

# Gaussian Sensor Networks with Adversarial Nodes

Emrah Akyol, Kenneth Rose  
 {eakyol, rose}@ece.ucsb.edu  
 UC, Santa Barbara

Tamer Başar  
 basar1@illinois.edu  
 University of Illinois at Urbana-Champaign

**Abstract**—This paper studies a particular sensor network model which involves one single Gaussian source observed by many sensors, subject to additive independent Gaussian observation noise. Sensors communicate with the receiver over an additive Gaussian multiple access channel. The aim of the receiver is to reconstruct the underlying source with minimum mean squared error. The scenario of interest here is one where some of the sensors act as adversary (jammer): they strive to maximize distortion. We show that the ability of transmitter sensors to secretly agree on a random event, that is “coordination”, plays a key role in the analysis. Depending on the coordination capability of sensors and the receiver, we consider two problem settings. The first setting involves transmitters with “coordination” capabilities in the sense that all transmitters can use identical realization of randomized encoding for each transmission. In this case, the optimal strategy for the adversary sensors also requires coordination, where they all generate the same realization of independent and identically distributed Gaussian noise. In the second setting, the transmitter sensors are restricted to use fixed, deterministic encoders and this setting, which corresponds to a Stackelberg game, does not admit a saddle-point solution. We show that the optimal strategy for all sensors is uncoded communications where encoding functions of adversaries and transmitters are in opposite directions. For both settings, digital compression and communication is strictly suboptimal.

**Index Terms**—Sensor networks, game theory, uncoded communication, analog mappings, coordinated transmission

## I. INTRODUCTION

Distributed compression and communication over sensor networks has been an important problem, see e.g. [1] for an overview. Joint source-channel coding (JSCC) has certain advantages over separate source and channel coding, and several specific aspects of JSCC over sensor networks have been studied in previous work; see e.g., [2] and the references therein. In this paper, we extend the game theoretic analysis of the Gaussian test channel [3]–[6] to Gaussian sensor networks introduced by [7]. A particular extension of [7] to asymmetric sensor networks was studied in [8], [9]. Communication games within general multiple input-multiple output settings were considered in [10], [11].

In this paper, we consider two settings for the sensor network model illustrated in Figure 1 and explained in detail in Section II. The first  $M$  sensors (i.e., the transmitters) and the receiver constitute Player 1 (minimizer) and the remaining  $K$  sensors (i.e., the adversaries) constitute Player 2 (maximizer). This zero-sum game does not admit a saddle-point in pure strategies (fixed encoding functions), but admits one in mixed strategies (randomized functions). In the first setting,

This work was supported in part by the NSF under grants CCF-1016861, CCF-1118075 and CCF-1111342.

the transmitter sensors can use randomized encoders, i.e., all transmitters and the receiver agree on some (pseudo)random sequence, denoted as  $\{\gamma\}$  in the paper. We coin the term “coordination” for this capability and show that it has a pivotal role in the analysis and the implementation of optimal strategies for both the transmitter and adversarial sensors. In the first setting, we consider the more general case of mixed strategies and present the saddle-point solution in Theorem 1. In the second setting, encoding functions of transmitters are limited to the fixed mappings. This setting can be viewed as a Stackelberg game where Player 1 is the leader, restricted to pure strategies, and Player 2 is the follower, who observes Player 1’s choice of pure strategies and plays accordingly. We present in Theorem 2 the optimal strategies for this Stackelberg game, whose cost is strictly higher than the cost associated with the first setting. The sharp contrast between the two settings underlines the importance of “coordination” in sensor networks with adversarial nodes.

## II. PROBLEM DEFINITION

In general, lowercase letters (e.g.,  $x$ ) denote scalars, boldface lowercase (e.g.,  $\mathbf{x}$ ) vectors, uppercase (e.g.,  $U, X$ ) matrices and random variables, and boldface uppercase (e.g.,  $\mathbf{X}$ ) random vectors.  $\mathbb{E}(\cdot)$ ,  $\mathbb{P}(\cdot)$  and  $\mathbb{R}$  denote the expectation and probability operators, and the set of real numbers respectively.  $Bern(p)$  denotes the Bernoulli random variable, taking values 1 with probability  $p$  and  $-1$  with  $1-p$ . Gaussian distribution with mean  $\boldsymbol{\mu}$  and covariance matrix  $R$  is denoted as  $\mathcal{N}(\boldsymbol{\mu}, R)$ .

