

On the Optimal Power-Distortion Region for Asymmetric Gaussian Sensor Networks with Fading*

Hamid Behroozi, Fady Alajaji and Tamás Linder

Department of Mathematics and Statistics, Queen's University, Kingston, Ontario, Canada, K7L 3N6

Email: {behroozi, fady, linder}@mast.queensu.ca

Abstract—We consider the estimation of a Gaussian source by a Gaussian sensor network where L distributed sensors transmit noisy observations of the source through a fading Gaussian multiple access channel (MAC) to a fusion center (FC). Since sensor power is usually limited, our goal is to characterize the optimal tradeoff between the transmission cost, i.e., the power vector $\mathbf{P} = (P_1, P_2, \dots, P_L)$, and the average estimation distortion, D . We focus on asymmetric fading sensor networks in which the sensors have differing signal to noise ratios and transmission powers. We present necessary and sufficient conditions for the achievability of $(L + 1)$ -tuples $(P_1, P_2, \dots, P_L, D)$. For a symmetric Gaussian sensor network with deterministic and equal-magnitude fading, we derive the optimal power-distortion tradeoff. We also provide an achievable power-distortion region for the asymmetric sensor network with deterministic fading by analyzing the transmission of scaled versions of vector-quantized observations. We show that some of the power-distortion tuples achievable by this scheme are not achievable via an uncoded system.

I. INTRODUCTION

We consider the estimation of a memoryless Gaussian source by a Gaussian sensor network where L sensors observe the source signal X corrupted by additive independent noise. The overall system is depicted in Fig. 1. The sensors communicate information about their observations through a fading Gaussian MAC to a single FC. The fading coefficients are not known by the encoders but are available at the FC. The encoders are distributed and cannot cooperate to exploit the correlation between their inputs. Each encoder is subject to a transmission cost constraint. The FC aims to reconstruct the main source, X at the smallest cost in the communication link. Our interest lies in determining the optimal power-distortion region, with the fidelity of estimation at the FC measured by the mean squared-error (MSE) distortion.

In the recent work of Gastpar [1], it is proved that uncoded transmission is exactly optimal for symmetric Gaussian sensor networks with a finite number of sensors and no fading. By uncoded transmission we mean scaling the encoder input subject to the channel power constraint and transmitting without explicit channel coding. In the case of deterministic fading, lower and upper bounds on the minimum distortion are presented in [1], and for random fading, bounds are also presented in [2]. The optimality of uncoded transmission in some other multi-user communication systems was recently shown in [3], [4]. For the sensor network we consider, a joint source-channel coding (JSCC) approach similar to that of [3] is presented in [5] where each sensor transmits directly

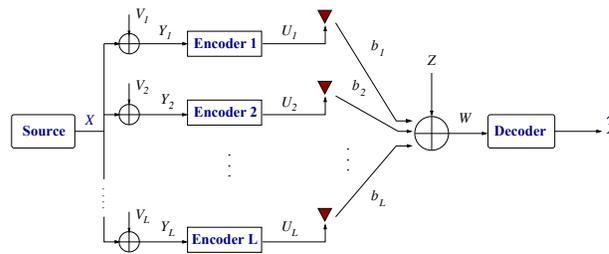


Figure 1. A Gaussian sensor network with fading.

a scaled version of the encoder input. In particular, the authors of [5] compare the minimum total power consumptions of three transmission schemes: separate source and channel coding, uncoded and JSCC. They demonstrate that the optimality of the uncoded scheme in symmetric sensor networks does not always carry over to asymmetric sensor networks. In [6], Bross *et al.* study the optimality of sending a Gaussian source over an average-power limited additive white Gaussian noise (AWGN) channel by transmitting a linear combination of the source sequence and its quantized version using a high dimensional Gaussian vector quantizer. It is shown that the proposed scheme is asymptotically optimal as the quantizer's dimension tends to infinity. However, at least three important issues remain unknown: (i) What is the optimal power-distortion tradeoff in an asymmetric Gaussian sensor network with deterministic fading? (ii) What is the optimal coding strategy, in the sense of achieving the optimal power-distortion tradeoff? (iii) In the presence of random fading, what is the optimal power-distortion tradeoff and what is the optimal transmission scheme?

In this work, we provide partial answers to the above questions by establishing a necessary condition for the achievability of all transmission cost-distortion tuples $(P_1, P_2, \dots, P_L, D)$ in the asymmetric Gaussian sensor network with random fading. We also analyze the uncoded transmission scheme and provide a sufficient condition for the achievability of $(P_1, P_2, \dots, P_L, D)$. Finally, by analyzing the transmission of vector-quantized sequences in an asymmetric sensor network with deterministic fading [3], an achievable power-distortion region is provided under deterministic fading. By examining some asymmetric cases, we show that there are some points in this achievable region that are not achievable by the uncoded transmission scheme.

