Lossy Network Correlated Data Gathering with High-Resolution Coding

Răzvan Cristescu
Caltech, 136-93 Pasadena, CA 91125, USA
razvanc@caltech.edu

Baltasar Beferull-Lozano
EPFL, Lausanne CH-1015, Switzerland
Baltasar.Beferull@epfl.ch

***Abstract*— We consider a sensor network measuring correlated data, where the task is to gather all data from the network nodes to a sink. We consider the case where data at nodes is lossy coded with high-resolution, and the information measured by the nodes should be available at the sink within certain total and individual distortion bounds. First, we consider the problem of finding the optimal transmission structure and the rate-distortion allocations at the various spatially located nodes, such as to minimize the total power consumption cost of the network. We prove that the optimal transmission structure is the shortest path tree and that the problems of rate and distortion allocation separate in the high-resolution case, namely, we first find the distortion allocation as a function of the transmission structure, and then the rate allocation is computed. Then, we also study the case when the node positions can be chosen, by finding the optimal node placement when two different targets of interest are considered, namely total power minimization and network lifetime extension.**

I. Introduction

A. Lossy Correlated Data Gathering

Consider a network of sensors that measures a correlated data field. An important task in such a scenario is data gathering, where the goal is to gather data from all the nodes to the sink via a subset of the links of the network. An example is shown in Fig. 1, where there are N nodes with sources $X_1, \ldots, X_N$, a sink denoted by S, and a representation graph of connectivity with edges connecting certain nodes. The transmission topology in our model is assumed to be an undirected connected graph with nodes connected within a certain distance with point-to-point links. Data at nodes have to be encoded, and for proper reconstruction of the field, they need to be available at the sink within certain prescribed total distortion (in Fig. 1, we denote the distortion and rate at node i by D_i and R_i, respectively). In specific scenarios where a minimum accuracy is necessary for measurements across network nodes, there are also individual distortion upper bounds imposed on the accuracy of data transmitted from the nodes.

In addition to the encoding of the data by the nodes, these data also need to be transmitted over the network from the sources to the sinks, which results in certain transmission costs (in Fig. 1, the total weight of the path from node i to the sink is denoted by c_i). It is thus important to study the interplay between the rate and distortion allocation at nodes, and the transmission structure used to transport the data. We consider a joint treatment of the rate allocation and the chosen transmission structure, by means of cost functions that are functions of both of them. The cost functions usually found in practice separate the rate term from the path weight term [3]. For instance, the [rate] $\times$ [path weight] cost function measures the power consumption in sensor networks, where [rate] is the number of transmitted bits, and [path weight] is a supraunitary power κ (usually $\kappa = 2$ or $\kappa = 4$) of the Euclidean distance over which the transmission is done. Typical goals in such settings are (a) the minimization of the total network power consumption, defined as the total sum of energies consumed by nodes, and (b) maximization of the network lifetime, defined as the time until the first node fails. In this work, we consider jointly the optimization of both rate and distortion allocation at nodes, transmission structure, and node placement, in the context of sensor networks with correlated data at nodes.

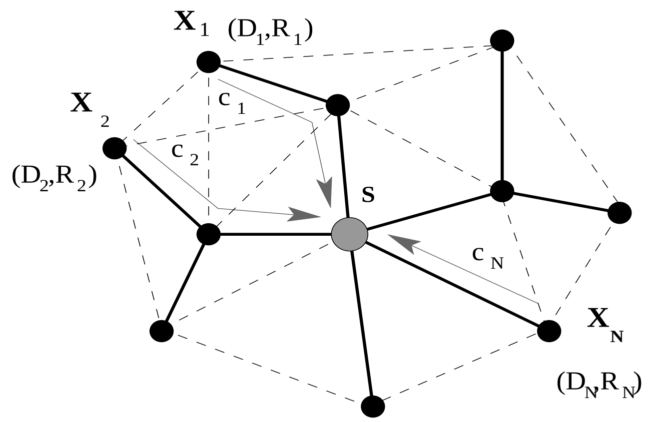


Fig. 1. In this example, data from nodes $X_1, X_2, \ldots, X_N$ need to arrive at sink S. A rate supply R_i is allocated to each node X_i, and the distortion at that node is D_i. In thick solid lines, a chosen tree transmission structure is shown. In thin dashed lines, the other possible links are shown. The weight of the path from node i to the sink is c_i, and the corresponding path is shown in gray line.

Consider a network of N nodes (see Fig. 1). Let $\mathbf{X} = (X_1, \ldots, X_N)$ be the vector formed by the random variables representing the sources measured at nodes $1, \ldots, N$. The samples taken at nodes are spatially correlated. We assume that the random variables are continuous and that there is a high-resolution quantizer in each sensor. A rate/distortion allocation $\{(R_1, D_1), \ldots, (R_N, D_N)\}$ (bits) has to be assigned at the nodes so that the quantized measured information samples are described with certain total D and potentially individual $D_i^{max}, i = 1, \ldots, N$ distortions. That information has to be transmitted through the links of the network to the designated base station (see Fig. 1). In [6], the problem of lossless coding of the data at nodes with Slepian-Wolf coding [13] has been studied. In this work, we extend that problem by allowing nodes to code their data in a *lossy* manner, such that the information received by the sink describes the data at nodes within the prescribed distortion constraints. Namely, the rate allocation at nodes in our lossy data gathering scenario essentially depends on the differential entropy of the non-quantized data that is measured and on the distortion levels for quantization that are assigned to each of the sources located at nodes.

