

Rate-Distortion Problem for Physics Based Distributed Sensing*

Baltasar Beferull-Lozano
School of Computer and
Communication Sciences
Swiss Federal Institute of
Technology-EPFL
CH-1015, Lausanne,
Switzerland

Baltasar.Beferull@epfl.ch

Robert L. Konsbruck
School of Computer and
Communication Sciences
Swiss Federal Institute of
Technology-EPFL
CH-1015, Lausanne,
Switzerland

Robert.Konsbruck@epfl.ch

Martin Vetterli[†]
School of Computer and
Communication Sciences
Swiss Federal Institute of
Technology-EPFL
CH-1015, Lausanne,
Switzerland

Martin.Vetterli@epfl.ch

ABSTRACT

We consider the rate-distortion problem for sensing the continuous space-time physical temperature in a circular ring on which a heat source is applied over space and time, and which is also allowed to cool by radiation or convection to its surrounding medium. The heat source is modelled as a continuous space-time stochastic process which is bandlimited over space and time. The temperature field is the result of a circular convolution over space and a continuous-time causal filtering over time of the heat source with the Green's function corresponding to the heat equation, which is space and time invariant. The temperature field is sampled at uniform spatial locations by a set of sensors and it has to be reconstructed at a base station. The goal is to minimize the mean-square-error per second, for a given number of nats per second, assuming ideal communication channels between sensors and base station. We find a) the centralized $R^c(D)$ function of the temperature field, where all the space-time samples can be observed and encoded jointly. Then, we obtain b) the $R^{s-i}(D)$ function, where each sensor, independently, encodes its samples optimally over time and c) the $R^{st-i}(D)$ function, where each sensor is constrained to encode also independently over time. We also study two distributed prediction-based approaches: a) with perfect feedback from the base station, where temporal prediction is performed at the base station and each sensor performs differential en-

*The work presented in this paper was supported (in part) by the National Competence Center in Research on Mobile Information and Communications Systems (NCCR-MICS), a center supported by the Swiss National Science Foundation under grant number 5005-67322.

[†]Martin Vetterli is also with the Department of EECS, University of California, Berkeley, CA 94720, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IPSN'04, April 26–27, 2004, Berkeley, California, USA.
Copyright 2004 ACM 1-58113-846-6/04/0004 ...\$5.00.

coding, and b) without feedback, where each sensor locally performs temporal prediction.

Categories and Subject Descriptors

B.4.1 [Input/Output and Data Communications]: Data Communication Devices—*receivers, transmitters*; C.2.4 [Computer-Communication Networks]: Distributed Systems; E.4 [Coding and Information Theory]: Data compaction and compression, Formal models of communication.

General Terms

Theory, Design, Performance.

Keywords

Sensor networks, distributed sampling, heat equation, temperature field, Green's function, spatio-temporal correlation, rate-distortion, centralized coding, local coding, distributed coding, prediction, feedback.

1. INTRODUCTION

In sensor networks, some continuous space-time physical phenomenon is sampled by a set of remote sensors [1], which are limited in power, and an estimate of this physical phenomenon has to be obtained at a base station. Due to the restricted processing and communication power in the sensors, it is usually not allowed to have communication between sensors and at the same time, it is very important to reduce as much as possible the rate at which the information is encoded. There has been important recent research work going in this direction [2, 5]. However, in all this previous work, the physics that describes the phenomenon, has not been taken into account. As an example, it is usually assumed in the literature that the process that is sampled is i.i.d. over time, while for real physical processes, such as the temperature field in a heat conducting medium, there is a very particular structure over time which is described completely by physical laws. In this paper, we incorporate the physics into a sensor network rate-distortion problem. There are two main reasons why this is important: first, in practice, sensor networks sample actually in space and time real physical phenomena (e.g. temperature), which have a

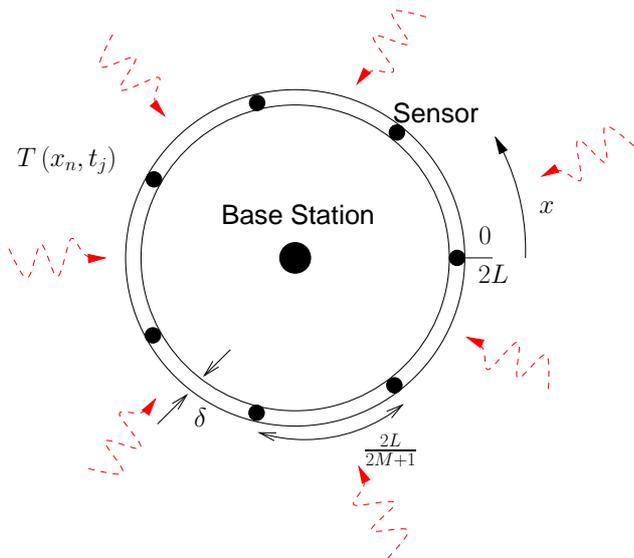


Figure 1: Ring structure.

very particular spatio-temporal structure, and second, this particular structure described by the physics, can be very useful in terms of coding efficiency if it is exploited properly.

We consider the rate-distortion problem for sensing the continuous space-time physical temperature on a closed circular ring on which a heat source (e.g. electric heating) is applied over space and time, and which is also allowed to cool by electromagnetic radiation or forced convection to its surrounding medium (see Fig. 1). The heat source is allowed to both produce (additive source) and withdraw (sink) heat and is modelled by a continuous Gaussian space-time stochastic process which is periodic and bandlimited over space (finite set of harmonics), and bandlimited over time. Using Green's Theorem for the heat equation, the temperature field over the ring is the result of a certain continuous-space circular convolution and a certain continuous-time causal filtering of the heat source with the corresponding Green's function of the system, which turns out to be space and time invariant. The temperature field is sampled at uniform spatial locations by a set of sensors and it has to be reconstructed at a base station. The goal is to minimize the distortion measured in mean-square-error per second, for a given number of nats per second, assuming that there are ideal communication channels between the sensors and the base station.

We first obtain the exact closed-form expression for the centralized $R^c(D)$ function of the temperature field, which corresponds to the idealized case where all the spatio-temporal (unquantized) samples together can be observed simultaneously and encoded jointly. Then, we obtain the $R^{s-i}(D)$ function, where each sensor, independently from the other sensors, encodes optimally its samples over time, thus taking full advantage of the correlation over time, and the $R^{st-i}(D)$ function, where each sensor individually encodes also independently over time. We also study two distributed approaches: a) a prediction based system with a perfect feedback channel from the base station to the sensors, where temporal prediction is performed at the base station and

each sensor performs differential encoding; thus, this scheme is equivalent to a closed-loop DPCM system where the encoders are the sensors, the decoder is the base station and, most importantly, the prediction operation at the base station makes use of the heat equation and the heat conductivity properties of the heat conductor ring, and b) each sensor independently performs temporal prediction without using feedback information from the base station. We compare the different approaches. Finally, we also explain briefly how to use nested code based constructions [9] for our problem¹.

This paper is structured as follows. In Section 2, we introduce the heat problem by explaining the physics governing the temperature field in a ring, and the fundamental concept of the Green's function of the system. In Section 3, we explain the rate-distortion problem. In Section 4, we obtain the centralized $R^c(D)$ function, and in Section 5, we find the $R^{s-i}(D)$ and $R^{st-i}(D)$ functions. Finally, in Section 6, we study the two distributed prediction-based approaches and compare the different approaches.