The remainder of this paper is organized as follows. In Section II, we present the system model and problem statement. Section III provides the necessary and sufficient conditions for the achievability of the power-distortion tuples

*This work was supported in part by NSERC of Canada.

$(P_1, P_2, \dots, P_L, D)$. Special cases of the system are also analyzed. Conclusions are presented in Section IV.

II. PROBLEM STATEMENT

We consider a simple Gaussian sensor network, illustrated in Fig. 1, where a team of L sensors observe independent noisy versions of the memoryless Gaussian source $\{X(t)\}_{t=1}^{\infty}$, represented by the sequences $\{Y_i(t)\}_{t=1}^{\infty}$ for $i = 1, 2, \dots, L$. For each observation time $t = 1, 2, 3, \dots$

$$Y_i(t) = X(t) + V_i(t), \quad i = 1, \dots, L \quad (1)$$

where $X(t) \sim \mathcal{N}(0, \sigma_X^2)$ are independent and identically distributed (i.i.d.) over t , $V_i(t) \sim \mathcal{N}(0, N_i)$, $i = 1, \dots, L$ are independently distributed over i and t , and are independent of $X(t)$. Thus, the random variables $Y_i(t)$ for $i = 1, \dots, L$ are conditionally independent given the source $X(t)$. We represent the first n instances of $\{X(t)\}_{t=1}^{\infty}$ and $\{Y_i(t)\}_{t=1}^{\infty}$ by the data sequences $X^n = (X(1), X(2), \dots, X(n))$ and $Y_i^n = (Y_i(1), Y_i(2), \dots, Y_i(n))$, respectively. The correlated sources Y_i are not co-located and their observers cannot cooperate to directly exploit their correlation. Instead, the sequences Y_i^n are separately encoded to $\varphi_i(Y_i^n) = U_i^n$ where the encoder functions are defined as

$$\varphi_i : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad i = 1, 2, \dots, L \quad (2)$$

The transmitted sequences U_i^n are each average-power limited to P_i , i.e.,

$$\frac{1}{n} \sum_{t=1}^n E[|U_i(t)|^2] \leq P_i, \quad i = 1, 2, \dots, L \quad (3)$$

The sensors communicate the coded sequences to the decoder through a multiple-access fading channel. In fact, each transmitted signal is multiplied by a real-valued fading random variable, b_i , for $i = 1, 2, \dots, L$ which are not known by the encoders but are available to the decoder. Equivalently, the channel output consists of the pair (W^n, \mathbf{b}^n) where \mathbf{b}^n denotes $(b_1^n, b_2^n, \dots, b_L^n)$, and the distribution of \mathbf{b}^n is known at the transmitter. The fading coefficients have non-zero mean and are independent of each other, of the U_i random variables and of the channel noise Z . The time- t output of the channel is given by

$$W(t) = \sum_{i=1}^L b_i(t)U_i(t) + Z(t), \quad (4)$$

where $Z(t) \sim \mathcal{N}(0, \sigma_Z^2)$ is i.i.d. over t and is independent of $U_i(t)$ and $V_i(t)$. Based on the channel output $W^n \in \mathbb{R}^n$, the FC makes an estimate of the main source X^n as \hat{X}^n . The measure of fidelity between X^n and \hat{X}^n is the average distortion criterion, i.e., $\Delta = \frac{1}{n} E[\sum_{j=1}^n d(X(j), \hat{X}(j))]$ where $d(X(j), \hat{X}(j))$ is the MSE distortion measure. The reconstructed signal can be described by $\hat{X}^n = \psi(W^n, \mathbf{b}^n)$, where the decoder function is a mapping

$$\psi : \mathbb{R}^n \times \mathbb{R}^{nL} \rightarrow \mathbb{R}^n. \quad (5)$$