Moreover, we study the problem of optimal node placement for the task of minimizing the total network power consumption and network lifetime. Note that the placement solutions resulting from the corresponding optimizations are different, since optimizing the lifetime of the network means equalizing power consumption at nodes, which does not necessarily coincide with the optimal placement solution when the sum of power consumptions is considered. Interestingly, when total network power consumption is minimized, the energy bottleneck node in the optimal positioning is the one at the *extremity* of the network, rather than the node closest to the sink. This is both due to the coding strategy, that assigns low rates to nodes far from the sink, and to the inter-node distance optimization.

The general distributed rate-distortion region for coding correlated data at general quantization resolution levels (e.g. low rates) is not known even for the Gaussian case [11]. Thus, we will restrict this study to the case of lossy *high-resolution* data gathering [1], [14], for which the region is known. However, the results we obtain for the data gathering scenario in the high-resolution case provide a lower bound on the performance of *any* rate-distortion allocation strategy, in terms of power efficiency, for a network that performs data gathering of correlated data. Moreover, our results on the problem of joint rate allocation and transmission structure optimization can be generalized should good bounds for low resolution rate-distortion curves become available.

B. Related Work

The rate-distortion region of coding with high-resolution for arbitrarily correlated sources has been found in [14]. In general (e.g. low resolution) this region provides an outer bound which becomes tighter as the resolution increases. A setting related to the problem we consider in this paper, namely optimizing correlated data gathering with communication costs for the case of lossless coding of finely quantized data with the same scalar quantizer at each node, has been studied in [6], where Slepian-Wolf [4], [13] bounds are used, and in [7], where a simple coding technique by explicit communication is considered; in this work, while assuming high-rate quantization, lattice quantizers are used where the resolutions of the quantizers can be adapted across nodes [14]. A similar problem of data compression by opportunistic aggregation was considered in [10]. In this work we consider *lossy* coding of data at nodes, by providing the optimal allocation within the rate-distortion region given in [14], that minimizes the transport cost to perform the data gathering to a sink. Moreover, we consider the effect of this coding technique on the optimal placement of the nodes for gathering lossy coded data when both total network power efficiency and lifetime optimization are considered.

C. Main Contributions and Outline

We solve the power efficiency optimization problem by finding the transmission structure and rate-distortion allocation in a closed form, for high-resolution coding of correlated data. Our result also provides a general lower bound in terms of power efficiency for the cost of data gathering of lossy coded data in sensor networks. Further, when the positions of the nodes can be chosen, we determine the optimal node placing for optimizing two metrics of interest, namely total power consumption and network lifetime.

In Section II, we formulate the optimization problem we consider in this paper, and briefly introduce a family of Gaussian random fields that we use as an example of correlation fields. In Section III, we show that the optimization problem separates, and we sequentially find the optimal transmission structure, rate allocation, and distortion allocation under various settings with total and individual distortion constraints. Further, in Section IV, we find the optimal placement in an one-dimensional scenario for two optimization goals, namely network total power and the network lifetime, and we present numerical simulations to illustrate our results. We conclude with Section V.

II. Problem Setting

A. Optimization Problem

For the optimization of rate-distortion allocation, and transmission structure, we will use a family of cost functions that separate the rate and the total path coefficients. These cost functions are relevant for various problems in sensor networks related to power efficiency optimization [3].

The problem considered in this paper is a generalization of the scenario in [6], where the problem of data gathering with joint rate allocation and transmission structure optimization was considered in the case of *lossless* coding of data at nodes, under Slepian-Wolf constraints. Namely here we consider the case of *lossy* coding of the data at nodes. This implies an additional number of optimization variables, namely the values of the *distortions* $\{D_i\}_{i=1}^N$ at nodes, which increases importantly the complexity of the problem. Also, additional constraints are introduced by the rate-distortion region. In the case of high resolution coding ($D_i \rightarrow 0$, for any $i \in V$), that we consider in this paper, the rate-distortion region derived in [14] becomes tight. In general, this region provides an outer bound which becomes tighter as the rates become large. The expressions for the rate/distortion region for high-resolution data rates found in [14] are similar in form to the Slepian-Wolf region [13]. Also, as in the Slepian-Wolf scenario, an important feature of the rate-distortion region is that no communication is allowed among the nodes for the data coding, although of course communication is needed to transport the coded data. As a result, the problems of transmission structure optimization, rate and distortion allocation separate, as we will see further in Section III.