2. TEMPERATURE PROBLEM IN THE RING

The sensed physical field that we consider in this work is the continuous space-time temperature field $T(x, t)$ in a closed circular (heat conductor) ring of length $2L$, where $x \in [0, 2L[$ indicates the spatial position in the ring (see Fig. 1), and $t \in \mathbb{R}$. It is assumed that the ring has a very small cross-section, so that the temperature at all points of the cross-section may be taken to be the same. A space-time varying heat source (e.g. electric heating), which we denote by $g(x, t)$ and which is measured in (Watts/meter³), is applied on the ring. Moreover, the ring is also allowed to cool by electromagnetic radiation or forced convection to its surrounding medium.

2.1 Heat Source and Thermal Properties

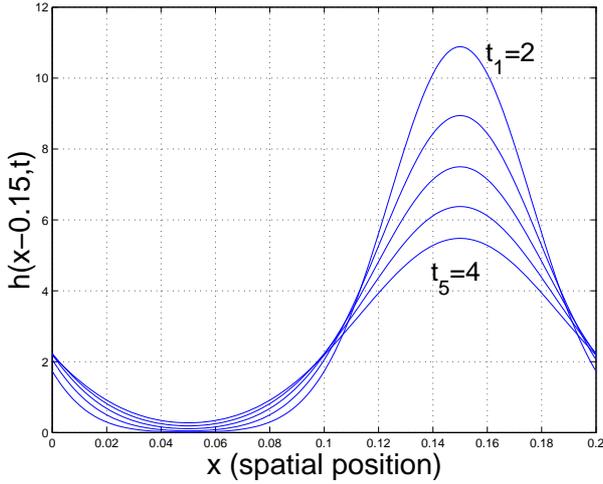
The heat source $g(x, t)$ is allowed to both produce (additive source) and withdraw (sink) heat², and we model it by a continuous space-time Gaussian stochastic process, which is periodic and bandlimited over space, and which is bandlimited over time. That is, $g(x, t)$ is given by:

$$g(x, t) = \frac{g_0(t)}{\sqrt{2}} + \sum_{m=1}^M (g_m(t) \cos(\lambda_m x) + g_{m+M}(t) \sin(\lambda_m x)) \quad (1)$$

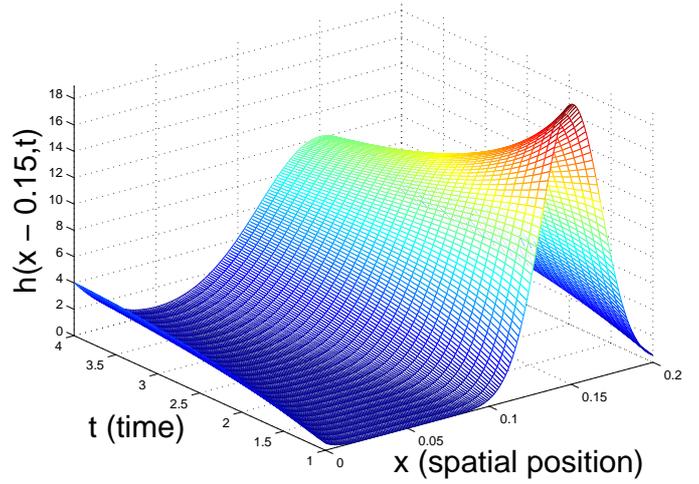
where $M+1$ is the number of harmonics (spatial bandwidth of $2M+1$), $\lambda_m = \frac{m\pi}{L}$, $m \geq 0$ (thus the fundamental spatial period is $2L$, which is equal to the length of the ring), and $\{g_m(t)\}_{m=0}^{2M}$ is a set of $2M+1$ continuous-time real wide-sense stationary (WSS) Gaussian stochastic processes with zero mean and which are assumed to be independent, that is, the cross-correlation $R_{g_{m_1} g_{m_2}}(\gamma) = E[g_{m_1}(t+\gamma)g_{m_2}(t)] = 0$ for $m_1 \neq m_2$, $\forall \gamma \in \mathbb{R}$. Moreover, each continuous-time process $g_m(t)$, $m = 0, \dots, 2M$, is assumed to be bandlimited to $[-\pi, \pi]$ and to have a constant gain $G_m > 0$, that is, each process has a power spectral density (PSD) given by $S_{g_m}(\Omega) = G_m 1_{[-\pi, \pi]}(\Omega)$, where Ω denotes the continuous-time angular frequency. We also assume that $G_m = G_{m+M}$, for $m = 1, \dots, M$, that is, for a given harmonic m , the even and odd terms have the same power. It can be easily

¹This is part of our current research work.

²It can be shown that this is physically realizable [4].



(a)



(b)

Figure 2: (a) Green's function $h(x - 0.15, t)$, which represents the temperature caused by a heat source $g(x, t) = \delta(x - 0.15)\delta(t)$. The material of the ring is Silver, with properties: $\alpha = 17.004 \cdot 10^{-5}$, $\mu = 100$ and $L = 0.1$. Each curve corresponds to a different value of time t , more specifically, $t_1 = 2$, $t_2 = 2.5$, $t_3 = 3$, $t_4 = 3.5$ and $t_5 = 4$ sec., (b) same Green's function in 3D.

shown that this condition is necessary to make the resulting temperature process $T(x, t)$ stationary over space.

The thermal properties of the ring play also an important role in the rate-distortion properties that we analyze in Sections 4, 5 and 6. The thermal parameters of the ring are: a) its thermal conductivity κ in (Watts/(meter Kelvin)), b) its thermal diffusivity $\alpha = \frac{\kappa}{\rho C_p}$ in (meter²/second), where ρ is the density in (kg/meter³) and C_p is the specific heat in (Joules/(kg Kelvin)), and c) its dissipation parameter $\mu = \frac{2h}{\kappa\delta}$ in (meter⁻²), where h , given in (Watts/(meter² Kelvin)), is the heat transfer coefficient, and δ is the thickness of the ring, which is assumed to be very small. The dissipation parameter μ is related to the loss of heat energy from the ring to its surrounding medium either by electromagnetic radiation or by forced convection through contact with a fluid flow [4].

2.2 Heat Differential Equation for the Ring

Given a source $g(x, t)$ and the different thermal parameters of the ring, the differential equation describing the resulting temperature $T(x, t)$ in the ring, is given by:

$$\frac{\partial^2 T(x, t)}{\partial x^2} + \frac{g(x, t)}{\kappa} - \mu T(x, t) = \frac{1}{\alpha} \frac{\partial T(x, t)}{\partial t}$$

$$\left. \begin{aligned} T(0, t) &= T(2L, t) \\ \frac{\partial T(x, t)}{\partial x} \Big|_{x=0} &= \frac{\partial T(x, t)}{\partial x} \Big|_{x=2L} \end{aligned} \right\} \text{(periodic boundary conditions)}$$

$$T(x, -\infty) = 0 \quad \text{(initial condition)}$$

(2)

where it can be seen that the periodic boundary conditions enforce continuity of both the temperature value and its

spatial derivative at the point³ 0 (see Fig. 1).

