$ as candidates for halo centres.

   Because the grid resolution is finite (one cell can contain thousands of particles even at the finest levels for very cuspy haloes), the implementation of a recentring scheme (using information at finer levels, and ultimately using the particle distribution to iteratively relocate the local maximum of the density field) is crucial to avoid centre offsets. Preventing this erroneous centring is critical to describe the properties of the inner regions of haloes.
2. While the local escape velocity unbinding that is implemented in many halo finders is successful in removing physically unbound particles, the fact that a conservative threshold (e.g. $2v_{\rm esc}$) is normally used can make it insufficient in some occasions. This can be improved by an additional unbinding scheme in pure velocity space, which proves useful to disentangle dynamically unrelated particles from the halo (see the examples in Sect. 3.3, especially Fig. 7).
3. The new version of ASOHF implements an entirely new scheme for searching for substructure that is different from the search for isolated haloes. In our experiments, we find that delimiting substructures by their density profile (e.g. placing a radial cut when the spherically averaged density profile increases) leads to too generous radii, which therefore contaminate the mean velocity of the halo and can degrade the performance of unbinding and even miss the halo. In ASOHF, density peaks located within the volume of previously identified haloes, at finer levels of refinement, are characterised by their Jacobi radius, which approximately delimits the region in which the gravitational force towards the substructure centre is dominates that of the host. In addition to its clear physical motivation, this (typically more stringent) radius together with the improved unbinding scheme helps to avoid the bias in the halo properties.
4. Using the DM haloes, ASOHF is able to identify and characterise stellar haloes, that is, galaxies. While the starting point for this identification is the catalogue of DM haloes and subhaloes, the stellar haloes are subsequently treated as independent objects. In particular, they can have a different centre, radial extent, dynamics, and so on than the DM halo they are built onto.
5. We have implemented a domain decomposition for ASOHF that enables the analysis of large simulations that would otherwise not fit in memory. Incidentally, as discussed in Sects. 2.6.1 and 3.4.1, the domain decomposition also strongly reduces the CPU time. Furthermore, the procedure is implemented by external, automated scripts, and each domain can be treated as a separate job (requiring no communication). This means that different domains can be either run concurrently (so that the wall time is reduced by a factor equal to the number of cores) or sequentially. The resulting outputs for each domain can be merged by an independent script, so that the whole process is transparent to the user.
6. By saving the particle IDs of all the bound particles of each halo, the merger tree can easily be built once the catalogues have been obtained for (at least) a pair of code outputs.

This process is dramatically accelerated by pre-sorting the lists of particles by ID, and thus intersecting them in $O(N)$ operations, instead of $O(N^2)$. Building the merger tree as a separate task from the halo-finding procedure has several benefits, such as avoiding the need to overload the memory with all the information of previous iterations, or the ability to connect a halo with a progenitor skipping iterations, if it has not been found in the immediately previous iteration. There are also drawbacks of this approach: for instance, the inability to incorporate the merger tree information into the halo finding process itself, such as other modern finders do (e.g. HBT+; Han et al. 2018).

Through a battery of tests described in Sect. 3, we have shown the ASOHF capabilities of recovering virtually all haloes and subhaloes in a broad range of masses in idealised but complex set-ups. When compared to other publicly available halo finders, such as AHF, Rockstar, or Subfind, ASOHF produces comparable results in terms of halo mass functions and properties of haloes. It provides excellent results in terms of substructure identification (in the sense that it is able to identify a large number of substructures, while using a tight, physically motivated definition of substructure extent). This is commonly a tough task for SO halo finders when compared to 3D-FoF, but especially when compared to 6D-FoF finders.

Regarding computational performance, the new version of ASOHF has been rewritten to be extremely efficient in terms of speed, parallel scaling, and memory usage. As discussed in Sect. 3.4.2, the parallel behaviour of the code scales nearly ideally, which allows a large performance gain using shared-memory platforms with a large number of cores. Moreover, the optimisation of the code allows analysing individual snapshots from simulations of $\gtrsim 10^7$ particles within tens of seconds and using a few GB of RAM, and simulations with $\gtrsim 10^8$ particles within tens of minutes and using less than 30 GB of RAM in any case.

When large domains with an even larger number of particles are involved, the time taken by ASOHF can be further improved by the domain-decomposition strategy, which involves no communication between different domains until a last step at which catalogues are merged. As an example, for a simulation with $\sim 10^{10}$ particles within a cubic domain with a side length of 500 Mpc/h, where $\sim 15 \times 10^6$ haloes are expected to form by $z = 0$, scaling the benchmarks characterised in Sect. 3.4.1 we estimate that by performing a decomposition in $d = 4^3$ subdomains, ASOHF could run over all the domains (sequentially i.e. one domain at a time) in ∼93 h of wall time, in a 24-core processor like the one used for the Tests in Sect. 3.4.1. If 8 of such nodes were concurrently available in a distributed memory architecture, this would take $\lesssim$12 h. The implementation of ASOHF, together with the several utilities and scripts mentioned in this paper, is publicly available and documented[13].

Acknowledgements. We thank the anonymous referee for his/her constructive comments, which have improved the presentation of this work. This work has been supported by the Spanish Ministerio de Ciencia e Innovación (MICINN, grant PID2019-107427GB-C33) and by the Generalitat Valenciana (grant PROMETEO/2019/071). D.V. acknowledges support from Universitat de València through an Atracció de Talent fellowship. Part of the tests have been carried out with the supercomputer Lluís Vives at the Servei d'Informàtica of the Universitat de València. This research has made use of the following open-source packages: NumPy (Oliphant 2006), SciPy (Virtanen et al. 2020), Matplotlib (Hunter 2007), and Colossus (Diemer 2018).

[13] See links to the code repository and documentation at the beginning of Sect. 2.

References

Angulo, R. E., & Hahn, O. 2022, Liv. Rev. Comput. Astrophys., 8, 1
Bagla, J. S. 2002, J. Astrophys. Astron., 23, 185
Barnes, J., & Hut, P. 1986, Nature, 324, 446
Behroozi, P. S., Wechsler, R. H., & Wu, H.-Y. 2013, ApJ, 762, 109
Binney, J., & Tremaine, S. 1987, Galactic dynamics, Princeton Series in Astrophysics (Princeton: Princeton University Press)
Borgani, S. 2008, in A Pan-Chromatic View of Clusters of Galaxies and the Large-Scale Structure, eds. M. Plionis, O. López-Cruz, & D. Hughes (Berlin: Springer), 740, 24
Bryan, G. L., & Norman, M. L. 1998, ApJ, 495, 80
Cañas, R., Elahi, P. J., Welker, C., et al. 2019, MNRAS, 482, 2039
Clerc, N., & Finoguenov, A. 2022, ArXiv e-prints [arXiv:2203.11906]
Cole, S., & Lacey, C. 1996, MNRAS, 281, 716
Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, ApJ, 292, 371
Diemand, J., Kuhlen, M., & Madau, P. 2006, ApJ, 649, 1
Diemer, B. 2018, ApJS, 239, 35
Dolag, K., Borgani, S., Murante, G., & Springel, V. 2009, MNRAS, 399, 497
Eisenstein, D. J., & Hut, P. 1998, ApJ, 498, 137
Elahi, P. J., Thacker, R. J., & Widrow, L. M. 2011, MNRAS, 418, 320
Elahi, P. J., Cañas, R., Poulton, R. J. J., et al. 2019, PASA, 36, e021
Gill, S. P. D., Knebe, A., & Gibson, B. K. 2004, MNRAS, 351, 399
Han, J., Jing, Y. P., Wang, H., & Wang, W. 2012, MNRAS, 427, 2437
Han, J., Cole, S., Frenk, C. S., Benitez-Llambay, A., & Helly, J. 2018, MNRAS, 474, 604
Hoffmann, K., Planelles, S., Gaztañaga, E., et al. 2014, MNRAS, 442, 1197
Hunter, J. D. 2007, Comput. Sci. Eng., 9, 90
Iannuzzi, F., & Dolag, K. 2012, MNRAS, 427, 1024
Ishiyama, T., Prada, F., Klypin, A. A., et al. 2021, MNRAS, 506, 4210
Kaviraj, S., Laigle, C., Kimm, T., et al. 2017, MNRAS, 467, 4739
Klypin, A., Gottlöber, S., Kravtsov, A. V., & Khokhlov, A. M. 1999, ApJ, 516, 530
Knebe, A., Knollmann, S. R., Muldrew, S. I., et al. 2011, MNRAS, 415, 2293
Knebe, A., Pearce, F. R., Lux, H., et al. 2013, MNRAS, 435, 1618
Knollmann, S. R., & Knebe, A. 2009, ApJS, 182, 608
Martin-Alvarez, S., Planelles, S., & Quilis, V. 2017, Ap&SS, 362, 91
Murante, G., Monaco, P., Giovalli, M., Borgani, S., & Diaferio, A. 2010, MNRAS, 405, 1491
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Navarro-González, J., Ricciardelli, E., Quilis, V., & Vazdekis, A. 2013, MNRAS, 436, 3507
Oliphant, T. E. 2006, A Guide to NumPy (USA: Trelgol Publishing), 1
Onions, J., Knebe, A., Pearce, F. R., et al. 2012, MNRAS, 423, 1200
Onions, J., Ascasibar, Y., Behroozi, P., et al. 2013, MNRAS, 429, 2739
Pillepich, A., Vogelsberger, M., Deason, A., et al. 2014, MNRAS, 444, 237
Pillepich, A., Springel, V., Nelson, D., et al. 2018, MNRAS, 473, 4077
Planelles, S., & Quilis, V. 2010, A&A, 519, A94
Planelles, S., & Quilis, V. 2013, MNRAS, 428, 1643
Planelles, S., Schleicher, D. R. G., & Bykov, A. M. 2015, Space Sci. Rev., 188, 93
Planelles, S., Mimica, P., Quilis, V., & Cuesta-Martínez, C. 2018, MNRAS, 476, 4629
Press, W. H., & Schechter, P. 1974, ApJ, 187, 425
Quilis, V. 2004, MNRAS, 352, 1426
Quilis, V., Planelles, S., & Ricciardelli, E. 2017, MNRAS, 469, 80
Quilis, V., Martí, J.-M., & Planelles, S. 2020, MNRAS, 494, 2706
Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521
Schechter, P. 1976, ApJ, 203, 297
Springel, V. 2010, MNRAS, 401, 791
Springel, V., White, S. D. M., Tormen, G., & Kauffmann, G. 2001, MNRAS, 328, 726
Tinker, J., Kravtsov, A. V., Klypin, A., et al. 2008, ApJ, 688, 709
Tormen, G., Moscardini, L., & Yoshida, N. 2004, MNRAS, 350, 1397
Vallés-Pérez, D., Planelles, S., & Quilis, V. 2020, MNRAS, 499, 2303
Vallés-Pérez, D., Planelles, S., & Quilis, V. 2021, MNRAS, 504, 510
Villaescusa-Navarro, F., Anglés-Alcázar, D., Genel, S., et al. 2021, ApJ, 915, 71
Villaescusa-Navarro, F., Genel, S., Anglés-Alcázar, D., et al. 2022, ArXiv e-prints [arXiv:2201.01300]
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nat. Methods, 17, 261
Vogelsberger, M., Genel, S., Springel, V., et al. 2014, Nature, 509, 177
Vogelsberger, M., Marinacci, F., Torrey, P., & Puchwein, E. 2020, Nat. Rev. Phys., 2, 42
Weinberger, R., Springel, V., & Pakmor, R. 2020, ApJS, 248, 32

Appendix A: Solution of Poisson's equation in spherical symmetry

The local unbinding procedure, described in Sect. 2.3, relies on the computation of the gravitational potential in spherical symmetry. Here we describe our implementation of the procedure.

In general, the gravitational potential, Φ, of any continuous or discrete density distribution ρ is obtained by solving Poisson's equation,

$$\nabla_{\boldsymbol{r}}^2\Phi = \frac{1}{a^2}\nabla_{\boldsymbol{x}}^2\Phi = 4\pi G\rho, \tag{A.1}$$

which is an elliptic partial derivative equation, where $\boldsymbol{r}$ and $\boldsymbol{x} \equiv \boldsymbol{r}/a$ are the physical and the comoving position vectors, respectively, related by the expansion factor $a(t)$ that solves Friedmann equations for a given cosmology. Under the assumption of spherical symmetry in the density distribution (and thus in the potential), the Poisson equation reduces to a non-linear second-order ordinary differential equation,

$$\frac{\mathrm{d}}{\mathrm{d}r}\left[r^2\frac{\mathrm{d}\Phi}{\mathrm{d}r}\right] = 4\pi G\rho(r)r^2. \tag{A.2}$$

If the mass distribution is bound, it is always possible to take $\lim_{r\to\infty}\Phi(r) = 0$ and show that

$$\Phi(r) = \int_{\infty}^{r}\frac{GM(<r')}{r'^2}\mathrm{d}r', \tag{A.3}$$

where $M(<r) \equiv \int_0^r 4\pi r'^2\rho(r')\mathrm{d}r'$ is the mass enclosed in a sphere of radius r. Assuming the mass distribution to be bound within a maximum radius r_{max}, it is then straightforward to rewrite this expression in a more convenient form for its numerical integration,

$$\Phi(r) = \int_0^r\frac{GM(<r')}{r'^2}\mathrm{d}r' - \frac{GM(<r_{\mathrm{max}})}{r_{\mathrm{max}}} - \int_0^{r_{\mathrm{max}}}\frac{GM(<r')}{r'^2}\mathrm{d}r'. \tag{A.4}$$

Only the first term in Eq. A.4 is a function of r. Operationally, we use the following recipe to efficiently perform the integration.

1. Start from a list of n_{part} particles ($i = 1, \ldots, n_{\mathrm{part}}$), with masses $\{m_i\}_{i=1}^{n_{\mathrm{part}}}$, sorted in increasing comoving distance to the halo centre (density peak), $\{r_i \,/\, r_j \le r_k \forall j \le k\}_{i=1}^{n_{\mathrm{part}}}$. Denote $r_{\mathrm{max}} = r_{n_{\mathrm{part}}}$.
2. Compute the cumulative mass profile, which we denote $M_i \equiv \sum_{j=1}^{i} m_j$.
3. Identify $\mathcal{J}$ as the smallest integer such that $r_{\mathcal{J}} > 0.01 r_{\mathrm{max}}$, and define $\tilde{\Phi}_{j\le\mathcal{J}} = \frac{M_{\mathcal{J}}}{r_{\mathcal{J}}}$.
4. Compute iteratively, for $j = \mathcal{J}+1, \ldots, n_{\mathrm{part}}$, the value of $\tilde{\Phi}$ at the position of the next particle as $\tilde{\Phi}_j = \tilde{\Phi}_{j-1} + \frac{M_j(r_j - r_{j-1})}{r_j^2}$.
5. Define the constant $C = \tilde{\Phi}_{n_{\mathrm{part}}} + \frac{M_{n_{\mathrm{part}}}}{r_{n_{\mathrm{part}}}}$.
6. The spherically symmetric gravitational potential at comoving radius r_i is $\Phi_i = \frac{G}{a}(C - \tilde{\Phi}_i) \le 0 \;\forall\; i$.

We perform step 3, that is, we assume a constant gravitational potential in the inner 1% of the radial space, to mitigate numerical errors since r_i can be arbitrarily small. This has a negligible impact on the unbinding procedure because it is rare to find unbound particles near halo centres.

Appendix B: Computation of the gravitational binding energy by sampling

The gravitational binding (or potential) energy of a system of N particles is explicitly given by the sum over all particle pairs:

$$E_{\mathrm{grav}} = -G\sum_{i=1}^{N}\sum_{j=i+1}^{N}\frac{m_i m_j}{|\boldsymbol{r_i} - \boldsymbol{r_j}|}, \tag{B.1}$$

which therefore implies and $O(N^2)$ calculation, which becomes prohibitively expensive for large enough N. In `ASOHF`, we perform a direct summation over the $N(N-1)/2$ pairs for haloes with $N \le N_{\mathrm{max}}^{\mathrm{energy}}$ particles (see the discussion of Fig. B.1 for the dependence of the accuracy of the results on this parameter), while for larger haloes, we use a sampling estimate of this quantity, which we describe and test below.

In order to estimate the gravitational binding energy of large haloes by sampling, we consider $n_{\mathrm{sample}}^2 = \max(\lfloor N_{\mathrm{max}}^{\mathrm{energy}}/\sqrt{2}\rfloor, 0.01N)^2$ pairs of particles and compute the contribution of each to the potential energy, that is, $E_{\mathrm{grav,ij}} = -G\frac{m_i m_j}{|\boldsymbol{r_i}-\boldsymbol{r_j}|}$ for the pair of particles (ij). The procedure is as follows:

First, we randomly select n_{sample} particles (with replacement), which we denote $i = 1, \ldots, n_{\mathrm{sample}}$. For each of these particles, we select a new set of n_{sample} particles (with replacement), which we denote $j = 1, \ldots, n_{\mathrm{sample}}$, and compute for particle i its contribution to the gravitational energy, $E_{\mathrm{grav,i}}^{\mathrm{sample}} = \sum_{j=1}^{n_{\mathrm{sample}}} E_{\mathrm{grav,ij}}$. The gravitational energy per pair of particles therefore is

$$\langle E_{\mathrm{grav,pairs}}\rangle = \frac{1}{n_{\mathrm{sample}}^2}\sum_{i=1}^{n_{\mathrm{sample}}} E_{\mathrm{grav,i}}^{\mathrm{sample}} = \frac{1}{n_{\mathrm{sample}}^2}\sum_{i=1}^{n_{\mathrm{sample}}}\sum_{j=1}^{n_{\mathrm{sample}}} E_{\mathrm{grav,ij}}. \tag{B.2}$$

Since there are $N(N-1)/2$ pairs of particles in the halo, we estimate its gravitational energy as

$$E_{\mathrm{grav}} \cong \frac{N(N-1)}{2}\langle E_{\mathrm{grav,pairs}}\rangle. \tag{B.3}$$

We note that $E_{\mathrm{grav}} \propto \langle E_{\mathrm{grav,pairs}}\rangle$ and thus the relative error in E_{grav} is equal to the relative error in $\langle E_{\mathrm{grav,pairs}}\rangle$, which in turn is proportional to $1/\sqrt{n_{\mathrm{sample}}}$ and independent of N. This is graphically exemplified in Fig. B.1.

The gravitational binding energy per unit mass of particle i is proportional to $E_{\mathrm{grav,i}}/m_i$. Thus, we report as the most bound particle the particle with the most negative value of $E_{\mathrm{grav,i}}/m_i$, which can incidentally be used as an estimation of the location of the minimum of gravitational potential. The uncertainty in determining this position can be estimated as the sphere, with centre in the most-bound particle, that encloses $\lceil N/n_{\mathrm{sample}}\rceil$ particles.

To assess the convergence of this method, we generated NFW haloes with an arbitrarily fixed concentration of $c = 10$ and radius $R = 1.5$ (in arbitrary units), realised using different numbers of particles, N, from 5×10^4 to 10^6. We then computed the gravitational energy by direct $O(N^2)$ summation as in Eq. B.1, and by the sampling method varying n_{sample} between 10^2 and 10^4. For each pair of values of (N, n_{sample}), we performed $n_{\mathrm{boots}} = 100$ bootstrap iterations of the calculation by sampling, so that we can estimate the mean relative error in the gravitational potential energy by comparing with the result by direct summation, and its (16 – 84)% confidence intervals.

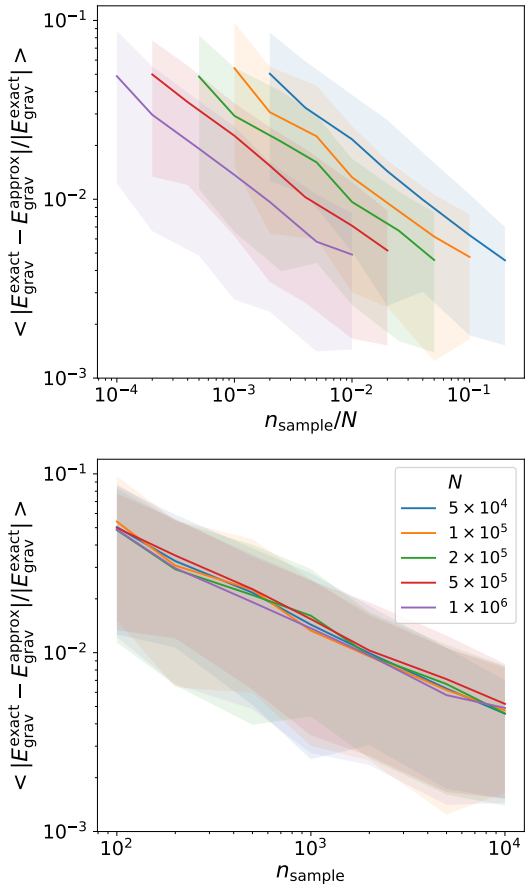


Fig. B.1. Dependence of the error in obtaining the gravitational binding energy by sampling, rather than by direct summation over all the pairs of particles, with the fraction $n_{\rm sample}/N$ (*upper panel*) and with just the number of sample particles $n_{\rm sample}$ (*lower panel*). The number of evaluated pairs of particles is $n^2_{\rm sample}$. Different colours, according to the legend in the lower panel, represent different number of particles within the halo. The shadowed regions correspond to (16 – 84)% confidence regions.

This is shown in Fig. B.1, whose upper panel shows the mean of the relative error, computed over the 100 bootstrap iterations, in computing the gravitational energy by sampling. Each colour represents a different number of particles in the halo according to the legend in the lower panel, while on the horizontal axis, we show the fraction of $n_{\rm sample}$ to N. For each N, the error clearly decreases as $1/\sqrt{n_{\rm sample}}$. These curves can be brought to match each other when we represent the relative error as a function of $n_{\rm sample}$, as shown in the lower panel of Fig. B.1. In this case, regardless the number of particles of the halo, the mean errors fall below 1% when more than $n_{\rm sample} \approx 2000$ are used, while $n_{\rm sample}$ must be increased to $\sim 5000 - 6000$ to achieve an error below 1% at the 84 percentile.

It would be reasonable to ask why the error scales as $1/\sqrt{n_{\rm sample}}$, instead of $1/\sqrt{n_{\rm pairs}} = 1/n_{\rm sample}$, which would be expected from averaging $n_{\rm pairs}$ values of the energy per pair of particles. This is due to the fact that we are considering $n_{\rm sample}$ particles, and estimating the gravitational binding energy of each of these with a new set of $n_{\rm sample}$ particles. This is fundamentally different from choosing $n_{\rm pairs}$ completely independent pairs, in which case it is easy to verify that the error indeed scales as $\propto 1/n_{\rm sample}$. However, this latter method, although much faster in terms of convergence (lower values of $n_{\rm sample}$ are required, and the computational cost is therefore greatly reduced), does not allow us to estimate the most-bound particles of haloes (and therefore the position of the gravitational potential minima). This is so because to achieve this aim, we require a good estimate of the energies of a subsample of individual particles. This cannot be achieved by randomly picking pairs of particles because each particle will be considered a few times as much. We therefore opt in `ASOHF` for this more slowly converging method.

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **519,** 6111–6125 (2023)
Advance Access publication 2023 January 10

https://doi.org/10.1093/mnras/stad059

On the choice of the most suitable indicator for the assembly state of dark matter haloes through cosmic time

David Vallés-Pérez [1]★ Susana Planelles [1,2] Óscar Monllor-Berbegal [1] and Vicent Quilis [1,2]

[1]*Departament d'Astronomia i Astrofísica, Universitat de València, E-46100 Burjassot, València, Spain*
[2]*Observatori Astronòmic, Universitat de València, E-46980 Paterna, València, Spain*

Accepted 2023 January 4. Received 2022 December 13; in original form 2022 October 14

ABSTRACT

The dynamical state and morphological features of galaxies and galaxy clusters, and their high-redshift precursors, are tightly connected with their assembly history, encoding crucial information about the formation and evolution of such cosmic structures. As a first step towards finding an optimal indicator of the assembly state of observed structures, we use a cosmological simulation of a moderate volume to critically examine the best definition of an indicator that is able to discriminate dark matter haloes undergoing mergers and/or strong accretion from haloes experimenting a relaxed evolution. Using a combination of centre offset, virial ratio, mean radial velocity, sparsity, and ellipticity of the dark matter halo, we study how the thresholds on these parameters, as well as their relative weights, should evolve with redshift to provide the best classification possible. This allows us to split a sample of haloes in a totally relaxed, a marginally relaxed and an unrelaxed subsamples. The resulting classification strongly correlates with the merging activity obtained from the analysis of complete merger trees extracted from whole simulation data. The results on how the different indicators depend on redshift and halo mass, and their optimal combination to better match the true assembly history of haloes, could constitute relevant hints to find a suitable set of indicators applicable to observational data.

Key words: methods: numerical – galaxies: clusters: general – dark matter – large-scale structure of Universe.

1 INTRODUCTION

Deeply interwoven through a complex network of filaments and sheets, dark matter (DM) haloes are bound, diffuse structures which result from the gravitational collapse of the primordial density fluctuations and a hierarchical merging history (Zel'dovich 1970; Press & Schechter 1974; Gott & Rees 1975). DM haloes constitute the fundamental building blocks of the large-scale structure (LSS) of the Universe, and host their baryonic counterparts that we observe over the electromagnetic spectrum (see, for instance, Planelles, Schleicher & Bykov 2015, for a review). At the galactic scale, the current theories of galaxy formation typically assume DM haloes to be virialized (e.g. White & Rees 1978), although this does not necessarily hold for each galactic DM halo; while, at larger masses (at the galaxy cluster scale), most DM haloes are still expected to be in the process of virialization, since they are the latest objects to have assembled (e.g. Kravtsov & Borgani 2012, for a review on galaxy cluster formation).

However, the dynamical state of individual haloes is tightly connected to their assembly history and, in particular, to the presence of mergers and the accretion rates in the last one or few dynamical times. A merger or a period of intense accretion usually triggers many morphological and dynamical disturbances in the halo (asphericity, higher velocity dispersions, abundance of substructures, changes to the internal structure, etc.), which gradually fade away once the assembly episode is over (see e.g. Poole et al. 2006 for a thorough analysis of the disturbances and the subsequent relaxation after a merger event at cluster scales).

Since dynamically relaxed and disturbed structures often present fundamentally different properties, a characterization of the dynamical state of the sample of cosmic structures is often a necessary procedure in many analyses of very different natures, such as in studies about the geometry of the cosmic web (Gouin, Bonnaire & Aghanim 2021), statistical properties of the population of galaxy clusters (scaling relations, mass functions, etc.; e.g. Chen et al. 2019; Seppi et al. 2021), hydrostatic mass bias (Nelson et al. 2014; Biffi et al. 2016; Angelinelli et al. 2020), turbulence (Vazza et al. 2017; Valdarnini 2019; Vallés-Pérez, Planelles & Quilis 2021a,b; Simonte et al. 2022), or galactic environments (Kuchner et al. 2022), just to mention a few.