Assume a given node placement such as the one illustrated in Fig. 1. Denote R_i and D_i, $i = 1 \dots N$, the allocated rate and distortion at node i. The most general form of our optimization problem is given as follows:

$$\{R_i^*, D_i^*, ST^*\}_{i=1}^N = \arg \min_{\{R_i, D_i, ST\}_{i=1}^N} \sum_{i=1}^N c_i R_i \quad (1)$$

under constraints

$$\sum_{i \in \mathcal{X}} R_i \geq h(\mathcal{X}|V \backslash \mathcal{X}) - \log(2\pi e)^{|\mathcal{X}|} \Pi_{i \in \mathcal{X}} D_i, \quad \forall \mathcal{X} \subset V \quad (2)$$

$$\sum_{i=1}^N D_i \leq D \quad (3)$$

$$D_i \leq D_i^{\max}, i = 1 \dots N. \quad (4)$$

where ST is the spanning tree to be found, which defines the transmission structure; $c_i, i = 1 \dots N$ are the total weights of the path from node i to the sink on the spanning tree ST, thus $c_i = \sum_{e \in \mathcal{E}_i} w_e^\kappa$, where $e \in \mathcal{E}_i$ is the set of edges that links node i to the sink S on ST, and w_e the Euclidean distance of edge e. Constraints (2) express the rate-distortion region constraints given in [14], where the assumption of high-resolution is considered, namely that the sum of rates for any given subset of nodes has to be larger than the entropy of the random variables measured at those nodes, conditioned on the random variables measured at all other nodes. In constraint (3), D is the maximum total distortion and in (4), $D_i^{\max}, i = 1 \dots N$ are the maximum individual constraints.

The optimization problem (1) assumes a fixed placement of nodes. However, in practice, a well chosen placement of nodes (namely, allowing adjustable values of c_i, such that the measured field is fully covered by the nodes, under spatial distortion constraints) can drastically increase the total efficiency of the network [9]. In the general two-dimensional setting, the problem (1) is hard to formalize and solve when the node placement is introduced as a variable in the optimization. Thus, in Section III, where we study the general optimization for a two-dimensional network, we will consider that the node placement is fixed. However, in Section IV, we will particularize the optimization (1) for the one-dimensional case, for which we solve the placement problem too. Namely, we provide optimal node placements for minimization of (a) the total power consumed by the network, and (b) the lifetime of the network. The study of node placement for the two-dimensional case is a subject of our current research.

B. Example: Correlated Gaussian Random Field

For the sake of clarity, in our numerical simulations we assume a *jointly Gaussian model* for the spatial data $\mathbf{X}$ measured at nodes, with an N-dimensional multivariate normal distribution $\mathbf{X} \sim \mathcal{N}^N(\mu, \mathbf{K})$:

$$f(\mathbf{X}) = \frac{1}{\sqrt{2\pi}\det(\mathbf{K})^{1/2}} e^{-\left(\frac{1}{2}(\mathbf{X}-\mu)^T\mathbf{K}^{-1}(\mathbf{X}-\mu)\right)}$$

where $\mathbf{K}$ is the covariance matrix of $\mathbf{X}$, and μ the mean vector. The diagonal elements of $\mathbf{K}$ are the variances $K_{ii} = \sigma_i^2$. The rest of K_{ij} depend on the distance between the corresponding nodes (e.g. $K_{ij} = \sigma_{ii}\sigma_{ij}\exp(-ad_{i,j}^{\beta}), \beta \in \{1, 2\}$), where $d_{i,j}$ is the distance between nodes i and j [5], [8]. Without loss of generality, we will restrict our analysis to unit variance $\sigma_{ii} = 1, i = 1 \ldots N$ and zero-mean $\mu = 0$ processes $\mathbf{X} \sim \mathcal{N}^N(0, \mathbf{K})$.

Notice that although the numerical evaluations will be performed using the Gaussian random field model, our results are valid for any spatially correlated random processes, whose correlation decreases with the distance.

III. Optimization

A. The Optimal Communication Tree Structure is SPT

Constraints (2) imply that nodes can code (asymptotically at high resolution) with any rate that obeys the rate-distortion region *without* explicitly communicating data with each other. As a consequence of the fact that it is possible to perform joint source coding without nodes communicating among them, we can state the following Proposition:

Proposition 1 *– Separation of the joint optimization of source coding and transmission structure:*

The overall joint optimization (1) can be achieved by first optimizing the transmission structure with respect to only *the link weights c_i, and then optimizing the distortion and rate allocation for the given transmission structure under the remaining constraints.*

Proof: The joint cost function we consider is separable as the product of a function that depends only on the rate and another function that depends only on the link weights of the transmission structure. Once the rate allocation is *fixed*, the best way (least cost) to transport any amount of data from a given node i to the sink S does not depend on the value of the rate R_i. Since this holds for any rate allocation, it is also true for the minimizing rate allocation and the result follows.■ ■

Proposition 1 implies that the optimal transmission structure that optimizes (1) is the shortest path tree (ST^*=*SPT*). Denote by c_i^* the total weight of the path from node i to the sink S on the *SPT*. Suppose without loss of generality that nodes are ordered in a list with increasing values of the weights corresponding to the shortest paths from each node to the sink, that is, $c_1^* \le c_2^* \le \cdots \le c_N^*$. Then, by using Proposition 1, the optimization (1) becomes:

$$\begin{aligned} \{R_i^*, D_i^*\}_{i=1}^N &= \arg\min_{\{R_i, D_i\}_{i=1}^N} \sum_{i=1}^N c_i^* R_i \\ &\text{under constraints} \\ &(2), (3), (4). \end{aligned} \tag{5}$$

B. Optimal Rate Allocation

First, we show that, regardless of constraints (3), (4), the optimal rate allocation has to obey only constraints (2). Moreover, for any set of distortion values $\{D_i\}_{i=1}^N$, the rate allocation is given by:

Theorem 1 *Optimal rate allocation:*

$$\begin{aligned} R_1^* &= h(X_1) - \log 2\pi e D_1 \\ R_2^* &= h(X_2|X_1) - \log 2\pi e D_2 \\ &\cdots \\ R_N^* &= h(X_N|X_{N-1}, \ldots, X_1) - \log 2\pi e D_N. \end{aligned} \tag{6}$$

Proof: First, we prove that (6) is a feasible solution for (5). Consider any constraint from (2), for some subset $Y \subset V$. Denote by $M = |Y|$ the number of elements in Y. Order the indices of $X_i \in Y$ as $i_1, i_2, i_3, \ldots i_M$, with i_1 indexing the closest node and i_M the furthest node from the sink on the SPT.

If we rewrite the left-hand-side of (2) in terms of the solutions that we provide in the statement of the theorem, we have:

$$\begin{aligned} \sum_{i \in Y} R_i &= h(X_{i_M}|X_{i_M-1}, \ldots, X_1) - \log 2\pi e D_{i_M} + \\ &\quad h(X_{i_{M-1}}|X_{i_{M-1}-1}, \ldots, X_1) - \log 2\pi e D_{i_{M-1}} + \\ &\quad \cdots + h(X_{i_1}|X_{i_1-1}, \ldots, X_1) - \log(2\pi e) D_{i_1}. \end{aligned} \tag{7}$$

Expanding the right-hand-side terms with the chain law for conditional entropies, we obtain:

$$\begin{aligned} h(Y|Y^C) &= h(X_{i_M}|Y^C \cup \{Y - \{X_{i_M}\}\}) + \\ &\quad h(X_{i_{M-1}}|Y^C \cup \{Y - \{X_{i_M}, X_{i_{M-1}}\}\}) + \\ &\quad \cdots + h(X_{i_1}|Y^C \cup \{Y - \{X_{i_M}, \ldots, X_{i_1}\}\}) \\ &\quad - \log(2\pi e)^{|Y|} D_{i_1} D_{i_2} \ldots D_{i_M} \\ &= h(X_{i_M}|V - \{X_{i_M}\}) - \log 2\pi e D_{i_M} + \\ &\quad h(X_{i_{M-1}}|V - \{X_{i_M}, X_{i_{M-1}}\}) - \\ &\quad - \log 2\pi e D_{i_{M-1}} + \\ &\quad \cdots + h(X_{i_1}|V - \{X_{i_M}, X_{i_{M-1}}, \ldots, X_{i_1}\}) \\ &\quad - \log 2\pi e D_{i_1}. \end{aligned} \tag{8}$$

Consider the terms on the right-hand-side of (7) and (8). For any $i_k \in Y$, the term corresponding to X_{i_k} in (8) is at most equal to its counterpart in (7), because the set of nodes on which the entropy conditioning is done for each term in (7) is a subset of its counterpart in (8). Since the choice of Y was arbitrary, then any constraint in (2) is satisfied by the assignment (6). Note also that the rate allocation in (6) satisfies with equality the constraint on the total sum of rates:

$$\sum_{i \in V} R_i \ge h(X_1, \ldots, X_N) - \log(2\pi e)^N \Pi_{i=1}^N D_i. \tag{9}$$

We prove now by induction that the assignment in (6) makes the expression to be minimized in (5) smaller than any other valid assignment. Note first that we cannot allocate to X_N less than $h(X_N|X_{N-1}, X_{N-2}, \dots, X_1) - \log 2\pi e D_N$ bits, due to the particular constraint in (2) corresponding to $Y = \{X_N\}$. Assume now that a rate allocation solution with $h(X_N|X_{N-1}, X_{N-2}, \dots, X_1) - \log 2\pi e D_N$ bits to X_N is not optimal, and X_N is assigned $h(X_N|X_{N-1}, \dots, X_1) - \log 2\pi e D_N + b$ bits. Due to (9), at most b bits in total can be extracted from the rates assigned to some of the other nodes. But since c_N^* is the largest coefficient in the optimization problem (5), it is straightforward to see that any such change in rate allocation increases the cost function in (5). Thus assigning $R_N = h(X_N|X_{N-1}, \dots, X_1) - \log 2\pi e D_N$ bits to X_N is optimal.

Consider now the rate assigned to X_{N-1}. From the rate constraint corresponding to $Y = \{X_{N-1}, X_N\}$, it follows that:

$$\begin{aligned} R_N + R_{N-1} &\geq h(X_N, X_{N-1}|X_{N-2}, \dots, X_1) \\ &\quad - \log(2\pi e)^2 D_N D_{N-1} \\ &= h(X_N|X_{N-1}, X_{N-2}, \dots, X_1) - \\ &\quad - \log 2\pi e D_N + \\ &\quad h(X_{N-1}|X_{N-2}, \dots, X_1) - \\ &\quad - \log 2\pi e D_{N-1}. \end{aligned}$$