Using Green's Theorem [3], it can be shown that the solution to this differential equation is given by:

$$T(x, t) = \frac{\alpha}{\kappa} \int_{\tau=-\infty}^{\tau=t} \int_{\xi=0}^{\xi=2L} g(\xi, \tau) h(x - \xi, t - \tau) d\xi d\tau \quad (3)$$

where $h(x, t; \xi, \tau) = h(x - \xi, t - \tau)$ is the Green's function of the system described by (2), which turns out to be time-invariant and space-invariant⁴. This means that the operator corresponding to the differential equations in (2) is equivalent to a linear time-invariant space-invariant system described by an impulse response $h(x, t)$. This impulse response can be shown to be [3]:

$$h(x, t) = u(t) \left(\frac{1}{2L} e^{-\mu\alpha t} + \frac{1}{L} \sum_{m=1}^{\infty} e^{-(\lambda_m^2 + \mu)\alpha t} \cos(\lambda_m x) \right)$$

where $u(t)$ is the step function, that is, $u(t) = 1$ for $t \geq 0$ and $u(t) = 0$ for $t < 0$. Thus, the impulse response is causal over time, which is intuitively expected because a physical system cannot respond before an input has been applied to it. Moreover, it is periodic over space with a period of $2L$. This implies that the filtering over space in (3) consists of a continuous-space circular convolution, which in the Fourier domain, corresponds to the product of the corresponding Fourier coefficients. After performing the circular convolution over space, the final temperature process $T(x, t)$ can be

³It can be shown that this condition actually implies also continuity of the temperature and its spatial derivative at any position x , $0 < x < 2L$.

⁴In general, the Green's function $h(x, t; \xi, \tau)$ of a system is not space-invariant [3] and represents the response of the system to a Dirac $\delta(x - \xi)\delta(t - \tau)$ applied at position ξ and time τ that is observed at position x and time t .

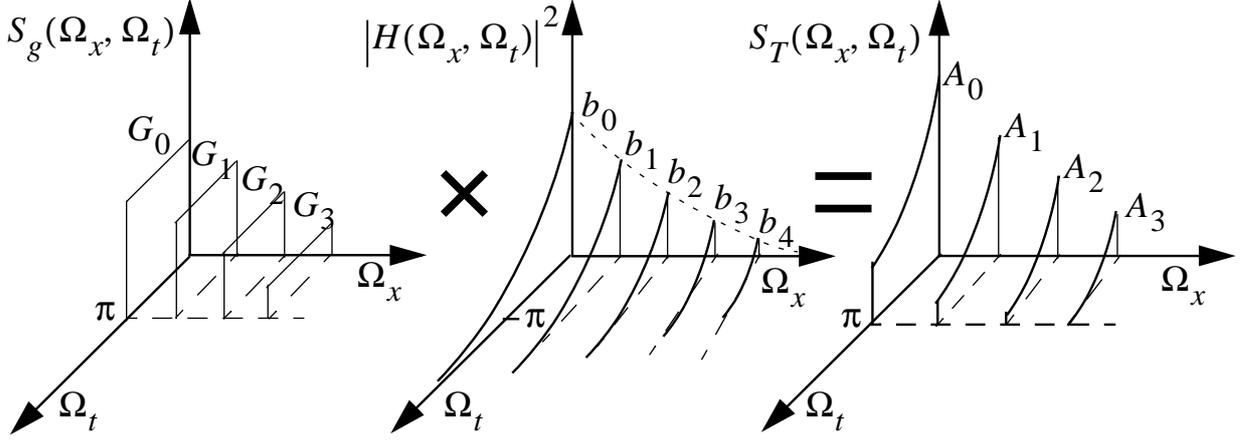


Figure 3: Space-time convolution in the frequency domain.

written as:

$$T(x, t) = \frac{\beta_0(t)}{\sqrt{2}} + \sum_{m=1}^M (\beta_m(t) \cos(\lambda_m x) + \beta_{m+M}(t) \sin(\lambda_m x))$$

$$\beta_m(t) = g_m(t) * h_m(t), \quad h_m(t) = u(t) \left(\frac{\alpha}{\kappa} e^{-(\lambda_m^2 + \mu)\alpha t} \right) \quad (4)$$

where $\lambda_m = \lambda_{m+M}$ for $m = 1, \dots, M$, and the $2M + 1$ independent Gaussian processes $\{\beta_m(t)\}_{m=0}^{2M}$ are still bandlimited to $[-\pi, \pi]$, but their spectral densities are given by:

$$S_{\beta_m}(\Omega) = |\mathcal{F}(h_m(t))|^2 G_m \mathbf{1}_{[-\pi, \pi]}(\Omega) = \frac{A_m c_m^2}{\Omega^2 + c_m^2}, \quad |\Omega| \leq \pi,$$

$$A_m = \frac{G_m}{\kappa^2(\lambda_m^2 + \mu)^2}, \quad c_m = (\lambda_m^2 + \mu)\alpha \quad (5)$$

where $\mathcal{F}(\cdot)$ denotes the continuous-time Fourier transform. Fig. 3 illustrates, in the frequency domain, the space-time convolution performed by Green's Theorem. It is observed, as expected, that the temperature process $T(x, t)$ is periodic (with period $2L$) and bandlimited over space with a spatial bandwidth of $2M + 1$.

3. DISTRIBUTED SAMPLING AND RATE-DISTORTION PROBLEM

The temperature process given in (4) is sampled uniformly in space and time at the corresponding spatial and temporal Nyquist sampling frequencies, respectively. This means that the sampling period in space is $T_s = \frac{2L}{2M+1}$ meters and the sampling period in time is $T_t = 1$ seconds. The sensing task is performed by $2M + 1$ sensors which are located uniformly at $x_n = \frac{2L}{2M+1}n$, $n = 0, \dots, 2M$. Thus, each sensor samples the temperature process every second (temporal Nyquist sampling rate), that is, the n -th sensor takes the samples $\{T(x_n, -\infty), \dots, T(x_n, t_j), T(x_n, t_{j+1}), \dots, T(x_n, \infty)\}$, where $t_{j+1} - t_j = 1$.

The continuous space-time temperature process $T(x, t)$ has to be reconstructed at a base station (BS), as illustrated in Fig. 4. Given a reconstructed temperature process $\hat{T}(x, t)$, the distortion, which we denote by D , is measured

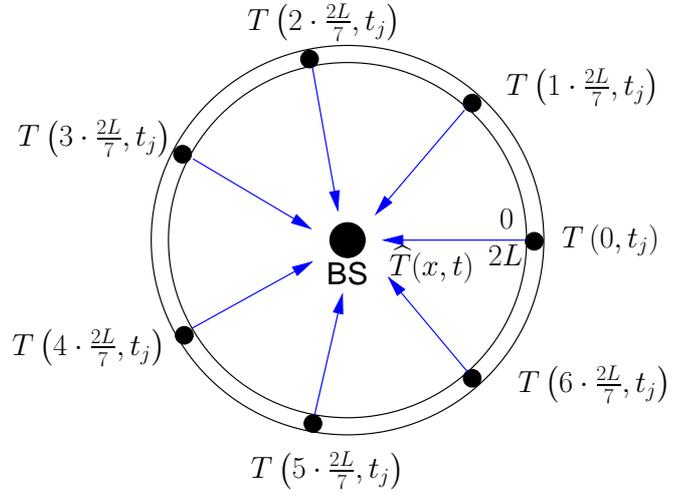


Figure 4: Rate-Distortion problem in the ring.

in mean-square-error (MSE) per second, and is defined by:

$$D = \frac{1}{2L} \int_0^{2L} \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T E[(T(x, t) - \hat{T}(x, t))^2] dt dx \quad (6)$$

The goal is to minimize the distortion at the BS, for a given number of nats per second, denoted by R . Regarding the communication model, it is assumed throughout this work that:

1. From the sensors to the BS, there are ideal channels, that is, if the n -th sensor has to deliver R_n nats per second to the BS, the sensor will spend the necessary power to transmit them without error⁵.
2. Sensors communicate using some orthogonal communication method (e.g. TDMA, FDMA) so that there is no interference.

⁵The consideration of the channel in the heat problem, which involves analyzing the behavior of distortion as a function of the power is a subject of our current investigation.