The sensor network model is illustrated in Figure 1. The underlying source  $\{S(i)\}$  is a sequence of i.i.d. real valued Gaussian random variables with zero mean and variance  $\sigma_S^2$ . Sensor  $m \in [1 : M+K]$  observes a sequence  $\{U_m(i)\}$  defined as

$$U_m(i) = S(i) + W_m(i), \quad (1)$$

where  $\{W_m(i)\}$  is a sequence of i.i.d. Gaussian random variables with zero mean and variance  $\sigma_{W_m}^2$ , independent of  $\{S(i)\}$ . Sensor  $m \in [1 : M+K]$  can apply arbitrary Borel measurable function  $g_m^N : \mathbb{R}^N \rightarrow \mathbb{R}$  to the observation sequence of length  $N$ ,  $\mathbf{U}_m$  so as to generate sequence of channel inputs  $X_m(i) = g_m^N(\mathbf{U}_m)$  under power constraint:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \mathbb{E}\{X_m^2(i)\} \leq P_m \quad (2)$$

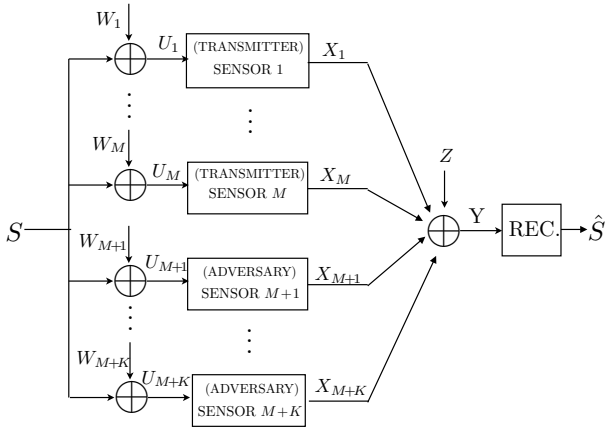


Fig. 1. The sensor network model.

The channel output is then given as

$$Y(i) = Z(i) + \sum_{j=1}^{M+K} X_j(i) \quad (3)$$

where  $\{Z(i)\}$  is a sequence of i.i.d. Gaussian random variables of zero mean and variance  $\sigma_Z^2$ , independent of  $\{S(i)\}$  and  $\{W_m(i)\}$ . The receiver applies a Borel measurable function  $h^N : \mathbb{R}^N \rightarrow \mathbb{R}$  to the received sequence  $\{Y(i)\}$  to minimize the cost, which is measured as mean squared error (MSE) between the underlying source  $S$  and the estimate at the receiver  $\hat{S}$  as

$$J(g_m^N(\cdot), h^N(\cdot)) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \mathbb{E}\{(S(i) - \hat{S}(i))^2\} \quad (4)$$

for  $m = 1, 2, \dots, M + K$ .

The transmitters  $g_m^N(\cdot)$  for  $m \in [1 : M]$  and the receiver  $h^N(\cdot)$  seek to minimize the cost (4) while the adversaries aim to maximize (4) by properly choosing  $g_k^N(\cdot)$  for  $k \in [M+1 : M+K]$ . We focus on the symmetric sensor and symmetric source where  $P_m = P_T$  and  $\sigma_{W_m}^2 = \sigma_{W_T}^2$ ,  $\forall m \in [1 : M]$  and  $\sigma_{W_k}^2 = \sigma_{W_T}^2$  and  $P_k = P_A$ ,  $\forall k \in [M+1 : M+K]$ .

A transmitter-receiver-adversarial policy  $(g_m^{N*}, g_k^{N*}, h^{N*})$  constitutes a saddle-point solution if it satisfies the pair of inequalities

$$J(g_m^{N*}, g_k^N, h^N) \leq J(g_m^{N*}, g_k^{N*}, h^{N*}) \leq J(g_m^N, g_k^{N*}, h^N) \quad (5)$$

### III. PROBLEM SETTING I

The first scenario is concerned with the setting where the transmitter sensors have the ability to *coordinate*, i.e., all transmitters and the receiver can agree on an i.i.d. sequence of random variables  $\{\gamma(i)\}$  generated, for example, by a side channel, the output of which is, however, not available to the adversarial sensors<sup>1</sup>. The ability of coordination allows transmitters and the receiver to agree on randomized encoding

<sup>1</sup>An alternative practical method to coordinate is to generate the identical pseudo-random numbers at each sensor, based on pre-determined seed.

mappings. Surprisingly, in this setting, the adversarial sensors also need to coordinate, i.e., agree on an i.i.d. random sequence, denoted as  $\{\theta(i)\}$ , to generate the optimal jamming strategy.