Let $\mathcal{F}^{(n)}(P_1, P_2, \dots, P_L)$ denote all encoder and decoder functions $(\varphi_1, \dots, \varphi_L, \psi)$ that satisfy (2)-(5). For a particular coding scheme $(\varphi_1, \dots, \varphi_L, \psi)$, the performance is determined by the cost vector $\mathbf{P} = (P_1, P_2, \dots, P_L)$ and the incurred distortion Δ . For any target distortion $D \geq 0$, the power-distortion region is defined in [7] as

$$\mathcal{P}(D) = \{(P_1, P_2, \dots, P_L) \mid (\mathbf{P}, D) \text{ is achievable}\}.$$

where power-distortion pair (\mathbf{P}, D) is achievable if for any $\delta > 0$ and any $n \geq n_0(\delta)$ there exists a sequence of $(\varphi_1, \dots, \varphi_L, \psi) \in \mathcal{F}^{(n)}(P_1, P_2, \dots, P_L)$ such that $\Delta \leq D + \delta$. Our aim is to investigate the power-distortion region of this fading Gaussian sensor network and present lower and upper bounds on the minimum achievable distortion, D_{min} , for a fixed \mathbf{P} .

III. INFORMATION-THEORETIC BOUNDS ON THE POWER-DISTORTION TRADEOFF

We present necessary and sufficient conditions for the achievability of $(P_1, P_2, \dots, P_L, D)$. Our necessary condition is based on analyzing the remote source coding scenario, where the sensor observations are given to one common encoder. Two sufficient conditions are provided based on analyzing two coding schemes: 1) uncoded transmission which is the transmission of scaled versions of the sensors observations and 2) transmission of scaled versions of vector-quantized observation sequences. Both schemes are JSCC schemes in which the scaling factors are used so that the required average transmission powers are satisfied.

A. Necessary Condition

Theorem 1: A necessary condition for the achievability of $(P_1, P_2, \dots, P_L, D)$ is that

$$D \geq D_l, \quad (6)$$

where

$$D_l = D_0^* \left(1 + \frac{\sigma_X^2 \sigma_Z^2 \sum_{i=1}^L \frac{1}{N_i}}{\sigma_Z^2 + \Omega_r} \right), D_0^* = \left(\frac{1}{\sigma_X^2} + \sum_{i=1}^L \frac{1}{N_i} \right)^{-1},$$

$$\Omega_r = \sum_{i=1}^L P_i E[|b_i|^2] + 2\sigma_X^2 \sum_{i=1}^L \sum_{j>i}^L E[b_i] E[b_j] \sqrt{\frac{P_i P_j}{(\sigma_X^2 + N_i)(\sigma_X^2 + N_j)}}, \quad (7)$$

and the expectations are taken with respect to the distribution of the fading random variables b_i .

Proof: By considering a (\mathbf{P}, D) achievable code $(\varphi_1, \dots, \varphi_L, \psi)$, this necessary condition follows from two information-theoretic concepts: the data processing inequality and remote source coding. By using the data processing inequality, the mutual information between the observation vector and the vector of the source estimate, $I(Y_1^n, Y_2^n, \dots, Y_L^n; \hat{X}^n) = I(\mathbf{Y}^n; \hat{X}^n)$, can be upper bounded by the mutual information between the transmitted waveforms and the received waveform, $I(U_1^n, U_2^n, \dots, U_L^n; W^n) = I(\mathbf{U}^n; W^n)$. Specifically, since the decoder knows the fading coefficients, the channel output is the pair (W^n, \mathbf{b}^n) , and thus

$$\begin{aligned} I(\mathbf{U}^n; (W^n, \mathbf{b}^n)) &= I(\mathbf{U}^n; W^n \mid \mathbf{b}^n) \\ &= E_b [I(\mathbf{U}^n; W^n \mid \mathbf{b}^n = b^n)]. \end{aligned} \quad (8)$$

The conditional mutual information can be obtained as

$$\frac{1}{n} I(\mathbf{U}^n; W^n \mid \mathbf{b}^n = b^n) \stackrel{(a)}{=} \frac{1}{n} h(W^n \mid \mathbf{b}^n = b^n) - \frac{1}{n} h(Z^n), \quad (9)$$

where (a) follows from the fact that U_i^n ($i = 1, \dots, L$) are independent of Z^n and $h(\cdot)$ denotes differential entropy. Let $\underline{b}(k)$ denotes the L -tuple $(b_1(k), b_2(k), \dots, b_L(k))$. We will upper bound $\frac{1}{n}h(W^n | \mathbf{b}^n = \mathbf{b}^n)$ using the bound

$$\begin{aligned} \text{Var}(W(k) | \underline{b}(k)) &= \text{Var}\left(\sum_{i=1}^L b_i(k)U_i(k) + Z(k)\right) \\ &\stackrel{(c)}{\leq} \left\{ \sum_{i=1}^L |b_i|^2 \sigma_{i,k}^2 + 2 \sum_{i=1}^L \sum_{j>i}^L b_i b_j \rho_{Y_i(k)Y_j(k)} \sqrt{\sigma_{i,k}^2 \sigma_{j,k}^2} + \sigma_Z^2 \right\} \end{aligned} \quad (10)$$