Even though we usually define haloes using the virial radius prescription of Eke, Cole & Frenk (1996) and Bryan & Norman (1998), based on the spherical collapse model, this does not imply that, in general, 3D haloes (non-necessarily spherical, in a full-cosmological, i.e. not isolated environment) defined this way are necessarily in virial equilibrium. While in simulations one can access the whole temporal evolution of the objects, and thus recover the assembly history of the halo under study in order to assess the dynamical state, this is not possible in observations. Thus, for the sake of a more direct comparison with observational works, simple schemes for characterizing the dynamical state using halo properties at a given time are usually involved in many analyses.

For the time being, most works have relied on placing a threshold on some halo property expected to correlate with the dynamical state, in order to split the relaxed and unrelaxed subsamples. Perhaps, the most direct of such indicators is the *virial ratio*, usually defined

★ E-mail: david.valles-perez@uv.es

© 2023 The Author(s)
Published by Oxford University Press on behalf of Royal Astronomical Society

as $\eta \equiv 2T/|W|$, where T is the intrinsic kinetic energy of the halo and W is its gravitational potential energy. η would be expected to be 1 for an isolated system in a steady state. However, different works have found different thresholds to best suite their particular classification (e.g. Shaw et al. 2006; Neto et al. 2007; Knebe & Power 2008, see also the discussion in Cui et al. 2017). Similarly, there is debate about the necessity of including a surface tension term to account for the fact that haloes are not isolated (Poole et al. 2006; Shaw et al. 2006; Knebe et al. 2011). Another frequently used indicator, both in simulations and observations, is the centre offset, which quantifies the departure from smoothness and spherical symmetry of the matter distribution, and serves as an indicator of substructure (Crone, Evrard & Richstone 1996). In practice, however, there are many possibilities regarding the choice of centres (see Cui et al. 2016) and how to set the thresholds (cf. D'Onghia & Navarro 2007; Macciò et al. 2007). Additionally, in observations the centre offset may depend crucially on the orientation, posing an additional challenge. Last, other authors have used the fraction of mass in substructures as a measure of the dynamical unrelaxedness of a DM halo (e.g. Neto et al. 2007; cf. other recently suggested approaches, e.g. Kimmig et al. 2022). While the election of this magnitude is well-motivated, the mass contained in substructures in simulated haloes depends critically on numerical resolution and the precise definition of the substructure extent (see e.g, the discussion in Vallés-Pérez, Planelles & Quilis 2022), making this criterion less comparable.

Since it is difficult that a single property can reflect the complex picture of the dynamical state of a halo, many recent studies have used combinations of these indicators, either by considering as relaxed the haloes which simultaneously fulfill several relaxation criteria (Neto et al. 2007; Biffi et al. 2016), or by defining some combined indicator (Haggar et al. 2020; De Luca et al. 2021; Zhang et al. 2021a). Finally, other metrics of the dynamical state are based on the X-ray morphology, such as the centroid shift ω (Mohr, Fabricant & Geller 1993), or the power ratio, P_3/P_0 (Buote & Tsai 1995, see also the review of Rasia, Meneghetti & Ettori 2013 on X-ray morphological estimators for galaxy clusters); or more sophisticated ones such as those involving Fourier analyses of the fluctuations in mass and X-ray maps (Cerini, Cappelluti & Natarajan 2022), or the expansion of the Compton y-maps in Zernike polynomials (Capalbo et al. 2021).

However, in most of the previous works, the parameters being used and, especially, the thresholds imposed on them have been tuned in a somewhat empirical way. This has lead to variations in the criteria from work to work, even though the underlying idea is kept. Furthermore, a possible redshift evolution of these thresholds or of their very relevance has been devoted marginal attention, either because the studies were focused on a particular cosmic epoch or because it had been implicitly assumed that these criteria should not evolve with redshift.

In this work, we intend to critically examine a set of possible indicators of the assembly state, all of which can be obtained from the complete 3D information in simulations, and develop a criterion which accommodates redshift-dependent thresholds and the possibility that different indicators have more or less relevance at different cosmic epochs. We note the reader that, while in the following we may refer to the *dynamical* state of haloes, our main focus is oriented towards the dynamical disturbances associated with the assembly history of haloes (i.e. the presence of merger events or episodes of strong accretion; rather than a more general sense of dynamical unrelaxedness which could include, e.g. the presence of substructures even when they are not associated with a merger episode, since they have an impact on properties such as the hydrostatic equilibrium).

The rest of the manuscript is organized as follows. In Section 2, we introduce our simulation, halo sample, and the methodology that we employ for setting the thresholds and relative weights of the different dynamical state indicators. Our resulting criterion is presented in Section 3, including the analysis of the mass dependence of our results and a validation of our method with a different simulation. Finally, we discuss the applicability of our results in Section 4. Appendix A contains the fitting formulae for the thresholds and weights applicable for massive haloes.

2 METHODS

The results reported in this paper have been extracted from the analysis of a Lambda cold dark matter (ΛCDM) cosmological simulation tracking the coupled evolution of baryons and DM. We describe the relevant details of the simulation in Section 2.1, then cover the halo catalogues and merger tree elaboration in Section 2.2, and discuss how do we compute the dynamical state indicators in Section 2.3. Finally, we introduce our classification strategy in Section 2.4.

2.1 The simulation

The haloes we analyse in this paper are extracted from a numerical simulation run with MASCLET (Quilis 2004; Quilis, Martí & Planelles 2020), a (magneto-)hydrodynamics and N-body code primarily designed for cosmological applications. For evolving the DM component, which is the primary focus of this work, MASCLET implements a multilevel Particle-Mesh (PM) scheme (Hockney & Eastwood 1988), which takes advantage of the adaptive-mesh refinement (AMR) strategy (Berger & Colella 1989) to gain spatial, temporal and force resolution.

We have simulated a periodic, cubic ($L = 100\,h^{-1}\,\mathrm{Mpc}$) domain, under the assumption of a flat, ΛCDM cosmology specified by the matter density parameter $\Omega_\mathrm{m} = 0.31$ ($\Omega_\Lambda = 1 - \Omega_\mathrm{m}$), baryon density parameter $\Omega_\mathrm{b} = 0.048$, and Hubble parameter $h \equiv H_0/(100\,\mathrm{km\,s^{-1}}) = 0.678$. The initial conditions stem from a realization of the primordial Gaussian random field assuming a spectral index $n_s = 0.96$ and an amplitude yielding $\sigma_8 = 0.82$, and are set up at redshift $z_\mathrm{ini} = 100$ using a CDM transfer function (Eisenstein & Hu 1998). The values selected for the cosmological parameters are consistent with the latest results reported by Planck Collaboration (2020).

A first simulation is run at low resolution, using a fix grid of $N_x^3 = 256^3$ cells and the same number of equal-mass particles. This is used to identify the Lagrangian regions in the initial conditions which will evolve into dense structures by $z = 0$, and mapping them with enhanced numerical resolution already at z_ini. We use this approach to establish three nested levels of initial conditions, resulting in a best mass resolution of $1.48 \times 10^7\,\mathrm{M}_\odot$.

Using these high-resolution initial conditions, the simulation is evolved again using AMR based on gas/DM overdensities, converging flows, and Jeans length criteria, achieving a peak resolution of $\Delta x_8 = 2.3\,\mathrm{kpc}$ at the maximum ($\ell = n_\ell \equiv 8$) level of refinement. While the baryonic component is not the primary focus of this work, the simulation includes gas cooling, but no other baryonic effect or feedback mechanism.

2.2 Halo catalogue and merging history

For each snapshot of the simulation, we have identified the DM haloes using the public halo finder ASOHF (Planelles & Quilis 2010;

Knebe et al. 2011; Vallés-Pérez et al. 2022),[1] which is based on the spherical-overdensity definition and uses the virial radius (according to the prescription of Bryan & Norman 1998) to delimit the extent of the haloes that are not substructure.

After determining the halo catalogues, these are linked in between snapshots using the merger tree code presented by Vallés-Pérez et al. (2022, their section 2.6.2), which identifies all the haloes at a given code output which have contributed to an object in a following one, allowing to skip an arbitrary number of snapshots, if necessary. Using it, we determine the main evolutionary line of each halo, as well as the presence and characterization of mergers.

Following Planelles & Quilis (2009), Chen et al. (2019), and Vallés-Pérez, Planelles & Quilis (2020), we have classified each merger event in the sample as either a major merger (if the mass ratio, $M_{\rm min}/M_{\rm max}$, between the two haloes involved exceeds 1/3), or a minor merger ($1/3 > M_{\rm min}/M_{\rm max} \geq 1/10$). Mergers below a mass ratio of 1/10 are disregarded. The merger time is determined as the moment in which the centre of the infalling (the least massive) halo crosses the virial boundary of the host (the most massive) halo.

2.2.1 *Fiducial classification: assembly history of the haloes*

In order to determine the optimal thresholds on the dynamical state indicators (see below, Section 2.3 and therein), we compare with a reference, or *fiducial*, classification of the dynamical state based on the full assembly history of haloes (i.e. the presence of past or ongoing mergers, as well as the accretion rates).

As a tentative classification of the unrelaxedness induced by a merger event, we will assume that a typical halo remains in a disturbed state for one dynamical time after a major merger, or half a dynamical time after a minor merger, with the dynamical time $\tau_{\rm dyn}$ being defined as

$$\tau_{\rm dyn}(z) \equiv \frac{1}{\sqrt{G\rho}} = \frac{1}{\sqrt{G\rho_B(z)\Delta_{\rm vir}(z)}}, \tag{1}$$

with G the gravitational constant, ρ the density of the halo, $\rho_B(z)$ the background matter density, and $\Delta_{\rm vir}(z)$ the virial overdensity (Bryan & Norman 1998).

While the choice of the time-span is a crude approximation, it responds to the fact that many works have shown that the disturbance triggered by a minor merger is, in general terms, smaller than the effect of a major merger, both for the dark and for the baryonic components (Planelles & Quilis 2009; Yu et al. 2015; Vallés-Pérez et al. 2020; Zhang et al. 2021b). In practical terms, since $\tau_{\rm dyn}(z)$ varies strongly with redshift and reaches considerable fractions of the age of the Universe, especially at low redshift, we choose to define the number of dynamical times between two moments, t_1 and t_2, as in Jiang & van den Bosch (2016) and Wang et al. (2020):

$$N_\tau(t_1, t_2) = \int_{t_1}^{t_2} \frac{{\rm d}t}{\tau_{\rm dyn}(z)}. \tag{2}$$

Additionally, it might be the case that a halo is accreting strongly, but without undergoing any significant merger (either physically or due to the finite resolution of a simulation). Thus, we also consider as unrelaxed, for the purpose of the fiducial classification, any halo which has assembled more than 50 per cent of their mass in the last dynamical time.

For the analyses within this work, all the 28 snapshots of the simulation since redshift $z = 5$ are considered. We select the 1000 most massive haloes at each epoch, and discard all those which cannot be reliably traced back in time for at least one $\tau_{\rm dyn}(z)$. We show, in Fig. 1, the redshift evolution of the median mass in the sample (solid line), together with shaded regions enclosing the confidence intervals corresponding to the 16–84 per cent (dark blue) and 2.5–97.5 per cent (light blue) percentiles of the distribution of masses. The dotted lines mark the maximum mass (upper line) and the minimum mass, or mass limit (lower line) in the sample at each time. Thus, the mass limit in our sample evolves from $\sim 10^{12}\,M_\odot$ at $z = 5$ to $\sim 4.5 \times 10^{12}$ M$_\odot$ at $z = 0$. The wide redshift interval considered in this study includes from the cluster-, group-, and massive galaxy-sized haloes at $z \simeq 0$, to the DM counterpart of galaxies and the progenitors of low-redshift clusters at the high-redshift end.

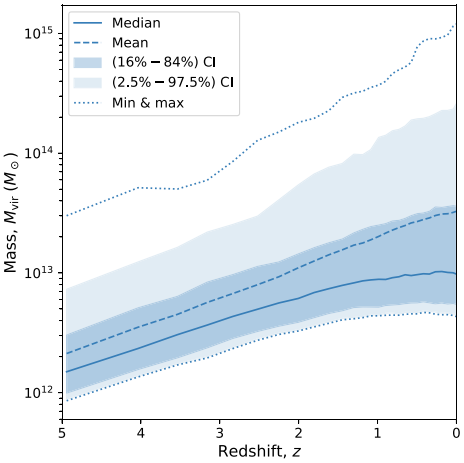


Figure 1. Evolution of the distribution of halo masses in our sample. The solid line indicates the median mass of the sample, with the dark and light shaded regions enclosing the 16–84 per cent (dark blue) and 2.5–97.5 per cent (light blue) confidence intervals (CIs) around it, respectively. The dashed line corresponds to the mean mass, while the dotted lines correspond to the maximum and minimum masses.

The results of the fiducial classification are summarized in Fig. 2, where we show the number of haloes which are finally considered at each snapshot (blue line, referring to the axis on the left). Only at high redshift ($z \gtrsim 3$), a large fraction (10–25 per cent) of the preliminary haloes get discarded because they cannot be tracked back in time for at least one dynamical time. The fraction of unrelaxed haloes according to the fiducial classification (green, dashed line; referring to the axis on the right) varies from ~ 80 to ~ 30 per cent through the considered redshift interval. Purple and orange, dotted lines show the number of haloes, as a fraction of the total, which are unrelaxed due to either the condition on recent mergers or the condition on the accretion rate, respectively. Most of the low-redshift haloes that are labelled unrelaxed are undergoing mergers, while at high redshift the cause for unrelaxedness is more usually a high level of smooth accretion. This may be due to several reasons, amongst which we can mention the higher density in the vincinity of haloes at high redshift, or resolution limitations of the simulation (i.e. at high redshift, a halo may be accreting small, under-resolved structures, which are therefore not accounted as mergers).

[1] https://github.com/dvallesp/ASOHF.

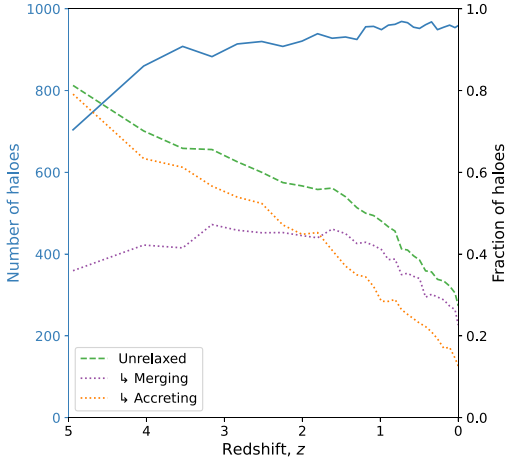


Figure 2. Fiducial classification of dynamical states of the halo sample. The blue line represents the number of haloes selected per snapshot, according to the left axis. The dashed/dotted lines, according to the right axis, describe the evolution with redshift of the fraction of unrelaxed haloes (green), which can have been labelled as such due to mergers (purple) or strong accretion/mass growth (orange).

2.3 Indicators for the assembly state

Many possible proxies for the dynamical and assembly state of a DM halo, or their corresponding baryonic structure (e.g. a galaxy or a galaxy cluster) have been proposed in the literature (see for instance, Cole & Lacey 1996; Crone et al. 1996; Shaw et al. 2006; Haggar et al. 2020; Zhang et al. 2021a, just to cite a few). While simulations allow to access the complete 3D picture, the lack of the whole information in observations (due to, e.g. projection or the inability to observe the dark component, or even the plasma out to large radii) requires that, generally, different quantities are used for assessing the dynamical state in simulations and in observations (Rasia et al. 2013; De Luca et al. 2021; Yuan, Han & Wen 2022). While comparison with observations is crucial and will be dealt with in future work, here we shall focus on the dynamical state indicators extracted from the complete, 3D data in simulations as a first step. Unless otherwise specified, all quantities below are referred to the virial volume.

2.3.1 *Centre offset.*

The centre offset is usually defined as the distance between two different choices of centre, in units of some aperture radius (typically, the virial radius of the halo, $R_{\rm vir}$). Many examples for the choices of centre pair exist in the literature, such as centre of mass (CM) versus density peak (Baldi et al. 2017) or CM versus potential minimum (Biffi et al. 2016), extracted from the 3D description in simulations, or the morphological offset of the BCG location versus X-ray surface brightness peak (Rossetti et al. 2016), amongst many others. We address the interested reader to Cui et al. (2016), who compare many different choices of observable for defining the centre of galaxy clusters.

In this work, we have tested the three possible combinations between the minimum of gravitational potential (defined as the location of the most-bound DM particle, as obtained by ASOHF and described in detail in Vallés-Pérez et al. 2022), the DM density peak, and the DM centre of mass. We find that the most robust results are obtained for the Peak–CM pair. Therefore, we defined the centre offset parameter as

$$\Delta_r = \frac{\left|\boldsymbol{r}_{\rm peak,DM} - \boldsymbol{r}_{\rm CM,DM}\right|}{R_{\rm vir}}. \tag{3}$$

2.3.2 *Virial ratio*

For a gravitational system in steady state, the virial theorem predicts $2T + W - E_s = 0$, where T is the kinetic energy, W is the gravitational binding energy, and E_s is the surface energy term (Chandrasekhar 1961). Neglecting the surface term, the virial ratio is usually defined as

$$\eta \equiv \frac{2T}{|W|}, \tag{4}$$

and it is expected that $\eta \to 1$ for isolated systems. However, haloes are not generally isolated systems, and therefore there is not a good *a priori* reason to drop the surface term in the virial theorem. Thus, many works define the virial ratio as $\eta' = (2T - E_s)/|W|$ (Shaw et al. 2006), while others claim that the surface term overcorrects the virial ratio (Power, Knebe & Knollmann 2012). As the latter, we find that correcting the virial ratio by the surface term wipes out the correlation with merging activity, and thus we shall use the definition in equation (4) in the remainder of this work.

2.3.3 *Mean radial velocity*

In a relaxed object, we do not expect important changes in the radial structure, while an unrelaxed system will experience significant disturbances as it settles down to equilibrium. This motivates the consideration of the mean radial velocity of DM particles

$$\langle v_r \rangle_{\rm DM} = \frac{\sum_i m_i v_{r,i}}{\sum_i m_i}, \tag{5}$$

being m_i the mass of the ith DM particle, and $v_{r,i}$ its radial velocity relative to the halo reference frame. In practical terms, we scale this quantity by the circular velocity at the virial radius, $V_{\rm circ,vir} \equiv \sqrt{GM_{\rm vir}/R_{\rm vir}}$, and define the corresponding normalized indicator as

$$\langle \widetilde{v_r} \rangle_{\rm DM} \equiv \frac{|\langle v_r \rangle_{\rm DM}|}{V_{\rm circ,vir}}. \tag{6}$$

2.3.4 *Sparsity*

Systems which have experienced recent significant mergers tend to display shallower central density profiles due to the disturbance caused by the infalling halo, and thus are less concentrated. Many works (e.g. Neto et al. 2007; Wang et al. 2020) have pointed out the relation between the time spanned since the last major merger and halo concentration, $c_{\rm vir} = R_{\rm vir}/R_{\rm s}$, being $R_{\rm s}$ the scale radius of the Navarro, Frenk & White (1997) profile (or the radius where the logarithmic slope of the DM density profile equals -2).

More recently, sparsity has been suggested as a non-parametric alternative to concentration, which reduces the scatter with halo mass (Balmès et al. 2014; Corasaniti et al. 2018), and has also been found to correlate with the timing since the last relevant merger (Richardson & Corasaniti 2022). While sparsity is generally defined as the quotient

between the masses at different spherical overdensities, we find that the one maximizing the correlation with merging activity is

$$s_{200c,500c} \equiv \frac{M_{200c}}{M_{500c}}. \tag{7}$$

2.3.5 Ellipticity

DM haloes are generally triaxial (Frenk et al. 1988; Knebe & Wießner 2006), with significant scatter in halo shape at a given mass and redshift. Many recent studies have pointed out at the correlation between triaxiality and/or ellipticity of the halo shape and the formation history of a halo, with relaxed haloes tending to be rounder (Chen et al. 2019; Lau et al. 2021).

We define the overall shape of the DM halo by finding the eigenvalues of the shape tensor, defined as

$$S_{\alpha\beta} = \sum_i m_i \frac{r_{i,\alpha} r_{i,\beta}}{r_i^2}, \tag{8}$$

which are proportional to the semiaxes squared. The positions, $\boldsymbol{r}_i$, are relative to the cluster centre (defined as the location of the density peak), and we choose to normalize them to be unit length to prevent the shape to be dominated by the particles in the outskirts of the halo. Note this corresponds to the E2 method introduced by Zemp et al. (2011). If a, b, and c are the semiaxes sorted in non-increasing order, we define the ellipticity of the halo, ϵ, as

$$\epsilon = 1 - \frac{c}{a}. \tag{9}$$

2.3.6 Other indicators not considered in this work

Amongst the most widely used proxies for the dynamical state of DM haloes in the literature, we have not included the fraction of substructure, $f_{\rm sub}$, in this study (neither defined as the mass in substructures as a fraction of the host mass, nor as the ratio between the mass of the heaviest substructure and the host mass, as in Cialone et al. 2018). While $f_{\rm sub}$ should naturally correlate with the assembly state (especially, with the merging state), its interpretation is very subtle due to several factors. First, there is not a unique way to define the extent of a subhalo, and differences amongst halo finders have a dramatic impact on the recovered masses of substructures (Vallés-Pérez et al. 2022, see their figs 5 and 10). In second place, the amount of substructure produced in simulations depends strongly, not only on resolution, but also on the numerical scheme employed to solve gravity. This introduces strong mass biases (while the most massive haloes in our simulation may host well-resolved substructure, haloes with less than a few ten thousands particles are likely to be substructure-deficient). These obscure dependencies with mass, resolution, and numerical scheme limit our ability to consistently incorporate this indicator in our work. Simulations with enhanced resolution, capable of fully resolving rich substructure in our wide range of masses, could be able to overcome this limitation of the present work.

Regarding the indicators describing the shape of the mass distribution, while ϵ alone does not fully characterize the shape of an ellipsoid, we have not considered any additional parameter, such as triaxiality $T \equiv \frac{a^2-b^2}{a^2-c^2}$ (Franx, Illingworth & de Zeeuw 1991). While ellipticity measures directly the deviation from sphericity, which is expected during assembly episodes, the same is not true for triaxiality. As a matter of fact, triaxiality is undefined for spherical objects, and we do not find a clear reason to have a preference towards either prolateness/oblateness during mergers or strong accretion periods.

Table 1. Summary of the redshift binning considered for the subsequent analyses. Each bin contains the haloes extracted from the $N_{\rm snaps}$ available with $z \in [z_{\rm min}, z_{\rm max}]$. The mean redshift of the $N_{\rm haloes}$ haloes in the bin is $\bar{z}$, with a fraction $f_{\rm unrelaxed}$ of them being unrelaxed (either merging or experiencing intense accretion) according to the fiducial classification. Note we report $\bar{z}$, instead of the median, because z is not continuously distributed (at each redshift bin, there are only $N_{\rm snaps}$ different values of z).

$z_{\rm min}$	$z_{\rm max}$	$N_{\rm snaps}$	$\bar{z}$	$N_{\rm haloes}$	$f_{\rm unrelaxed}$
0	0.2	4	0.084	3828	0.309
0.2	0.5	4	0.381	3830	0.360
0.5	0.75	3	0.651	2890	0.406
0.75	1.0	3	0.897	2871	0.468
1.0	1.5	4	1.253	3769	0.512
1.5	2.0	3	1.808	2788	0.562
2.0	3.0	3	2.536	2742	0.600
3.0	4.0	2	3.350	1791	0.657
4.0	5.0	2	4.443	1564	0.751

2.4 Classification strategy

2.4.1 Redshift binning

A total of 28 snapshots of the simulation, since $z = 5$, are saved and used in this analysis. To augment the statistics, we have grouped the snapshots in several redshift bins, which are described in Table 1.[2]

2.4.2 Optimizing the thresholds

In a first step, we place a threshold, $X_i^{\rm thr}$, for each of the dynamical state indicators, X_i, described in the previous section ($i = 1, \ldots, 5$, for the five dynamical state indicators). This is performed independently at each redshift bin. To do so, we vary $X_i^{\rm thr}$ from the minimum to the maximum value of X_i through the sample, and identify how well does $X_i^{\rm thr}$ separate the relaxed and the unrelaxed samples of the fiducial classification.

For each value of $X_i^{\rm thr}$, we compute two complementary metrics of the goodness of the classification,[3] namely the efficiency in discriminating the unrelaxed haloes,

$$\epsilon_{\rm unrelaxed}(X_i^{\rm thr}) = \frac{\#\text{ of unrelaxed haloes properly identified}}{\#\text{ of unrelaxed haloes (fiducial)}} \tag{10}$$

and the efficiency in discriminating the relaxed haloes,

$$\epsilon_{\rm relaxed}(X_i^{\rm thr}) = \frac{\#\text{ of relaxed haloes properly identified}}{\#\text{ of relaxed haloes (fiducial)}}. \tag{11}$$

Out of all the possible values of $X_i^{\rm thr}$, we choose the one which maximizes the product of both metrics, that is to say:

$$\hat{X}_i^{\rm thr} = \text{argmax}_{X_i^{\rm thr}} \left[\epsilon_{\rm unrelaxed}(X_i^{\rm thr}) \cdot \epsilon_{\rm relaxed}(X_i^{\rm thr})\right]. \tag{12}$$

Since $\epsilon_{\rm unrelaxed}$ ($\epsilon_{\rm relaxed}$) can be thought, in a frequentist approach, as the probability of correctly identifying an unrelaxed (relaxed) halo as

[2] Not all bins contain the same number of snapshots (or haloes): higher redshift bins comprise less snapshots. While this may increase the scatter in our results at high redshift, grouping more snapshots together at high redshift would increase the systematic uncertainty due to stacking objects of more different epochs.

[3] Note that the metrics introduced in equations (10) and (11) also correspond, respectively, to the True Positive Rate (TPR) or sensitivity, and the True Negative Rate (TNR) or specificity in the usual jargon of binary classifications (e.g. Fawcett 2006). However, we choose this notation here for better readability.

such, our choice of threshold in equation (12) corresponds to picking the one which enhances the likelihood of correctly classifying both an unrelaxed and a relaxed halo, and thus serves as a compromise between too generous and too stringent thresholds.

2.4.3 *Totally relaxed, marginally relaxed and disturbed haloes*

Once the final (redshift-dependent) thresholds, $\left\{X_i^{\rm thr}(z)\right\}_{i=1}^{5}$, are established, any halo will be regarded as *totally relaxed* if

$$X_i < X_i^{\rm thr}(z) \qquad \forall i = 1, \ldots, 5, \tag{13}$$

that is, if it has a low value of all the dynamical state indicators (low centre offset, mean radial velocity and ellipticity, virial ratio and sparsity close to unity). This allows a very conservative definition of the most relaxed haloes.