3. From the BS to the sensors there is a perfect channel where the BS can spend a large amount of power and hence, it can be assumed that the BS can transmit at an infinite rate to the sensors. Thus, we assume that if feedback from the BS to the sensors is used, it is perfect.

Our objective in this paper is to study the rate-distortion problem under these communication channel idealizations.

As explained in the following sections, the number of nats per second R required to achieve a given distortion D depends on how the coding is performed.

4. $R^c(D)$ FOR CENTRALIZED CODING

In this section, we obtain the Rate-Distortion function for the most idealized (genie-aided) case where the exact (un-quantized) space-time temperature samples can be observed simultaneously by the sensors, that is, all the space-time samples $\{T(x_n, -\infty), \dots, T(x_n, t_j), T(x_n, t_{j+1}), \dots, T(x_n, \infty)\}_{n=0}^{2M}$ can be encoded jointly before being sent to the BS. We denote this function by $R^c(D)$.

In order to find the $R^c(D)$ function for this idealized scenario, we first express the distortion D , as defined in (6), as a function of the distortions corresponding to the processes $\{\beta_m(t)\}_{m=0}^{2M}$, which completely determine the temperature process $T(x, t)$. Let \mathbf{F} be the $(2M+1) \times (2M+1)$ matrix with m -th row $(\mathbf{F})_m = [\frac{1}{\sqrt{2}}, \cos(\frac{2\pi m}{2M+1}), \dots, \cos(\frac{2\pi m M}{2M+1}), \sin(\frac{2\pi m}{2M+1}), \dots, \sin(\frac{2\pi m M}{2M+1})]$, $\beta(t) = [\beta_0(t), \dots, \beta_{2M}(t)]^T$ and $\mathbf{T}(t) = [T(0, t), \dots, T(2M \frac{2L}{2M+1}, t)]^T$, which is the spatial vector of temperature samples. Then, since $\mathbf{F}^{-1} = \frac{2}{2M+1} \mathbf{F}^T$, the BS can find $\beta(t)$ from $\mathbf{T}(t)$ and viceversa by:

$$\mathbf{T}(t) = \mathbf{F}\beta(t) \iff \beta(t) = \frac{2}{2M+1} \mathbf{F}^T \mathbf{T}(t) \quad (7)$$

Thus, to provide a set of reconstructed processes $\{\hat{\beta}_m(t)\}_{m=0}^{2M}$ is equivalent to providing a reconstruction $\hat{T}(x, t)$. Because of the orthogonality property of the Fourier series over space, and the fact that $\{\beta_m(t)\}_{m=0}^{2M}$ are independent random processes, it can be easily shown that

$$D = \frac{1}{2} \sum_{m=0}^{2M} D_m, \quad R = \sum_{m=0}^{2M} R_m \quad (8)$$

where D_m is the MSE per second associated to the temporal process $\beta_m(t)$, which is given by:

$$D_m = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T E[(\beta_m(t) - \hat{\beta}_m(t))^2] dt \quad (9)$$

and R_m is the corresponding number of nats per second. These additivity properties allow us to use the well-known waterfilling (equal-slope) technique [8] over the set of processes $\{\beta_m(t)\}_{m=0}^{2M}$, or equivalently, over the set of spatial harmonics. In order to do this, we first need to calculate the rate-distortion functions $\{R_m(D_m)\}_{m=0}^{2M}$. Notice that since each process $\beta_m(t)$ is bandlimited, each $R_m(D_m)$ can be calculated by considering the discrete-time sampled process $\beta_m(t_j)$ and performing waterfilling over time [8], that is, $R_m(D_m)$ is given in parametric form by [8]:

$$D_m(\theta_m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \min[\theta_m, S_{\beta_m}(\omega)] d\omega \quad (10)$$

$$R_m(\theta_m) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \max\left[0, \log\left(\frac{S_{\beta_m}(\omega)}{\theta_m}\right)\right] d\omega$$

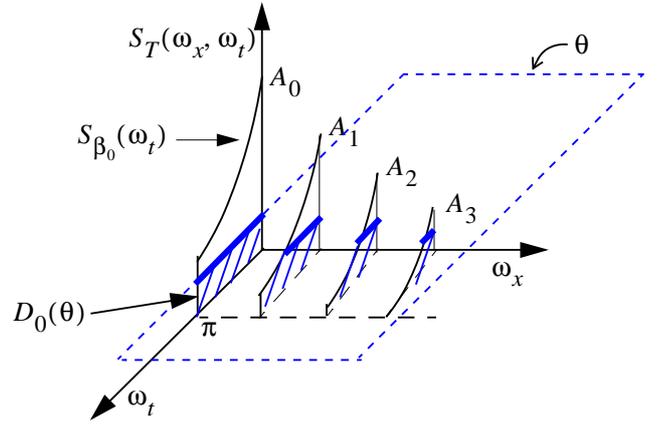


Figure 5: Waterfilling technique: only the energy above θ is retained.

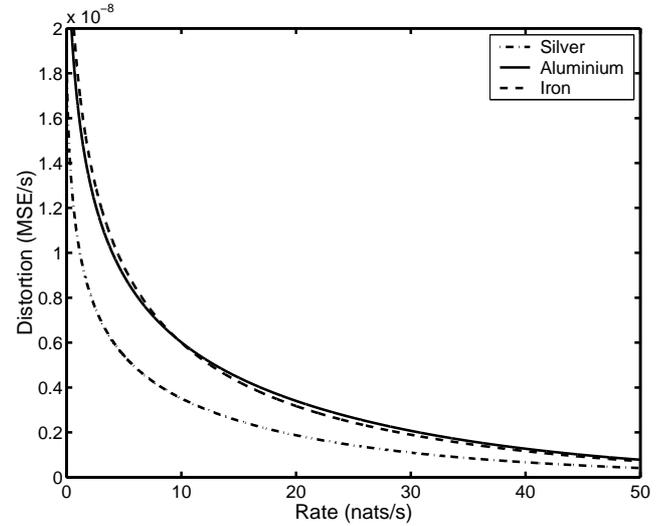


Figure 6: Centralized $R^c(D)$ function for Silver, Aluminium and Iron with $G_m = 1$, $m = 0, \dots, 2M$, $L = 0.1$ and $M = 20$. Time scale: 0.1 sec.

where $0 \leq \theta_m \leq S_{\beta_m}(0)$ is the waterfilling parameter, and ω denotes the discrete-time angular frequency. Performing this calculation, we get the following parameterized expression for $R_m(D_m)$:

$$D_m(\phi_m) = \frac{A_m c_m}{\pi} \left(\arctan\left(\frac{\pi}{c_m}\right) + \frac{\phi_m}{1+\phi_m^2} - \arctan(\phi_m) \right)$$

$$R_m(\phi_m) = \frac{c_m}{\pi} (\phi_m - \arctan(\phi_m))$$

for $0 \leq \phi_m \leq \frac{\pi}{c_m}$

$$D_m(\phi_m) = \frac{A_m}{1+\phi_m^2}$$

$$R_m(\phi_m) = \frac{1}{2} \log\left(\frac{(1+\phi_m^2)c_m^2}{c_m^2 + \pi^2}\right) + 1 - \frac{c_m}{\pi} \arctan\left(\frac{\pi}{c_m}\right)$$

for $\phi_m \geq \frac{\pi}{c_m}$

where $\phi_m = \sqrt{\frac{A_m - \theta_m}{\theta_m}}$, and the parameters A_m and c_m are as defined in (5).