The saddle-point solution of this problem is presented in the following theorem.

*Theorem 1:* Under scenario 1, the optimal encoding function for the transmitters is randomized uncoded transmission:

$$X_m(i) = \gamma(i)\alpha_T U_m(i), \quad M \geq m \geq 1 \quad (6)$$

where  $\gamma(i)$  is i.i.d. Bernoulli ( $\frac{1}{2}$ ) over the alphabet  $\{-1, 1\}$

$$\gamma(i) \sim \text{Bern}\left(\frac{1}{2}\right).$$

The optimal jamming function (for adversarial sensors) is to generate i.i.d. Gaussian output

$$X_k(i) = \theta(i), \quad M + K \geq k \geq M + 1$$

where

$$\theta(i) \sim \mathcal{N}(0, P_A),$$

and is independent of the adversarial sensor input  $U_k(i)$ . The optimal receiver is the Bayesian estimator of  $S$  given  $Y$ , i.e.,

$$h(Y(i)) = \frac{M\alpha_T\sigma_S^2}{M\alpha_T^2\sigma_S^2 + M^2\alpha_T^2\sigma_{W_T}^2 + K^2P_A + \sigma_Z^2} Y(i). \quad (7)$$

Cost at this saddle-point is

$$J_2 = \sigma_S^2 \frac{M^2\alpha_T^2\sigma_{W_T}^2 + K^2P_A + \sigma_Z^2}{M\alpha_T^2\sigma_S^2 + M^2\alpha_T^2\sigma_{W_T}^2 + K^2P_A + \sigma_Z^2} \quad (8)$$

where  $\alpha_T = \sqrt{\frac{P_T}{\sigma_S^2 + \sigma_{W_T}^2}}$ .

*Proof:* We prove this result by verifying that the mappings in this theorem satisfy the saddle-point criterion given in (5), following the approach in [5].

RHS of (5): Suppose the policy of the adversarial sensors is given as in Theorem 1. Then, the communication system at hand becomes identical to the problem considered in [7], whose solution is uncoded communication with deterministic, linear encoders, i.e.,  $X_m(i) = \alpha_T U_m(i)$ . Any probabilistic encoder, given in the form of (6) (irrespective of the density of  $\gamma$ ) yield the same cost (8) with deterministic encoders and hence is optimal.

LHS of (5): Note that all the adversarial sensors must use the same jamming strategy to maximize the overall cost. Let us derive the overall cost conditioned on the realization of the transmitter mappings (i.e.,  $\gamma = 1$  and  $\gamma = -1$ ) used in conjunction with optimal linear decoders given in (7). If  $\gamma = 1$

$$D_1 = J_1 + \xi \mathbb{E}\{SX_k\} + \psi \mathbb{E}\{ZX_k\} \quad (9)$$

for some constants  $\xi, \psi$ , and similarly if  $\gamma = -1$

$$D_2 = J_1 - \xi \mathbb{E}\{SX_k\} - \psi \mathbb{E}\{ZX_k\} \quad (10)$$

where the overall cost is

$$D(i) = \mathbb{P}(\gamma(i) = 1)D_1 + \mathbb{P}(\gamma(i) = -1)D_2. \quad (11)$$

Clearly, for  $\gamma(i) \sim \text{Bern}(\frac{1}{2})$  the overall cost  $J_1$  is only a function of the second-order statistics of the adversarial outputs, irrespective of the distribution of  $\{\theta(i)\}$ , and hence the solution presented here is indeed a saddle-point. ■

*Corollary 1:* The solution in Theorem 1 is (almost surely) the unique solution for the transmitters, the adversaries (jammer) and the receiver.

*Proof:* We start by restating the fact that the optimal solution for transmitter sensors must be in the randomized form given in (6). Let us prove the properties which were not covered by the proof of the saddle-point.