However, it may be the case that a halo has a high value of one of the parameters, but is relaxed according to the rest. This might be the case for a variety of reasons, ranging from physical (e.g. a halo with high ellipticity due to a strong tidal field generated by the surrounding large-scale structure; Chen et al. 2016) to numerical (e.g. under-resolved haloes with higher sparsities, misidentification of the centre, etc.). Thus, we deal with all haloes not falling into the *totally relaxed* category by defining a combined *relaxedness* indicator, in the manner of Haggar et al. (2020; see also Kuchner et al. 2020; Zhang et al. 2021a; Gouin, Gallo & Aghanim 2022) but adding weights which account for the fact that some dynamical state indicators can be more insightful than others at any given particular epoch.

$$\chi = \left[w_1 \left(\frac{\Delta_r}{\Delta_r^{\rm thr}} \right)^2 + w_2 \left(\frac{\eta - 1}{\eta^{\rm thr} - 1} \right)^2 + w_3 \left(\frac{\langle \widetilde{v_r} \rangle_{\rm DM}}{\langle \widetilde{v_r} \rangle_{\rm DM}^{\rm thr}} \right)^2 + w_4 \left(\frac{s_{200c,500c} - 1}{s_{200c,500c}^{\rm thr} - 1} \right)^2 + w_5 \left(\frac{\epsilon}{\epsilon^{\rm thr}} \right)^2 \right]^{-1/2}. \tag{14}$$

The weights, $\{w_i\}_{i=1}^5$, are normalized so that $\sum_{i=1}^5 w_i = 1$, and are fixed at each redshift bin to be proportional to the performance of their corresponding indicator in splitting the merging and non-merging subsamples of the fiducial classifications. In particular, we set $w_i \propto \epsilon_{\rm relaxed}\epsilon_{\rm unrelaxed} - 0.25$ (the absolute values being set by the closure relation $\sum_{i=1}^5 w_i = 1$). If, at a given redshift bin, $\epsilon_{\rm relaxed}\epsilon_{\rm unrelaxed} \leq 0.25$, we consider that the particular indicator is not meaningful and its weight is set to $w_i = 0$.

A particular halo that does not belong to the *totally relaxed category* will be classified as *marginally relaxed* if $\chi \geq 1$, and *disturbed* whenever $\chi < 1$. Additionally, this classification scheme can naturally handle missing data. For instance, if $s_{200c,\,500c}$ is missing (e.g. due to a low resolution not enabling to resolve R_{500c}), one can simply evaluate χ neglecting the sparsity term (and multiplying χ by a factor $\sqrt{1 - w_4}$; or, alternatively, renormalizing the weights after setting $w_4 = 0$).

2.4.4 *Redshift evolution of the thresholds and weights*

With the procedure outlined in Sections 2.4.2 and 2.4.3, we obtain a threshold and a weight for each dynamical state indicator at each of the redshift bins specified in Table 1. In order to obtain a continuous trend for each of these parameters, we fit them to polynomial functions of arbitrary degree.

First, we estimate the uncertainties in the thresholds ($X_i^{\rm thr}$) and weights (w_i) by computing the standard deviation of the distribution of these parameters obtained in 1000 bootstrap resamplings (Efron 1979). Then, we fit the redshift evolution of the given parameter to polynomial functions of increasing degree, until the p-value of the highest degree coefficient falls above $p = 0.046$ (low significance), or the reduced chi-squared falls below 1 (indicating possible overfitting of the model). Fits are performed using least squares weighted to the inverse of the variance of each data point.

3 RESULTS

Following the procedure described in Sections 2.4.2 and 2.4.4 over the whole sample, we have found the optimal thresholds for the dynamical state indicators, and fitted them to the best possible polynomial models. The results are shown in Fig. 3, from top to bottom, for the centre offset, virial ratio, mean radial velocity, sparsity, and ellipticity thresholds.

Most of the thresholds on the assembly state indicators present a clear redshift evolution. At earlier times, the thresholds on the dynamical state indicators tend to take higher values, reflecting the fact that haloes at earlier times were more irregular or exhibited more disturbed features, even when not having experienced any relevant merging activity or growth during the last dynamical time.

The evolution of the thresholds ranges from very mild or almost non-existent (e.g. $\Delta_r^{\rm thr}$, $\varepsilon^{\rm thr}$) to noticeable (and definitely worth taking into account; e.g. $\eta^{\rm thr}$, $s_{200c,500c}^{\rm thr}$, $\langle \widetilde{v_r} \rangle^{\rm thr}$). This unequivocally evidences that fixed, set thresholds on certain parameters may not be able to correctly discriminate relaxed from merging haloes through the whole evolutionary history of the objects, especially when delving into the realm of high-redshift haloes.

The thresholds can be fitted by the following equations (solid lines in Fig. 3, whose uncertainties are represented by the shaded regions), valid for $0 \leq z \leq 5$, where the figures in parentheses correspond to the uncertainty in the two last digits of each coefficient:

$$\Delta_r^{\rm thr}(z) = 0.0849(13) \tag{15}$$

$$\eta^{\rm thr}(z) = 1.3383(56) + 0.197(11)z - 0.0276(32)z^2 \tag{16}$$

$$\langle \widetilde{v_r} \rangle_{\rm DM}^{\rm thr}(z) = 0.0718(22) + 0.0056(14)z \tag{17}$$

$$s_{200c,500c}^{\rm thr}(z) = 1.491(16) + 0.064(37)z - 0.031(22)z^2 + 0.0060(35)z^3 \tag{18}$$

$$\varepsilon^{\rm thr}(z) = 0.2696(27). \tag{19}$$

Based on the performance of each assembly state indicator in matching the fiducial classification, we fix the weights of each indicator in equation (14) as described in Section 2.4.3. The results are summarized in Fig. 4, which is analogous to Fig. 3 but this time showing the weights instead of the thresholds. Note that, if all indicators were equivalently important, $w_i = 0.2\,\forall i$. Thus, $w_i > 0.2$ ($w_i < 0.2$) implies above-average (below-average) performance for the given dynamical state indicator at the given epoch.

Interestingly, the importance of each indicator in determining the dynamical state of DM haloes varies strongly with redshift. For example, one of the most widely used indicators, the centre offset Δ_r, is exceedingly effective in discriminating the disturbed haloes at high redshift, but its effectiveness declines steeply with decreasing redshift and has slightly below-average performance at $z \simeq 0$. As an example of the opposite trend, the virial ratio, η, appears to be irrelevant at high redshift ($z \gtrsim 2$), and is only useful at low redshifts ($\lesssim 1$). This dissimilar behaviour between centre offset and virial ratio

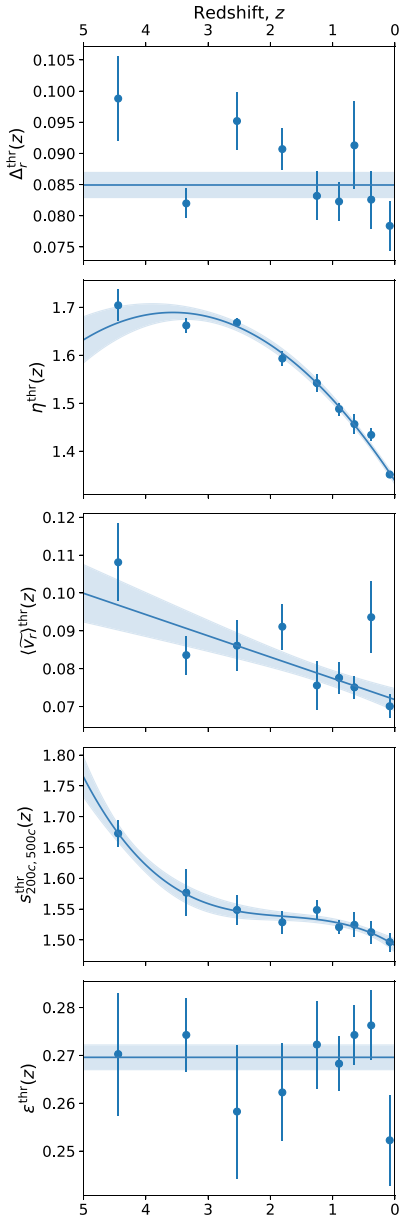


Figure 3. Redshift evolution of the thresholds on the dynamical state indicators. From top to bottom, the panels refer to the centre offset ($\Delta_r^{\rm thr}$), virial ratio ($\eta^{\rm thr}$), mean radial velocity ($\langle\widetilde{v_r}\rangle^{\rm thr}_{\rm DM}$), sparsity ($s^{\rm thr}_{200c,500c}$), and ellipticity ($\varepsilon^{\rm thr}$) thresholds. Dots correspond to the optimal threshold obtained within the redshift bin, with the error bars obtained by means of bootstrap resampling. Solid lines correspond to the best polynomial fits, with their (16–84) per cent confidence interval as the shaded region.

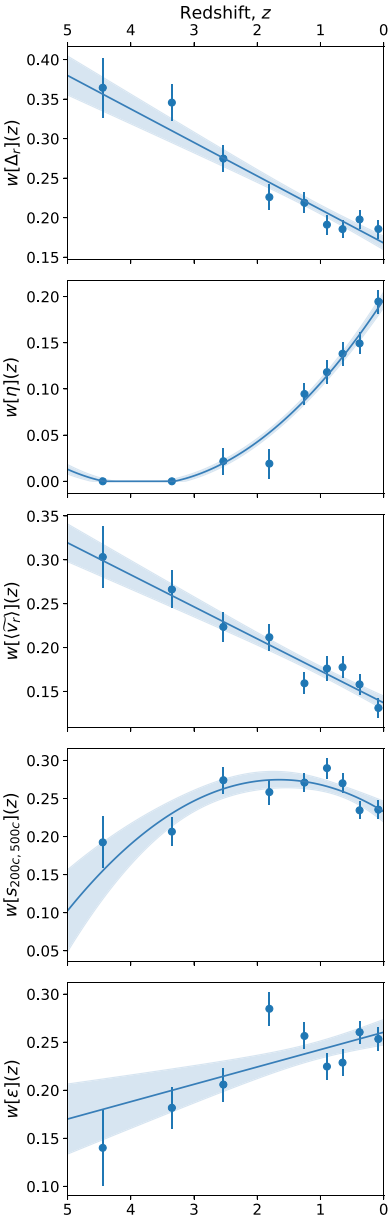


Figure 4. Redshift evolution of the weights on the dynamical state indicators. From top to bottom, the panels refer to the centre offset ($w[\Delta_r]$), virial ratio ($w[\eta]$), mean radial velocity ($w[\langle\widetilde{v_r}\rangle]$), sparsity ($w[s_{200c,\,500c}]$), and ellipticity ($w[\varepsilon]$) weights. Dots correspond to the weights obtained within the redshift bin, with the error bars obtained by means of bootstrap resampling. Solid lines correspond to the best polynomial fits, with their (16–84) per cent confidence interval as the shaded region.

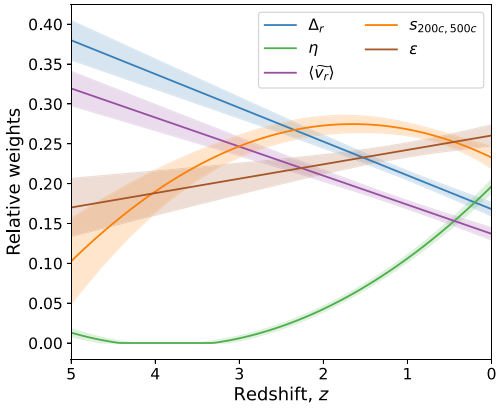


Figure 5. Overall fitted redshift evolution of the weights of the dynamical state indicators, using the complete sample.

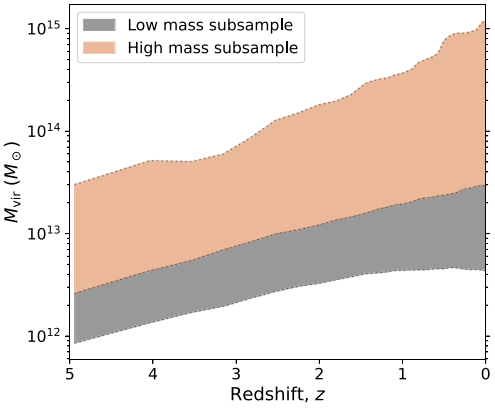


Figure 6. Definition of the mass subsamples in terms of mass, as a function of redshift. The grey (salmon) shaded regions correspond to the low-mass (high-mass) subsamples. The dotted lines mark the mass limits.

is also reported by the analyses at high redshift of Davis, D'Aloisio & Natarajan (2011).

As a purely dynamical parameter, the mean radial velocity $\langle\widetilde{v_r}\rangle$ is especially relevant at high redshift, likely due to the fact that smooth (nearly radial) accretion could be more important at these stages given the relatively higher density in the surroundings of the halo. Sparsity, as well as ellipticity, are especially correlated with the fiducial dynamical state classification at more recent redshifts, although they cannot generally be neglected at any epoch. As a matter of fact, at low redshift, ε is the most relevant indicator of the dynamical state of haloes.

With the same procedure as above, we have fitted the weights to polynomial functions capturing their evolution (solid lines in Fig. 4, whose uncertainties are represented by the shaded regions), valid for $0 \leq z \leq 5$:

$$w[\Delta_r](z) \propto 0.1679(70) + 0.0423(50)z \tag{20}$$

$$w[\eta](z) \propto 0.1965(78) - 0.1037(60)z + 0.0134(11)z^2 \tag{21}$$

$$w[\langle\widetilde{v_r}\rangle_{\rm DM}](z) \propto 0.1370(70) + 0.0364(48)z \tag{22}$$

$$w[s_{200c,500c}](z) \propto 0.2327(97) + 0.051(14)z - 0.0153(38)z^2 \tag{23}$$

$$w[\varepsilon](z) \propto 0.2603(75) - 0.0181(51)z \tag{24}$$

We note that, while at any epoch the data points fulfilled $\sum_{i=1}^{5} w_i = 1$, this is not necessarily true for the fitting polynomials evaluated at any arbitrary redshift (although it holds to a few per cents). Thus, they must be normalized by their sum before plugging them into equation (14).

For better comparison of the relative importance of each of the indicators, Fig. 5 presents the fits of the weights on each dynamical state indicator, as a function of decreasing redshift. The behaviour can be roughly summarized as

(i) At high redshift ($z \in [2, 5]$), the centre offset provides the most insightful information about the recent assembly activity. This can be primarily complemented by the mean radial velocity (especially at $z \gtrsim 3$), or sparsity and ellipticity (at $z \lesssim 3$). The virial ratio does not seem to provide any insight on the dynamical state at high redshift.

(ii) At intermediate redshifts, ($z \in [1, 2]$), sparsity, ellipticity, and centre offset provide similarly useful information about the dynamical state. The relevance of the virial ratio is still limited at this epoch.

(iii) At low redshifts ($z \lesssim 1$), the ellipticity of the DM halo correlates exceptionally well with the dynamical state, as well as sparsity does. While no dynamical state indicator is negligible at this stage, centre offset and virial ratio also present reasonable performances, while mean radial velocity is the least useful indicator at this time.

3.1 Dependence on halo mass

The previous analyses have considered all haloes on an equal footing, despite their broad distribution in masses. Invoking self-similarity (see e.g. Navarro et al. 2010 for a thorough analysis on the level of self-similarity of haloes), it may be argued that the same thresholds and weights could be used for all halo masses. However, the fact that many halo properties (related to the dynamical state indicators in our work) scale with mass demands to explicitly check how our results depend on the mass scale of the haloes being considered.

We have split the complete sample, introduced in Section 2.4.3, in two subsamples, namely a low-mass and a high-mass subsample. The high-mass subsample contains, at each redshift, the 20 per cent most massive haloes in the complete sample. This is chosen so as to contain, at redshift 0, all the haloes associated with massive groups and clusters ($M_{\rm DM} > 3 \times 10^{13}$ M$_\odot$). Fig. 6 shows the evolution of the mass limits on each subsample. Note that, therefore, our mass groups do not correspond to fixed-mass ranges, but rather to two sub-populations with a redshift-dependent mass threshold. While it would definetely be interesting to explore the dependence of our thresholds and weights with actual mass ranges, our limited statistics prevent us from this goal and we may defer this for future work.

Repeating the analyses above separately for each mass subsample, we can infer the mass dependence of the thresholds on the dynamical state indicators, and that of their corresponding weights in equation (14). The redshift evolution of the thresholds for each mass subsample is presented in Fig. 7, which is analogous to Fig. 3 but displaying only the fits for clarity. The same colour coding as in Fig. 6 is used here.

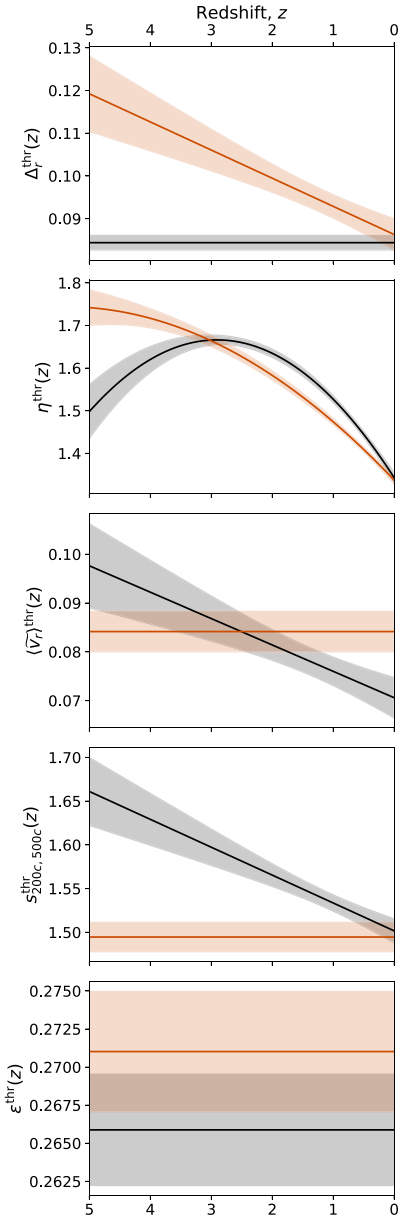


Figure 7. Mass dependence of the redshift evolution of the thresholds on the dynamical state indicators. The figure is analogous to Fig. 3, with each line corresponding to the fit performed with a mass subsample (the same colour coding as in Fig. 6 is used: gray [salmon] lines correspond to the low-mass [high-mass] subsamples), and the shaded regions enclosing 1σ confidence intervals.

For some indicators, such as ellipticity, there is no hint for any significant trend of the evolution of the threshold with mass, at least within the statistical uncertainties given by our sample size. That is to say, at least within the mass range considered in this work (roughly, $[10^{12} - 10^{15}]\,M_\odot$), the relaxation criteria based on this indicator can be used regardless of the scale of the objects (as customarily done with many indicators, e.g. Power et al. 2012).

However, the rest of indicators of the dynamical state do present significant dependence on halo mass. In particular, the threshold on Δ_r remains constant with redshift for the low-mass subsample, while it increases linearly with increasing redshift for group and cluster-sized DM haloes. This would suggest that imposing a constant threshold on the centre offset may be too conservative and could, for massive haloes at high redshift, artificially increase in excess the number of disturbed haloes.

Regarding virial ratio, there is a minor trend with mass at intermediate and low redshifts ($z \lesssim 3$), with higher-mass haloes preferring a slightly more stringent threshold to separate dynamically relaxed and unrelaxed haloes, but this difference corresponds to a small variation on the value of the parameter ($\Delta\eta^{\rm thr} \sim 0.05$). The most striking difference appears at high redshift, but is not relevant since we have found that η itself is not meaningful at high redshift (see the second panel in Fig. 4, as well as Fig. 8 below).

The mass dependence of $\langle\widetilde{v_r}\rangle$ is moderate, with massive haloes preferring a constant threshold around $\langle\widetilde{v_r}\rangle \approx 0.085$ and low-mass systems displaying a decreasing trend with decreasing redshift.

Last, the two mass subsamples present a different behaviours with respect to the threshold on halo sparsity, $s^{\rm thr}_{200c,500c}$. In this case, lower mass haloes require consistently larger thresholds on sparsity to discriminate relaxed and merging objects. While smaller haloes tend to be more concentrated (see for instance, Dutton & Macciò 2014), the mass-dependence of sparsity is much more contained (Corasaniti et al. 2018; Corasaniti & Rasera 2019). It might be the case that, both for physical (e.g. stronger influence of the environment) and numerical (e.g. less mass resolution elements leading to more unresolved central regions) reasons, low-mass haloes present a broader distribution in sparsities (see e.g. fig. 10 in Balmès et al. 2014) and thus a larger sparsity threshold is required at the low-mass end.

Finally, we show in Fig. 8 the evolution of the weights on each of the dynamical state indicators (as they appear in equation 14), for each of the mass subsamples. Besides the general trends already analysed for the whole sample when Fig. 4 was presented, several differences emerge between the high-mass and the low-mass subsamples, especially at intermediate and high redshifts. At low redshift, however, the weights are essentially compatible amongst the two subsamples, with only a small hint of virial ratio and centre offset being – comparatively – more effective in higher-mass haloes.

At intermediate redshifts, $z \sim 2$–3, high-mass haloes find ellipticity and mean radial velocity to be better indicators of their dynamical state, and centre offset and sparsity comparatively worse ones, when confronted to the low-mass sample. At high redshifts, $z \sim 5$, the performance of sparsity gets more penalized for lower mass haloes and, in these objects, centre offset can be relatively more important than in higher mass haloes. This further highlights that, besides not being able to put fix criteria (in the sense of them not evolving with redshift) for assessing the dynamical state of dark matter haloes, they have to be carefully chosen depending on the scale of the object being studied.

We provide fits for the thresholds and weights for the high-mass (group and cluster-sized) subsample in Appendix A.

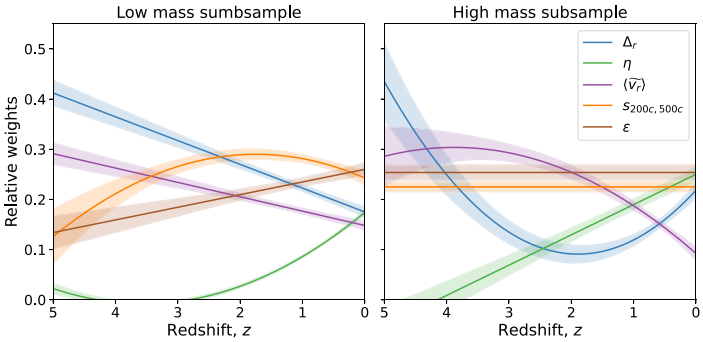


Figure 8. Mass dependence of the redshift evolution of the weights on the dynamical state indicators. Each panel is analogous to Fig. 5 (also using the same colour coding), showing the results for each of the mass samples defined in Fig. 6: low-mass (left-hand panel) and high-mass (right-hand panel) subsamples.

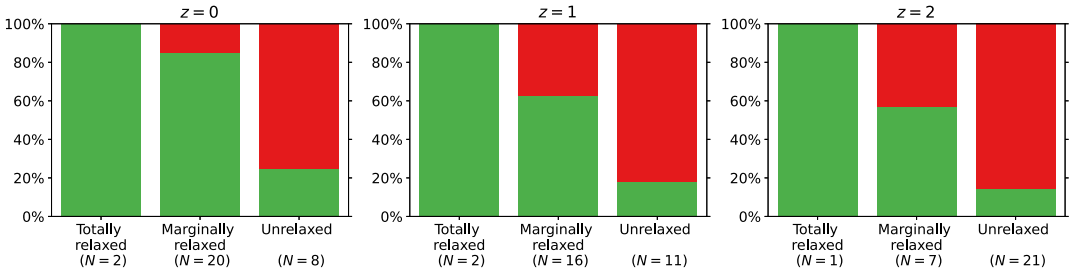


Figure 9. Classification summary at redshifts $z = 0$ (left-hand panel), $z = 1$ (middle panel), and $z = 2$ (right-hand panel). Within each panel, each of the columns corresponds to one of the possible classifications (totally relaxed, if $X_i < X_i^{\rm thr}\ \forall i$; marginally relaxed, if the previous condition fails but $\chi \geq 1$; or unrelaxed otherwise). Within each bar, the green portion represents the fraction of the haloes which have not suffered any mergers nor strong accretion, while the red portion corresponds to the fraction of haloes having suffered mergers or strong accretion (i.e. the colour encodes the fiducial classification). Below each column, N indicates the number of objects falling into each category.

3.2 Classification assessment

The aim of this section is to validate to which extent the dynamical state classification introduced in this work, which only uses information at a given time-step, is capable of predicting the merging state of the DM halo. That is to say, whether we can predict the fiducial classification (Section 2.2.1) based on the dynamical state indicators, when confronting our method with haloes from a different simulation (corresponding to different resolution, gravity solver, etc.).

We use public simulation data from the suite `CAMELS` (Villaescusa-Navarro et al. 2021, 2022), which contains over 4000 simulations of $25h^{-1}$ Mpc cubic, periodic volumes run with different physics, cosmological and astrophysical parameters, and numerical codes. In particular, we have analysed the haloes in the `IllustrisTNG-DM CV-0` simulation, which corresponds to a DM-only simulation run with `Arepo` (Springel 2010; Weinberger, Springel & Pakmor 2020). `Arepo` implements a Tree + Particle-Mesh approach (Bagla 2002) for solving the evolution of DM, thus providing a high dynamical range even though the number of particles ($N_{\rm part} = 256^3$) is rather small. The `CV-0` simulation corresponds to a background cosmology with $h = 0.6711$, $\Omega_{\rm m} = 0.3$, $n_s = 0.9624$, and $\sigma_8 = 0.8$, the initial conditions having been set up at $z_{\rm ini} = 127$.

We have extracted the halo catalogues and merger trees with `ASOHF` (Vallés-Pérez et al. 2022) by following the exact same procedure described in Section 2.2, and determined the dynamical state indicators (Section 2.3). For our analyses, we have considered the 30 most massive haloes at each time, which corresponds to a similar mass limit as in the main analysis (cf. Fig. 1). Out of these 30 haloes, we have dropped the ones that we are not able to trace back in time for at least one dynamical time (which is most usually none or one halo, at the considered epochs).

In Fig. 9, we assess the performance of our classification scheme at three cosmological epochs ($z = 0$, 1, and 2, for the panels left to right) by computing, at each time, the fraction of haloes in each class (totally relaxed, marginally relaxed and unrelaxed) which have recently suffered mergers or strong accretion (red), or has undergone a quiet evolution (green). We note that, when evaluating the dynamical state criteria, at any given z we only apply the indicators with weight $w_i(z) > 0.05$. Otherwise, we consider the given dynamical state indicator as not meaningful at that particular epoch. While this particular threshold is arbitrary, it is a sensible choice and the results do not depend strongly on variations around this value. According to Fig. 5, this only removes the virial ratio, η, at $z \gtrsim 1.9$.