Using these results, we can show the following Theorem:

THEOREM 1. The $R^c(D)$ function is given by the follow-

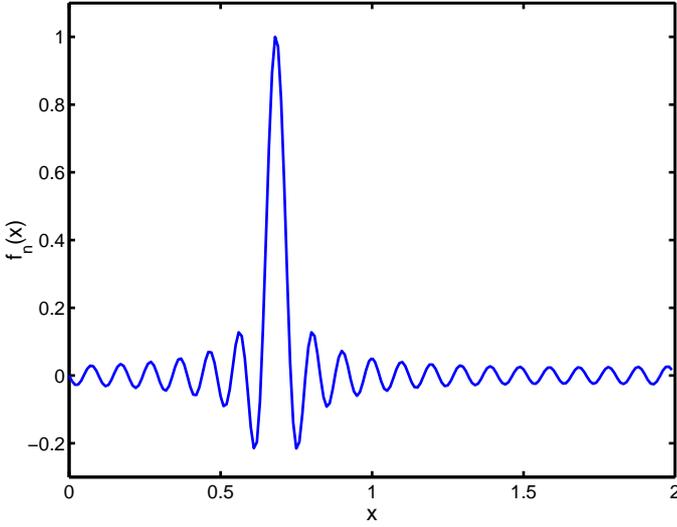


Figure 7: Interpolation function $f_n(x)$ for a particular position n .

ing parametric expression:

$$D^c(\eta) = \frac{1}{2} \sum_{m=0}^{2M} D_m(f_m(\eta)), \quad R^c(\eta) = \sum_{m=0}^{2M} R_m(f_m(\eta))$$

where $f_m(\eta) = \sqrt{2A_m\eta - 1}$ and $\eta \geq \min_{\{0 \leq m \leq 2M\}} \frac{1}{2A_m}$.

Proof: The proof follows easily by performing waterfilling (see Fig. 5) over the set of Rate-Distortion functions $\{R_m(D_m)\}_{m=0}^{2M}$. This is done by solving the Lagrangian optimization problem with functional $J = R + \eta D = \sum_{m=0}^{2M} R_m + \frac{\eta}{2} \sum_{m=0}^{2M} D_m$, and finding for each value of η the operating points in each of the curves $\{R_m(D_m)\}_{m=0}^{2M}$. ■

Notice that $R^c(D)$ provides a lower bound for any potential distributed lossy coding⁶ approach for correlated data [5]. This is because, as shown in Section 5, the vector $\mathbf{T}(t)$ is spatially correlated, so clearly, any distributed coding approach, as considered in Section 6, can only be worse than the centralized coding approach considered in this section.

Fig. 6 shows the $R^c(D)$ functions corresponding to three different heat conductor materials for the ring, namely, Silver, Aluminium and Iron. It can be seen that there is a clearly different rate-distortion curve for each of the materials, which shows the fact that the conductivity properties of the heat conductor have a very important impact on the rate-distortion performance. Interestingly, the lowest distortion is achieved for the Silver material, which is a better heat conductor than Aluminium and Iron.

5. RATE DISTORTION FOR LOCAL CODING

In this section, we consider the scenario in which the source coding is performed independently by each sensor. We obtain the rate distortion function for the case where

⁶Notice that although there exist binning techniques [5] that can be used, the solution to the general problem of distributed lossy coding for correlated data has not been solved yet [6].

each sensor, independently, exploits optimally the data correlation over time, as well as for the case where each sensor is additionally constrained to encode each time sample separately, disregarding the correlation over time.

5.1 $R^{s-i}(D)$ Function for Spatially Independent Coding

First, we consider the case where each sensor at its position $x_n = 2nL/(2M+1)$ observes the process $T(x_n, t)$, encodes it independently from the other sensors, and transmits the encoded process to the BS. The original temperature $T(x, t)$ is completely determined by the spatial samples according to :

$$T(x, t) = \sum_{n=0}^{2M} T(x_n, t) f_n(x), \quad (11)$$

where

$$f_n(x) = \frac{2}{2M+1} \left[\frac{1}{2} + \sum_{m=1}^M \left(\cos\left(mn \frac{2\pi}{2M+1}\right) \cos(\lambda_m x) + \sin\left(mn \frac{2\pi}{2M+1}\right) \sin(\lambda_m x) \right) \right]. \quad (12)$$

The interpolation function $f_n(x)$ is illustrated in Fig. 7. Receiving the quantized spatial samples of the temperature process, the BS reconstructs an estimate $\hat{T}(x, t)$ of the original temperature process $T(x, t)$. We call $R^{s-i}(D)$ the rate distortion function corresponding to this scenario. Because of the orthogonality of the family $\{f_n(x)\}_{n=0}^{2M}$, and because of the constraint of the source coding to be performed locally, it can be easily shown that:

$$D^{s-i} = \frac{1}{2M+1} \sum_{n=0}^{2M} D_n, \quad R^{s-i} = \sum_{n=0}^{2M} R_n, \quad (13)$$

where D_n is the MSE per second associated to the temporal process $T(x_n, t)$ and R_n is the rate used by the encoder of the sensor located at position x_n . It is important to notice that since $G_m = G_{m+M}$, the power spectral density of each locally observed process $T(x_n, t)$ does not depend on the location x_n . Let $S(\omega)$ denote this power spectral density. Because of the whole symmetry, it is clear that an equal rate allocation among the sensors is optimal, assuming that each sensor is coding independently. The rate distortion function $R^{s-i}(D)$ is then given by the following Theorem:

THEOREM 2. *The $R^{s-i}(D)$ function for the scenario of spatially independent coding is given in the following parameterized form:*

$$D^{s-i}(\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \min(\theta, S(\omega)) d\omega, \\ R^{s-i}(\theta) = \frac{2M+1}{4\pi} \int_{-\pi}^{\pi} \max\left(0, \log\left(\frac{S(\omega)}{\theta}\right)\right) d\omega,$$

where

$$S(\omega) = \frac{1}{2} \frac{A_0 c_0^2}{\omega^2 + c_0^2} + \sum_{m=1}^M \frac{A_m c_m^2}{\omega^2 + c_m^2}, \quad |\omega| \leq \pi \quad (14)$$

and $0 \leq \theta \leq S(0)$ is the waterfilling parameter [8].

Proof: The proof follows immediately by performing waterfilling over time and from the optimality of the equal rate allocation over space due to the circular symmetry. ■

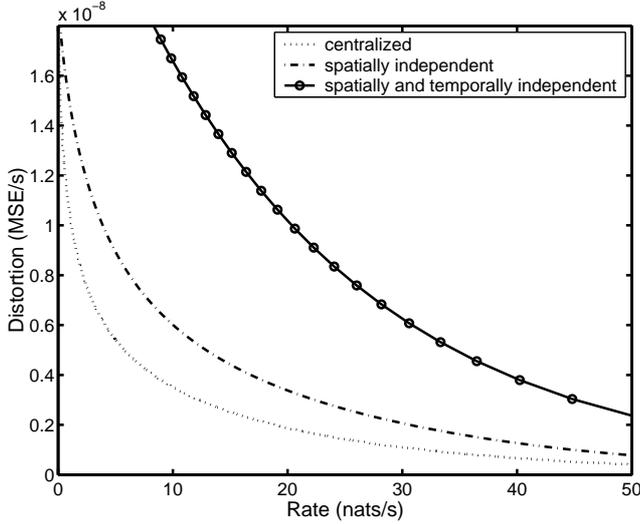


Figure 8: Comparison between $R^c(D)$, $R^{s-i}(D)$ and $R^{st-i}(D)$ for $G_m = 1$, $m = 0, \dots, 2M$, $L = 0.1$, $M = 20$, and the material is Silver: $\alpha = 17.004 \cdot 10^{-5}$, $\mu = 10^4$. Time scale: 0.1 sec.