**Gaussianity of  $\{\theta(i)\}$ :** The choice  $\theta(i) \sim \mathcal{N}(0, P_A)$  maximizes (4) since it renders the simple uncoded linear mappings asymptotically optimal, i.e., the transmitters cannot improve on the zero-delay performance by utilizing asymptotically high delays. Moreover, the optimal zero-delay performance is always lower bounded by the performance of the linear mappings, which is imposed by the adversarial choice of  $\theta(i) \sim \mathcal{N}(0, P_A)$ .

**Independence of  $\{\theta(i)\}$  of  $\{S(i)\}$  and  $\{W(i)\}$ :** If the adversarial sensors introduce some correlation, i.e., if  $\mathbb{E}\{SX_k\} \neq 0$  or  $\mathbb{E}\{WX_k\} \neq 0$ , the transmitter can adjust its Bernoulli parameter to decrease the distortion. Hence, the optimal adversarial strategy is setting  $\mathbb{E}\{SX_k\} = \mathbb{E}\{WX_k\} = 0$  which implies independence since all variables are jointly Gaussian.

**Choice of Bernoulli parameter:** Note that the optimal choice of the Bernoulli parameter for the transmitters is  $\frac{1}{2}$  since other choices will not cancel the cross terms in (9) and (10), i.e.,  $\mathbb{E}\{SX_k\}$  and  $\mathbb{E}\{WX_k\}$ . These cross terms can be exploited by the adversary to increase the cost, hence optimal strategy for transmitter is to set  $\gamma = \text{Bern}(1/2)$ . ■

*Corollary 2:* Coordination is essential for adversarial sensors in the case of coordinating transmitters and receiver, in the sense that lack of adversarial coordination strictly decreases the overall cost.

*Proof:* Note that coordination enables adversarial sensors to create a Gaussian noise with variance  $K^2 P_A$  yielding the cost in (8). However, without coordination, each sensor can only generate independent Gaussian random variables, yielding an overall Gaussian noise with variance  $K P_A$  and the total cost

$$J_2 = \sigma_S^2 \frac{M^2 \alpha_T^2 \sigma_{W_T}^2 + K P_A + \sigma_Z^2}{M \alpha_T^2 \sigma_S^2 + M^2 \alpha_T^2 \sigma_{W_T}^2 + K P_A + \sigma_Z^2} < J_1 \quad (12)$$

Hence, coordination of adversarial sensors strictly increases the overall cost, i.e., coordination is essential for adversarial sensors in this setting. ■

*Remark 1:* Note that the optimality of this jamming strategy does not depend on the “symmetry” assumption for the adversaries. Hence, it is straightforward to show that in the more general setting of  $\sigma_{W_{k_1}}^2 \neq \sigma_{W_{k_2}}^2$  and  $P_{k_1} \neq P_{k_2}$ , for  $(k_1, k_2) \in [M+1 : M+K]$  our results hold. We do not include such generalizations for brevity.

*Remark 2:* We note that the optimal strategies do not depend on the sensor index  $m$ , hence the implementation of the

optimal strategy, for both transmitter and adversarial sensors, requires “coordination” among the sensors. This highlights the need for coordination in game theoretic settings in sensor networks. Note that this coordination requirement arises purely from the game theoretic considerations, i.e., the presence of adversarial sensors. In the case where no adversarial node exists, transmitters do not need to “coordinate”. Moreover, as we will show in Theorem 2 if the transmitters cannot coordinate, then adversarial sensors do not need to coordinate.

#### IV. PROBLEM SETTING II

In this section, we focus on the problem, where the transmitters do not have the ability to secretly agree on a “coordination” random variable to generate their transmission function  $X_k$ . In this case, it turns out that the optimal transmitter strategy, which is almost surely unique, is uncoded transmission with linear mappings, while the adversarial optimal strategy for the (jamming) sensors is uncoded, linear mappings with the opposite sign of the transmitter functions.

The following theorem captures this result.