The totally relaxed subsample is, naturally, the smallest one, since it is defined rather conservatively as the set of haloes simultaneously fulfilling all five relaxation criteria. It typically contains ~ 10 per cent of the haloes (slightly lower in this case; nevertheless, the statistics are small). Within this test, all haloes being classified as totally relaxed have not suffered any major (minor) merger within one (half) $\tau_{\rm dyn}$ or built up more than half of their mass in the last

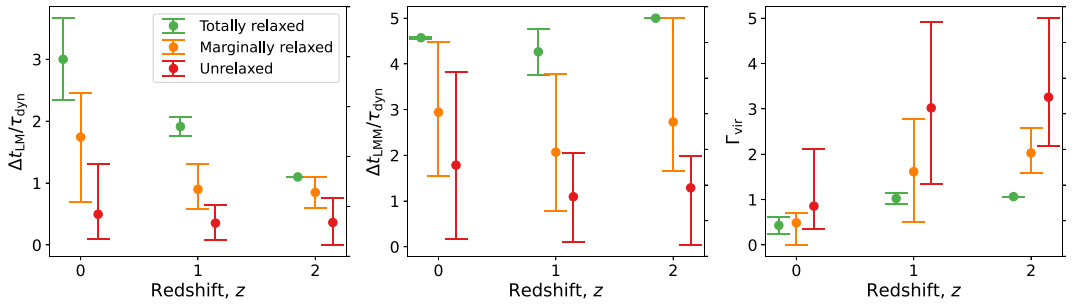


Figure 10. Trends of the evolutionary properties of the haloes according to their dynamical state classification. In each panel, red, orange, and green dots correspond, respectively, to the unrelaxed, marginally relaxed and totally relaxed subsamples. At each redshift (encoded in the horizontal axis) the dot and the error bars represent, respectively, the mean and the 1σ dispersion of the distribution of the given variable within the subsample. The left-hand panel presents the time since the last merger (either major or minor) in units of the dynamical time ($\Delta t_{\rm LM}/\tau_{\rm dyn}$), the central panel corresponds to the time since the last major merger ($\Delta t_{\rm LMM}/\tau_{\rm dyn}$), and the right-hand panel shows the accretion rate $\Gamma_{\rm vir}$.

dynamical time, thus proving to be a selection of the relaxed sample with high specificity.

The marginally relaxed sample is the most numerous at low redshifts ($z \lesssim 1$) and mostly contains objects which have not experienced any relevant merging or accretion activity, although the fraction of objects having experienced it increases with increasing redshift (from ~ 15 per cent at $z \simeq 0$ to ~ 40 per cent at $z \simeq 2$). The unrelaxed subsample, which is especially numerous at high redshift when merger rates are higher (see e.g. Wetzel, Cohn & White 2009), contains mostly merging objects, although a small fraction (10–25 per cent) of objects not experiencing mergers or accretion seem to fall into this category. This may happen because a halo appears to be disturbed, even when not merging or accreting intensely, due to environmental effects (e.g. strong tidal field due to the presence of another nearby massive halo, for instance in a pre-merger state), or even numerical effects (mainly associated to low resolution). Naturally, it may also be the case that unrelaxedness after a major merger is last for longer than $1\tau_{\rm dyn}(z)$, since the fiducial classification in Section 2.2.1 was only a rough estimation.

Expanding upon the previous figure, in Fig. 10 we focus on some quantities tied to the merger and accretion history of the haloes. In particular, the left-hand panel represents, at three redshifts ($z = 0$, 1, and 2, respectively, from left to right) and for the three subsamples, the distribution of the time since the last merger (either major or minor) in units of the dynamical time. Haloes classified as unrelaxed have usually suffered some merger recently while, on the other hand, the totally relaxed sample has typically not experienced any merging activity since several dynamical times ago. A similar situation is seen for the major mergers (middle panel), although naturally not all unrelaxed haloes have suffered a major merger (the unrelaxedness can be due to one or several minor mergers, or smooth accretion, as well). However, the same trend holds, with major mergers having occurred a longer time ago as we move from unrelaxed to marginally relaxed, and to totally relaxed haloes.

Finally, the third panel presents, in a similar way, the distribution of accretion rates $\Gamma_{\rm vir}$, which are computed according to the prescription of Diemer & Kravtsov (2014),

$$\Gamma_{\rm vir} = \frac{\Delta \log M_{\rm vir}}{\Delta \log a} \tag{25}$$

with a being the scale factor of the Universe, and the increments computed over a dynamical time following its definition in equation (2). The figure shows, in line with the previous results, an increasing trend of the accretion rate when moving from the relaxed to the more disturbed subsamples, within a wide redshift interval. This reflects how the dynamical state classification presented in this work, which only uses information at a fixed time, can offer insight on the temporal evolution of the systems over the last dynamical time.

3.2.1 *Does a smaller set of indicators provide similar insight?*

Lastly, it might be interesting to assess whether a single dynamical state indicator, or a combination of them, is capable of providing a similarly accurate classification; i.e. to motivate why it is important to involve a high number of indicators. This is briefly exemplified in Table 2, where we show the classification summary for each indicator (or combination). In particular, for each classification class we give the number of haloes falling into this class (and its percentage with respect to the total), and the fraction of them which is unrelaxed according to the fiducial classification ($f_{\rm merging}$). Ideally, this fraction would be 0 for the totally relaxed class and 1 for the unrelaxed class. Naturally, when using only one indicator, there is no *marginally relaxed* or intermediate category, since the value of the dynamical state indicator can only be above or below the threshold. In the case of using two indicators, we have defined the marginally relaxed sample as the set of haloes fulfilling only one of the two relaxedness conditions, as it is often done in the literature (e.g. Biffi et al. 2016; Planelles et al. 2017).

Generally speaking, involving only one dynamical state indicator leads to far poorer results, since the relaxed sample gets often contaminated (around ~ 40 per cent) by haloes which have suffered mergers. Likewise, the unrelaxed sample may end up containing a high fraction of haloes undergoing quiescent evolution for some indicators (e.g. ϵ; although the particular results have to be considered carefully due to the reduced statistics).

Interestingly, when using a combination of virial ratio and centre offset, which is a common option in the literature (e.g. Power et al. 2012), we are still not able to pick out all merging haloes with these criteria and even the totally relaxed subsample gets contaminated with ~ 40 per cent of merging haloes. Other common options in the literature are the combination of mass ratio and centre offset (De Luca et al. 2021), or centre offset, virial ratio and mass ratio (Cui et al. 2017; Haggar et al. 2020). We have also tested these combinations, taking $f_{\rm sub}^{\rm thr} = 0.1$ from the afore-

Table 2. Classification properties using only one or a combination of dynamical state indicators, exemplified at $z = 1$. Each row corresponds to one dynamical indicator or combination, and for each classification class we give the number of objects falling into the class (N, and the percentage with respect to the total) and the fraction of them which is unrelaxed according to the fiducial classification ($f_{\rm merging}$). When only one indicator is used, there is no intermediate (*marginally relaxed*) class. The first block corresponds to the individual indicators involved in this work. The second block contains several combinations widely used in the literature, with the fitted thresholds in this work (for $f_{\rm sub}$, not involved in this work, we use $f_{\rm sub}^{\rm thr} = 0.1$). The last row corresponds to the complete method introduced here, using the five indicators (therefore, these results are the same shown in the central panel of Fig. 9).

Indicator(s)	Totally relaxed		Marginally relaxed		Unrelaxed	
	N	$f_{\rm merging}$	N	$f_{\rm merging}$	N	$f_{\rm merging}$
Δ_r	20 (69%)	0.40	–	–	9 (31%)	0.78
η	25 (86%)	0.48	–	–	4 (14%)	0.75
$\langle\widetilde{v}_r\rangle$	21 (72%)	0.43	–	–	8 (28%)	0.75
$s_{200c,\,500c}$	17 (59%)	0.47	–	–	12 (41%)	0.58
ϵ	10 (34%)	0.50	–	–	19 (66%)	0.53
Δ_r and η	19 (66%)	0.42	7 (24%)	0.57	3 (10%)	1.00
$f_{\rm sub}$	23 (79%)	0.48	–	–	6 (21%)	0.67
η and $f_{\rm sub}$	18 (62%)	0.44	7 (24%)	0.43	4 (14%)	1.00
Δ_r, η, and $f_{\rm sub}$	17 (59%)	0.47	3 (10%)	0.33	9 (31%)	0.67
Full set of indicators	2 (7%)	0.00	16 (55%)	0.38	11 (38%)	0.82

mentioned references, since we have not involved this indicator in our previous analyses. In these cases, the results are similar to the Δ_r and η combination. This highlights the necessity of involving and combining as many indicators of the dynamical state as possible. When using the full set of indicators derived in this work, the totally relaxed subsample is rather small, due to its conservative definition. However, even our marginally relaxed subsample is purer (contains a smaller fraction of merging/accreting haloes) than the totally relaxed sample of the previous combinations, proving to provide a robust splitting of haloes according to their dynamical state.

4 DISCUSSION AND CONCLUSIONS

To fully exploit the capabilities of ongoing surveys (e.g. eROSITA; Ghirardini et al. 2022), and in the advent of upcoming instruments over the electromagnetic spectrum (from X-ray, e.g. ATHENA, Nandra et al. 2013; to radio, e.g. SKA, Acosta-Pulido et al. 2015; going through the optical, e.g. EUCLID, Sartoris et al. 2016; Euclid Collaboration 2019), which will provide samples of galaxies and galaxy clusters unprecedented in size and depth, it remains crucial to provide reliable indicators of the dynamical state (which in turns is a fast proxy of the – recent – evolution of the system). As a first step towards that aim, using N-body + hydrodynamics simulations, in this work we have systematically analysed how to best combine a series of quantities which can be measured from simulation data at a given time in order to be able to detect the presence of mergers and/or ongoing strong accretion.

As a result, we have built an algorithm that combines a series of different indicators of the dynamical state of a DM halo (namely, its centre offset Δ_r, the virial ratio η, the mean radial velocity $\langle\widetilde{v}_r\rangle$, the sparsity $s_{200c,\,500c}$, and the ellipticity ε) in order to classify haloes within three classes. The *totally relaxed* haloes, comprising the objects simultaneously fulfilling all relaxedness conditions (which are redshift-dependent, in general), is a conservatively defined subsample which, therefore, only contains around ~ 10 per cent of the haloes at a given time. Haloes where some relaxedness condition may fail, but are remarkably relaxed according to the rest of indicators may be categorized in the *marginally relaxed* class, using a criterion similar to Haggar et al. (2020), but allowing different indicators to have different (redshift-dependent) weights, which are tuned based on the performance of each indicator on telling relaxed and unrelaxed haloes apart. Thus, we defined a *relaxedness parametre* (χ), which tells *marginally relaxed* ($\chi \geq 1$) and *unrelaxed* ($\chi < 1$) apart. The fits for the redshift dependence of the thresholds and weights are given in equations (15)–(24), while equivalent results for *massive* haloes are provided in Appendix A.

Furthermore, we have confronted our classification scheme against an independent DM-only simulation from the CAMELS suite (Villaescusa-Navarro et al. 2021, 2022), corresponding to different input physics, initial conditions, and numerical solvers. Using it, we find that our algorithm performs a clean splitting of relaxed and unrelaxed haloes across a wide cosmic time interval, and that this classification improves upon the usage of any single indicator or some widely used combinations (Δ_r and η; $f_{\rm sub}$ and η; or Δ_r, η and $f_{\rm sub}$).

As a qualitative summary of the main highlights of the classification scheme, we can mention:

(i) Placing fix thresholds (which do not evolve with redshift) is generally undesirable. While some indicators do not show strong evolution of their optimal thresholds with redshift (e.g. ellipticity, centre offset), others do (e.g. sparsity, mean radial velocity; all tending to increase with redshift). This has important consequences, since it implies that classification schemes for the dynamical state of haloes that are set at $z = 0$ cannot be directly used at high redshifts.

- At high halo mass (see the precise definition of the high-mass subsample in Fig. 6), however, the results are slightly changed: in particular, the redshift dependence of the thresholds on $\langle\widetilde{v}_r\rangle$ and $s_{200c,\,500c}$ is not significant anymore, while the classification based on the Δ_r benefits from an increasing trend with increasing redshift. This warns us that the classification cannot be universal, and that haloes on different mass scales may need slightly modified criteria.

(ii) At low redshift ($z \lesssim 1$), even though all indicators offer insight into the merging state of the halo, it is sparsity and ellipticity of the DM halo the ones which provide the most valuable information, well beyond other, more widely used indicators such as centre offset or virial ratio. Nevertheless, the fact that all relative weights are not

very dissimilar at this epoch (see Fig. 5) means that the classification scheme can importantly benefit from combining as many indicators as possible.

- The difference in weights amongst the different observables (except $\langle\widetilde{v_r}\rangle$) is importantly reduced when looking at the high-mass sample (right-hand panel in Fig. 8), reinforcing that, for group- and cluster-sized haloes at low redshift, it may be important to combine all indicators suggested in this work.

(iii) At high redshifts ($z \gtrsim 3$), η becomes irrelevant for the determination of the assembly state of the halo, while centre offset and mean radial velocity become, by far, the dominant indicators.

- Again, the differences are lower for the high-mass subsample, but the prevalence of Δ_r and $\langle\widetilde{v_r}\rangle$ still holds.

In this work, we have focused on the determination of the assembly state of DM haloes using the full information contained in a snapshot of a numerical simulation. The motivation for this is two-fold. On the one hand, it is important to devise efficient methods to classify large samples of simulated haloes, especially given the ever-growing trend of simulations, both in size and resolution (see e.g. Angulo & Hahn 2022, their table 1), made possible by the increasing computational power available. On the other hand, it serves as a first step, which can be further connected to observations using projected data or, more realistically, mock multiwavelength observations (e.g. Planelles et al. 2018).

Much of the information comprised in the dynamical state indicators we involve in this work can be lost, or at least hindered, when moving from the 3D description to the 2D observed data. The first, most natural consequence is the effect of projection on any geometrical indicator, such as the centre offset, ellipticity or the mean radial velocity. For the case of centre offset and ellipticity, the measured values will only be a lower limit, with the actual 3D value depending on the inclination between the direction of the offset, or the plane containing the major and minor axis, with the line of sight.

Regarding the mean radial velocity, which is especially important for determining the dynamical state at high redshift, besides the difficulty induced by projection (only velocities along the line of sight, and distances on the plane of the sky, can be measured), future kinetic Sunyaev–Zel'dovich (kSZ) observations could be able to provide some constraints on proper velocities of the intracluster medium (ICM; see for instance, the estimates of Baldi et al. 2018 about the kSZ effect due to the coherent rotation of the ICM), even for high-redshift objects since the SZ effect is essentially distance-independent (e.g. Voit 2005). Even though the dynamics of the ICM, especially in the inner regions of haloes, may differ significantly from those of the DM halo, probing the velocity field of the diffuse gas in haloes could supply useful insight onto the dynamical state of haloes at high redshift.

Lastly, sparsity may be a suitable option for observations, given its good performance shown across the whole redshift span considered here (especially, for high-mass haloes). However, care must be taken when using this quantity: here, we have defined sparsity from the DM masses obtained from the full, 3D information. However, in observations, masses can be obtained from several methods (e.g. hydrostatic, lensing, and caustic masses), and biases amongst them are non-negligible (see for instance, Lovisari et al. 2020). Moreover, mass biases tend to correlate with the merging state (Bennett & Sijacki 2022; cf. Gianfagna et al. 2022) and, while the quotient of two masses at different apertures derived from the same method may cancel out part of these biases, the fact that the bias itself depends on the aperture and the large object-to-object scatter still make the interpretation non-trivial and deserve further attention themselves.

This work provides a motivated definition of a scheme for classifying DM haloes according to their dynamical status, based on simple properties which can be readily extracted from the outputs of typical halo finders. Future work will need to deal with the connection of these dynamical and morphological properties of the DM halo with the baryonic component, as well as the application to observations, in order to being able to extract the largest possible amount of information about the assembly state of haloes from future observational campaigns.

ACKNOWLEDGEMENTS

We gratefully thank the anonymous referee for their valuable feedback, which has helped us to improve the quality of this manuscript. This work has been supported by the Agencia Estatal de Investigación Española (AEI; grant PID2019-107427GB-C33), by the Ministerio de Ciencia e Innovación (MCIN) within the Plan de Recuperación, Transformación y Resiliencia del Gobierno de España through the project ASFAE/2022/001, with funding from European Union NextGenerationEU (PRTR-C17.I1), and by the Generalitat Valenciana (grant PROMETEO/2019/071). DV acknowledges support from Universitat de València through an Atracció de Talent fellowship, and gratefully thanks the hospitality of the Dipartimento di Fisica e Astronomia of the Università di Bologna, where part of this work was done during a research stay funded by Universitat de València. We also thank F. Vazza and A. Ragagnin for fruitful scientific conversations. Simulations have been carried out with the supercomputer Lluís Vives at the Servei d'Informàtica of the Universitat de València. This research has made use of the following open-source packages: `NUMPY` (Harris et al. 2020), `SCIPY` (Virtanen et al. 2020), `MATPLOTLIB` (Hunter 2007), `statsmodels` (Seabold & Perktold 2010), `scikit-learn` (Pedregosa et al. 2011), and `Colossus` (Diemer 2018).

DATA AVAILABILITY

The data underlying this article will be shared upon reasonable request to the corresponding author.

REFERENCES

Acosta-Pulido J. A. et al., 2015, preprint (arXiv:1506.03474)
Angelinelli M., Vazza F., Giocoli C., Ettori S., Jones T. W., Brunetti G., Brüggen M., Eckert D., 2020, MNRAS, 495, 864
Angulo R. E., Hahn O., 2022, Liv. Rev. Comput. Astrophys., 8, 1
Bagla J. S., 2002, J. Astrophys. Astron., 23, 185
Baldi A. S., De Petris M., Sembolini F., Yepes G., Lamagna L., Rasia E., 2017, MNRAS, 465, 2584
Baldi A. S., De Petris M., Sembolini F., Yepes G., Cui W., Lamagna L., 2018, MNRAS, 479, 4028
Balmès I., Rasera Y., Corasaniti P. S., Alimi J. M., 2014, MNRAS, 437, 2328
Bennett J. S., Sijacki D., 2022, MNRAS, 514, 313
Berger M. J., Colella P., 1989, J. Computat. Phys., 82, 64
Biffi V. et al., 2016, ApJ, 827, 112
Bryan G. L., Norman M. L., 1998, ApJ, 495, 80
Buote D. A., Tsai J. C., 1995, ApJ, 452, 522
Capalbo V., De Petris M., De Luca F., Cui W., Yepes G., Knebe A., Rasia E., 2021, MNRAS, 503, 6155
Cerini G., Cappelluti N., Natarajan P., 2022, preprint (arXiv:2209.06831)
Chandrasekhar S., 1961, Hydrodynamic and Hydromagnetic Stability. Clarendon, Oxford

Chen S., Wang H., Mo H. J., Shi J., 2016, ApJ, 825, 49
Chen H., Avestruz C., Kravtsov A. V., Lau E. T., Nagai D., 2019, MNRAS, 490, 2380
Cialone G., De Petris M., Sembolini F., Yepes G., Baldi A. S., Rasia E., 2018, MNRAS, 477, 139
Cole S., Lacey C., 1996, MNRAS, 281, 716
Corasaniti P. S., Rasera Y., 2019, MNRAS, 487, 4382
Corasaniti P. S., Ettori S., Rasera Y., Sereno M., Amodeo S., Breton M. A., Ghirardini V., Eckert D., 2018, ApJ, 862, 40
Crone M. M., Evrard A. E., Richstone D. O., 1996, ApJ, 467, 489
Cui W. et al., 2016, MNRAS, 456, 2566
Cui W., Power C., Borgani S., Knebe A., Lewis G. F., Murante G., Poole G. B., 2017, MNRAS, 464, 2502
D'Onghia E., Navarro J. F., 2007, MNRAS, 380, L58
Davis A. J., D'Aloisio A., Natarajan P., 2011, MNRAS, 416, 242
De Luca F., De Petris M., Yepes G., Cui W., Knebe A., Rasia E., 2021, MNRAS, 504, 5383
Diemer B., 2018, ApJS, 239, 35
Diemer B., Kravtsov A. V., 2014, ApJ, 789, 1
Dutton A. A., Macciò A. V., 2014, MNRAS, 441, 3359
Efron B., 1979, Ann. Stat., 7, 1
Eisenstein D. J., Hu W., 1998, ApJ, 496, 605
Eke V. R., Cole S., Frenk C. S., 1996, MNRAS, 282, 263
Euclid Collaboration, 2019, A&A, 627, A23
Fawcett T., 2006, Pattern Recognit. Lett., 27, 861
Franx M., Illingworth G., de Zeeuw T., 1991, ApJ, 383, 112
Frenk C. S., White S. D. M., Davis M., Efstathiou G., 1988, ApJ, 327, 507
Ghirardini V. et al., 2022, A&A, 661, A12
Gianfagna G., Rasia E., Cui W., De Petris M., Yepes G., 2022, in mm Universe @ NIKA2 - Observing the mm Universe with the NIKA2 Camera. p. 00020
Gott J. R. I., Rees M. J., 1975, A&A, 45, 365
Gouin C., Bonnaire T., Aghanim N., 2021, A&A, 651, A56
Gouin C., Gallo S., Aghanim N., 2022, A&A, 664, A198
Haggar R., Gray M. E., Pearce F. R., Knebe A., Cui W., Mostoghiu R., Yepes G., 2020, MNRAS, 492, 6074
Harris C. R. et al., 2020, Nature, 585, 357
Hockney R. W., Eastwood J. W., 1988, Computer Simulation Using Particles. Hilger, Bristol
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Jiang F., van den Bosch F. C., 2016, MNRAS, 458, 2848
Kimmig L. C., Remus R.-S., Dolag K., Biffi V., 2022, preprint (arXiv:2209.09916)
Knebe A., Power C., 2008, ApJ, 678, 621
Knebe A., Wießner V., 2006, PASA, 23, 125
Knebe A. et al., 2011, MNRAS, 415, 2293
Kravtsov A. V., Borgani S., 2012, ARA&A, 50, 353
Kuchner U. et al., 2020, MNRAS, 494, 5473
Kuchner U. et al., 2022, MNRAS, 510, 581
Lau E. T., Hearin A. P., Nagai D., Cappelluti N., 2021, MNRAS, 500, 1029
Lovisari L., Ettori S., Sereno M., Schellenberger G., Forman W. R., Andrade-Santos F., Jones C., 2020, A&A, 644, A78
Macciò A. V., Dutton A. A., van den Bosch F. C., Moore B., Potter D., Stadel J., 2007, MNRAS, 378, 55
Mohr J. J., Fabricant D. G., Geller M. J., 1993, ApJ, 413, 492
Nandra K. et al., 2013, preprint (arXiv:1306.2307)
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Navarro J. F. et al., 2010, MNRAS, 402, 21
Nelson K., Lau E. T., Nagai D., Rudd D. H., Yu L., 2014, ApJ, 782, 107
Neto A. F. et al., 2007, MNRAS, 381, 1450
Pedregosa F. et al., 2011, J. Mach. Learn. Res., 12, 2825
Planck Collaboration VI, 2020, A&A, 641, A6
Planelles S., Quilis V., 2009, MNRAS, 399, 410
Planelles S., Quilis V., 2010, A&A, 519, A94
Planelles S., Schleicher D. R. G., Bykov A. M., 2015, Space Sci. Rev., 188, 93
Planelles S. et al., 2017, MNRAS, 467, 3827
Planelles S., Mimica P., Quilis V., Cuesta-Martínez C., 2018, MNRAS, 476, 4629
Poole G. B., Fardal M. A., Babul A., McCarthy I. G., Quinn T., Wadsley J., 2006, MNRAS, 373, 881
Power C., Knebe A., Knollmann S. R., 2012, MNRAS, 419, 1576
Press W. H., Schechter P., 1974, ApJ, 187, 425
Quilis V., 2004, MNRAS, 352, 1426
Quilis V., Martí J.-M., Planelles S., 2020, MNRAS, 494, 2706
Rasia E., Meneghetti M., Ettori S., 2013, Astron. Rev., 8, 40
Richardson T. R. G., Corasaniti P. S., 2022, MNRAS, 513, 4951
Rossetti M. et al., 2016, MNRAS, 457, 4515
Sartoris B. et al., 2016, MNRAS, 459, 1764
Seabold S., Perktold J., 2010, in 9th Python in Science Conference
Seppi R. et al., 2021, A&A, 652, A155
Shaw L. D., Weller J., Ostriker J. P., Bode P., 2006, ApJ, 646, 815
Simonte M., Vazza F., Brighenti F., Brüggen M., Jones T. W., Angelinelli M., 2022, A&A, 658, A149
Springel V., 2010, MNRAS, 401, 791
Valdarnini R., 2019, ApJ, 874, 42
Vallés-Pérez D., Planelles S., Quilis V., 2020, MNRAS, 499, 2303
Vallés-Pérez D., Planelles S., Quilis V., 2021a, Comput. Phys. Commun., 263, 107892
Vallés-Pérez D., Planelles S., Quilis V., 2021b, MNRAS, 504, 510
Vallés-Pérez D., Planelles S., Quilis V., 2022, A&A, 664, A42
Vazza F., Jones T. W., Brüggen M., Brunetti G., Gheller C., Porter D., Ryu D., 2017, MNRAS, 464, 210
Villaescusa-Navarro F. et al., 2021, ApJ, 915, 71
Villaescusa-Navarro F. et al., 2022, preprint (arXiv:2201.01300)
Virtanen P. et al., 2020, Nat. Methods, 17, 261
Voit G. M., 2005, Rev. Mod. Phys., 77, 207
Wang K., Mao Y.-Y., Zentner A. R., Lange J. U., van den Bosch F. C., Wechsler R. H., 2020, MNRAS, 498, 4450
Weinberger R., Springel V., Pakmor R., 2020, ApJS, 248, 32
Wetzel A. R., Cohn J. D., White M., 2009, MNRAS, 395, 1376
White S. D. M., Rees M. J., 1978, MNRAS, 183, 341
Yu L., Nelson K. L., Nagai D., Lau E., 2015, ApJ, 807, 12
Yuan Z. S., Han J. L., Wen Z. L., 2022, MNRAS, 513, 3013
Zel'dovich Y. B., 1970, A&A, 5, 84
Zemp M., Gnedin O. Y., Gnedin N. Y., Kravtsov A. V., 2011, ApJS, 197, 30
Zhang B., Cui W., Dave R., De Petris M., 2021a, MNRAS, 516, 26
Zhang C., Zhuravleva I., Kravtsov A., Churazov E., 2021b, MNRAS, 506, 839

APPENDIX A: EVOLUTION OF THE THRESHOLDS AND WEIGHTS ON THE DYNAMICAL STATE INDICATORS FOR THE HIGH-MASS SAMPLE

The high-mass subsample, defined in Section 3.1 and shown in Fig. 6, corresponds to a redshift-dependent mass limit, which can be parametrized by

$$\log_{10}\frac{M_{\rm lim}(z)}{M_\odot} = 13.49 - 0.21z \tag{A1}$$

within $\lesssim 0.05$ dex. Thus, the sample corresponds to massive groups and clusters at $z \sim 0$; to objects above $10^{13}\,{\rm M}_\odot$ at $z \sim 2$; and to a mass limit of $\sim 3 \times 10^{12}\,{\rm M}_\odot$ at $z \sim 5$, which might most often be the progenitors of the massive haloes we find at $z \sim 0$.