where $U_i(k)$ is the k -th component of $U_i^n = \varphi_i(Y_i^n)$, $\sigma_{i,k}^2 = E[|U_i(k)|^2]$ and ρ_{XY} denotes the correlation coefficient between X and Y . Inequality (c) follows from two known results in the literature: 1) The maximum correlation coefficient between any two finite-variance functions of two jointly Gaussian random variables is equal to the correlation coefficient between the original random variables [8]. 2) The correlation coefficient between any two finite-variance functions of the correlated data sequences (with i.i.d. elements) cannot be greater than the maximum correlation coefficient between any two elements of both sequences [9], [10]. From (1), we have $\rho_{Y_i Y_j} = \frac{\sigma_X^2}{\sqrt{(\sigma_X^2 + N_i)(\sigma_X^2 + N_j)}}$. Substituting this correlation coefficient in (10) and using the power constraints of (3) for the transmitted coded sequences U_i^n , and also using the Cauchy-Schwarz inequality, we obtain

$$\frac{1}{n} \sum_{k=1}^n \text{Var}(W(k) | \underline{b}(k)) \leq \Omega_{r|b} + \sigma_Z^2,$$

where

$$\Omega_{r|b} = 2\sigma_X^2 \sum_{i=1}^L \sum_{j>i}^L b_i b_j \sqrt{\frac{P_i P_j}{(\sigma_X^2 + N_i)(\sigma_X^2 + N_j)}} + \sum_{i=1}^L P_i |b_i|^2. \quad (11)$$

Since i.i.d. random variables with normal distribution maximize the differential entropy subject to the sum of the variances [11], we obtain that $\frac{1}{n}h(W^n | \mathbf{b}^n = \mathbf{b}^n) \leq \frac{1}{2} \log \{2\pi e (\Omega_{r|b} + \sigma_Z^2)\}$. Thus, (9) yields the bound $\frac{1}{n}I(\mathbf{U}^n; W^n | \mathbf{b}^n = \mathbf{b}^n) \leq \frac{1}{2} \log_2 \left(1 + \frac{\Omega_{r|b}}{\sigma_Z^2}\right)$, and then using (8), we obtain

$$\frac{1}{n}I(\mathbf{U}^n; (W^n, \mathbf{b}^n)) = E_b \left[\frac{1}{2} \log_2 \left(1 + \frac{\Omega_{r|b}}{\sigma_Z^2}\right) \right]. \quad (12)$$

By Jensen's inequality [11], (12) is upper bounded by $\frac{1}{2} \log_2 \left(1 + \frac{E_b[\Omega_{r|b}]}{\sigma_Z^2}\right)$. Using the data processing inequality,

$$\frac{1}{n}I(\mathbf{Y}^n; \widehat{X}^n) \leq \frac{1}{2} \log_2 \left(1 + \frac{\Omega_r}{\sigma_Z^2}\right), \quad (13)$$

where $\Omega_r = E_b[\Omega_{r|b}]$ which is given in (7) using the fact that the b_i are independent.

On the other hand, a lower bound on $\frac{1}{n}I(\mathbf{Y}^n; \widehat{X}^n)$ can be obtained by considering the "remote source coding problem" introduced in [12], where the sensors are allowed to collaborate, i.e., the L -tuple $\underline{Y} = (Y_1, \dots, Y_L)$ is observed by one common encoder. For any coding system that attains an average distortion D , applying the converse of the rate-distortion theorem yields that $\frac{1}{n}I(\mathbf{Y}^n; \widehat{X}^n) \geq \min \frac{1}{n}I(\mathbf{Y}^n; \widehat{X}^n)$ where the mini-

um is over all conditional distributions $p(\widehat{x} | y_1, \dots, y_L)$ satisfying $E[(X - \widehat{X})^2] \leq D$. We can calculate $\min \frac{1}{n}I(\mathbf{Y}^n; \widehat{X}^n)$ [13], [14] and obtain the following rate-distortion function:

$$R_X^{rem}(D) = \frac{1}{2} \log_2 \left(\frac{\sigma_X^2 \sum_{i=1}^L \frac{1}{N_i}}{D \frac{1}{D_0^*} - \frac{1}{D}} \right), \quad D_0^* < D < \sigma_X^2.$$