Since for optimality of the solution R_N must be given by $R_N = h(X_N|X_{N-1}, X_{N-2}, \dots, X_1) - \log 2\pi e D_N$, thus $R_{N-1} \geq h(X_{N-1}|X_{N-2}, \dots, X_1) - \log 2\pi e D_{N-1}$. Following a similar argument as for X_N, the optimal solution allocates $R_{N-1} = h(X_{N-1}|X_{N-2}, \dots, X_1) - \log 2\pi e D_{N-1}$. The rest of the proof follows similarly by considering successively the constraints corresponding to subsets
$Y = \{X_i, X_{i+1}, \dots, X_N\}$, with $i = N-2, N-3, \dots 1$. ▪ ■

C. Total Distortion Constraint

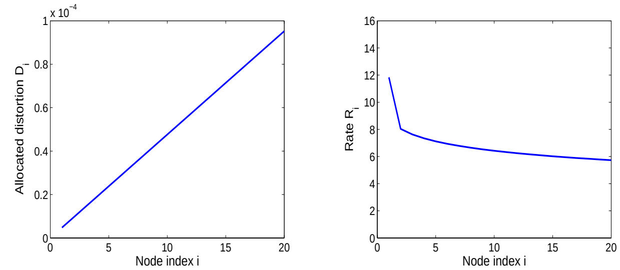
In this section, we consider optimization of (1) in the case when we assume the constraints (4) are not active. By Theorem 1, if the positions of nodes are fixed, $\{R_i^*\}_{i=1}^N$ *only* depends on $\{D_i\}_{i=1}^N$. Therefore, at this point, we can insert in (5) the values obtained for $\{R_i^*\}_{i=1}^N$, and thus obtain an optimization problem having as argument the set of distortions $\{D_i\}_{i=1}^N$ only:

$$\begin{aligned} \{D_i^*\}_{i=1}^N &= \arg \min_{\{D_i\}_{i=1}^N} \sum_{i=1}^N c_i^* \cdot \\ &\quad \cdot (h(X_i|X_{i-1}, \dots, X_1) - \log 2\pi e D_i) \\ \text{under constraint} & \qquad\qquad\qquad\qquad (10) \\ \sum_{i=1}^N D_i &\leq D. \end{aligned}$$

Note that the differential entropy terms in (10) do not depend on the distortions D_i. Thus, (10) can be equivalently written as:

$$\begin{aligned} \{D_i^*\}_{i=1}^N &= \arg \max_{\{D_i\}_{i=1}^N} \sum_{i=1}^N c_i^* \log 2\pi e D_i \\ \text{under constraint} & \qquad\qquad\qquad\qquad (11) \\ \sum_{i=1}^N D_i &\leq D. \end{aligned}$$

Denote $\sum_{i=1}^N c_i^* = C$. The solution of the optimization problem (11) is easily obtained, using Lagrange multipliers, giving a distribution of distortions:



(a) Distortion allocation as a function of node index. (b) Rate allocation as a function of node index.

Fig. 2. One-dimensional grid with constraint on the total distortion. Nodes are labelled with increasing indexes, with node 1 closest to the sink and node N farthest.

Fig. 3. One-dimensional grid network.

$$D_i^* = D \cdot \frac{c_i^*}{C}, i = 1 \dots N. \tag{12}$$

By combining (6) and (12), we obtain that the rate-distortion allocation at nodes is given by:

$$R_i^* = h(X_i|X_{i-1}, \dots, X_1) - \log\left(\frac{2\pi e D c_i^*}{C}\right), i = 1 \dots N. \tag{13}$$

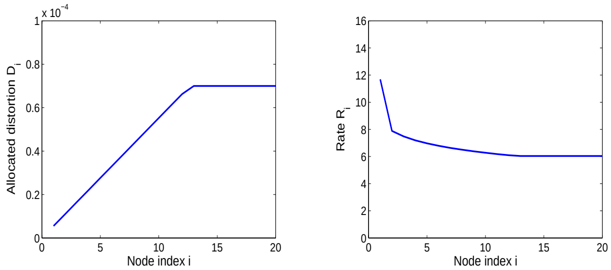
Moreover, in the case of correlated Gaussian random fields, (13) can be written as:

$$R_i^* = \log \frac{\det(\mathbf{K}(1, \dots i))}{\det(\mathbf{K}(1, \dots, i-1))} \frac{C}{c_i^* \cdot D}, \ i = 1 \dots N, \tag{14}$$

where $\mathbf{K}(1, \dots, i)$ is the correlation matrix corresponding to nodes $1, \dots, i$. In Fig. 2 we illustrate the distortion and rate allocations provided by (13) for the one-dimensional grid with uniform internode distances in Fig. 3 measuring a correlated Gaussian random field with $\beta = 1$ (which corresponds to a Gauss-Markov process). The same analysis holds for arbitrary two-dimensional networks. For clarity in our numerical experiments, we keep the values of the model parameters constant along this paper: $N = 20$, $D = 10^{-3}$, $D_i^{\max} = 0.7 \cdot 10^{-3}$, $\alpha = 10^{-3}$, $\kappa = 2$, $\beta = 1$. We have observed similar results are obtained for many other network parameter settings. For instance, for $\beta = 2$, the rate decay in Fig. 2(b) is less abrupt, due to the smooth decay of differential entropy in (13) with the number of nodes on which conditioning is done.