Within this sample, the evolution with redshifts of the thresholds on the dynamical state indicators, shown in Fig. 7, can be given by the following polynomial fits:

$$\Delta_r^{\rm thr}(z)\big|_{\rm massive} = 0.0863(39) + 0.0066(23)z \tag{A2}$$

$$\eta^{\rm thr}(z)\big|_{\rm massive} = 1.3371(88) + 0.151(14)z - 0.0139(37)z^2 \tag{A3}$$

$$\langle\widetilde{v_r}\rangle_{\rm DM}^{\rm thr}(z)\big|_{\rm massive} = 0.0842(32) \tag{A4}$$

$$s_{200c,500c}^{\rm thr}(z)\big|_{\rm massive} = 1.495(10) \tag{A5}$$

$$\varepsilon^{\rm thr}(z)\big|_{\rm massive} = 0.2710(33) \tag{A6}$$

The weights on these indicators, as they appear on the relaxedness parameter (equation 14), are fitted by

$$w[\Delta_r](z)\big|_{\rm massive} \propto 0.218(16) - 0.134(26)z + 0.0356(69)z^2 \tag{A7}$$

$$w[\eta](z)\big|_{\rm massive} \propto 0.250(11) - 0.0603(66)z \tag{A8}$$

$$w[\langle\widetilde{v_r}\rangle_{\rm DM}](z)\big|_{\rm massive} \propto 0.092(17) + 0.109(27)z - 0.0141(71)z^2 \tag{A9}$$

$$w[s_{200c,500c}](z)\big|_{\rm massive} \propto 0.2251(87) \tag{A10}$$

$$w[\varepsilon](z)\big|_{\rm massive} \propto 0.2537(86) \tag{A11}$$

This paper has been typeset from a TEX/LATEX file prepared by the author.

APPENDIX B

Summary of co-authored contributions

This Appendix contains a brief summary of the contributions done to works led by other authors during the course of this PhD, where I have contributed and I am a co-author. These topics are not part of the Publications included in Appendix A, which constitute the compendium of publications of this Thesis. The list is presented in chronological order of the submission date of the corresponding publication.

The imprints of galaxy cluster internal dynamics on the Sunyaev-Zeldovich effect

This section contains a brief summary of Monllor-Berbegal, Vallés-Pérez, Planelles & Quilis (2024)[1].

Overview

Context. The morphological features of galaxy groups and clusters, in several bands, could be connected with the dynamical and thermodynamic state of the intracluster medium. In the near future, measurements of the Sunyaev-Zeldovich

[1] Ó. Monllor-Berbegal et al. "The imprints of galaxy cluster internal dynamics on the Sunyaev-Zeldovich effect." In: *Astron. Astrophys.* 686, A243 (June 2024), A243. DOI: 10.1051/0004-6361/202348967

(SZ) effect in groups and clusters will become a powerful tool to study these properties.

Aims. We aim to study the impact of the dynamics of galaxy clusters (characterised through the presence of merger events and dynamical state indicators) on the thermal and kinetic SZ signals (tSZ and kSZ, respectively).

Methods. With a sample of simulated galaxy clusters, we produce synthetic maps of the SZ effect (both thermal and kinetic) in several projections. We use the correlations between properties extracted from these SZ maps, and the dynamical history extracted from the full three-dimensional description of the simulation, to look for the imprints of the evolutionary events, mainly mergers, on the SZ signals.

Results. The integrated tSZ signal does not show any noticeable dependence with dynamical state, although its radial distribution does. The kinetic effect, however, shows a remarkable segregation amongst dynamical state classes. Unrelaxed clusters present a higher signal at all masses and radii. Studying the individual evolution of a few systems, we find that galaxy cluster mergers are the main responsible agents for this enhancement on the kSZ signal. Additionally, performing a harmonic expansion of the kSZ signal, we find that the dipolar component of the signal is correlated with the projected angular momentum of the cluster, but this correlation is only significant in relaxed systems.

Conclusions. Even though the scatter in the different relations is large, the results presented in this work suggest that the SZ signal can yield valuable information about the dynamical history of galaxy clusters.

My contribution

I have participated in this work within several aspects. First, I have contributed to the simulation (in the form of the different improvements and optimisations to MASCLET), generated the halo catalogues and merger trees, and the corresponding dynamical state classification. Subsequently, I have participated in the analysis of the results: in particular, in several aspects regarding the study of the kSZ signal and its harmonic decomposition, as well as the statistical analysis. Finally, I have contributed to the discussion of the results and the revision of the manuscript.

Unravelling the relativistic magnetised jet dynamics of the microquasar SS 433

This section contains a brief summary of the work by López-Miralles, Perucho, Vallés-Pérez, et al., which is currently in the last stages of preparation for submission to the journal Nature Astronomy.

Overview

Context. The high-mass X-ray binary SS 433 is one of the most powerful and exotic star systems in our Galaxy, which has exerted a huge influence in the field of high-energy Astrophysics. Discovered more than four decades ago, numerous multi-wavelength observations of this source have unveiled unconventional physical processes, seemingly unique among other galactic binaries. In particular, the precessing jet synchrotron emission, along with its polarization signature at the sub-parsec scale, suggests a complex outflow morphology and a highly specific magnetic field configuration, parallel to the motion of the SS 433 jet components.

Aims. Using large-scale numerical simulations, in this work we explore the way a precessing jet based on the inferred physical parameters of SS 433, together with an initial toroidal magnetic field configuration, can develop a component aligned with the local velocity direction after the first precession period, and specifically, the role of discrete blob collisions into this process. We also address the effect of these interactions on the overall jet morphology, deceleration and long-term stability. This information is also essential for improving our understanding of particle acceleration and high-energy emission.

Methods. We developed large-scale three-dimensional relativistic magnetohydrodynamics (RMHD) simulations of the SS 433 precessing jets at the scale of the sub-parsec, including for the first time the dynamical evolution of magnetic fields. These simulations were performed with the code Lóstrego (López-Miralles et al. 2022, López-Miralles, Martí, and Perucho 2023) in the Picasso supercomputer (Red Española de Supercomputación), using 4096 cores and a total of four million hours of HPC computing time.

Results. These numerical simulations show that a discrete flow morphology drives a magnetic field orientation aligned with the local velocity after one precession period, in good agreement with observations. The interaction of the jet discrete components in the sub-parsec scale leads to elongated bullets with

no significant evidence of flow deceleration. A continuous stream is by contrast prone to the development of different types of hydrodynamical instabilities.

Conclusions. The results of the RMHD simulations suggest that the observed characteristics of the source are compatible with the existence of discrete structures, rather than with jets based on a continuous flow. This was previously suggested by radio and X-ray polarization studies. Further work is also required to probe the existence of ultra-relativistic outflows (as previously inferred from X-ray data).

My contribution

My contribution to the project focused on the post-processing stage of the simulations, developing and implementing algorithms to extract the information that can be compared with the polarization data. For instance, these algorithms allowed us to integrate the jet velocity and magnetic field vectors along the line of sight, extract the kinematic jet trace, and derive the projected angles between all these variables, what ultimately enabled us to show how the discrete flow generates an aligned magnetic field configuration after one precession period. In the first stages of the project, I also contributed to the simulation setup, by deriving the necessary transformations for the injection of the precessing jet. In a more secondary level, I have also participated in the discussion of the results, as well as the revision of the manuscript.

Bibliography

The references within the texts of the articles in Appendix A are not reproduced here for the sake of conciseness, and can be found at the end of each paper.

S. J. Aarseth. "Dynamical evolution of clusters of galaxies, I." In: *Mon. Not. R. Astron. Soc.* 126 (Jan. 1963), p. 223. DOI: 10.1093/mnras/126.3.223.

G. O. Abell. "The Distribution of Rich Clusters of Galaxies." In: *Astrophys. J. Suppl. Ser.* 3 (May 1958), p. 211. DOI: 10.1086/190036.

G. O. Abell, J. Corwin Harold G., and R. P. Olowin. "A Catalog of Rich Clusters of Galaxies." In: *Astrophys. J. Suppl. Ser.* 70 (May 1989), p. 1. DOI: 10.1086/191333.

O. Agertz, R. Teyssier, and B. Moore. "The formation of disc galaxies in a ΛCDM universe." In: *Mon. Not. R. Astron. Soc.* 410.2 (Jan. 2011), pp. 1391–1408. DOI: 10.1111/j.1365-2966.2010.17530.x.

O. Agertz et al. "Fundamental differences between SPH and grid methods." In: *Mon. Not. R. Astron. Soc.* 380.3 (Sept. 2007), pp. 963–978. DOI: 10.1111/j.1365-2966.2007.12183.x.

R. A. Alpher, H. Bethe, and G. Gamow. "The Origin of Chemical Elements." In: *Physical Review* 73.7 (Apr. 1948), pp. 803–804. DOI: 10.1103/PhysRev.73.803.

M. Angelinelli et al. "Turbulent pressure support and hydrostatic mass bias in the intracluster medium." In: *Mon. Not. R. Astron. Soc.* 495.1 (June 2020), pp. 864–885. DOI: 10.1093/mnras/staa975.

M. A. Aragon-Calvo and A. S. Szalay. "The hierarchical structure and dynamics of voids." In: *Mon. Not. R. Astron. Soc.* 428.4 (Feb. 2013), pp. 3409–3424. DOI: 10.1093/mnras/sts281.

M. Arnaud and A. E. Evrard. "The $L_X - T$ relation and intracluster gas fractions of X-ray clusters." In: *Mon. Not. R. Astron. Soc.* 305.3 (May 1999), pp. 631–640. DOI: 10.1046/j.1365-8711.1999.02442.x.

A. Arth et al. "Anisotropic thermal conduction in galaxy clusters with MHD in Gadget." In: *arXiv e-prints* (Dec. 2014). arXiv: 1412.6533 [astro-ph.CO].

C. Avestruz et al. "Non-equilibrium Electrons in the Outskirts of Galaxy Clusters." In: *Astrophys. J.* 808.2, 176 (Aug. 2015), p. 176. DOI: 10.1088/0004-637X/808/2/176.

J. S. Bagla. "TreePM: A Code for Cosmological N-Body Simulations." In: *Journal of Astrophysics and Astronomy* 23.3-4 (Dec. 2002), pp. 185–196. DOI: 10.1007/BF02702282.

D. S. Balsara. "von Neumann stability analysis of smooth particle hydrodynamics–suggestions for optimal algorithms." In: *Journal of Computational Physics* 121.2 (Jan. 1995), pp. 357–372. DOI: 10.1016/S0021-9991(95)90221-X.

D. S. Balsara and J. Kim. "A Comparison between Divergence-Cleaning and Staggered-Mesh Formulations for Numerical Magnetohydrodynamics." In: *Astrophys. J.* 602.2 (Feb. 2004), pp. 1079–1090. DOI: 10.1086/381051.

S. Banfi, F. Vazza, and C. Gheller. "On the alignment of haloes, filaments and magnetic fields in the simulated cosmic web." In: *Mon. Not. R. Astron. Soc.* 503.3 (May 2021), pp. 4016–4031. DOI: 10.1093/mnras/stab655.

S. Banfi, F. Vazza, and D. Wittor. "Shock waves in the magnetized cosmic web: the role of obliquity and cosmic ray acceleration." In: *Mon. Not. R. Astron. Soc.* 496.3 (Aug. 2020), pp. 3648–3667. DOI: 10.1093/mnras/staa1810.

J. M. Bardeen et al. "The Statistics of Peaks of Gaussian Random Fields." In: *Astrophys. J.* 304 (May 1986), p. 15. DOI: 10.1086/164143.

J. Barnes and P. Hut. "A hierarchical O(N log N) force-calculation algorithm." In: *Nature* 324.6096 (Dec. 1986), pp. 446–449. DOI: 10.1038/324446a0.

M. Bartelmann. "The dark Universe." In: *Reviews of Modern Physics* 82.1 (Jan. 2010), pp. 331–382. DOI: 10.1103/RevModPhys.82.331.

C. M. Baugh, E. Gaztanaga, and G. Efstathiou. "A comparison of the evolution of density fields in perturbation theory and numerical simulations - II. Counts-in-cells analysis." In: *Mon. Not. R. Astron. Soc.* 274.4 (June 1995), pp. 1049–1070. DOI: 10.1093/mnras/274.4.1049.

A. N. Baushev. "The central region of a void: an analytical solution." In: *Mon. Not. R. Astron. Soc.* 504.1 (June 2021), pp. L56–L60. DOI: 10.1093/mnrasl/slab036.

A. M. Beck, K. Dolag, and J. M. F. Donnert. "Geometrical on-the-fly shock detection in smoothed particle hydrodynamics." In: *Mon. Not. R. Astron. Soc.* 458.2 (May 2016), pp. 2080–2087. DOI: 10.1093/mnras/stw487.

A. M. Beck et al. "An improved SPH scheme for cosmological simulations." In: *Mon. Not. R. Astron. Soc.* 455.2 (Jan. 2016), pp. 2110–2130. DOI: 10.1093/mnras/stv2443.

J. S. Bennett and D. Sijacki. "A disturbing FABLE of mergers, feedback, turbulence, and mass biases in simulated galaxy clusters." In: *Mon. Not. R. Astron. Soc.* 514.1 (July 2022), pp. 313–328. DOI: 10.1093/mnras/stac1216.

A. Beresnyak and F. Miniati. "Turbulent Amplification and Structure of the Intracluster Magnetic Field." In: *Astrophys. J.* 817.2, 127 (Feb. 2016), p. 127. DOI: 10.3847/0004-637X/817/2/127.

M. J. Berger and P. Colella. "Local Adaptive Mesh Refinement for Shock Hydrodynamics." In: *Journal of Computational Physics* 82.1 (May 1989), pp. 64–84. DOI: 10.1016/0021-9991(89)90035-1.

M. J. Berger and J. Oliger. "Adaptive Mesh Refinement for Hyperbolic Partial Differential Equations." In: *Journal of Computational Physics* 53.3 (Mar. 1984), pp. 484–512. DOI: 10.1016/0021-9991(84)90073-1.

E. Bertschinger. "Self-similar secondary infall and accretion in an Einstein-de Sitter universe." In: *Astrophys. J. Suppl. Ser.* 58 (May 1985), pp. 39–65. DOI: 10.1086/191028.

E. Bertschinger. "Multiscale Gaussian Random Fields and Their Application to Cosmological Simulations." In: *Astrophys. J. Suppl. Ser.* 137.1 (Nov. 2001), pp. 1–20. DOI: 10.1086/322526.

S. P. Bhavsar and E. N. Ling. "II. Large-Scale Distribution of Galaxies: Filamentary Structure and Visual Bias." In: *Publ. Astron. Soc. Pac.* 100 (Nov. 1988), p. 1314. DOI: 10.1086/132325.

V. Biffi et al. "On the Nature of Hydrostatic Equilibrium in Galaxy Clusters." In: *Astrophys. J.* 827.2, 112 (Aug. 2016), p. 112. DOI: 10.3847/0004-637X/827/2/112.

M. Birkinshaw. "Observations of the Sunyaev-Zeldovich Effect." In: *The Cosmic Microwave Backround: 25 Years Later*. Ed. by N. Mandolesi and N. Vittorio. Vol. 164. Astrophysics and Space Science Library. Jan. 1990, p. 77. DOI: 10.1007/978-94-009-0655-6_6.

A. Biviano. "From Messier to Abell: 200 Years of Science with Galaxy Clusters." In: *Constructing the Universe with Clusters of Galaxies*. Ed. by F. Durret and D. Gerbal. Jan. 2000, 1, p. 1. arXiv: astro-ph/0010409 [astro-ph].

R. D. Blandford and J. P. Ostriker. "Particle acceleration by astrophysical shocks." In: *Astrophys. J. Lett.* 221 (Apr. 1978), pp. L29–L32. DOI: 10.1086/182658.

L. E. Bleem et al. "Galaxy Clusters Discovered via the Sunyaev-Zel'dovich Effect in the 2500-Square-Degree SPT-SZ Survey." In: *Astrophys. J. Suppl. Ser.* 216.2, 27 (Feb. 2015), p. 27. DOI: 10.1088/0067-0049/216/2/27.

G. R. Blumenthal et al. "Formation of galaxies and large-scale structure with cold dark matter." In: *Nature* 311 (Oct. 1984), pp. 517–525. DOI: 10.1038/311517a0.

P. Bode, J. P. Ostriker, and G. Xu. "The Tree Particle-Mesh N-Body Gravity Solver." In: *Astrophys. J. Suppl. Ser.* 128.2 (June 2000), pp. 561–569. DOI: 10.1086/313398.

H. Böhringer and N. Werner. "X-ray spectroscopy of galaxy clusters: studying astrophysical processes in the largest celestial laboratories." In: *Astronomy and Astrophysics Reviews* 18.1-2 (Feb. 2010), pp. 127–196. DOI: 10.1007/s00159-009-0023-3.

A. Bonafede et al. "LOFAR discovery of radio emission in MACS J0717.5+3745." In: *Mon. Not. R. Astron. Soc.* 478.3 (Aug. 2018), pp. 2927–2938. DOI: 10.1093/mnras/sty1121.

A. Bonafede et al. "The Coma cluster magnetic field from Faraday rotation measures." In: *Astron. Astrophys.* 513, A30 (Apr. 2010), A30. DOI: 10.1051/0004-6361/200913696.

J. R. Bond et al. "Excursion Set Mass Functions for Hierarchical Gaussian Fluctuations." In: *Astrophys. J.* 379 (Oct. 1991), p. 440. DOI: 10.1086/170520.

J. R. Bond, L. Kofman, and D. Pogosyan. "How filaments of galaxies are woven into the cosmic web." In: *Nature* 380.6575 (Apr. 1996), pp. 603–606. DOI: 10.1038/380603a0.

H. Bondi. "On spherically symmetrical accretion." In: *Mon. Not. R. Astron. Soc.* 112 (Jan. 1952), p. 195. DOI: 10.1093/mnras/112.2.195.

H. Bondi. "Spherically symmetrical models in general relativity." In: *Mon. Not. R. Astron. Soc.* 107 (Jan. 1947), p. 410. DOI: 10.1093/mnras/107.5-6.410.

V. Bonjean et al. "Gas and galaxies in filaments between clusters of galaxies. The study of A399-A401." In: *Astron. Astrophys.* 609, A49 (Jan. 2018), A49. DOI: 10.1051/0004-6361/201731699.

C. M. Booth and J. Schaye. "Cosmological simulations of the growth of supermassive black holes and feedback from active galactic nuclei: method and tests." In: *Mon. Not. R. Astron. Soc.* 398.1 (Sept. 2009), pp. 53–74. DOI: 10.1111/j.1365-2966.2009.15043.x.

S. Borgani et al. "X-ray properties of galaxy clusters and groups from a cosmological hydrodynamical simulation." In: *Mon. Not. R. Astron. Soc.* 348.3 (Mar. 2004), pp. 1078–1096. DOI: 10.1111/j.1365-2966.2004.07431.x.

S. Borgani and A. Kravtsov. "Cosmological Simulations of Galaxy Clusters." In: *Advanced Science Letters* 4.2 (Feb. 2011), pp. 204–227. DOI: 10.1166/asl.2011.1209.

L. M. Böss et al. "CRESCENDO: an on-the-fly Fokker-Planck solver for spectral cosmic rays in cosmological simulations." In: *Mon. Not. R. Astron. Soc.* 519.1 (Feb. 2023), pp. 548–572. DOI: 10.1093/mnras/stac3584.

J. Bovy. *Dynamics and Astrophysics of Galaxies*. Ed. by Princeton University Press (in preparation). 2023.

G. E. P. Box and M. E. Muller. "A Note on the Generation of Random Normal Deviates." In: *The Annals of Mathematical Statistics* 29.2 (1958), pp. 610 –611. DOI: 10.1214/aoms/1177706645. URL: https://doi.org/10.1214/aoms/1177706645.

M. Boylan-Kolchin, J. S. Bullock, and M. Kaplinghat. "Too big to fail? The puzzling darkness of massive Milky Way subhaloes." In: *Mon. Not. R. Astron. Soc.* 415.1 (July 2011), pp. L40–L44. DOI: 10.1111/j.1745-3933.2011.01074.x.

H. Bradt et al. "Evidence for X-Radiation from the Radio Galaxy M87." In: *Astrophys. J. Lett.* 150 (Dec. 1967), p. L199. DOI: 10.1086/180125.

A. Brandenburg and Å. Nordlund. "Astrophysical turbulence modeling." In: *Reports on Progress in Physics* 74.4, 046901 (Apr. 2011), p. 046901. DOI: 10.1088/0034-4885/74/4/046901.

M. Brüggen and F. Vazza. "Turbulence in the Intracluster Medium." In: *Magnetic Fields in Diffuse Media*. Ed. by A. Lazarian, E. M. de Gouveia Dal Pino, and C. Melioli. Vol. 407. Astrophysics and Space Science Library. Jan. 2015, p. 599. DOI: 10.1007/978-3-662-44625-6_21.

M. Brüggen et al. "Magnetic Fields, Relativistic Particles, and Shock Waves in Cluster Outskirts." In: *Space Sci. Rev.* 166.1-4 (May 2012), pp. 187–213. DOI: 10.1007/s11214-011-9785-9.

G. L. Bryan and M. L. Norman. "Statistical Properties of X-Ray Clusters: Analytic and Numerical Comparisons." In: *Astrophys. J.* 495.1 (Mar. 1998), pp. 80–99. DOI: 10.1086/305262.

G. L. Bryan et al. "ENZO: An Adaptive Mesh Refinement Code for Astrophysics." In: *Astrophys. J. Suppl. Ser.* 211.2, 19 (Apr. 2014), p. 19. DOI: 10.1088/0067-0049/211/2/19.

T. Buchert. "Lagrangian theory of gravitational instability of Friedman-Lemaitre cosmologies and the 'Zel'dovich approximation'." In: *Mon. Not. R. Astron. Soc.* 254 (Feb. 1992), pp. 729–737. DOI: 10.1093/mnras/254.4.729.

J. S. Bullock and M. Boylan-Kolchin. "Small-Scale Challenges to the ΛCDM Paradigm." In: *Annu. Rev. Astron. Astrophys.* 55.1 (Aug. 2017), pp. 343–387. DOI: 10.1146/annurev-astro-091916-055313.

D. A. Buote and J. C. Tsai. "Quantifying the Morphologies and Dynamical Evolution of Galaxy Clusters. I. The Method." In: *Astrophys. J.* 452 (Oct. 1995), p. 522. DOI: 10.1086/176326.

A. M. Bykov, K. Dolag, and F. Durret. "Cosmological Shock Waves." In: *Space Sci. Rev.* 134.1-4 (Feb. 2008), pp. 119–140. DOI: 10.1007/s11214-008-9312-9.

E. T. Byram, T. A. Chubb, and H. Friedman. "Cosmic X-ray Sources, Galactic and Extragalactic." In: *Science* 152.3718 (Apr. 1966), pp. 66–71. DOI: 10.1126/science.152.3718.66.

R. Cassano and G. Brunetti. "Cluster mergers and non-thermal phenomena: a statistical magneto-turbulent model." In: *Mon. Not. R. Astron. Soc.* 357.4 (Mar. 2005), pp. 1313–1329. DOI: 10.1111/j.1365-2966.2005.08747.x.

M. Cautun et al. "Evolution of the cosmic web." In: *Mon. Not. R. Astron. Soc.* 441.4 (July 2014), pp. 2923–2973. DOI: 10.1093/mnras/stu768.

A. G. Cavaliere, H. Gursky, and W. H. Tucker. "Extragalactic X-ray Sources and Associations of Galaxies." In: *Nature* 231.5303 (June 1971), pp. 437–438. DOI: 10.1038/231437a0.

R. Cen. "A Hydrodynamic Approach to Cosmology: Methodology." In: *Astrophys. J. Suppl. Ser.* 78 (Feb. 1992), p. 341. DOI: 10.1086/191630.

D. Chapon, L. Mayer, and R. Teyssier. "Hydrodynamics of galaxy mergers with supermassive black holes: is there a last parsec problem?" In: *Mon. Not. R. Astron. Soc.* 429.4 (Mar. 2013), pp. 3114–3122. DOI: 10.1093/mnras/sts568.

H. Chen et al. "Imprints of mass accretion history on the shape of the intracluster medium and the T_X-M relation." In: *Mon. Not. R. Astron. Soc.* 490.2 (Dec. 2019), pp. 2380–2389. DOI: 10.1093/mnras/stz2776.

E. Churazov et al. "Measuring the non-thermal pressure in early-type galaxy atmospheres: a comparison of X-ray and optical potential profiles in M87 and NGC 1399." In: *Mon. Not. R. Astron. Soc.* 388.3 (Aug. 2008), pp. 1062–1078. DOI: 10.1111/j.1365-2966.2008.13507.x.

E. Churazov et al. "Tempestuous life beyond R_{500}: X-ray view on the Coma cluster with SRG/eROSITA. I. X-ray morphology, recent merger, and radio halo connection." In: *Astron. Astrophys.* 651, A41 (July 2021), A41. DOI: 10.1051/0004-6361/202040197.

E. Churazov et al. "X-ray surface brightness and gas density fluctuations in the Coma cluster." In: *Mon. Not. R. Astron. Soc.* 421.2 (Apr. 2012), pp. 1123–1135. DOI: 10.1111/j.1365-2966.2011.20372.x.

J. M. Colberg et al. "The Aspen-Amsterdam void finder comparison project." In: *Mon. Not. R. Astron. Soc.* 387.2 (June 2008), pp. 933–944. DOI: 10.1111/j.1365-2966.2008.13307.x.

P. Colella and H. M. Glaz. "Efficient Solution Algorithms for the Riemann Problem for Real Gases." In: *Journal of Computational Physics* 59.2 (June 1985), pp. 264–289. DOI: 10.1016/0021-9991(85)90146-9.

P. Colella and P. R. Woodward. "The Piecewise Parabolic Method (PPM) for Gas-Dynamical Simulations." In: *Journal of Computational Physics* 54 (Sept. 1984), pp. 174–201. DOI: 10.1016/0021-9991(84)90143-8.

M. Colless et al. "The 2dF Galaxy Redshift Survey: spectra and redshifts." In: *Mon. Not. R. Astron. Soc.* 328.4 (Dec. 2001), pp. 1039–1063. DOI: 10.1046/j.1365-8711.2001.04902.x.

S. Colombi et al. "Vlasov versus N-body: the Hénon sphere." In: *Mon. Not. R. Astron. Soc.* 450.4 (July 2015), pp. 3724–3741. DOI: 10.1093/mnras/stv819.

E. Contini et al. "On the formation and physical properties of the intracluster light in hierarchical galaxy formation models." In: *Mon. Not. R. Astron. Soc.* 437.4 (Feb. 2014), pp. 3787–3802. DOI: 10.1093/mnras/stt2174.

J. W. Cooley and J. W. Tukey. "An algorithm for the machine calculation of complex Fourier series." In: *Mathematics of Computation* 19 (1965). URL: http://cr.yp.to/bib/entries.html#1965/cooley, pp. 297–301. ISSN: 0025–5718.