*Theorem 2:* Under scenario 2, the optimal encoding function for the transmitters is uncoded transmission, i.e.,

$$X_m(i) = \alpha_T U_m(i), \quad M \geq m \geq 1$$

The optimal jamming function (for adversarial sensors) is uncoded transmission with the opposite sign of the transmitters, i.e.,

$$X_k(i) = \alpha_A U_k(i), \quad M+K \geq k \geq M+1$$

The optimal decoding function is the Bayesian estimator of  $S$  given  $Y$ , i.e.,

$$h(Y(i)) = \frac{[(M\alpha_T + K\alpha_A)\sigma_S^2] Y(i)}{(M\alpha_T + K\alpha_A)\sigma_S^2 + M^2\alpha_T^2\sigma_{W_T}^2 + K^2\alpha_A^2\sigma_{W_A}^2 + \sigma_Z^2}.$$

Cost at this saddle-point is

$$J_3 = \sigma_S^2 \frac{M^2\alpha_T^2\sigma_{W_T}^2 + K^2\alpha_A^2\sigma_{W_A}^2 + \sigma_Z^2}{(M\alpha_T + K\alpha_A)\sigma_S^2 + M^2\alpha_T^2\sigma_{W_T}^2 + K^2\alpha_A^2\sigma_{W_A}^2 + \sigma_Z^2} \quad (13)$$

where  $\alpha_T = \sqrt{\frac{P_T}{\sigma_S^2 + \sigma_{W_T}^2}}$  and  $\alpha_A = -\sqrt{\frac{P_A}{\sigma_S^2 + \sigma_{W_A}^2}}$ .

*Proof:* First, we note that adversarial sensors have the knowledge of the transmitter encoding functions, and hence the adversarial encoding functions will be in the same form as the transmitters functions but with a negative sign i.e.,  $\mathbf{g}_A(\cdot) = -\sqrt{\frac{P_A}{P_T}} \mathbf{g}_T(\cdot)$  since outputs are sent over an additive channel (see e.g., [5], [6] for a proof of this result). We next proceed to find the optimal encoding functions for the transmitters, given  $\mathbf{g}_A(\cdot) = -\sqrt{\frac{P_A}{P_T}} \mathbf{g}_T(\cdot)$ . From the data processing theorem, we must have

$$I(\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_{M+K}; \hat{\mathbf{S}}) \leq I(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{M+K}; \mathbf{Y}) \quad (14)$$

The left hand side can be lower bounded as:

$$I(\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_{M+K}; \hat{\mathbf{S}}) \geq R(D) \quad (15)$$

where  $R(D)$  is derived in Appendix A. The right hand side can be upper bounded by

$$I(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{M+K}; \mathbf{Y}) \quad (16)$$

$$\stackrel{(a)}{\leq} \sum_{i=1}^N I(X_1(i), \dots, X_{M+K}(i); Y(i)) \quad (17)$$

$$\leq \max \sum_{i=1}^N I(X_1(i), \dots, X_{M+K}(i); Y(i)) \quad (18)$$

$$= \frac{1}{2} \sum_{i=1}^N \log\left(1 + \frac{1}{\sigma_Z^2} \mathbf{1}^T R_X(i) \mathbf{1}\right) \quad (19)$$

where  $R_X(i)$  is defined as

$$\{R_X(i)\}_{p,r} \triangleq \mathbb{E}\{X_p(i)X_r(i)\} \quad \forall p, r \in [1 : M+K]. \quad (20)$$

Note that (a) follows from the memoryless property of the channel and the maximum in (18) is over the joint density over  $X_1(i), \dots, X_{M+K}(i)$  given the structural constraints on  $R_X(i)$  due to the power constraints and also the fact that  $g_A(\cdot) = -\sqrt{\frac{P_A}{P_T}} g_T(\cdot)$ . It is well known that the maximum is achieved by the jointly Gaussian density for a given fixed covariance structure [12], yielding (19). Since logarithm is a monotonically increasing function, hence the optimal encoding functions  $g_m^N(\cdot), m \in [1 : M]$  equivalently maximize  $\sum_{p,r} \mathbb{E}\{X_p(i)X_r(i)\}$ . Note that

$$X_m(i) = g_m^N(\mathbf{U}_m) \quad (21)$$

and hence  $g_m^N(\cdot), m \in [1 : M]$  that maximize

$$\sum_{p=1}^{p=M+K} \sum_{r=1}^{r=M+K} \mathbb{E}\{g_p^N(\mathbf{U}_p)g_r^N(\mathbf{U}_r)\} \quad (22)$$

can be found by invoking Witsenhausen's lemma (given in Appendix B) as  $g_m^N(\mathbf{U}_m) = \alpha_N \mathbf{U}_m$  where  $\alpha_N = [\alpha_T, \dots, \alpha_T]$ . Finally, we obtain  $J_3$  as an outer bound by equating the left and right hand sides of (14). The linear mappings in Theorem 2 achieve this outer bound and hence are optimal. ■

*Corollary 3:* Source-channel separation, based on digital compression and communications is strictly suboptimal for this setting.