5.2 $R^{st-i}(D)$ for Spatially and Temporally Independent Coding

In addition to the constraint of local processing, we now also require that each sensor encodes the observed time samples separately, disregarding the correlation over time. We call $R^{st-i}(D)$ the rate distortion function associated to this scenario. Notice that in this case we are constrained to perform quantization only on one temporal sample at a time, and thus, for any optimal entropy-constrained one-dimensional quantizer, there is a quantization shaping loss [8]. The following Theorem directly follows:

THEOREM 3. *For the scenario of spatially and temporally independent coding, the achievable $R^{st-i}(D)$ function is characterized by:*

$$(2M+1)R_G(D) \leq R^{st-i}(D) \leq (2M+1) \left(R_G(D) + \frac{1}{2} \log \left(\frac{\pi e}{6} \right) \right)$$

where $R_G(D) = \frac{1}{2} \max \left[0, \log \left(\frac{\sigma^2}{D} \right) \right]$, $\sigma^2 = \frac{1}{\pi} \int_0^\pi S(\omega) d\omega$ is the variance of each sample, and $S(\omega)$ is as given in Theorem 2.

Fig. 8 shows the comparison between the centralized rate distortion function and the rate distortion functions resulting from spatially independent as well as from spatially and temporally independent coding. In addition to the obvious suboptimality of the latter strategy, it is important to note that spatially independent coding is also suboptimal because the spatial vector of temperature samples $\mathbf{T}(t)$ has a nondiagonal correlation matrix. This can be easily seen as follows. Since the vector $\mathbf{T}(t)$ is related to the vector $\boldsymbol{\beta}(t)$ by the equation $\mathbf{T}(t) = \mathbf{F}\boldsymbol{\beta}(t)$, then, if $\boldsymbol{\Sigma}_{\mathbf{T}}$ denotes the auto-correlation matrix of $\mathbf{T}(t)$ and $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$ denotes the auto-correlation matrix of $\boldsymbol{\beta}(t)$, the following holds:

$$\boldsymbol{\Sigma}_{\mathbf{T}} = \mathbf{F}\boldsymbol{\Sigma}_{\boldsymbol{\beta}}\mathbf{F}^T. \quad (15)$$

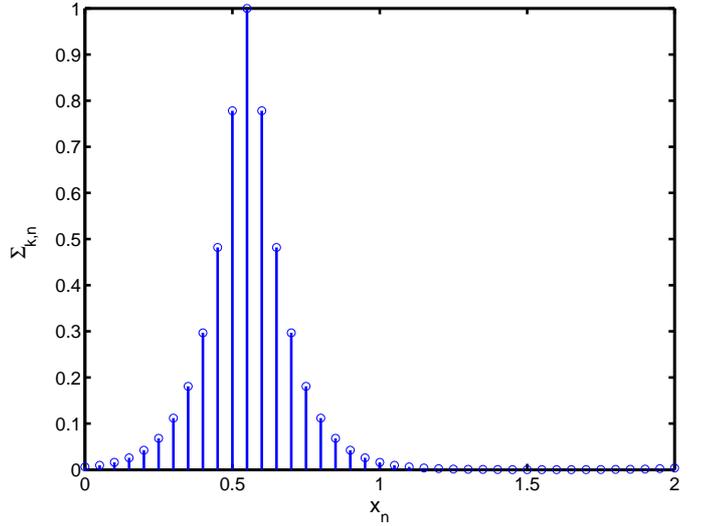


Figure 9: Spatial correlation.

Since the components of $\boldsymbol{\beta}(t)$, although mutually independent, do not have the same variance, the matrix $\boldsymbol{\Sigma}_{\mathbf{T}}$ is not diagonal and the entries of $\mathbf{T}(t)$ are correlated. This is illustrated in Fig. 9. Therefore, $R^{s-i}(D) > R^c(D)$.

6. DISTRIBUTED PREDICTION-BASED CODING

In this section, we consider two prediction based systems. In the first system, we assume that there is a perfect feedback channel from the BS to the sensors (see Fig. 12), that is, the BS can transmit at an infinite rate, while in the second system, no feedback is available.

6.1 Prediction with Feedback from Base Station

In this case, the idea is to perform prediction over time at the BS, and to send the prediction to the sensors, which then perform differential coding, that is, the sensors encode the prediction error. Over time, the system works by sampling at the temporal Nyquist rate, which under our assumptions is 1. The overall processing is illustrated in Fig. 10 and is performed as follows. Given an estimate (reconstruction) of the temperature vector $\hat{\mathbf{T}}(t_j) = [\hat{T}(x_0, t_j), \dots, \hat{T}(x_{2M}, t_j)]^T$ at time t_j , the BS performs a temporal prediction for the temperature field $T(x, t_{j+1})$ at time t_{j+1} . We call this prediction $\tilde{T}(x, t_{j+1})$. Notice that since $\tilde{\boldsymbol{\beta}}(t) = \frac{2}{2M+1} \mathbf{F}^T \tilde{\mathbf{T}}(t)$, we only really need to get the prediction $\tilde{\boldsymbol{\beta}}(t_{j+1})$, that is, we just need to get independently each prediction $\tilde{\beta}_m(t_{j+1})$, $m = 0, \dots, 2M$. In order to perform this prediction, we make use of the Green's Theorem for the heat equation [3], which, applied to our system, establishes that if $\hat{\boldsymbol{\beta}}(t_j)$ were the actual vector at time t_j , then the temperature field $T(x, t_{j+1})$ for the next time slot t_{j+1} is given by $\tilde{\mathbf{T}}(t_{j+1}) = \mathbf{F}\tilde{\boldsymbol{\beta}}(t_{j+1})$, where each component is given by:

$$\tilde{\beta}_m(t_{j+1}) = \hat{\beta}_m(t_j)e^{-c_m} + \frac{\alpha}{\kappa} e^{-c_m} \int_0^1 g_m(t_j + \tau) e^{c_m \tau} d\tau \quad (16)$$