R. Courant, K. Friedrichs, and H. Lewy. "Über die partiellen Differenzengleichungen der mathematischen Physik." In: *Mathematische Annalen* 100 (Jan. 1928), pp. 32–74. DOI: 10.1007/BF01448839.

R. Courant and K. O. Friedrichs. *Supersonic flow and shock waves*. 1948.

R. A. Crain et al. "Galaxies-intergalactic medium interaction calculation - I. Galaxy formation as a function of large-scale environment." In: *Mon. Not. R. Astron. Soc.* 399.4 (Nov. 2009), pp. 1773–1794. DOI: 10.1111/j.1365-2966.2009.15402.x.

C. Cuesta-Martínez, M. A. Aloy, and P. Mimica. "Numerical models of blackbody-dominated gamma-ray bursts - I. Hydrodynamics and the origin of the thermal emission." In: *Mon. Not. R. Astron. Soc.* 446.2 (Jan. 2015), pp. 1716–1736. DOI: 10.1093/mnras/stu2186.

W. Cui et al. "On the dynamical state of galaxy clusters: insights from cosmological simulations - II." In: *Mon. Not. R. Astron. Soc.* 464.2 (Jan. 2017), pp. 2502–2510. DOI: 10.1093/mnras/stw2567.

C. Dalla Vecchia and J. Schaye. "Simulating galactic outflows with thermal supernova feedback." In: *Mon. Not. R. Astron. Soc.* 426.1 (Oct. 2012), pp. 140–158. DOI: 10.1111/j.1365-2966.2012.21704.x.

G. B. Dalton et al. "The APM Galaxy Survey - V. Catalogues of galaxy clusters." In: *Mon. Not. R. Astron. Soc.* 289.2 (Aug. 1997), pp. 263–284. DOI: 10.1093/mnras/289.2.263.

S. De Grandi and S. Molendi. "Metallicity Gradients in X-Ray Clusters of Galaxies." In: *Astrophys. J.* 551.1 (Apr. 2001), pp. 153–159. DOI: 10.1086/320098.

V. de Lapparent, M. J. Geller, and J. P. Huchra. "A Slice of the Universe." In: *Astrophys. J. Lett.* 302 (Mar. 1986), p. L1. DOI: 10.1086/184625.

A. Dedner et al. "Hyperbolic Divergence Cleaning for the MHD Equations." In: *Journal of Computational Physics* 175.2 (Jan. 2002), pp. 645–673. DOI: 10.1006/jcph.2001.6961.

W. Dehnen and J. I. Read. "N-body simulations of gravitational dynamics." In: *European Physical Journal Plus* 126, 55 (May 2011), p. 55. DOI: 10.1140/epjp/i2011-11055-3.

W. Dehnen. "A Very Fast and Momentum-conserving Tree Code." In: *Astrophys. J. Lett.* 536.1 (June 2000), pp. L39–L42. DOI: 10.1086/312724.

N. Deruelle and J.-P. Uzan. "The Lambda-CDM model of the hot Big Bang." In: *Relativity in Modern Physics*. Oxford University Press, Aug. 2018. ISBN: 9780198786399. DOI: 10.1093/oso/9780198786399.003.0059. URL: https://doi.org/10.1093/oso/9780198786399.003.0059.

T. Di Matteo, V. Springel, and L. Hernquist. "Energy input from quasars regulates the growth and activity of black holes and their host galaxies." In: *Nature* 433.7026 (Feb. 2005), pp. 604–607. DOI: 10.1038/nature03335.

A. Diaferio and M. J. Geller. "Infall Regions of Galaxy Clusters." In: *Astrophys. J.* 481.2 (May 1997), pp. 633–643. DOI: 10.1086/304075.

R. H. Dicke et al. "Cosmic Black-Body Radiation." In: *Astrophys. J.* 142 (July 1965), pp. 414–419. DOI: 10.1086/148306.

B. Diemer. "COLOSSUS: A Python Toolkit for Cosmology, Large-scale Structure, and Dark Matter Halos." In: *Astrophys. J. Suppl. Ser.* 239.2, 35 (Dec. 2018), p. 35. DOI: 10.3847/1538-4365/aaee8c.

B. Diemer and A. V. Kravtsov. "Dependence of the Outer Density Profiles of Halos on Their Mass Accretion Rate." In: *Astrophys. J.* 789.1, 1 (July 2014), p. 1. DOI: 10.1088/0004-637X/789/1/1.

K. Dolag et al. "Simulation Techniques for Cosmological Simulations." In: *Space Sci. Rev.* 134.1-4 (Feb. 2008), pp. 229–268. DOI: 10.1007/s11214-008-9316-5.

K. Dolag et al. "Substructures in hydrodynamical cluster simulations." In: *Mon. Not. R. Astron. Soc.* 399.2 (Oct. 2009), pp. 497–514. DOI: 10.1111/j.1365-2966.2009.15034.x.

K. Dolag et al. "Thermal Conduction in Simulated Galaxy Clusters." In: *Astrophys. J. Lett.* 606.2 (May 2004), pp. L97–L100. DOI: 10.1086/420966.

K. Dolag et al. "Turbulent gas motions in galaxy cluster simulations: the role of smoothed particle hydrodynamics viscosity." In: *Mon. Not. R. Astron. Soc.* 364.3 (Dec. 2005), pp. 753–772. DOI: 10.1111/j.1365-2966.2005.09630.x.

J. Domínguez-Gómez et al. "Galaxies in voids assemble their stars slowly." In: *Nature* 619.7969 (July 2023), pp. 269–271. DOI: 10.1038/s41586-023-06109-1.

A. Drlica-Wagner et al. "Eight Ultra-faint Galaxy Candidates Discovered in Year Two of the Dark Energy Survey." In: *Astrophys. J.* 813.2, 109 (Nov. 2015), p. 109. DOI: 10.1088/0004-637X/813/2/109.

J. Dubinski and R. G. Carlberg. "The Structure of Cold Dark Matter Halos." In: *Astrophys. J.* 378 (Sept. 1991), p. 496. DOI: 10.1086/170451.

A. A. Dutton and A. V. Macciò. "Cold dark matter haloes in the Planck era: evolution of structural parameters for Einasto and NFW profiles." In: *Mon. Not. R. Astron. Soc.* 441.4 (July 2014), pp. 3359–3374. DOI: 10.1093/mnras/stu742.

D. Eckert et al. "On the Connection between Turbulent Motions and Particle Acceleration in Galaxy Clusters." In: *Astrophys. J. Lett.* 843.2, L29 (July 2017), p. L29. DOI: 10.3847/2041-8213/aa7c1a.

G. Efstathiou et al. "Numerical techniques for large cosmological N-body simulations." In: *Astrophys. J. Suppl. Ser.* 57 (Feb. 1985), pp. 241–260. DOI: 10.1086/191003.

J. Einasto, M. Joeveer, and E. Saar. "Superclusters and galaxy formation." In: *Nature* 283.5742 (Jan. 1980), pp. 47–48. DOI: 10.1038/283047a0.

B. Einfeldt. "On Godunov-Type Methods for Gas Dynamics." In: *SIAM Journal on Numerical Analysis* 25.2 (Apr. 1988), pp. 294–318. DOI: 10.1137/0725021.

D. J. Eisenstein and W. Hu. "Baryonic Features in the Matter Transfer Function." In: *Astrophys. J.* 496.2 (Mar. 1998), pp. 605–614. DOI: 10.1086/305424.

Euclid Collaboration et al. "Euclid preparation. III. Galaxy cluster detection in the wide photometric survey, performance and algorithm selection." In: *Astron. Astrophys.* 627, A23 (July 2019), A23. DOI: 10.1051/0004-6361/201935088.

A. C. Fabian et al. "Cooling flows and the X-ray luminosity-temperature relation for clusters." In: *Mon. Not. R. Astron. Soc.* 267 (Apr. 1994), pp. 779–784. DOI: 10.1093/mnras/267.3.779.

A. Faltenbacher et al. "Supersonic motions of galaxies in clusters." In: *Mon. Not. R. Astron. Soc.* 358.1 (Mar. 2005), pp. 139–148. DOI: 10.1111/j.1365-2966.2005.08769.x.

B. Famaey and S. S. McGaugh. "Modified Newtonian Dynamics (MOND): Observational Phenomenology and Relativistic Extensions." In: *Living Reviews in Relativity* 15.1, 10 (Sept. 2012), p. 10. DOI: 10.12942/lrr-2012-10.

J. E. Felten et al. "X-Rays from the Coma Cluster of Galaxies." In: *Astrophys. J.* 146 (Dec. 1966), pp. 955–958. DOI: 10.1086/148972.

L. Feretti et al. "Clusters of galaxies: observational properties of the diffuse radio emission." In: *Astron. Astrophys. Rev.* 20, 54 (May 2012), p. 54. DOI: 10.1007/s00159-012-0054-z.

C. Ferrari et al. "Observations of Extended Radio Emission in Clusters." In: *Space Sci. Rev.* 134.1-4 (Feb. 2008), pp. 93–118. DOI: 10.1007/s11214-008-9311-x.

C. S. Frenk et al. "The Formation of Dark Halos in a Universe Dominated by Cold Dark Matter." In: *Astrophys. J.* 327 (Apr. 1988), p. 507. DOI: 10.1086/166213.

M. Frigo and S. G. Johnson. "The Design and Implementation of FFTW3." In: *Proceedings of the IEEE* 93.2 (2005). Special issue on "Program Generation, Optimization, and Platform Adaptation", pp. 216–231.

G. Fritz et al. "Discovery of an X-Ray Source in Perseus." In: *Astrophys. J. Lett.* 164 (Mar. 1971), p. L81. DOI: 10.1086/180697.

Y. Fujita, M. Takizawa, and C. L. Sarazin. "Nonthermal Emissions from Particles Accelerated by Turbulence in Clusters of Galaxies." In: *Astrophys. J.* 584.1 (Feb. 2003), pp. 190–202. DOI: 10.1086/345599.

F. Gagliardi et al. "The international race towards Exascale in Europe." In: *CCF Transactions on High Performance Computing* 1.1 (May 2019), pp. 3–13. ISSN: 2524-4930. DOI: 10.1007/s42514-019-00002-y. URL: https://doi.org/10.1007/s42514-019-00002-y.

G. Gamow. "The Evolution of the Universe." In: *Nature* 162.4122 (Oct. 1948), pp. 680–682. DOI: 10.1038/162680a0.

J. García-Bellido. "The origin of matter and structure in the universe." In: *Philosophical Transactions of the Royal Society of London Series A* 357.1763 (Dec. 1999), p. 3237. DOI: 10.1098/rsta.1999.0492.

L. H. Garrison et al. "Improving initial conditions for cosmological N-body simulations." In: *Mon. Not. R. Astron. Soc.* 461.4 (Oct. 2016), pp. 4125–4145. DOI: 10.1093/mnras/stw1594.

M. Gaspari. "Shaping the X-ray spectrum of galaxy clusters with AGN feedback and turbulence." In: *Mon. Not. R. Astron. Soc.* 451 (July 2015), pp. L60–L64. DOI: 10.1093/mnrasl/slv067.

M. Gaspari et al. "The relation between gas density and velocity power spectra in galaxy clusters: High-resolution hydrodynamic simulations and the role of conduction." In: *Astron. Astrophys.* 569, A67 (Sept. 2014), A67. DOI: 10.1051/0004-6361/201424043.

E. Gatuzz et al. "Chemical enrichment of the ICM within the virgo cluster - I. Radial profiles." In: *Mon. Not. R. Astron. Soc.* 520.3 (Apr. 2023), pp. 4793–4800. DOI: 10.1093/mnras/stad447.

T. Gehren et al. "Host galaxies of quasars and their association with galaxy clusters." In: *Astrophys. J.* 278 (Mar. 1984), pp. 11–27. DOI: 10.1086/161763.

C. Gheller et al. "Evolution of cosmic filaments and of their galaxy population from MHD cosmological simulations." In: *Mon. Not. R. Astron. Soc.* 462.1 (Oct. 2016), pp. 448–463. DOI: 10.1093/mnras/stw1595.

C. Gheller et al. "Numerical cosmology on the GPU with Enzo and Ramses." In: *Journal of Physics Conference Series*. Vol. 640. Journal of Physics Conference Series. Sept. 2015, 012058, p. 012058. DOI: 10.1088/1742-6596/640/1/012058.

R. A. Gingold and J. J. Monaghan. "Smoothed particle hydrodynamics: theory and application to non-spherical stars." In: *Mon. Not. R. Astron. Soc.* 181 (Nov. 1977), pp. 375–389. DOI: 10.1093/mnras/181.3.375.

S. Giodini et al. "Scaling Relations for Galaxy Clusters: Properties and Evolution." In: *Space Sci. Rev.* 177.1-4 (Aug. 2013), pp. 247–282. DOI: 10.1007/s11214-013-9994-5.

M. D. Gladders and H. K. C. Yee. "A New Method For Galaxy Cluster Detection. I. The Algorithm." In: *Astron. J.* 120.4 (Oct. 2000), pp. 2148–2162. DOI: 10.1086/301557.

N. Y. Gnedin and A. A. Kaurov. "Cosmic Reionization on Computers. II. Reionization History and Its Back-reaction on Early Galaxies." In: *Astrophys. J.* 793.1, 30 (Sept. 2014), p. 30. DOI: 10.1088/0004-637X/793/1/30.

S. K. Godunov and I. Bohachevsky. "Finite difference method for numerical computation of discontinuous solutions of the equations of fluid dynamics." In: *Matematičeskij sbornik* 47(89).3 (1959), pp. 271–306. URL: https://hal.science/hal-01620642.

S. Gottloeber, Y. Hoffman, and G. Yepes. "Constrained Local UniversE Simulations (CLUES)." In: *arXiv e-prints* (May 2010). arXiv: 1005.2687 [astro-ph.CO].

S. A. Gregory and L. A. Thompson. "The Coma/A1367 supercluster and its environs." In: *Astrophys. J.* 222 (June 1978), pp. 784–799. DOI: 10.1086/156198.

E. H. B. M. Gronenschild and R. Mewe. "Calculated X-radiation from optically thin plasmas. III. Abundance effects on continuum emission." In: *Astron. Astroph. Suppl.* 32 (May 1978), pp. 283–305.

F. Groth et al. "The cosmological simulation code OPENGADGET3 - implementation of meshless finite mass." In: *Mon. Not. R. Astron. Soc.* 526.1 (Nov. 2023), pp. 616–644. DOI: 10.1093/mnras/stad2717.

J. E. Gunn and I. Gott J. Richard. "On the Infall of Matter Into Clusters of Galaxies and Some Effects on Their Evolution." In: *Astrophys. J.* 176 (Aug. 1972), p. 1. DOI: 10.1086/151605.

H. Gursky et al. "A Strong X-Ray Source in the Coma Cluster Observed by UHURU." In: *Astrophys. J. Lett.* 167 (Aug. 1971), p. L81. DOI: 10.1086/180765.

A. H. Guth and S. Y. Pi. "Fluctuations in the New Inflationary Universe." In: *Phys. Rev. Lett.* 49.15 (Oct. 1982), pp. 1110–1113. DOI: 10.1103/PhysRevLett.49.1110.

A. H. Guth. "Inflationary universe: A possible solution to the horizon and flatness problems." In: *Phys. Rev. D* 23.2 (Jan. 1981), pp. 347–356. DOI: 10.1103/PhysRevD.23.347.

A. H. Guth and E. J. Weinberg. "Cosmological consequences of a first-order phase transition in the SU_5 grand unified model." In: *Phys. Rev. D* 23.4 (Feb. 1981), pp. 876–885. DOI: 10.1103/PhysRevD.23.876.

F. Haardt and P. Madau. "Radiative Transfer in a Clumpy Universe. II. The Ultraviolet Extragalactic Background." In: *Astrophys. J.* 461 (Apr. 1996), p. 20. DOI: 10.1086/177035.

O. Hahn, T. Abel, and R. Kaehler. "A new approach to simulating collisionless dark matter fluids." In: *Mon. Not. R. Astron. Soc.* 434.2 (Sept. 2013), pp. 1171–1191. DOI: 10.1093/mnras/stt1061.

O. Hahn et al. "The evolution of dark matter halo properties in clusters, filaments, sheets and voids." In: *Mon. Not. R. Astron. Soc.* 381.1 (Oct. 2007), pp. 41–51. DOI: 10.1111/j.1365-2966.2007.12249.x.

N. Hamaus, P. M. Sutter, and B. D. Wandelt. "Universal Density Profile for Cosmic Voids." In: *Phys. Rev. Lett.* 112.25, 251302 (June 2014), p. 251302. DOI: 10.1103/PhysRevLett.112.251302.

J. C. Hamilton. "What have we learned from observational cosmology?" In: *Studies in the History and Philosophy of Modern Physics* 46 (May 2014), pp. 70–85. DOI: 10.1016/j.shpsb.2013.02.002.

N. Hand et al. "nbodykit: An Open-source, Massively Parallel Toolkit for Large-scale Structure." In: *Astron. J.* 156.4, 160 (Oct. 2018), p. 160. DOI: 10.3847/1538-3881/aadae0.

A. Harten and J. M. Hyman. "Self-Adjusting Grid Methods for One-Dimensional Hyperbolic Conservation Laws." In: *Journal of Computational Physics* 50.2 (May 1983), pp. 235–269. DOI: 10.1016/0021-9991(83)90066-9.

L. Hernquist and N. Katz. "TREESPH: A Unification of SPH with the Hierarchical Tree Method." In: *Astrophys. J. Suppl. Ser.* 70 (June 1989), p. 419. DOI: 10.1086/191344.

W. Herschel. "On the Construction of the Heavens." In: *Philosophical Transactions of the Royal Society of London Series I* 75 (Jan. 1785), pp. 213–266.

R. W. Hockney and J. W. Eastwood. *Computer Simulation Using Particles*. 1981.

Y. Hoffman and E. Ribak. "Constrained Realizations of Gaussian Fields: A Simple Algorithm." In: *Astrophys. J. Lett.* 380 (Oct. 1991), p. L5. DOI: 10.1086/186160.

F. Hofmann et al. "eROSITA cluster cosmology forecasts: Cluster temperature substructure bias." In: *Astron. Astrophys.* 606, A118 (Oct. 2017), A118. DOI: 10.1051/0004-6361/201730742.

E. Holmberg. "On the Clustering Tendencies among the Nebulae. II. a Study of Encounters Between Laboratory Models of Stellar Systems by a New Integration Procedure." In: *Astrophys. J.* 94 (Nov. 1941), p. 385. DOI: 10.1086/144344.

P. F. Hopkins. "A new class of accurate, mesh-free hydrodynamic simulation methods." In: *Mon. Not. R. Astron. Soc.* 450.1 (June 2015), pp. 53–110. DOI: 10.1093/mnras/stv195.

P. F. Hopkins et al. "FIRE-2 simulations: physics versus numerics in galaxy formation." In: *Mon. Not. R. Astron. Soc.* 480.1 (Oct. 2018), pp. 800–863. DOI: 10.1093/mnras/sty1690.

P. F. Hopkins et al. "Galaxies on FIRE (Feedback In Realistic Environments): stellar feedback explains cosmologically inefficient star formation." In: *Mon. Not. R. Astron. Soc.* 445.1 (Nov. 2014), pp. 581–603. DOI: 10.1093/mnras/stu1738.

F. Hoyle and R. A. Lyttleton. "On the accretion theory of stellar evolution." In: *Mon. Not. R. Astron. Soc.* 101 (Jan. 1941), p. 227. DOI: 10.1093/mnras/101.4.227.

W. Hu and S. Dodelson. "Cosmic Microwave Background Anisotropies." In: *Annu. Rev. Astron. Astrophys.* 40 (Jan. 2002), pp. 171–216. DOI: 10.1146/annurev.astro.40.060401.093926.

E. Hubble. "A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae." In: *Proceedings of the National Academy of Science* 15.3 (Mar. 1929), pp. 168–173. DOI: 10.1073/pnas.15.3.168.

L. Iapichino and J. C. Niemeyer. "Hydrodynamical adaptive mesh refinement simulations of turbulent flows - II. Cosmological simulations of galaxy clusters." In: *Mon. Not. R. Astron. Soc.* 388.3 (Aug. 2008), pp. 1089–1100. DOI: 10.1111/j.1365-2966.2008.13518.x.

V. Icke. "Voids and filaments." In: *Mon. Not. R. Astron. Soc.* 206 (Jan. 1984), 1P–3P. DOI: 10.1093/mnras/206.1.1P.

T. Ito et al. "A special-purpose N-body machine GRAPE-1." In: *Computer Physics Communications* 60.2 (Sept. 1990), pp. 187–194. DOI: 10.1016/0010-4655(90)90003-J.

N. Itoh, Y. Kohyama, and S. Nozawa. "Relativistic Corrections to the Sunyaev-Zeldovich Effect for Clusters of Galaxies." In: *Astrophys. J.* 502.1 (July 1998), pp. 7–15. DOI: 10.1086/305876.

M. Jõeveer, J. Einasto, and E. Tago. "Spatial distribution of galaxies and of clusters of galaxies in the southern galactic hemisphere." In: *Mon. Not. R. Astron. Soc.* 185 (Nov. 1978), pp. 357–370. DOI: 10.1093/mnras/185.2.357.

A. Jenkins. "Second-order Lagrangian perturbation theory initial conditions for resimulations." In: *Mon. Not. R. Astron. Soc.* 403.4 (Apr. 2010), pp. 1859–1872. DOI: 10.1111/j.1365-2966.2010.16259.x.

Y. P. Jing and Y. Suto. "Triaxial Modeling of Halo Density Profiles with High-Resolution N-Body Simulations." In: *Astrophys. J.* 574.2 (Aug. 2002), pp. 538–553. DOI: 10.1086/341065.

M. E. Jones. "A High-Redshift Cluster Detected via its S-Z Effect." In: *Observational Cosmology with the New Radio Surveys*. Ed. by M. N. Bremer, N. Jackson, and I. Perez-Fournon. Vol. 226. Astrophysics and Space Science Library. Jan. 1998, p. 89. DOI: 10.1007/978-94-011-5238-9_16.

N. Kaiser. "Evolution and clustering of rich clusters." In: *Mon. Not. R. Astron. Soc.* 222 (Sept. 1986), pp. 323–345. DOI: 10.1093/mnras/222.2.323.

H. Kang, D. Ryu, and T. W. Jones. "Cluster Accretion Shocks as Possible Acceleration Sites for Ultra–High-Energy Protons below the Greisen Cutoff." In: *Astrophys. J.* 456 (Jan. 1996), p. 422. DOI: 10.1086/176666.

H. Kang et al. "Cosmological Shock Waves in the Large-Scale Structure of the Universe: Nongravitational Effects." In: *Astrophys. J.* 669.2 (Nov. 2007), pp. 729–740. DOI: 10.1086/521717.

R. Kannan et al. "Accurately simulating anisotropic thermal conduction on a moving mesh." In: *Mon. Not. R. Astron. Soc.* 458.1 (May 2016), pp. 410–424. DOI: 10.1093/mnras/stw294.

N. Katz, D. H. Weinberg, and L. Hernquist. "Cosmological Simulations with TreeSPH." In: *Astrophys. J. Suppl. Ser.* 105 (July 1996), p. 19. DOI: 10.1086/192305.

A. Kawai et al. "GRAPE-5: A Special-Purpose Computer for N-Body Simulations." In: *Publ. Astron. Soc. Jpn* 52 (Aug. 2000), pp. 659–676. DOI: 10.1093/pasj/52.4.659.

J. Kennicutt Robert C. "The Global Schmidt Law in Star-forming Galaxies." In: *Astrophys. J.* 498.2 (May 1998), pp. 541–552. DOI: 10.1086/305588.

R. P. Kirshner et al. "A million cubic megaparsec void in Bootes ?" In: *Astrophys. J. Lett.* 248 (Sept. 1981), pp. L57–L60. DOI: 10.1086/183623.

A. Klypin et al. "Resolving the Structure of Cold Dark Matter Halos." In: *Astrophys. J.* 554.2 (June 2001), pp. 903–915. DOI: 10.1086/321400.

A. Kolmogorov. "The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds' Numbers." In: *Akademiia Nauk SSSR Doklady* 30 (Jan. 1941), pp. 301–305.

E. Komatsu et al. "Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation." In: *Astrophys. J. Suppl. Ser.* 192.2, 18 (Feb. 2011), p. 18. DOI: 10.1088/0067-0049/192/2/18.

A. V. Kravtsov and S. Borgani. "Formation of Galaxy Clusters." In: *Annu. Rev. Astron. Astrophys.* 50 (Sept. 2012), pp. 353–409. DOI: 10.1146/annurev-astro-081811-125502.

K. Kreckel et al. "Only the Lonely: H I Imaging of Void Galaxies." In: *Astron. J.* 141.1, 4 (Jan. 2011), p. 4. DOI: 10.1088/0004-6256/141/1/4.

J. H. Krolik. *Active galactic nuclei : from the central black hole to the galactic environment*. 1999.

P. Kroupa, C. Theis, and C. M. Boily. "The great disk of Milky-Way satellites and cosmological sub-structures." In: *Astron. Astrophys.* 431 (Feb. 2005), pp. 517–521. DOI: 10.1051/0004-6361:20041122.

U. Kuchner et al. "An inventory of galaxies in cosmic filaments feeding galaxy clusters: galaxy groups, backsplash galaxies, and pristine galaxies." In: *Mon. Not. R. Astron. Soc.* 510.1 (Feb. 2022), pp. 581–592. DOI: 10.1093/mnras/stab3419.

W. Kutta. "Beitrag zur näherungsweisen Integration totaler Differentialgleichungen." In: *Zeitschrift für Mathematik und Physik* 46 (1901), pp. 435–453.

L. D. Landau and E. M. Lifshitz. *Fluid Mechanics*. 1987.

M. I. Large, D. S. Mathewson, and C. G. T. Haslam. "A High-Resolution Survey of the Coma Cluster of Galaxies at 408 Mc./s." In: *Nature* 183.4676 (June 1959), pp. 1663–1664. DOI: 10.1038/1831663a0.

R. B. Larson. "A model for the formation of a spherical galaxy." In: *Mon. Not. R. Astron. Soc.* 145 (Jan. 1969), p. 405. DOI: 10.1093/mnras/145.4.405.

E. T. Lau, A. V. Kravtsov, and D. Nagai. "Residual Gas Motions in the Intracluster Medium and Bias in Hydrostatic Measurements of Mass Profiles of Clusters." In: *Astrophys. J.* 705.2 (Nov. 2009), pp. 1129–1138. DOI: 10.1088/0004-637X/705/2/1129.

E. T. Lau et al. "Mass Accretion and its Effects on the Self-similarity of Gas Profiles in the Outskirts of Galaxy Clusters." In: *Astrophys. J.* 806.1, 68 (June 2015), p. 68. DOI: 10.1088/0004-637X/806/1/68.

G. Lavaux and B. D. Wandelt. "Precision Cosmography with Stacked Voids." In: *Astrophys. J.* 754.2, 109 (Aug. 2012), p. 109. DOI: 10.1088/0004-637X/754/2/109.