These results are intuitively expected: at the extremities of the network (where the cost c_i^* of the path from node i to the sink is large), small rates are allocated, meaning large allocated distortions and conditioning done on many nodes that are closer to the sink on the *SPT*; on the contrary, large rates, meaning small distortions and conditioning on fewer nodes closer to the sink on the *SPT*, are allocated at nodes near the sink (where the cost c_i^* is small). For the nodes in-between there is a monotonically increasing set of rates as we move towards the sink.

Also note that for $\beta = 1$, the resulting process is Gauss-Markov, which means that if the nodes are equally-spaced then the differential



(a) Distortion allocation as function of node index.
(b) Rate allocation as a function of node index.

Fig. 4. One-dimensional grid with constraints on the total *and* individual distortions. Nodes are labelled with increasing indexes, with node 1 closest to the sink and node N farthest.

entropies are equal in (13), and thus the decrease in rate in Fig. 2(b) is due only to the increase in allocated distortion. For a process with $\beta = 2$, the result is a smoother decay of the rate when the node index increases, as a function of both the decrease in differential entropy and the increase in allocated distortion in (13).

D. Individual Distortion Constraints

Let us now additionally introduce the individual constraints (4). For the sake of clarity, assume first that all individual constraints are equal, namely $D_i^{\max} = T$. Using again Lagrange optimization, it is simple to show that the solution in this case can be obtained as follows:

- Consider the solution given by (12) without introducing individual distortion constraints.
- Let $m+1$ be the smallest node index for which $D_i \geq T, i = m+1 \dots N$. Allocate $D_i^* = T, i = m+1, \dots, N$, that is, we make the last $N-m$ constraints alive. To find the rest of the distortion values $D_i, i = 1, \dots, m$, we solve:

$$\begin{aligned} \{D_i^*\}_{i=1}^m &= \arg \max_{\{D_i\}_{i=1}^m} \sum_{i=1}^m c_i^* \log 2\pi e D_i \\ \text{under constraint} & \\ \sum_{i=1}^m D_i &\leq D - (N-m)T. \end{aligned} \tag{15}$$

  which is exactly exactly the same type of problem as in (11). Thus, the complete solution is readily given by:

$$\begin{aligned} D_i &= T, i = m+1 \dots N \\ D_i &= c_i^* \frac{D - (N-m)T}{C_m} \end{aligned}$$

  where $C_m = \sum_{i=1}^m c_i^*$. Fig. 4 shows the distortion and rate allocations that are obtained in this case.

In Fig. 4(b) we see that when individual distortions are considered, the rate allocation at the nodes at the extremity of the network is equalized (as compared to Fig. 2(b)). As a result of this rate load added to nodes at the extremity of the network, the rates at the nodes closer to the sink can be correspondingly decreased.

In the case where the individual constraints are different, a similar procedure can be used, that is a subset of the constraints will be made active and then again a problem similar to (11) can be easily solved.

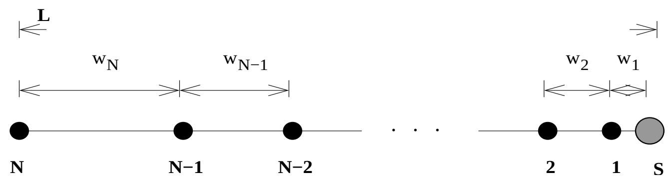


Fig. 5. Optimized placement for the one-dimensional network.

IV. Placement Optimization for the One-Dimensional Network

A. Total Power Minimization

In some scenarios, the positions of the nodes are not fixed in advance, but it is possible to place the nodes optimally so as to minimize various resources [9]. Since the study in this paper is concerned with power efficient scenarios, one possible task to be achieved when the node placement can be chosen is the total power efficiency. Namely, we consider the optimization (1) where the positions of the nodes are variable and the optimization is done additionally over the weights of the paths from the nodes to the sink:

$$\begin{aligned} \{R_i^*, D_i^*, c_i^*\}_{i=1}^N &= \arg \min_{\{R_i, D_i, c_i\}_{i=1}^N} \sum_{i=1}^N c_i R_i \\ \text{under constraints} & \quad (2), (3), (4), \end{aligned} \tag{16}$$

with additional constraints on $\max\{c_i\}_{i=1}^N$ (coverage constraint) and $\max\{c_i - c_{i-1}\}_{i=2}^N$ (inter-node space constraint).

For the sake of simplicity, in this section we consider a one-dimensional scenario (see Fig. 3). A similar analysis can be performed in the two-dimensional case, since the rate and distortion allocations only depend on the ordering of the nodes on the computed SPT; however, in the two-dimensional case, an additional model describing how nodes can be moved from their initial positions is required (e.g. on two-dimensional grids), which makes the problem more complicated.

Finding the analytically solution for (16) directly by partial derivatives is very complicated, since the expressions for the rates $\{R_i^*\}_{i=1}^N$ obtained in (13) depend on $\{D_i^*\}_{i=1}^N$ and $\{h(X_i|X_{i-1}, \dots, X_1)\}_{i=1}^N$; moreover, these sets of values depend themselves on the values of $\{c_i\}_{i=1}^N$. Instead, we use the following iterative algorithm for optimization:

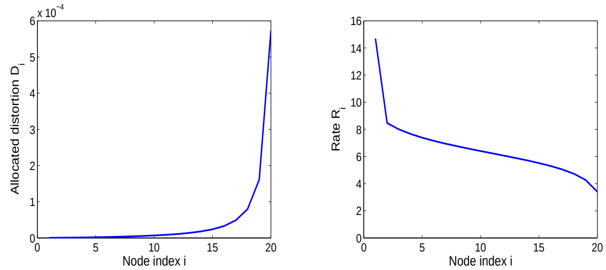
Algorithm 1: Optimal placement (see Fig. 5).