Combining this lower bound on $\frac{1}{n}I(\mathbf{Y}^n; \widehat{X}^n)$ with the upper bound of (13) (which is the capacity of the AWGN channel with a power constraint of Ω_r) yields the following result:

$$R_X^{rem}(D) \leq \frac{1}{n}I(\mathbf{Y}^n; \widehat{X}^n) \leq \frac{1}{n}I(\mathbf{U}^n; W^n) \leq \frac{1}{2} \log_2 \left(1 + \frac{\Omega_r}{\sigma_Z^2}\right). \quad \blacksquare$$

Therefore, using any transmission scheme, the incurred distortion must satisfy (6). As a result, the LHS of (6) provides a lower bound for the minimum achievable distortion, i.e., D_{min} . Comparing our general lower bound for the estimation distortion with the bound presented in [2], we can show in the following that our bound is tighter. The lower bound presented in [2], which is for the case where the sensors observations have the same noise level $N_1 = \dots = N_L = N$, can be expressed as

$$D_{lower}(L, P_{tot}) = D_0^* \left(1 + \frac{\sigma_X^2 \sigma_Z^2 \frac{L}{N}}{\sigma_Z^2 + P_{tot} \sum_{i=1}^L E[|b_i|^2]} \right), \quad (14)$$

where it is assumed that there is a total power constraint in the communication channel, i.e., $\frac{1}{n} \sum_{t=1}^n \sum_{i=1}^L E[|U_i(t)|^2] \leq P_{tot}$, and P_{tot} denotes the average total sensor power available per observation vector (U_1, U_2, \dots, U_L) . Let us denote $\sigma_1^2 = \sigma_Z^2 + \sum_{i=1}^L P_i E[|b_i|^2] + 2 \frac{\sigma_X^2}{\sigma_X^2 + N} \sum_{i=1}^L \sum_{j>i}^L E[b_i] E[b_j] \sqrt{P_i P_j}$ and $\sigma_2^2 = \sigma_Z^2 + P_{tot} \sum_{i=1}^L E[|b_i|^2]$. By comparing (6) and (14), we observe that the only difference is in the denominator, i.e., our lower bound is in the form $D_l = D_0^* \left(1 + \frac{\sigma_X^2 \sigma_Z^2 \frac{L}{N}}{\sigma_1^2}\right)$ and $D_{lower}(L, P_{tot}) = D_0^* \left(1 + \frac{\sigma_X^2 \sigma_Z^2 \frac{L}{N}}{\sigma_2^2}\right)$. We can show that the maximum of the denominator in (6), i.e., σ_1^2 , is strictly less than the denominator in (14), σ_2^2 , and therefore it is a tighter lower bound. Specifically, for a given total power, σ_1^2 is maximized when $P_1 = P_2 = \dots = P_L = \frac{P_{tot}}{L}$. Thus, the maximum of σ_1^2 can be expressed as

$$\sigma_Z^2 + \frac{P_{tot}}{L} \sum_{i=1}^L E[|b_i|^2] + P_{tot}(L-1) \frac{\sigma_X^2}{\sigma_X^2 + N} \bar{b}^2,$$

where $E[b_i] = \bar{b}$. It can be shown that $\sigma_2^2 - \sigma_1^2$ is equal to

$$\frac{P_{tot}(1 - \frac{1}{L})}{\sigma_X^2 + N} \left(N \sum_{i=1}^L E[|b_i|^2] + \sigma_X^2 \sum_{i=1}^L \text{Var}(b_i) \right),$$

which is strictly greater than zero. Therefore, our lower bound in (6) on D_{min} is tighter than (14); the improvement is mainly due to the maximization of the correlation coefficients as explained in the discussion following (10).

Examples:

We assume that $E[|b_1|^2] = E[|b_2|^2] = 4$, $E[b_1] = E[b_2] = 1$.

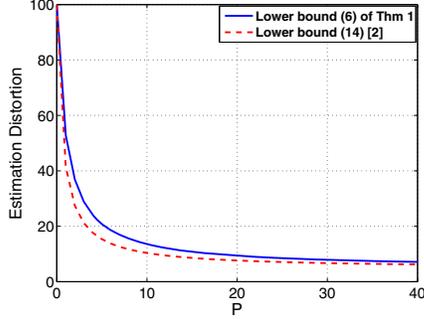


Figure 2. The estimation distortion v.s. the transmission power, $P_1 = P_2 = P$. We assume that $N_1 = N_2 = 10$, $\sigma_X^2 = 100$ and $\sigma_Z^2 = 10$.