R. J. LeVeque. *Numerical methods for conservation laws (2. ed.)*. Lectures in mathematics. Birkhäuser, 1992, pp. 1–214. ISBN: 978-3-7643-2723-1.

A. D. Linde. "A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems." In: *Physics Letters B* 108.6 (Feb. 1982), pp. 389–393. DOI: 10.1016/0370-2693(82)91219-9.

J. Liske et al. "Galaxy And Mass Assembly (GAMA): end of survey report and data release 2." In: *Mon. Not. R. Astron. Soc.* 452.2 (Sept. 2015), pp. 2087–2126. DOI: 10.1093/mnras/stv1436.

J. López-Miralles et al. "3D RMHD simulations of jet-wind interactions in high-mass X-ray binaries." In: *Astron. Astrophys.* 661, A117 (May 2022), A117. DOI: 10.1051/0004-6361/202142968.

J. López-Miralles, J. M. Martí, and M. Perucho. "On the application of Jacobian-free Riemann solvers for relativistic radiation magnetohydrodynamics under M1 closure." In: *Computer Physics Communications* 284, 108630 (Mar. 2023), p. 108630. DOI: 10.1016/j.cpc.2022.108630.

L. Lovisari and B. J. Maughan. "Scaling Relations of Clusters and Groups and Their Evolution." In: *Handbook of X-ray and Gamma-ray Astrophysics*. 2022, 65, p. 65. DOI: 10.1007/978-981-16-4544-0_118-1.

L. B. Lucy. "A numerical approach to the testing of the fission hypothesis." In: *Astron. J.* 82 (Dec. 1977), pp. 1013–1024. DOI: 10.1086/112164.

S. L. Lumsden et al. "The Edinburgh-Durham southern galaxy catalogue. IV. Thecluster catalogue." In: *Mon. Not. R. Astron. Soc.* 258 (Sept. 1992), pp. 1–22. DOI: 10.1093/mnras/258.1.1.

J. Makino et al. "GRAPE-4: A Massively Parallel Special-Purpose Computer for Collisional N-Body Simulations." In: *Astrophys. J.* 480.1 (May 1997), pp. 432–446. DOI: 10.1086/303972.

J. Makino et al. "GRAPE-6: Massively-Parallel Special-Purpose Computer for Astrophysical Particle Simulations." In: *Publ. Astron. Soc. Jpn* 55 (Dec. 2003), pp. 1163–1187. DOI: 10.1093/pasj/55.6.1163.

G. A. Mamon, A. Biviano, and G. Boué. "MAMPOSSt: Modelling Anisotropy and Mass Profiles of Observed Spherical Systems - I. Gaussian 3D velocities." In: *Mon. Not. R. Astron. Soc.* 429.4 (Mar. 2013), pp. 3079–3098. DOI: 10.1093/mnras/sts565.

A. Mantz et al. "The observed growth of massive galaxy clusters - I. Statistical methods and cosmological constraints." In: *Mon. Not. R. Astron. Soc.* 406.3 (Aug. 2010), pp. 1759–1772. DOI: 10.1111/j.1365-2966.2010.16992.x.

A. B. Mantz et al. "Cosmology and astrophysics from relaxed galaxy clusters - I. Sample selection." In: *Mon. Not. R. Astron. Soc.* 449.1 (May 2015), pp. 199–219. DOI: 10.1093/mnras/stv219.

M. Markevitch, A. Vikhlinin, and P. Mazzotta. "Nonhydrostatic Gas in the Core of the Relaxed Galaxy Cluster A1795." In: *Astrophys. J. Lett.* 562.2 (Dec. 2001), pp. L153–L156. DOI: 10.1086/337973.

M. Markevitch et al. "The Temperature Structure of 30 Nearby Clusters Observed with ASCA: Similarity of Temperature Profiles." In: *Astrophys. J.* 503.1 (Aug. 1998), pp. 77–96. DOI: 10.1086/305976.

J. M. Martí and E. Müller. "Extension of the Piecewise Parabolic Method to One-Dimensional Relativistic Hydrodynamics." In: *Journal of Computational Physics* 123.1 (Jan. 1996), pp. 1–14. DOI: 10.1006/jcph.1996.0001.

S. Mashchenko, J. Wadsley, and H. M. P. Couchman. "Stellar Feedback in Dwarf Galaxy Formation." In: *Science* 319.5860 (Jan. 2008), p. 174. DOI: 10.1126/science.1148666.

C. Maureira-Fredes and P. Amaro-Seoane. "GRAVIDY, a GPU modular, parallel direct-summation N-body integrator: dynamics with softening." In: *Mon. Not. R. Astron. Soc.* 473.3 (Jan. 2018), pp. 3113–3127. DOI: 10.1093/mnras/stx2468.

I. G. McCarthy et al. "Towards a holistic view of the heating and cooling of the intracluster medium." In: *Mon. Not. R. Astron. Soc.* 386.3 (May 2008), pp. 1309–1331. DOI: 10.1111/j.1365-2966.2008.13141.x.

I. G. McCarthy et al. "Models of the Intracluster Medium with Heating and Cooling: Explaining the Global and Structural X-Ray Properties of Clusters." In: *Astrophys. J.* 613.2 (Oct. 2004), pp. 811–830. DOI: 10.1086/423267.

R. McKinnon, P. Torrey, and M. Vogelsberger. "Dust formation in Milky Way-like galaxies." In: *Mon. Not. R. Astron. Soc.* 457.4 (Apr. 2016), pp. 3775–3800. DOI: 10.1093/mnras/stw253.

B. R. McNamara and P. E. J. Nulsen. "Heating Hot Atmospheres with Active Galactic Nuclei." In: *Annu. Rev. Astron. Astrophys.* 45.1 (Sept. 2007), pp. 117–175. DOI: 10.1146/annurev.astro.45.051806.110625.

F. Mernier and V. Biffi. "Chemical Enrichment in Groups and Clusters." In: *Handbook of X-ray and Gamma-ray Astrophysics*. 2022, 12, p. 12. DOI: 10.1007/978-981-16-4544-0_123-1.

J. C. Mihos et al. "The Burrell Schmidt Deep Virgo Survey: Tidal Debris, Galaxy Halos, and Diffuse Intracluster Light in the Virgo Cluster." In: *Astrophys. J.* 834.1, 16 (Jan. 2017), p. 16. DOI: 10.3847/1538-4357/834/1/16.

P. Mimica et al. "Spectral Evolution of Superluminal Components in Parsec-Scale Jets." In: *Astrophys. J.* 696.2 (May 2009), pp. 1142–1163. DOI: 10.1088/0004-637X/696/2/1142.

P. Mimica et al. "Numerical simulations of the jetted tidal disruption event Swift J1644+57." In: *Journal of Physics Conference Series*. Vol. 719. Journal of Physics Conference Series. May 2016, 012008, p. 012008. DOI: 10.1088/1742-6596/719/1/012008.

F. Miniati. "Intergalactic shock acceleration and the cosmic gamma-ray background." In: *Mon. Not. R. Astron. Soc.* 337.1 (Nov. 2002), pp. 199–208. DOI: 10.1046/j.1365-8711.2002.05903.x.

F. Miniati. "The Matryoshka Run: A Eulerian Refinement Strategy to Study the Statistics of Turbulence in Virialized Cosmic Structures." In: *Astrophys. J.* 782.1, 21 (Feb. 2014), p. 21. DOI: 10.1088/0004-637X/782/1/21.

F. Miniati et al. "Properties of Cosmic Shock Waves in Large-Scale Structure Formation." In: *Astrophys. J.* 542.2 (Oct. 2000), pp. 608–621. DOI: 10.1086/317027.

J. J. Mohr, D. G. Fabricant, and M. J. Geller. "An X-Ray Method for Detecting Substructure in Galaxy Clusters: Application to Perseus, A2256, Centaurus, Coma, and Sersic 40/6." In: *Astrophys. J.* 413 (Aug. 1993), p. 492. DOI: 10.1086/173019.

J. J. Monaghan. "SPH and Riemann Solvers." In: *Journal of Computational Physics* 136.2 (Sept. 1997), pp. 298–307. DOI: 10.1006/jcph.1997.5732.

J. J. Monaghan and R. A. Gingold. "Shock Simulation by the Particle Method SPH." In: *Journal of Computational Physics* 52.2 (Nov. 1983), pp. 374–389. DOI: 10.1016/0021-9991(83)90036-0.

J. J. Monaghan and J. C. Lattanzio. "A refined particle method for astrophysical problems." In: *Astron. Astrophys.* 149.1 (Aug. 1985), pp. 135–143.

Ó. Monllor-Berbegal et al. "The imprints of galaxy cluster internal dynamics on the Sunyaev-Zeldovich effect." In: *Astron. Astrophys.* 686, A243 (June 2024), A243. DOI: 10.1051/0004-6361/202348967.

M. Montes and I. Trujillo. "Intracluster light: a luminous tracer for dark matter in clusters of galaxies." In: *Mon. Not. R. Astron. Soc.* 482.2 (Jan. 2019), pp. 2838–2851. DOI: 10.1093/mnras/sty2858.

B. Moore et al. "Collisional versus Collisionless Dark Matter." In: *Astrophys. J. Lett.* 535.1 (May 2000), pp. L21–L24. DOI: 10.1086/312692.

J. P. Morris and J. J. Monaghan. "A Switch to Reduce SPH Viscosity." In: *Journal of Computational Physics* 136.1 (Sept. 1997), pp. 41–50. DOI: 10.1006/jcph.1997.5690.

T. Mroczkowski et al. "Astrophysics with the Spatially and Spectrally Resolved Sunyaev-Zeldovich Effects. A Millimetre/Submillimetre Probe of the Warm and Hot Universe." In: *Space Sci. Rev.* 215.1, 17 (Feb. 2019), p. 17. DOI: 10.1007/s11214-019-0581-2.

V. Mukhanov. "Inflation II: origin of the primordial inhomogeneities." In: *Physical Foundations of Cosmology*. Cambridge University Press, 2005, 322–355. DOI: 10.1017/CBO9780511790553.011.

G. Murante et al. "A subresolution multiphase interstellar medium model of star formation and supernova energy feedback." In: *Mon. Not. R. Astron. Soc.* 405.3 (July 2010), pp. 1491–1512. DOI: 10.1111/j.1365-2966.2010.16567.x.

S. Nadathur et al. "Beyond BAO: Improving cosmological constraints from BOSS data with measurement of the void-galaxy cross-correlation." In: *Phys. Rev. D* 100.2, 023504 (July 2019), p. 023504. DOI: 10.1103/PhysRevD.100.023504.

D. Nagai and E. T. Lau. "Gas Clumping in the Outskirts of ΛCDM Clusters." In: *Astrophys. J. Lett.* 731.1, L10 (Apr. 2011), p. L10. DOI: 10.1088/2041-8205/731/1/L10.

J. F. Navarro, C. S. Frenk, and S. D. M. White. "A Universal Density Profile from Hierarchical Clustering." In: *Astrophys. J.* 490.2 (Dec. 1997), pp. 493–508. DOI: 10.1086/304888.

J. F. Navarro et al. "The diversity and similarity of simulated cold dark matter haloes." In: *Mon. Not. R. Astron. Soc.* 402.1 (Feb. 2010), pp. 21–34. DOI: 10.1111/j.1365-2966.2009.15878.x.

D. Nelson et al. "The IllustrisTNG simulations: public data release." In: *Computational Astrophysics and Cosmology* 6.1, 2 (May 2019), p. 2. DOI: 10.1186/s40668-019-0028-x.

A. F. Neto et al. "The statistics of ΛCDM halo concentrations." In: *Mon. Not. R. Astron. Soc.* 381.4 (Nov. 2007), pp. 1450–1462. DOI: 10.1111/j.1365-2966.2007.12381.x.

J. A. Newman et al. "The DEEP2 Galaxy Redshift Survey: Design, Observations, Data Reduction, and Redshifts." In: *Astrophys. J. Suppl. Ser.* 208.1, 5 (Sept. 2013), p. 5. DOI: 10.1088/0067-0049/208/1/5.

J. Nickolls et al. "Scalable Parallel Programming with CUDA." en. In: *ACM Queue* 6.2 (Mar. 2008), pp. 40–53.

S. Nozawa, N. Itoh, and Y. Kohyama. "Relativistic corrections to the Sunyaev-Zeldovich effect for clusters of galaxies: effect of the motion of the observer." In: *Astron. Astrophys.* 440.1 (Sept. 2005), pp. 39–44. DOI: 10.1051/0004-6361:20052923.

B. K. Oh et al. "Calibration of a star formation and feedback model for cosmological simulations with ENZO." In: *Mon. Not. R. Astron. Soc.* 497.4 (Oct. 2020), pp. 5203–5219. DOI: 10.1093/mnras/staa2318.

T. Padmanabhan. *Theoretical Astrophysics - Volume 3, Galaxies and Cosmology*. Vol. 3. 2002. DOI: 10.2277/0521562422.

R. Pakmor et al. "Galactic Winds Driven by Isotropic and Anisotropic Cosmic-Ray Diffusion in Disk Galaxies." In: *Astrophys. J. Lett.* 824.2, L30 (June 2016), p. L30. DOI: 10.3847/2041-8205/824/2/L30.

P. J. E. Peebles. *Principles of Physical Cosmology*. 1993. DOI: 10.1515/9780691206721.

P. J. E. Peebles. "Structure of the Coma Cluster of Galaxies." In: *Astron. J.* 75 (Feb. 1970), p. 13. DOI: 10.1086/110933.

P. J. E. Peebles. *The large-scale structure of the universe*. 1980.

A. A. Penzias and R. W. Wilson. "A Measurement of Excess Antenna Temperature at 4080 Mc/s." In: *Astrophys. J.* 142 (July 1965), pp. 419–421. DOI: 10.1086/148307.

S. Perlmutter et al. "Discovery of a supernova explosion at half the age of the Universe." In: *Nature* 391.6662 (Jan. 1998), pp. 51–54. DOI: 10.1038/34124.

S. Perlmutter, M. S. Turner, and M. White. "Constraining Dark Energy with Type Ia Supernovae and Large-Scale Structure." In: *Phys. Rev. Lett.* 83.4 (July 1999), pp. 670–673. DOI: 10.1103/PhysRevLett.83.670.

C. Pfrommer et al. "Simulating cosmic ray physics on a moving mesh." In: *Mon. Not. R. Astron. Soc.* 465.4 (Mar. 2017), pp. 4500–4529. DOI: 10.1093/mnras/stw2941.

A. Pisani et al. “Cosmic voids: a novel probe to shed light on our Universe.” In: *Bull. Am. Astron. Soc.* 51.3, 40 (May 2019), p. 40. arXiv: 1903.05161 [astro-ph.CO].

Planck Collaboration et al. “Planck 2018 results. VI. Cosmological parameters.” In: *Astron. Astrophys.* 641, A6 (Sept. 2020), A6. DOI: 10.1051/0004-6361/201833910.

Planck Collaboration et al. “Planck early results. VIII. The all-sky early Sunyaev-Zeldovich cluster sample.” In: *Astron. Astrophys.* 536, A8 (Dec. 2011), A8. DOI: 10.1051/0004-6361/201116459.

S. Planelles and V. Quilis. “ASOHF: a new adaptive spherical overdensity halo finder.” In: *Astron. Astrophys.* 519, A94 (Sept. 2010), A94. DOI: 10.1051/0004-6361/201014214.

S. Planelles, D. R. G. Schleicher, and A. M. Bykov. “Large-Scale Structure Formation: From the First Non-linear Objects to Massive Galaxy Clusters.” In: *Space Sci. Rev.* 188.1-4 (May 2015), pp. 93–139. DOI: 10.1007/s11214-014-0045-7.

S. Planelles et al. “Baryon census in hydrodynamical simulations of galaxy clusters.” In: *Mon. Not. R. Astron. Soc.* 431.2 (May 2013), pp. 1487–1502. DOI: 10.1093/mnras/stt265.

S. Planelles et al. “Exploring the role of cosmological shock waves in the Dianoga simulations of galaxy clusters.” In: *Mon. Not. R. Astron. Soc.* 507.4 (Nov. 2021), pp. 5703–5719. DOI: 10.1093/mnras/stab2436.

S. Planelles et al. “On the role of AGN feedback on the thermal and chemodynamical properties of the hot intracluster medium.” In: *Mon. Not. R. Astron. Soc.* 438.1 (Feb. 2014), pp. 195–216. DOI: 10.1093/mnras/stt2141.

S. Planelles and V. Quilis. “Cosmological shock waves: clues to the formation history of haloes.” In: *Mon. Not. R. Astron. Soc.* 428.2 (Jan. 2013), pp. 1643–1655. DOI: 10.1093/mnras/sts142.

S. Planelles and V. Quilis. “Galaxy cluster mergers.” In: *Mon. Not. R. Astron. Soc.* 399.1 (Oct. 2009), pp. 410–424. DOI: 10.1111/j.1365-2966.2009.15290.x.

S. Planelles et al. “Multiwavelength mock observations of the WHIM in a simulated galaxy cluster.” In: *Mon. Not. R. Astron. Soc.* 476.4 (June 2018), pp. 4629–4648. DOI: 10.1093/mnras/sty527.

D. H. Porter, T. W. Jones, and D. Ryu. “Vorticity, Shocks, and Magnetic Fields in Subsonic, ICM-like Turbulence.” In: *Astrophys. J.* 810.2, 93 (Sept. 2015), p. 93. DOI: 10.1088/0004-637X/810/2/93.

C. Power et al. “The inner structure of ΛCDM haloes - I. A numerical convergence study.” In: *Mon. Not. R. Astron. Soc.* 338.1 (Jan. 2003), pp. 14–34. DOI: 10.1046/j.1365-8711.2003.05925.x.

V. Presotto et al. “Intracluster light properties in the CLASH-VLT cluster MACS J1206.2-0847.” In: *Astron. Astrophys.* 565, A126 (May 2014), A126. DOI: 10.1051/0004-6361/201323251.

W. H. Press and P. Schechter. “Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation.” In: *Astrophys. J.* 187 (Feb. 1974), pp. 425–438. DOI: 10.1086/152650.

W. H. Press et al. *Numerical recipes in FORTRAN. The art of scientific computing*. 1992.

D. J. Price. “Smoothed particle hydrodynamics and magnetohydrodynamics.” In: *Journal of Computational Physics* 231.3 (Feb. 2012), pp. 759–794. DOI: 10.1016/j.jcp.2010.12.011.

V. Quilis, J. M. Ibañez, and D. Sáez. “Modern high-resolution shock-capturing methods for structure evolution in cosmology.” In: *Astron. Astrophys.* 286 (June 1994), pp. 1–16.

V. Quilis, J. M. Ibañez, and D. Sáez. “Nonlinear Evolution of One Dimensional Cosmological Inhomogeneities.” In: *Rev. Mex. Astron. Astrofis.* 25 (Oct. 1993), p. 117.

V. Quilis. “A new multidimensional adaptive mesh refinement hydro + gravity cosmological code.” In: *Mon. Not. R. Astron. Soc.* 352.4 (Aug. 2004), pp. 1426–1438. DOI: 10.1111/j.1365-2966.2004.08040.x.

V. Quilis, R. G. Bower, and M. L. Balogh. “Bubbles, feedback and the intracluster medium: three-dimensional hydrodynamic simulations.” In: *Mon. Not. R. Astron. Soc.* 328.4 (Dec. 2001), pp. 1091–1097. DOI: 10.1046/j.1365-8711.2001.04927.x.

V. Quilis, J.-M. Martí, and S. Planelles. “Cosmic magnetic fields with MASCLET: an application to galaxy clusters.” In: *Mon. Not. R. Astron. Soc.* 494.2 (May 2020), pp. 2706–2717. DOI: 10.1093/mnras/staa877.

V. Quilis and I. Trujillo. "Expected Number of Massive Galaxy Relics in the Present Day Universe." In: *Astrophys. J. Lett.* 773.1, L8 (Aug. 2013), p. L8. DOI: 10.1088/2041-8205/773/1/L8.

V. Quilis et al. "On the Role of Shock Waves in Galaxy Cluster Evolution." In: *Astrophys. J.* 502.2 (Aug. 1998), pp. 518–530. DOI: 10.1086/305932.

E. Rasia et al. "Cool Core Clusters from Cosmological Simulations." In: *Astrophys. J. Lett.* 813.1, L17 (Nov. 2015), p. L17. DOI: 10.1088/2041-8205/813/1/L17.

E. Rasia, M. Meneghetti, and S. Ettori. "X-ray Morphological Estimators for Galaxy Clusters." In: *The Astronomical Review* 8 (Jan. 2013), pp. 40–70. DOI: 10.1080/21672857.2013.11519713.

J. C. Raymond and B. W. Smith. "Soft X-ray spectrum of a hot plasma." In: *Astrophys. J. Suppl. Ser.* 35 (Dec. 1977), pp. 419–439. DOI: 10.1086/190486.

J. I. Read et al. "The tidal stripping of satellites." In: *Mon. Not. R. Astron. Soc.* 366.2 (Feb. 2006), pp. 429–437. DOI: 10.1111/j.1365-2966.2005.09861.x.

A. E. Reines et al. "An actively accreting massive black hole in the dwarf starburst galaxy Henize2-10." In: *Nature* 470.7332 (Feb. 2011), pp. 66–68. DOI: 10.1038/nature09724.

T. H. Reiprich and H. Böhringer. "The Mass Function of an X-Ray Flux-limited Sample of Galaxy Clusters." In: *Astrophys. J.* 567.2 (Mar. 2002), pp. 716–740. DOI: 10.1086/338753.

A. Renzini. "Iron as a Tracer in Galaxy Clusters and Groups." In: *Astrophys. J.* 488.1 (Oct. 1997), pp. 35–43. DOI: 10.1086/304696.

Y. Rephaeli et al. "Nonthermal Phenomena in Clusters of Galaxies." In: *Space Sci. Rev.* 134.1-4 (Feb. 2008), pp. 71–92. DOI: 10.1007/s11214-008-9314-7.

Y. Rephaeli. "Comptonization Of The Cosmic Microwave Background: The Sunyaev-Zeldovich Effect." In: *Annu. Rev. Astron. Astrophys.* 33 (Jan. 1995), pp. 541–580. DOI: 10.1146/annurev.aa.33.090195.002545.

E. Ricciardelli, V. Quilis, and S. Planelles. "The structure of cosmic voids in a ΛCDM Universe." In: *Mon. Not. R. Astron. Soc.* 434.2 (Sept. 2013), pp. 1192–1204. DOI: 10.1093/mnras/stt1069.

E. Ricciardelli et al. "The star formation activity in cosmic voids." In: *Mon. Not. R. Astron. Soc.* 445.4 (Dec. 2014), pp. 4045–4054. DOI: 10.1093/mnras/stu2061.

L. F. Richardson. *Weather Prediction by Numerical Process*. Cambridge University Press, 1922.

A. G. Riess et al. "Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant." In: *Astron. J.* 116.3 (Sept. 1998), pp. 1009–1038. DOI: 10.1086/300499.

C. J. Riseley et al. "Radio fossils, relics, and haloes in Abell 3266: cluster archaeology with ASKAP-EMU and the ATCA." In: *Mon. Not. R. Astron. Soc.* 515.2 (Sept. 2022), pp. 1871–1896. DOI: 10.1093/mnras/stac1771.

P. L. Roe. "Approximate Riemann Solvers, Parameter Vectors, and Difference Schemes." In: *Journal of Computational Physics* 43.2 (Oct. 1981), pp. 357–372. DOI: 10.1016/0021-9991(81)90128-5.

E. Roediger et al. "Kelvin-Helmholtz Instabilities at the Sloshing Cold Fronts in the Virgo Cluster as a Measure for the Effective Intracluster Medium Viscosity." In: *Astrophys. J.* 764.1, 60 (Feb. 2013), p. 60. DOI: 10.1088/0004-637X/764/1/60.

J. Román and I. Trujillo. "Ultra-diffuse galaxies outside clusters: clues to their formation and evolution." In: *Mon. Not. R. Astron. Soc.* 468.4 (July 2017), pp. 4039–4047. DOI: 10.1093/mnras/stx694.

M. Roncarelli et al. "Measuring turbulence and gas motions in galaxy clusters via synthetic Athena X-IFU observations." In: *Astron. Astrophys.* 618, A39 (Oct. 2018), A39. DOI: 10.1051/0004-6361/201833371.

P. Rosati, S. Borgani, and C. Norman. "The Evolution of X-ray Clusters of Galaxies." In: *Annu. Rev. Astron. Astrophys.* 40 (Jan. 2002), pp. 539–577. DOI: 10.1146/annurev.astro.40.120401.150547.

J. Rosdahl et al. "The SPHINX cosmological simulations of the first billion years: the impact of binary stars on reionization." In: *Mon. Not. R. Astron. Soc.* 479.1 (Sept. 2018), pp. 994–1016. DOI: 10.1093/mnras/sty1655.

V. C. Rubin, J. Ford W. K., and N. Thonnard. "Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 (R=4kpc) to UGC 2885 (R=122kpc)." In: *Astrophys. J.* 238 (June 1980), pp. 471–487. DOI: 10.1086/158003.

C. Runge. "Ueber die numerische Auflsung von Differentialgleichungen." In: *Mathematische Annalen* 46.2 (June 1895), pp. 167–178. DOI: 10.1007/bf01446807. URL: https://doi.org/10.1007/bf01446807.

M. Ruszkowski and S. P. Oh. "Galaxy motions, turbulence and conduction in clusters of galaxies." In: *Mon. Not. R. Astron. Soc.* 414.2 (June 2011), pp. 1493–1507. DOI: 10.1111/j.1365-2966.2011.18482.x.

M. Ruszkowski et al. "Cosmological Magnetohydrodynamic Simulations of Cluster Formation with Anisotropic Thermal Conduction." In: *Astrophys. J.* 740.2, 81 (Oct. 2011), p. 81. DOI: 10.1088/0004-637X/740/2/81.

M. Ruszkowski, H. Y. K. Yang, and C. S. Reynolds. "Cosmic-Ray Feedback Heating of the Intracluster Medium." In: *Astrophys. J.* 844.1, 13 (July 2017), p. 13. DOI: 10.3847/1538-4357/aa79f8.

D. Ryu and H. Kang. "Accreting matter around clusters of galaxies: one- dimensional considerations." In: *Mon. Not. R. Astron. Soc.* 284.2 (Jan. 1997), pp. 416–424. DOI: 10.1093/mnras/284.2.416.

D. Ryu et al. "A Cosmological Hydrodynamic Code Based on the Total Variation Diminishing Scheme." In: *Astrophys. J.* 414 (Sept. 1993), p. 1. DOI: 10.1086/173051.

D. Ryu et al. "Cosmological Shock Waves and Their Role in the Large-Scale Structure of the Universe." In: *Astrophys. J.* 593.2 (Aug. 2003), pp. 599–610. DOI: 10.1086/376723.

E. E. Salpeter. "The Luminosity Function and Stellar Evolution." In: *Astrophys. J.* 121 (Jan. 1955), p. 161. DOI: 10.1086/145971.