- Initially, the values of c_i are chosen such that the nodes are equally spaced (as in Fig. 3)[1].
- **until** convergence **do**:
  1) Solve (16) with parameters $\{c_i\}_{i=1}^N$ to obtain $\{R_i\}_{i=1}^N$ as in (13).
  2) Solve (16) with parameters $\{R_i\}_{i=1}^N$ to obtain $\{c_i\}_{i=1}^N$.

The convergence of the algorithm is ensured by the convexity of the cost function and the constraint sets [2].

The optimization in Step 1 of the Algorithm has been presented in Section III. We describe now the optimization in Step 2. Denote $w_i^\kappa = c_i - c_{i-1}, i = 1 \dots N$, where w_i is the Euclidean distance between node i and node $i+1$, and $\kappa \in \{2, 4\}$ is the power coefficient of the

[1] Note that the placement uniquely determines the ordering of nodes. Namely, once an order is set among the nodes, the same order will hold along the optimization, while only the distances between the nodes will change.



(a) Distortion allocation as a function of node index.

(b) Rate allocation as a function of node index.

Fig. 6. Optimized placement for the one-dimensional network.

distance[2]. Thus, $c_i = \sum_{j=1}^{i} w_j^2$. The values of R_i^* are computed as in (13), using as initialization $w_i = \frac{L}{N}$, corresponding to equally spaced nodes. Notice that the form of the solution for R_i^* is independent of the placement. Therefore, we can now rewrite the optimization (16) including the unknowns $\{w_i\}_{i=1}^N$ as follows:

$$\begin{aligned} \{w_i^*\}_{i=1}^N &= \arg\min_{\{w_i\}_{i=1}^N} \sum_{i=1}^{N} (\sum_{j=i}^{N} R_j^*) w_i^\kappa \\ \text{under constraint:} & \qquad\qquad\qquad\qquad (17) \\ \sum_{i=1}^{N} w_i &= L. \end{aligned}$$

The solution of (17) is obtained again using Lagrange multipliers:

$$w_i^* = \frac{L}{\sum_{j=i}^N R_j^* \left(\sum_{l=1}^N \frac{1}{\sum_{j=l}^N R_j^*} \right)}, i = 1 \dots N, \qquad (18)$$

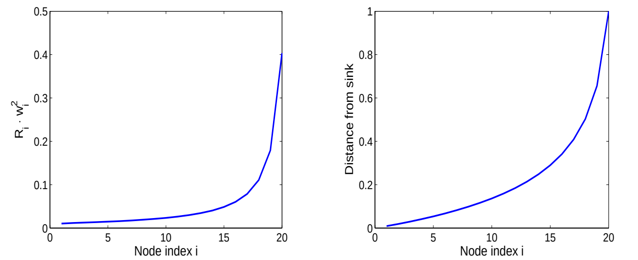
Note that by placing the nodes as given by (18) the rate allocation changes as a function of $\{w_i^*\}_{i=1}^N$, since both the differential entropies and optimal distortions in (13) depend on the values of c_i. Next, for these values of $\{c_i^*\}_{i=1}^N$, we optimize over $\{R_i^*\}_{i=1}^N$ (Step 1 of Algorithm 1). We repeat this iterative procedure until the algorithm converges. The resulting distortion and rate allocation are shown in Fig. 6.

The convergence of Algorithm 1 is insured by the convexity of the cost function and the constraint sets. For the instances we used ($N = 20$, $\alpha = 10^{-3}$, $\kappa = 2$, $\beta = 1$, no constraints on the inter-node distances), the algorithm converges in at most 4 steps. Our numerical results show that nodes at the extremity of the network are largely spread, and nodes near the sink are closer to each other (see Fig. 7(b)). Further constraints imposed on the maximum distance between nodes would make the distances between the nodes more even.

B. Network Lifetime Maximization

The plot in Fig. 7 shows the power consumption *per node* and distances from node to the sink for the scenario in Section IV-A as a result of solving problem (16). The per node power is a quantity inversely proportional to the lifetime of a node. Sometimes, in practice, a different target of interest is to maximize the lifetime of the network. So, in Fig. 7, the last node N will be the first to die. Let us now separately consider another set of constraints related to the lifetime of the network, namely we impose that all nodes consume the same amount of power. This new set of constraints are aimed

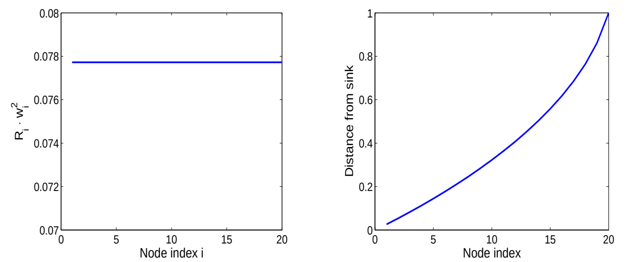
[2]For the sake of the simplicity of analysis, we will use $\kappa = 2$ in solving the optimization problems, however our results can be easily extended to arbitrary values of κ.