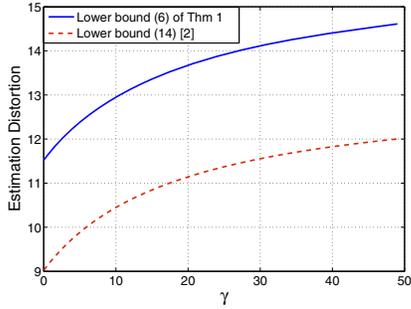


Figure 3. The estimation distortion v.s. the value of $\gamma = N_2 - N_1$. We assume that $P_1 = P_2 = 20$, $N_1 = 10$, $\sigma_X^2 = 100$ and $\sigma_Z^2 = 15$.

- We consider a *symmetric* sensor network consisting of two sensors and evaluate the resulting lower bounds numerically. The estimation distortion bound is plotted as a function of the power level $P_1 = P_2 = P$ in Fig. 2. We observe that the distortion is a decreasing function of P and that our lower bound is tighter than the lower bound presented in [2].
- We next consider an *asymmetric* sensor network. We assume that $N_2 = N_1 + \gamma$ and $P_1 = P_2 = 20$, and plot the estimation distortion bounds as a function of γ in Fig. 3. There is almost a fixed gap between our lower bound and the lower bound presented in [2].

B. Sufficient Condition: Uncoded Transmission

By analyzing the uncoded transmission in Gaussian sensor network, we next present a sufficient condition for the achievability of $(P_1, P_2, \dots, P_L, D)$. In this approach, each sensor transmits the scaled version of its observation, scaled to its power constraint, i.e., $U_i(t) = \sqrt{\frac{P_i}{\sigma_X^2 + N_i}} Y_i(t)$. The received signal at the FC is then given by

$$W(t) = \sum_{i=1}^L \left\{ \sqrt{\frac{P_i}{\sigma_X^2 + N_i}} b_i(t) (X(t) + V_i(t)) \right\} + Z(t).$$

Since the encoding is memoryless, the optimal estimator is the minimum mean squared error (MMSE) estimator of $X(t)$ from the received signal $W(t)$. By evaluating the resulting MSE distortion, we obtain a sufficient condition for the achievability of $(P_1, P_2, \dots, P_L, D)$, which is summarized in the next lemma.

Lemma 1: For the $(L+1)$ -tuple $(P_1, P_2, \dots, P_L, D)$ to be achievable it suffices that $D \geq D_{\text{uncoded}}$ where

$$D_{\text{uncoded}} = \sigma_X^2 E \left(\frac{\sigma_Z^2 + \sum_{i=1}^L \frac{P_i N_i |b_i|^2}{\sigma_X^2 + N_i}}{\sigma_Z^2 + \Omega_r |b|} \right), \quad (15)$$

$\Omega_r |b|$ is given in (11) and the expectation is with respect to fading random variables, b_i .

Since D_{uncoded} in (15) is the achievable distortion by uncoded transmission, it is an upper bound on D_{min} ; thus $D_{\text{min}} \leq D_{\text{uncoded}}$.

Using some algebraic manipulations, it is not hard to show that if $\text{Var}(b_i) = 0$, the lower bound in (6) and the upper bound in (15) agree. This means that we can obtain the optimal performance if the fading coefficients stay constant over the duration of transmission. This may apply to situations where the network conditions change very slowly. We next investigate the Gaussian sensor network with such deterministic channel gains.

C. Deterministic Fading System

Assume that the fading coefficients b_i are fixed known constants. Our lower and upper bounds can be expressed as follows:

$$D_{\text{min}} \geq D_l = D_0^* \left(1 + \frac{\sigma_X^2 \sigma_Z^2 \sum_{i=1}^L \frac{1}{N_i}}{\sigma_Z^2 + \Omega_r} \right), \quad (16)$$

$$D_{\text{min}} \leq D_{\text{uncoded}} = \left(\frac{\sigma_X^2}{\sigma_Z^2 + \Omega_r} \right) \left(\sigma_Z^2 + \sum_{i=1}^L \frac{P_i N_i |b_i|^2}{\sigma_X^2 + N_i} \right), \quad (17)$$

where $\Omega_r = \sum_{i=1}^L \frac{P_i N_i |b_i|^2}{\sigma_X^2 + N_i} + \left(\sum_{i=1}^L b_i \sqrt{\frac{P_i \sigma_X^2}{\sigma_X^2 + N_i}} \right)^2$. In the symmetric case, where $N_1 = N_2 = \dots = N_L = N$ and $P_1 = P_2 = \dots = P_L = P$, the lower bound (16) and the upper bound (17) coincide if $b_1 = b_2 = \dots = b_L = b$. Hence, we obtain the ‘‘optimal performance theoretically achievable’’ for the symmetric Gaussian sensor network under deterministic and identical fading. This is the same result as recently established by Gastpar in [1] where $b_1 = b_2 = \dots = b_L = 1$.