J. S. Sanders, A. C. Fabian, and R. K. Smith. "Constraints on turbulent velocity broadening for a sample of clusters, groups and elliptical galaxies using XMM-Newton." In: *Mon. Not. R. Astron. Soc.* 410.3 (Jan. 2011), pp. 1797–1812. DOI: 10.1111/j.1365-2966.2010.17561.x.

J. S. Sanders et al. "A direct limit on the turbulent velocity of the intracluster medium in the core of Abell 1835 from XMM-Newton." In: *Mon. Not. R. Astron. Soc.* 402.1 (Feb. 2010), pp. L11–L15. DOI: 10.1111/j.1745-3933.2009.00789.x.

A. J. R. Sanderson et al. "The Birmingham-CfA cluster scaling project - I. Gas fraction and the M-T_X relation." In: *Mon. Not. R. Astron. Soc.* 340.3 (Apr. 2003), pp. 989–1010. DOI: 10.1046/j.1365-8711.2003.06401.x.

K. Schaal and V. Springel. "Shock finding on a moving mesh - I. Shock statistics in non-radiative cosmological simulations." In: *Mon. Not. R. Astron. Soc.* 446.4 (Feb. 2015), pp. 3992–4007. DOI: 10.1093/mnras/stu2386.

M. Schaller et al. "A new framework for numerical simulations of structure formation." In: *Mon. Not. R. Astron. Soc.* 442.4 (Aug. 2014), pp. 3073–3095. DOI: 10.1093/mnras/stu1069.

J. Schaye. "Star Formation Thresholds and Galaxy Edges: Why and Where." In: *Astrophys. J.* 609.2 (July 2004), pp. 667–682. DOI: 10.1086/421232.

J. Schaye et al. "The EAGLE project: simulating the evolution and assembly of galaxies and their environments." In: *Mon. Not. R. Astron. Soc.* 446.1 (Jan. 2015), pp. 521–554. DOI: 10.1093/mnras/stu2058.

H.-Y. Schive et al. "Graphic-card cluster for astrophysics (GraCCA) - Performance tests." In: *New Astronomy* 13.6 (Aug. 2008), pp. 418–435. DOI: 10.1016/j.newast.2007.12.005.

M. Schmidt. "The Rate of Star Formation." In: *Astrophys. J.* 129 (Mar. 1959), p. 243. DOI: 10.1086/146614.

P. Schuecker et al. "Probing turbulence in the Coma galaxy cluster." In: *Astron. Astrophys.* 426 (Nov. 2004), pp. 387–397. DOI: 10.1051/0004-6361:20041039.

V. A. Semenov, A. V. Kravtsov, and N. Y. Gnedin. "The Physical Origin of Long Gas Depletion Times in Galaxies." In: *Astrophys. J.* 845.2, 133 (Aug. 2017), p. 133. DOI: 10.3847/1538-4357/aa8096.

S. F. Shandarin and Y. B. Zeldovich. "The large-scale structure of the universe: Turbulence, intermittency, structures in a self-gravitating medium." In: *Reviews of Modern Physics* 61.2 (Apr. 1989), pp. 185–220. DOI: 10.1103/RevModPhys.61.185.

S. Shandarin et al. "Shapes and sizes of voids in the Lambda cold dark matter universe: excursion set approach." In: *Mon. Not. R. Astron. Soc.* 367.4 (Apr. 2006), pp. 1629–1640. DOI: 10.1111/j.1365-2966.2006.10062.x.

L. D. Shaw et al. "Statistics of Physical Properties of Dark Matter Clusters." In: *Astrophys. J.* 646.2 (Aug. 2006), pp. 815–833. DOI: 10.1086/505016.

R. K. Sheth and R. van de Weygaert. "A hierarchy of voids: much ado about nothing." In: *Mon. Not. R. Astron. Soc.* 350.2 (May 2004), pp. 517–538. DOI: 10.1111/j.1365-2966.2004.07661.x.

C.-W. Shu and S. Osher. "Efficient Implementation of Essentially Non-oscillatory Shock-Capturing Schemes." In: *Journal of Computational Physics* 77.2 (Aug. 1988), pp. 439–471. DOI: 10.1016/0021-9991(88)90177-5.

D. Sijacki and V. Springel. "Physical viscosity in smoothed particle hydrodynamics simulations of galaxy clusters." In: *Mon. Not. R. Astron. Soc.* 371.3 (Sept. 2006), pp. 1025–1046. DOI: 10.1111/j.1365-2966.2006.10752.x.

D. Sijacki et al. "A unified model for AGN feedback in cosmological simulations of structure formation." In: *Mon. Not. R. Astron. Soc.* 380.3 (Sept. 2007), pp. 877–900. DOI: 10.1111/j.1365-2966.2007.12153.x.

J. D. Simon and M. Geha. "The Kinematics of the Ultra-faint Milky Way Satellites: Solving the Missing Satellite Problem." In: *Astrophys. J.* 670.1 (Nov. 2007), pp. 313–331. DOI: 10.1086/521816.

M. Simonte et al. "Exploring the relation between turbulent velocity and density fluctuations in the stratified intracluster medium." In: *Astron. Astrophys.* 658, A149 (Feb. 2022), A149. DOI: 10.1051/0004-6361/202141703.

S. W. Skillman et al. "Cosmological Shocks in Adaptive Mesh Refinement Simulations and the Acceleration of Cosmic Rays." In: *Astrophys. J.* 689.2 (Dec. 2008), pp. 1063–1077. DOI: 10.1086/592496.

B. D. Smith et al. "GRACKLE: a chemistry and cooling library for astrophysics." In: *Mon. Not. R. Astron. Soc.* 466.2 (Apr. 2017), pp. 2217–2234. DOI: 10.1093/mnras/stw3291.

G. F. Smoot et al. "Structure in the COBE Differential Microwave Radiometer First-Year Maps." In: *Astrophys. J. Lett.* 396 (Sept. 1992), p. L1. DOI: 10.1086/186504.

D. N. Spergel et al. "First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters." In: *Astrophys. J. Suppl. Ser.* 148.1 (Sept. 2003), pp. 175–194. DOI: 10.1086/377226.

E. A. Spiegel. *The Theory of Turbulence*. Vol. 810. 2011. DOI: 10.1007/978-94-007-0117-5.

R. J. Splinter et al. "Fundamental Discreteness Limitations of Cosmological N-Body Clustering Simulations." In: *Astrophys. J.* 497.1 (Apr. 1998), pp. 38–61. DOI: 10.1086/305450.

V. Springel et al. "The Aquarius Project: the subhaloes of galactic haloes." In: *Mon. Not. R. Astron. Soc.* 391.4 (Dec. 2008), pp. 1685–1711. DOI: 10.1111/j.1365-2966.2008.14066.x.

V. Springel. "E pur si muove: Galilean-invariant cosmological hydrodynamical simulations on a moving mesh." In: *Mon. Not. R. Astron. Soc.* 401.2 (Jan. 2010), pp. 791–851. DOI: 10.1111/j.1365-2966.2009.15715.x.

V. Springel. "The cosmological simulation code GADGET-2." In: *Mon. Not. R. Astron. Soc.* 364.4 (Dec. 2005), pp. 1105–1134. DOI: 10.1111/j.1365-2966.2005.09655.x.

V. Springel, T. Di Matteo, and L. Hernquist. "Modelling feedback from stars and black holes in galaxy mergers." In: *Mon. Not. R. Astron. Soc.* 361.3 (Aug. 2005), pp. 776–794. DOI: 10.1111/j.1365-2966.2005.09238.x.

V. Springel and L. Hernquist. "Cosmological smoothed particle hydrodynamics simulations: the entropy equation." In: *Mon. Not. R. Astron. Soc.* 333.3 (July 2002), pp. 649–664. DOI: 10.1046/j.1365-8711.2002.05445.x.

V. Springel and L. Hernquist. "The history of star formation in a Λ cold dark matter universe." In: *Mon. Not. R. Astron. Soc.* 339.2 (Feb. 2003), pp. 312–334. DOI: 10.1046/j.1365-8711.2003.06207.x.

V. Springel et al. "Populating a cluster of galaxies - I. Results at z=0." In: *Mon. Not. R. Astron. Soc.* 328.3 (Dec. 2001), pp. 726–750. DOI: 10.1046/j.1365-8711.2001.04912.x.

U. P. Steinwandel et al. "Towards cosmological simulations of the magnetized intracluster medium with resolved Coulomb collision scale." In: *arXiv e-prints* (June 2023).

G. S. Stinson et al. "Making Galaxies In a Cosmological Context: the need for early stellar feedback." In: *Mon. Not. R. Astron. Soc.* 428.1 (Jan. 2013), pp. 129–140. DOI: 10.1093/mnras/sts028.

G. Stinson et al. "Star formation and feedback in smoothed particle hydrodynamic simulations - I. Isolated galaxies." In: *Mon. Not. R. Astron. Soc.* 373.3 (Dec. 2006), pp. 1074–1090. DOI: 10.1111/j.1365-2966.2006.11097.x.

K. Subramanian, A. Shukurov, and N. E. L. Haugen. "Evolving turbulence and magnetic fields in galaxy clusters." In: *Mon. Not. R. Astron. Soc.* 366.4 (Mar. 2006), pp. 1437–1454. DOI: 10.1111/j.1365-2966.2006.09918.x.

R. A. Sunyaev and Y. B. Zeldovich. "Small-Scale Fluctuations of Relic Radiation." In: *Astrophys. Space Sci.* 7.1 (Apr. 1970), pp. 3–19. DOI: 10.1007/BF00653471.

R. S. Sutherland and M. A. Dopita. "Cooling Functions for Low-Density Astrophysical Plasmas." In: *Astrophys. J. Suppl. Ser.* 88 (Sept. 1993), p. 253. DOI: 10.1086/191823.

H. Tanimura, S. Zaroubi, and N. Aghanim. "Direct detection of the kinetic Sunyaev-Zel'dovich effect in galaxy clusters." In: *Astron. Astrophys.* 645, A112 (Jan. 2021), A112. DOI: 10.1051/0004-6361/202038846.

R. Teyssier. "Cosmological hydrodynamics with adaptive mesh refinement. A new high resolution code called RAMSES." In: *Astron. Astrophys.* 385 (Apr. 2002), pp. 337–364. DOI: 10.1051/0004-6361:20011817.

R. Teyssier et al. "Mass distribution in galaxy clusters: the role of Active Galactic Nuclei feedback." In: *Mon. Not. R. Astron. Soc.* 414.1 (June 2011), pp. 195–208. DOI: 10.1111/j.1365-2966.2011.18399.x.

T. Theuns et al. "P^3M-SPH simulations of the Lyα forest." In: *Mon. Not. R. Astron. Soc.* 301.2 (Dec. 1998), pp. 478–502. DOI: 10.1046/j.1365-8711.1998.02040.x.

R. C. Tolman. "The Present Status of Cosmology. II." In: *The Scientific Monthly* 44.1 (Jan. 1937), pp. 20–40.

K. Tomida and J. M. Stone. "The Athena++ Adaptive Mesh Refinement Framework: Multigrid Solvers for Self-gravity." In: *Astrophys. J. Suppl. Ser.* 266.1, 7 (May 2023), p. 7. DOI: 10.3847/1538-4365/acc2c0.

G. Tormen, F. R. Bouchet, and S. D. M. White. "The structure and dynamical evolution of dark matter haloes." In: *Mon. Not. R. Astron. Soc.* 286.4 (Apr. 1997), pp. 865–884. DOI: 10.1093/mnras/286.4.865.

L. Tornatore et al. "Chemical enrichment of galaxy clusters from hydrodynamical simulations." In: *Mon. Not. R. Astron. Soc.* 382.3 (Dec. 2007), pp. 1050–1072. DOI: 10.1111/j.1365-2966.2007.12070.x.

E. F. Toro, M. Spruce, and W. Speares. "Restoration of the contact surface in the HLL-Riemann solver." In: *Shock Waves* 4.1 (July 1994), pp. 25–34. DOI: 10.1007/BF01414629.

E. Toro. *Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction*. Springer Berlin Heidelberg, 2009. ISBN: 9783540498346. URL: https://books.google.es/books?id=SqEjX0um8o0C.

S. Tulin and H.-B. Yu. "Dark matter self-interactions and small scale structure." In: *Phys. Rep.* 730 (Feb. 2018), pp. 1–57. DOI: 10.1016/j.physrep.2017.11.004.

M. J. Turk et al. "yt: A Multi-code Analysis Toolkit for Astrophysical Simulation Data." In: *The Astrophysical Journal Supplement Series* 192, 9 (Jan. 2011), p. 9. DOI: 10.1088/0067-0049/192/1/9.

J. A. Tyson, F. Valdes, and R. A. Wenk. "Detection of Systematic Gravitational Lens Galaxy Image Alignments: Mapping Dark Matter in Galaxy Clusters." In: *Astrophys. J. Lett.* 349 (Jan. 1990), p. L1. DOI: 10.1086/185636.

K. Umetsu. "Cluster-galaxy weak lensing." In: *Astron. Astrophys. Rev.* 28.1, 7 (Dec. 2020), p. 7. DOI: 10.1007/s00159-020-00129-w.

V. Vacca et al. "The intracluster magnetic field power spectrum in Abell 665." In: *Astron. Astrophys.* 514, A71 (May 2010), A71. DOI: 10.1051/0004-6361/200913060.

D. Vallés-Pérez, S. Planelles, and V. Quilis. "On the accretion history of galaxy clusters: temporal and spatial distribution." In: *Mon. Not. R. Astron. Soc.* 499.2 (Dec. 2020), pp. 2303–2318. DOI: 10.1093/mnras/staa3035.

D. Vallés-Pérez, S. Planelles, and V. Quilis. "The halo-finding problem revisited: a deep revision of the ASOHF code." In: *Astron. Astrophys.* 664, A42 (Aug. 2022), A42. DOI: 10.1051/0004-6361/202243712.

D. Vallés-Pérez, S. Planelles, and V. Quilis. "Troubled cosmic flows: turbulence, enstrophy, and helicity from the assembly history of the intracluster medium." In: *Mon. Not. R. Astron. Soc.* 504.1 (June 2021), pp. 510–527. DOI: 10.1093/mnras/stab880.

D. Vallés-Pérez, S. Planelles, and V. Quilis. "Unravelling cosmic velocity flows: a Helmholtz-Hodge decomposition algorithm for cosmological simulations." In: *Computer Physics Communications* 263, 107892 (June 2021). DOI: 10.1016/j.cpc.2021.107892.

D. Vallés-Pérez, V. Quilis, and S. Planelles. "Cosmic accretion shocks as a tool to measure the dark matter mass of galaxy clusters." In: *Nat. Astron.* (2024, in press). DOI: 10.1038/s41550-024-02303-x.

D. Vallés-Pérez, V. Quilis, and S. Planelles. "Void Replenishment: How Voids Accrete Matter Over Cosmic History." In: *Astrophys. J. Lett.* 920.1, L2 (Oct. 2021), p. L2. DOI: 10.3847/2041-8213/ac2816.

D. Vallés-Pérez et al. "On the choice of the most suitable indicator for the assembly state of dark matter haloes through cosmic time." In: *Mon. Not. R. Astron. Soc.* 519.4 (Mar. 2023), pp. 6111–6125. DOI: 10.1093/mnras/stad059.

D. Vallés-Pérez et al. "VORTEX-P: A Helmholtz-Hodge+Reynolds decomposition algorithm for particle-based simulations." In: *Computer Physics Communications* (2024, in press).

R. van de Weygaert and E. Platen. "Cosmic Voids: Structure, Dynamics and Galaxies." In: *International Journal of Modern Physics Conference Series*. Vol. 1. International Journal of Modern Physics Conference Series. Jan. 2011, pp. 41–66. DOI: 10.1142/S2010194511000092.

F. Vazza, E. Roediger, and M. Brüggen. "Turbulence in the ICM from mergers, cool-core sloshing, and jets: results from a new multi-scale filtering approach." In: *Astron. Astrophys.* 544, A103 (Aug. 2012), A103. DOI: 10.1051/0004-6361/201118688.

F. Vazza et al. "Electron and proton acceleration efficiency by merger shocks in galaxy clusters." In: *Mon. Not. R. Astron. Soc.* 451.2 (Aug. 2015), pp. 2198–2211. DOI: 10.1093/mnras/stv1072.

F. Vazza et al. "Massive and refined: A sample of large galaxy clusters simulated at high resolution. I: Thermal gas and properties of shock waves." In: *New Astron.* 15.8 (Nov. 2010), pp. 695–711. DOI: 10.1016/j.newast.2010.05.003.

F. Vazza et al. "Massive and refined. II. The statistical properties of turbulent motions in massive galaxy clusters with high spatial resolution." In: *Astron. Astrophys.* 529, A17 (May 2011), A17. DOI: 10.1051/0004-6361/201016015.

F. Vazza et al. "Resolved magnetic dynamo action in the simulated intracluster medium." In: *Mon. Not. R. Astron. Soc.* 474.2 (Feb. 2018), pp. 1672–1687. DOI: 10.1093/mnras/stx2830.

F. Vazza et al. "The turbulent pressure support in galaxy clusters revisited." In: *Mon. Not. R. Astron. Soc.* 481.1 (Nov. 2018), pp. L120–L124. DOI: 10.1093/mnrasl/sly172.

F. Vazza et al. "Turbulence and vorticity in Galaxy clusters generated by structure formation." In: *Mon. Not. R. Astron. Soc.* 464.1 (Jan. 2017), pp. 210–230. DOI: 10.1093/mnras/stw2351.

F. Vazza et al. "Turbulent motions and shocks waves in galaxy clusters simulated with adaptive mesh refinement." In: *Astron. Astrophys.* 504.1 (Sept. 2009), pp. 33–43. DOI: 10.1051/0004-6361/200912535.

F. Vazza et al. "Simulating the Enrichment of Fossil Radio Electrons by Multiple Radio Galaxies." In: *Galaxies* 11.2 (Mar. 2023), p. 45. DOI: 10.3390/galaxies11020045.

T. Vernstrom et al. "Discovery of magnetic fields along stacked cosmic filaments as revealed by radio and X-ray emission." In: *Mon. Not. R. Astron. Soc.* 505.3 (Aug. 2021), pp. 4178–4196. DOI: 10.1093/mnras/stab1301.

F. Villaescusa-Navarro et al. "Cosmology with One Galaxy?" In: *Astrophys. J.* 929.2, 132 (Apr. 2022), p. 132. DOI: 10.3847/1538-4357/ac5d3f.

F. Villaescusa-Navarro et al. "The CAMELS Project: Cosmology and Astrophysics with Machine-learning Simulations." In: *Astrophys. J.* 915.1, 71 (July 2021), p. 71. DOI: 10.3847/1538-4357/abf7ba.

F. Villaescusa-Navarro et al. "The CAMELS Project: Public Data Release." In: *Astrophys. J. Suppl. Ser.* 265.2, 54 (Apr. 2023), p. 54. DOI: 10.3847/1538-4365/acbf47.

M. Vogelsberger et al. “A model for cosmological simulations of galaxy formation physics.” In: *Mon. Not. R. Astron. Soc.* 436.4 (Dec. 2013), pp. 3031–3067. DOI: 10.1093/mnras/stt1789.

M. Vogelsberger et al. “Cosmological simulations of galaxy formation.” In: *Nature Reviews Physics* 2.1 (Jan. 2020), pp. 42–66. DOI: 10.1038/s42254-019-0127-2.

M. Vogelsberger et al. “Dust in and around galaxies: dust in cluster environments and its impact on gas cooling.” In: *Mon. Not. R. Astron. Soc.* 487.4 (Aug. 2019), pp. 4870–4883. DOI: 10.1093/mnras/stz1644.

C. Vogt and T. A. Enßlin. “A Bayesian view on Faraday rotation maps Seeing the magnetic power spectra in galaxy clusters.” In: *Astron. Astrophys.* 434.1 (Apr. 2005), pp. 67–76. DOI: 10.1051/0004-6361:20041839.

G. M. Voit and G. L. Bryan. “Regulation of the X-ray luminosity of clusters of galaxies by cooling and supernova feedback.” In: *Nature* 414.6862 (Nov. 2001), pp. 425–427. DOI: 10.1038/35106523.

G. M. Voit, S. T. Kay, and G. L. Bryan. “The baseline intracluster entropy profile from gravitational structure formation.” In: *Mon. Not. R. Astron. Soc.* 364.3 (Dec. 2005), pp. 909–916. DOI: 10.1111/j.1365-2966.2005.09621.x.

S. von Hoerner. “Die numerische Integration des n-Körper-Problemes für Sternhaufen. I.” In: *Zeitschrift fuer Astrophysik* 50 (Jan. 1960), pp. 184–214.

M. G. Walker and J. Peñarrubia. “A Method for Measuring (Slopes of) the Mass Profiles of Dwarf Spheroidal Galaxies.” In: *Astrophys. J.* 742.1, 20 (Nov. 2011), p. 20. DOI: 10.1088/0004-637X/742/1/20.

S. Walker and E. Lau. “Cluster Outskirts and Their Connection to the Cosmic Web.” In: *Handbook of X-ray and Gamma-ray Astrophysics*. 2022, 13, p. 13. DOI: 10.1007/978-981-16-4544-0_120-1.

S. Walker et al. “The Physics of Galaxy Cluster Outskirts.” In: *Space Sci. Rev.* 215.1, 7 (Jan. 2019), p. 7. DOI: 10.1007/s11214-018-0572-8.

E. X. Wang et al. “The impact of black hole seeding in cosmological simulations.” In: *Mon. Not. R. Astron. Soc.* 483.4 (Mar. 2019), pp. 4640–4648. DOI: 10.1093/mnras/sty3491.

H. Wendland. “Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree.” In: *Advances in Computational Mathematics* 4.1 (Dec. 1995), pp. 389–396. DOI: 10.1007/BF02123482.

M. J. West and G. D. Bothun. “A Reanalysis of Substructure in Clusters of Galaxies and Their Surroundings.” In: *Astrophys. J.* 350 (Feb. 1990), p. 36. DOI: 10.1086/168360.

R. Weymann. “The Energy Spectrum of Radiation in the Expanding Universe.” In: *Astrophys. J.* 145 (Aug. 1966), p. 560. DOI: 10.1086/148795.

S. D. M. White. “Dynamical friction in spherical clusters.” In: *Mon. Not. R. Astron. Soc.* 174 (Jan. 1976), pp. 19–28. DOI: 10.1093/mnras/174.1.19.

S. D. M. White and M. J. Rees. “Core condensation in heavy halos: a two-stage theory for galaxy formation and clustering.” In: *Mon. Not. R. Astron. Soc.* 183 (May 1978), pp. 341–358. DOI: 10.1093/mnras/183.3.341.

L. M. Widrow and N. Kaiser. “Using the Schroedinger Equation to Simulate Collisionless Matter.” In: *Astrophys. J. Lett.* 416 (Oct. 1993), p. L71. DOI: 10.1086/187073.

R. P. C. Wiersma, J. Schaye, and B. D. Smith. “The effect of photoionization on the cooling rates of enriched, astrophysical plasmas.” In: *Mon. Not. R. Astron. Soc.* 393.1 (Feb. 2009), pp. 99–107. DOI: 10.1111/j.1365-2966.2008.14191.x.

M. A. G. Willson. “Radio observations of the cluster of galaxies in Coma Berenices - the 5C4 survey.” In: *Mon. Not. R. Astron. Soc.* 151 (Jan. 1970), p. 1. DOI: 10.1093/mnras/151.1.1.

H. A. Winther. *Cosmology II*. 2021. URL: https://cmb.wintherscoming.no/index.php.

D. Wittor et al. “Evolution of vorticity and enstrophy in the intracluster medium.” In: *Mon. Not. R. Astron. Soc.* 471.3 (Nov. 2017), pp. 3212–3225. DOI: 10.1093/mnras/stx1769.

G. Xu. “A New Parallel N-Body Gravity Solver: TPM.” In: *Astrophys. J. Suppl. Ser.* 98 (May 1995), p. 355. DOI: 10.1086/192166.

J. K. Yadav, J. S. Bagla, and N. Khandai. "Fractal dimension as a measure of the scale of homogeneity." In: *Mon. Not. R. Astron. Soc.* 405.3 (July 2010), pp. 2009–2015. DOI: 10.1111/j.1365-2966.2010.16612.x.

D. G. York et al. "The Sloan Digital Sky Survey: Technical Summary." In: *Astron. J.* 120.3 (Sept. 2000), pp. 1579–1587. DOI: 10.1086/301513.

K. Yoshikawa, N. Yoshida, and M. Umemura. "Direct Integration of the Collisionless Boltzmann Equation in Six-dimensional Phase Space: Self-gravitating Systems." In: *Astrophys. J.* 762.2, 116 (Jan. 2013), p. 116. DOI: 10.1088/0004-637X/762/2/116.

J. Zavala and C. S. Frenk. "Dark Matter Haloes and Subhaloes." In: *Galaxies* 7.4 (Sept. 2019), p. 81. DOI: 10.3390/galaxies7040081.

Y. B. Zeldovich. "Gravitational instability: An approximate theory for large density perturbations." In: *Astron. Astrophys.* 5 (Mar. 1970), pp. 84–89.

Q. Zhu, B. Smith, and L. Hernquist. "Gas cooling in hydrodynamic simulations with an exact time integration scheme." In: *Mon. Not. R. Astron. Soc.* 470.1 (Sept. 2017), pp. 1017–1025. DOI: 10.1093/mnras/stx1346.

I. Zhuravleva et al. "The Relation between Gas Density and Velocity Power Spectra in Galaxy Clusters: Qualitative Treatment and Cosmological Simulations." In: *Astrophys. J. Lett.* 788.1, L13 (June 2014), p. L13. DOI: 10.1088/2041-8205/788/1/L13.

J. A. ZuHone et al. "The Effect of Anisotropic Viscosity on Cold Fronts in Galaxy Clusters." In: *Astrophys. J.* 798.2, 90 (Jan. 2015), p. 90. DOI: 10.1088/0004-637X/798/2/90.

J. A. ZuHone et al. "What Do the Hitomi Observations Tell Us About the Turbulent Velocities in the Perseus Cluster? Probing the Velocity Field with Mock Observations." In: *Astrophys. J.* 853.2, 180 (Feb. 2018), p. 180. DOI: 10.3847/1538-4357/aaa4b3.

F. Zwicky. "Die Rotverschiebung von extragalaktischen Nebeln." In: *Helvetica Physica Acta* 6 (Jan. 1933), pp. 110–127.

F. Zwicky. "Nebulae as Gravitational Lenses." In: *Physical Review* 51.4 (Feb. 1937), pp. 290–290. DOI: 10.1103/PhysRev.51.290.

This PhD Thesis has been funded by an ***Atracció de Talent*** fellowship from the **Vicerrectorat d'Investigació** of **Universitat de València**.

Our research is additionally supported by the Spanish **Agencia Estatal de Investigación**, by the Spanish **Ministerio de Ciencia e Innovación**, and by the **Generalitat Valenciana**.

Referees: **Veronica Biffi** (INAF-OATS), **Elena Rasia** (University of Michigan & INAF-OATS), **Patricia Sánchez-Blázquez** (UCM).

Committee: **Stefano Borgani** (UniTS), **Franco Vazza** (UniBo), **José María Martí** (UV; chair).