(a) Power consumption as a function of node index.

(b) Distance from sink as a function of node index.

Fig. 7. Placement optimization for minimizing the total power.



(a) Power consumption as a function of node index.

(b) Distance from sink as a function of node index.

Fig. 8. Placement for lifetime optimization.

at maintaining a large number of nodes alive. Namely, our problem now amounts to solving the following set of equations:

$$\begin{aligned} \sum_{i=1}^{N} w_i &= L \qquad (19) \\ \left(\sum_{j=i}^{N} R_j^*\right) w_i^2 &= \left(\sum_{j=k}^{N} R_j^*\right) w_k^2, i, k = 1 \dots N \\ \sum_{i=1}^{N} D_i &= D. \end{aligned}$$

The solution is easily found to be:

$$w_i^* = \frac{L}{\sqrt{\sum_{j=i}^N R_j^*} \left(\sum_{l=1}^N \sqrt{\frac{1}{\sum_{j=l}^N R_j^*}} \right)}. \qquad (20)$$

Note that $\{R_i^*\}_{i=1}^N$ depend on $\{w_i^*\}_{i=1}^N$ as in (13), thus we can implement a similar iterative procedure to find the optimal weights $\{w_i^*\}_{i=1}^N$:

Algorithm 2: Lifetime optimization.

- **until** convergence **do**:
  1) Given $\{w_i\}_{i=1}^N$, express $\{R_i\}_{i=1}^N$ as (13).
  2) Solve (19) with parameters $\{R_i\}_{i=1}^N$ to obtain $\{w_i\}_{i=1}^N$ as (20).

Fig. 8 shows the distance from sink, and power consumption at each node, when the optimization of node placement is done as in (20).

For the same instance, the total power consumed with the optimized placement of Section IV-A is 1.17 [bit $\cdot$ m^2], whereas with the scenario in this section, the total power is 1.55 [bit $\cdot$ m^2]. We conclude that a placement that optimizes both the total power consumption of the network while maintaining the same individual consumption levels at nodes should be a tradeoff of the results of the two optimization problems.

V. Conclusions

We considered the problem of power efficient data gathering in sensor networks, with high-resolution coding, under distortion constraints. We found the rate and distortion allocations in a closed-form, and illustrated our results with numerical experiments on Gaussian correlated random fields. We also studied the problem of optimal placement for two power efficiency targets of interests, namely total power and network lifetime, and compared the tradeoffs involved. Our current work is focused on optimization problems where one of the two targets of interest is minimized under upper bound constraints on the other target.

References

[1] T. Berger: *Multiterminal source coding*, Lectures presented at CISM Summer School on the Information Theory Approach to Communications, 1977.

[2] D. Bertsekas: *Network Optimization: Continuous and Discrete Models*, Athena Scientific, 1998.

[3] C. Chong, S. P. Kumar: *Sensor Networks: Evolution, Opportunities and Challenges*, in Proc. IEEE, vol. 91, no. 8, 2003.

[4] T.M. Cover, J.A. Thomas: *Elements of Information Theory*, John Wiley and Sons, Inc., 1991.

[5] N. Cressie: *Statistics for Spatial Data*, Wiley 1991.

[6] R. Cristescu, B, Beferull-Lozano, M. Vetterli: *Networked Slepian-Wolf: Theory, Algorithms and Scaling Laws*, submitted to IEEE Trans. on Inf. Th., 2003.

[7] R. Cristescu, B. Beferull-Lozano, M. Vetterli, R. Wattenhofer: *Network Correlated Data Gathering with Explicit Communication: NP-Completeness and Algorithms*, submitted to IEEE Trans. on Networking, 2004.

[8] R. Cristescu, B. Beferull-Lozano, M. Vetterli: *On Network Correlated Data Gathering*, in Proc. INFOCOM 2004.

[9] D. Ganesan, R. Cristescu, B. Beferull-Lozano, *Power-Efficient Sensor Placement and Transmission Structure for Data Gathering under Distortion Constraints*, in Proc. IPSN 2004, submitted to ACM Trans. on Sensor Networks, 2004.

[10] A. Goel, D. Estrin: *Simultaneous Optimization for Concave Costs: Single Sink Aggregation or Single Source Buy-at-Bulk*, ACM-SIAM Symposium on Discrete Algorithms, 2003.

[11] A. Pandya, A. Kansal, G. Pottie, M. Srivastava: *Bounds on the Rate-Distortion of Multiple Cooperative Gaussian Sources*, CENS TR (UCLA), 2003.

[12] G. J. Pottie, W. J. Kaiser: *Wireless Integrated Sensor Networks*, Communications of ACM 2000.

[13] D. Slepian, J.K.Wolf: *Noiseless Coding of Correlated Information Sources*, IEEE Trans. Information Theory, IT-19, 1973, pp. 471-480.

[14] R.Zamir, T.Berger, *Multiterminal Source Coding with High Resolution*, IEEE Trans. on Information Theory, pp. 106-117, Jan. 1999.