D. Sufficient Condition with Deterministic Fading: Transmission of Vector-Quantized Observation Sequences

Another sufficient condition can be obtained by studying a JSCC scheme for the system with deterministic fading where the encoding functions are independent vector quantizers [3], [5]. We herein assume that $L = 2$ for the sake of simplicity.

Theorem 2: The triple (P_1, P_2, D) is achievable if there exist rates $R_1 > 0$ and $R_2 > 0$ such that the following conditions hold:

$$\begin{aligned} R_1 &\leq \frac{1}{2} \log_2 \left(\frac{P_1 |b_1|^2 (1 - \tilde{\rho}^2) + \sigma_Z^2}{\sigma_Z^2 (1 - \tilde{\rho}^2)} \right) \\ R_2 &\leq \frac{1}{2} \log_2 \left(\frac{P_2 |b_2|^2 (1 - \tilde{\rho}^2) + \sigma_Z^2}{\sigma_Z^2 (1 - \tilde{\rho}^2)} \right) \\ R_1 + R_2 &\leq \frac{1}{2} \log_2 \left(\frac{P_1 |b_1|^2 + P_2 |b_2|^2 + 2\tilde{\rho} b_1 b_2 \sqrt{P_1 P_2} + \sigma_Z^2}{\sigma_Z^2 (1 - \tilde{\rho}^2)} \right) \\ D &\geq \frac{D_1 D_2}{\sigma_X^2 (1 - \tilde{\rho}^2)} \end{aligned} \quad (18)$$

where

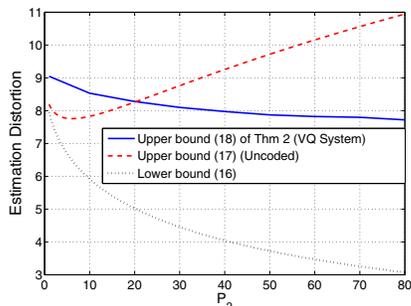


Figure 4. The estimation distortion versus the transmission power of the second sensor. We assume that $P_1 = 1$, $b_1 = 10$, $b_2 = 1$, $N_1 = 0.1$, $N_2 = 50$, $\sigma_X^2 = 100$, and $\sigma_Z^2 = 10$.

$$\tilde{\rho} = \sigma_X^2 \sqrt{\frac{(1 - 2^{-2R_1})(1 - 2^{-2R_2})}{(\sigma_X^2 + N_1)(\sigma_X^2 + N_2)}}, \quad (19)$$

and

$$D'_i = \frac{\sigma_X^2}{\sigma_X^2 + N_i} (N_i + \sigma_X^2 2^{-2R_i}), \quad i = 1, 2. \quad (20)$$

We only give a sketch of the proof, which is based on a scheme analyzed in [3]. The i^{th} encoder is a rate- R_i Gaussian vector quantizer with 2^{nR_i} codewords that lie on the sphere of approximate radius $\sqrt{n(\sigma_X^2 + N_i)(1 - 2^{-2R_i})}$. Encoder i uses nearest neighbor encoding and scales the resulting vector \mathbf{U}_i^* so that the power constraint P_i is satisfied. It is shown in [3] that the decoder can recover the pair $(\mathbf{U}_1^*, \mathbf{U}_2^*)$ with asymptotically vanishing error probability by using typicality decoding if R_1 and R_2 satisfy the conditions in (18). As a second step in reconstruction, the optimal linear estimator $\hat{\mathbf{X}} = \beta_1 \mathbf{U}_1^* + \beta_2 \mathbf{U}_2^*$ is formed. It is only a matter of some calculations to show that the resulting distortion can be arbitrarily close to $\frac{D'_1 D'_2}{\sigma_X^2 (1 - \tilde{\rho}^2)}$ where $\tilde{\rho}$ and D'_i are given in (19) and (20), respectively.