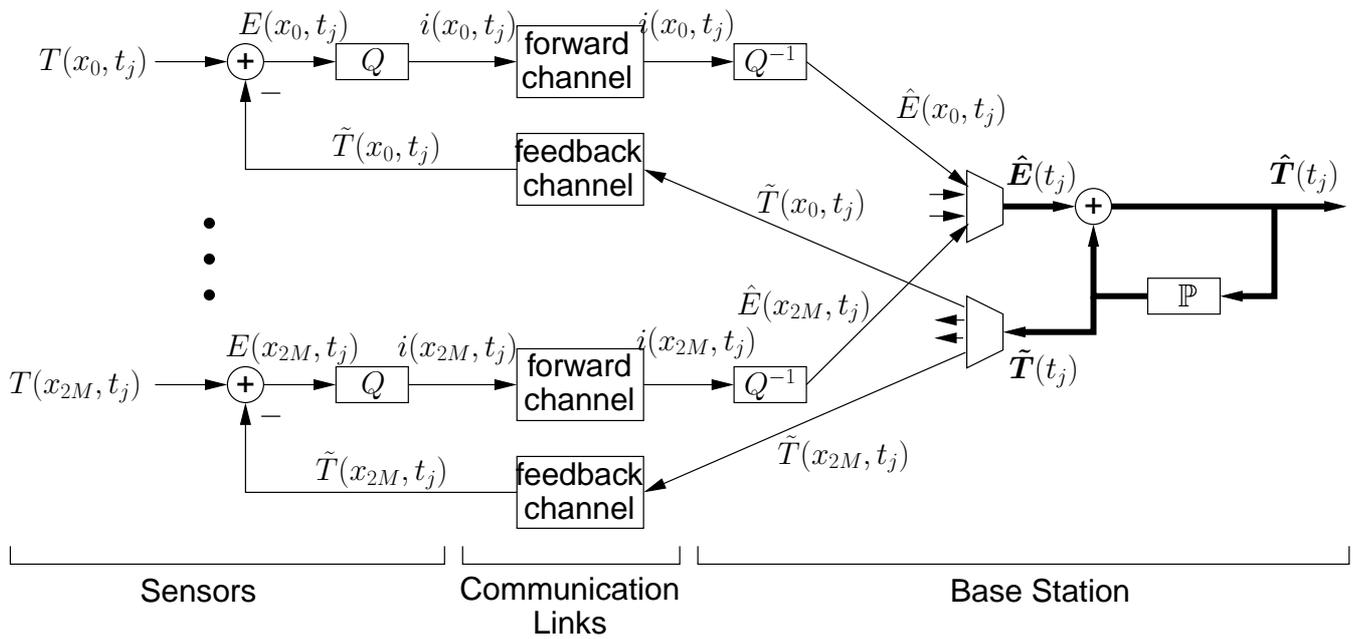


Figure 10: DPCM based system with feedback from BS.

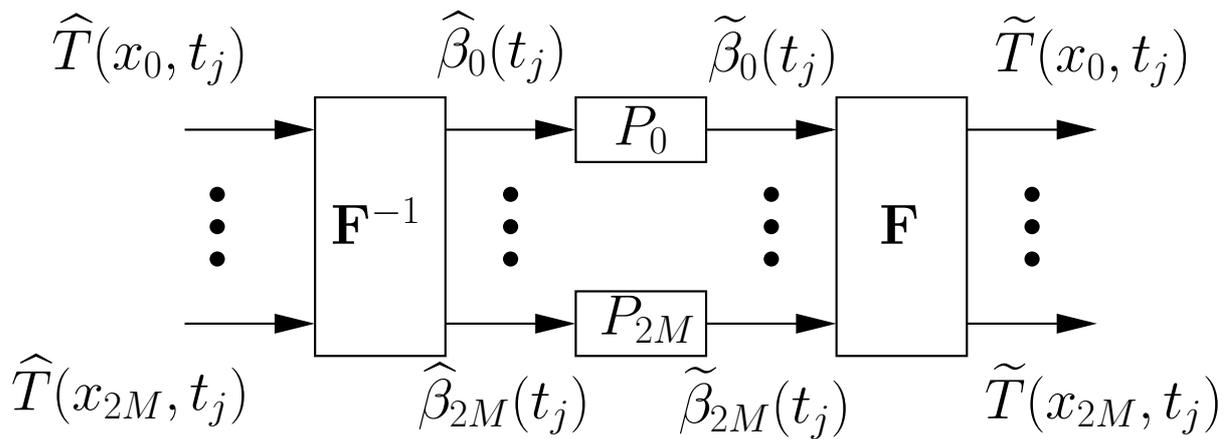


Figure 11: Prediction at the Base Station.

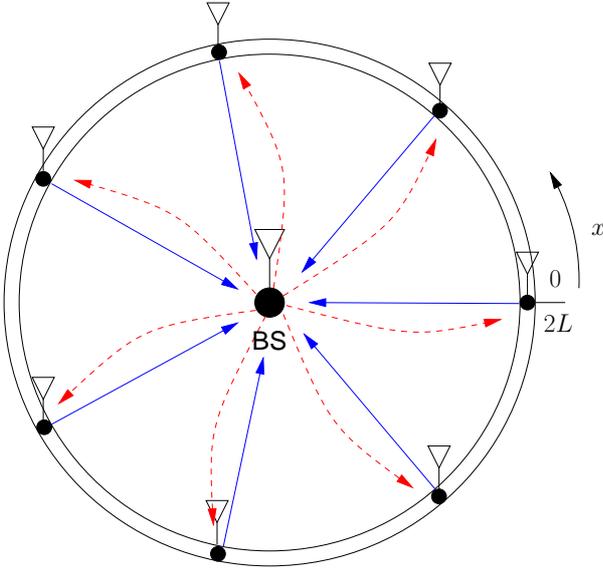


Figure 12: Feedback from Base Station.

We observe that there are two terms in (16), the first one consisting of a diffusion of the current state $\hat{\beta}_m(t_j)$ and the other one corresponding to an innovation term, which involves each source process $g_m(t)$ affecting the system during the interval $[t_j, t_{j+1}]$. Our approach is to base our prediction only on the first term, that is, our predicted $\tilde{\beta}_m(t_{j+1})$, is given by:

$$\tilde{\beta}_m(t_{j+1}) = \hat{\beta}_m(t_j)e^{-c_m}, \quad m = 0, \dots, 2M \quad (17)$$

Once this prediction is performed, the BS obtains the prediction for the temperature vector $\tilde{\mathbf{T}}(t_{j+1}) = \mathbf{F}\tilde{\beta}(t_{j+1})$, and sends to the n -th sensor the corresponding predicted value $\tilde{T}(x_n, t_{j+1})$. The prediction steps are illustrated in Fig. 11. Then, at time t_{j+1} , the n -th sensor reads the real temperature value $T(x_n, t_{j+1})$ and quantizes the prediction error $E(x_n, t_{j+1}) = T(x_n, t_{j+1}) - \tilde{T}(x_n, t_{j+1})$, getting the quantized output $\hat{E}(x_n, t_{j+1})$, with a scalar quantizer Q . Next, all the sensors send their quantized prediction errors to the BS, which obtains a reconstruction for each temperature value $\hat{T}(t_{j+1})$ by adding these quantized prediction errors to the previously predicted values, that is,

$$\hat{T}(x_n, t_{j+1}) = \tilde{T}(x_n, t_{j+1}) + \hat{E}(x_n, t_{j+1}), \quad n = 0, \dots, 2M \quad (18)$$

Then, the BS starts performing the prediction operation again and the whole process is repeated. Therefore, the system is very similar to that of a closed-loop DPCM where the encoders are the sensors and the decoder is the BS, and importantly, the prediction operation makes use of the heat equation and the conductivity properties of the heat conductor ring.

The following remarks are in order:

1. Notice that although the sensors and the BS are physically separated, we have a closed-loop DPCM system because of the existence of a perfect feedback channel from the BS to the sensors.

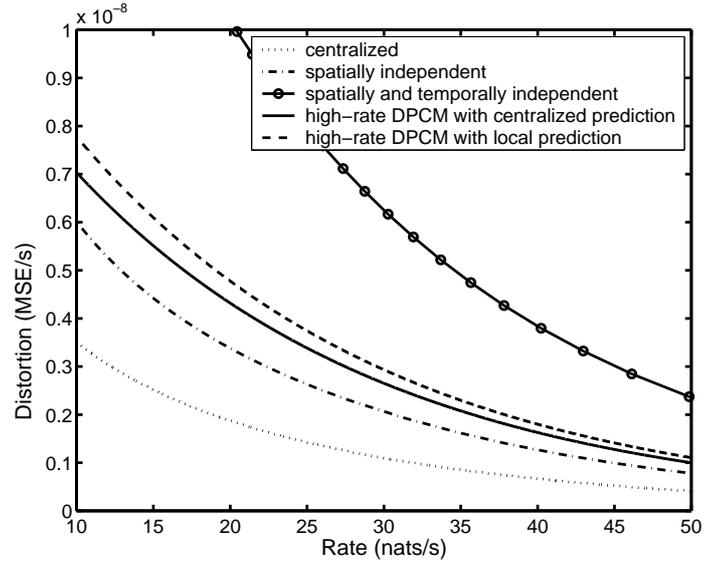


Figure 13: Comparison at high rates for $G_m = 1$, $m = 1, \dots, M$, $L = 0.1$, $M = 20$, and the material is Silver: $\alpha = 17.004 \cdot 10^{-5}$, $\mu = 10^4$. Time scale: 0.1 sec.