Remark 1: It is possible to show that power-distortion region of Theorem 2 contains points that are not achievable by the uncoded scheme. For example, consider a Gaussian sensor network with the parameters $\sigma_X^2 = 10$, $\sigma_Z^2 = 1$, $N_1 = 8$, $N_2 = 1$. The power-distortion triple $(P_1, P_2, D) = (20, 1, 3.188)$ is achievable by VQ-based JSCC. However, for these P_1 and P_2 , the best distortion that can be achieved by the uncoded scheme is 3.519. As another example, consider an asymmetric sensor network with fixed transmission power $P_1 = 1$ for the first sensor. The estimation distortion bounds are plotted against P_2 in Fig. 4. We observe that for small values of P_2 , uncoded transmission performs better than VQ-based transmission. In fact, since P_1 is fixed to 1, it is expected that uncoded scheme performs optimally in the symmetric case, where $P_1 = P_2$, and near optimally where $P_2 - P_1$ is small. But for larger values of P_2 , the VQ-based scheme achieves a lower estimation distortion at the decoder.

Consequently, uncoded transmission is not optimal for large values of P_2 . Since uncoded transmission is optimal in the symmetric network, we conjecture that the transmission of a linear combination of coded and uncoded sequences [6] is a close-to-optimal strategy in the asymmetric sensor network.

IV. CONCLUSIONS

In this paper, we considered distributed sensor networks, where L noisy observations of a memoryless Gaussian source are transmitted through a fading MAC to a decoder. The decoder wants to reconstruct the main source with an average distortion D at the smallest possible power consumption in the communication link. Our goal was to characterize the power-distortion region achievable by any coding strategy regardless of delay and complexity. We obtained a necessary condition for achievability of a given power-distortion tuple $(P_1, P_2, \dots, P_L, D)$. Also, by analyzing the uncoded transmission scheme we provided a sufficient condition for achievability of $(P_1, P_2, \dots, P_L, D)$. When specialized to the symmetric network with identical deterministic fading, we derived the optimal power-distortion tradeoff. We also provided an achievable power distortion region for the asymmetric network with deterministic fading by analyzing transmission of scaled versions of vector-quantized observations.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their helpful comments.

REFERENCES

- [1] M. Gastpar, "Uncoded transmission is exactly optimal for a simple Gaussian sensor network," in *Proc. 2007 Information Theory and Applications Workshop*, San Diego, CA, USA, Jan. 2007.
- [2] M. Gastpar and M. Vetterli, "Power, spatio-temporal bandwidth, and distortion in large sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 745–754, Apr. 2005.
- [3] A. Lapidoth and S. Tinguely, "Sending a bivariate Gaussian source over a Gaussian MAC," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Seattle, Washington, Jul. 2006, pp. 2124–2128.
- [4] S. Bross, A. Lapidoth, and S. Tinguely, "Broadcasting correlated Gaussians," arXiv:0710.5376v1[cs.IT], Oct. 2007.
- [5] S. Wei, R. Kannan, S. Iyengar, and N. Rao, "Energy efficient estimation of Gaussian sources over inhomogeneous Gaussian MAC channels," *IEEE Trans. Inf. Theory*, submitted for publication, Sept. 2007.
- [6] S. Bross, A. Lapidoth, and S. Tinguely, "Superimposed coded and uncoded transmissions of a Gaussian source over the Gaussian channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Seattle, Washington, Jul. 2006, pp. 2153–2155.
- [7] M. Gastpar and M. Vetterli, "Source-channel communication in sensor networks," *2nd International Workshop on Information Processing in Sensor Networks (IPSN'03)* (L. J. Guibas and F. Zhao, eds.), pp. 162–177, New York, NY: Lecture Notes in Computer Science, Springer, Apr. 2003.
- [8] H. O. Lancaster, "Some properties of the bivariate normal distribution considered in the form of a contingency table," *Biometrika*, vol. 44, pp. 289–292, 1957.
- [9] P. Csáki and J. Fisher, "On the general notion of maximal correlation," *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, vol. 8, pp. 27–51, 1963.
- [10] H. S. Witsenhausen, "On sequences of pairs of dependent random variables," *SIAM J. Appl. Math.*, vol. 28, no. 1, pp. 100–113, Jan. 1975.
- [11] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: 2nd Edition, John Wiley & Sons, 2006.
- [12] T. Berger, *Rate-distortion theory: A Mathematical Basis for Data Compression*. Englewood Cliffs, NJ: Prentice-Hall, 1971.
- [13] J. K. Wolf and J. Ziv, "Transmission of noisy information to a noisy receiver with minimum distortion," *IEEE Trans. Inf. Theory*, vol. 16, no. 4, pp. 406–411, Jul. 1970.
- [14] S. C. Draper, "Successive structuring of source coding algorithms for data fusion, buffering, and distribution in networks," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, Jun. 2002.