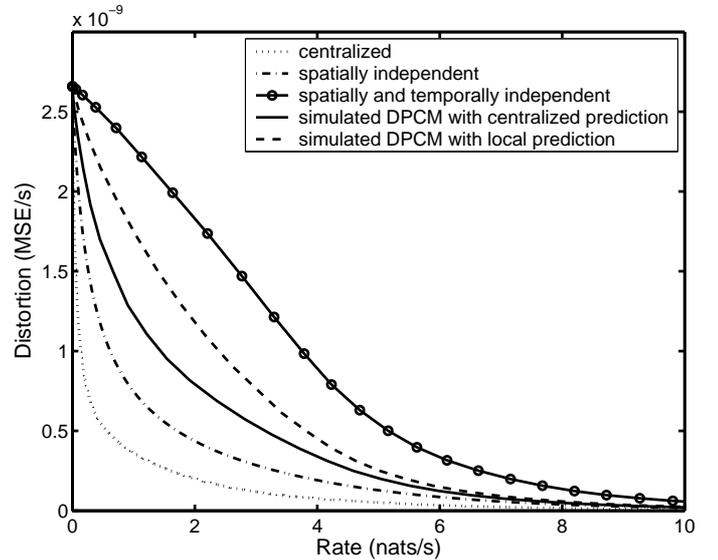


Figure 14: Comparison at low-to-middle rates for $M = 2$, $G_m = 1$, $m = 1, 2$, $L = 0.01$, and the material is Silver: $\alpha = 17.004 \cdot 10^{-5}$, $\mu = 10^4$. Time scale: 0.1 sec.

2. It can be shown that, if instead of using the prediction equation (17) inspired from Green's Theorem, one tries to obtain the optimal (infinite-length) discrete-time causal *LMMSE* predictor filter for the discrete-time sampled m -th process $\beta_m(t_j)$, which *theoretically* exists because the PSD $S_{\beta_m}(\omega)$ satisfies the Paley-Wiener criterion [12] (regular discrete-time process), the problem is that, in z -transform domain, $S_{\beta_m}(z)$ is a non-rational function in z , which results ultimately in a very complicated discrete-time filter to be built in terms of adders and delays [10].
3. The prediction given by (17) is actually the optimal continuous-time causal *LMMSE* predictor for a continuous-time (non-bandlimited) Gauss-Markov source having a PSD with the same shape as the PSD of $\beta_m(t)$, but without the bandlimitedness. Thus, the prediction given by the Green's Theorem makes use only of the immediate previous sample as in a Gauss-Markov process. Numerical results show that the performance obtained with the one-tap predictors is actually very close to the performance of the optimal predictors. This is because each PSD $S_{\beta_m}(\omega)$ is actually very close to the PSD of a first-order discrete-time autoregressive source.
4. As it happens in classical DPCM [10], at high rates, the closed-loop system becomes basically equivalent to the open-loop system, where the prediction is performed from the true past samples. In this case, the distortion only comes from the quantization of the innovation process given in (17) and it is possible to calculate exactly its PSD and the associated theoretical Rate-Distortion curve for the temperature process when encoding is performed with DPCM, using similar methods to the ones used in Section 5.

6.2 Local Prediction at the Sensors

In the second system, feedback from BS is not allowed and the n -th sensor, independently, performs a closed-loop DPCM system by predicting over time and where the corresponding observed sampled process $T(x_n, t_j)$ has a PSD $S(\omega)$ which is given as in Theorem 2, $\forall n = 0, \dots, 2M$. Again, it can also be shown [9] that the optimal one-tap prediction filter provides a prediction gain very close to that of the (infinite-length) optimal prediction filter. Once again, the reason stems from the fact that the PSD $S(\omega)$ is very similar to the PSD of a first-order discrete-time autoregressive source.

Fig. 13 shows the different analytical rate-distortion curves at high rates (see [9] for details), and Fig. 14 shows the performance obtained by simulation for both prediction-based systems and where the quantization has been performed with a simple uniform quantizer. For comparison, we also show $R^c(D)$, $R^{s-i}(D)$ and $R^{st-i}(D)$ in Fig. 13-14, where in Fig. 14, the curve corresponding to $R^{st-i}(D)$ has been also obtained by simulation with the same uniform scalar quantizer used for the DPCM systems. As expected, the local DPCM system provides better performance than $R^{s-i}(D)$. On the other hand, the DPCM system with centralized prediction, since it makes use of temporal information from all the sensors, or equivalently, it makes use of the spatial correlation among sensors, is superior to the local DPCM system. However, the DPCM system with centralized prediction is

still inferior to $R^{s-i}(D)$ which involves infinite complexity (although the difference is less than the quantization sphere shaping gain). Obviously, as expected, the best performance is given by $R^c(D)$.

6.3 Remark: Nested codes with Side Information based on Prediction

In [9], it is shown that it is possible to design an appropriate (Wyner-Ziv) nested code by using as side information the (spatial-temporal) physics-based predictions made at the BS, so that each sensor only needs to transmit coset (binning) indices. Notice that each $T(x_n, t_j)$ is correlated with the predicted vector $\tilde{T}(t_j)$. Thus, this is a Wyner-Ziv problem with multiple side-information ($2M + 1$ predictions). This scheme avoids the need of feedback from BS. For details, see [9]. Part of our future work includes also computing (inner and outer) bounds on the performance of distributed source coding for the heat problem, without the use of feedback.

7. REFERENCES

- [1] P. Ishwar, A. Kumar and K. Ramchandran. Distributed sampling for dense sensor networks: "a bit-conservation principle". IPSN 2003.
- [2] D. Marco, E. J. Duarte-Melo, M. Liu and D. L. Neuhoff. On the many-to-one transport capacity of a dense wireless sensor network and the compressibility of its data. IPSN 2003.
- [3] J. V. Beck, K. D. Cole, A. Haji-Sheikh and B. Litkouhi. Heat Conduction Using Green's Functions. Series in Computational and Physical Processes in Mechanics and Thermal Sciences, W. J. Minkowycz and E. M. Sparrow, Eds., Hemisphere Publishing Corporation, 1992.
- [4] H. S. Carslaw and J. C. Jaeger. Conduction of Heat in Solids. Oxford University Press, 1959.
- [5] S. Pradhan. Distributed Source Coding Using Syndromes (DISCUS). PhD Thesis, University of California, Berkeley, 2001.
- [6] T. Cover and J. Thomas. Elements of Information Theory. John Wiley & Sons, Inc., 1991.
- [7] T. Berger. Multiterminal source coding. Lecture Notes presented at CISM Summer School on the Information Theory Approach to Communications, 1997.
- [8] T. Berger. Rate Distortion Theory (A Mathematical Basis for Data Compression). Prentice-Hall, 1971.
- [9] B. Beferull-Lozano, R. L. Konsbruck and M. Vetterli. Rate-Distortion and Communication for Physics Based Distributed Sensing. To be submitted to IEEE Trans. on Signal Processing.
- [10] A. Gersho and R. Gray. Vector Quantization and Signal Compression. Kluwer Academic Publishers, 1992.
- [11] N. S. Jayant and P. Noll. Digital Coding of Waveforms. Prentice-Hall, 1984.
- [12] B. Porat. Digital Processing of Random Signals. Prentice-Hall, 1994.