*Proof:* We first note that the optimal adversarial encoding functions must be the negative of that of the transmitters to achieve the saddle-point solution derived in Theorem 2. But then, the problem at hand becomes equivalent to a problem with no adversary which was studied in [2], where source-channel separation was shown to be strictly suboptimal. Hence, separate source-channel coding has to be suboptimal for our problem. A more direct proof follows from the calculation of the separate source-channel coding performance. ■

*Corollary 4:* Coordination is essential for transmitter sensors, in the sense that lack of coordination strictly increases the overall cost.

*Proof:* Proof follows from the fact that  $J_1 < J_3$ . ■

## V. DISCUSSION AND CONCLUSION

This paper is an initial attempt to analyze game theoretic source-channel coding within sensor networks. We have conducted a game-theoretical analysis of a specific Gaussian sensor network, specialized to symmetric transmitter and adversarial sensors. Depending on the *coordination* capabilities of the sensors, we have analyzed two problem settings. The first setting allows coordination among the transmitter sensors. Coordination capability enables the transmitters to use randomized encoders. The saddle-point solution to this problem is randomized uncoded transmission for the transmitters and the coordinated generation of i.i.d. Gaussian noise for the adversarial sensors. In the second setting, transmitter sensors cannot coordinate, and hence they use fixed, deterministic mappings. The solution to this problem is shown to be uncoded communication with linear mappings for both the transmitter and the adversarial sensors, but with opposite signs. We note that coordination aspect of the problem is entirely due to game-theoretic considerations, i.e., if no adversarial sensors exist, optimal transmitter encoding functions do not need coordination.

Our analysis has uncovered an interesting result regarding coordination among transmitter nodes, and among adversarial nodes. If transmitter nodes can coordinate, then so must the adversaries, i.e., all must generate the identical realization of an i.i.d. Gaussian noise sequence. If transmitters cannot coordinate, adversarial sensors do not need to coordinate, and this saddle-point is at strictly higher cost than the one when transmitters can coordinate.

Several questions still remain open and are currently under investigation, including extensions of the analysis to vector sources and channels, which require a vector form of Witsenhausen's Lemma, an important research question in its own right; the asymptotic (in the number of sensors  $M$  and  $K$ ) analysis of the results presented here; and extension of our analysis to asymmetric sensor networks and non-Gaussian settings.

## APPENDIX A

### THE GAUSSIAN CEO PROBLEM

In the Gaussian CEO problem, an underlying Gaussian source  $S \sim \mathcal{N}(0, \sigma_S^2)$  is observed under additive noise  $\mathbf{W} \sim \mathcal{N}(\mathbf{0}, R_W)$  as  $\mathbf{U} = S + \mathbf{W}$ . These noisy observations, i.e.,  $\mathbf{U}$ , must be encoded in such a way that the decoder produces a good approximation to the original underlying source. This problem was proposed in [13] and solved in [14] (see also [15], [16]). A lower bound for this function for the non-Gaussian sources within the "symmetric" setting where all  $U$ 's have identical statistics was presented in [17]. Here, we simply extend the results in [15] to asymmetric settings, following the approach in [17], focusing on MSE distortion measure

$$D = \mathbb{E}\{(S - \hat{S})^2\} \quad (23)$$

and rate

$$R = \min I(\mathbf{U}, \hat{S}) \quad (24)$$

where  $\mathbf{U} = \mathbf{S} + \mathbf{W}$ ,  $\mathbf{W} \sim \mathcal{N}(\mathbf{0}, R_W)$ , and  $R_W$  is an  $(M + K) \times (M + K)$  diagonal matrix where the first  $M$  diagonal entries are  $\sigma_{W_T}^2$  and the remaining  $K$  are  $\sigma_{W_A}^2$ . The minimization in (24) is over all conditional densities  $p(\hat{s}|\mathbf{u})$  that satisfy (23). Distortion can be written as sum of two terms

$$D = \mathbb{E}\{(S - T + T - \hat{S})^2\}, \quad (25)$$

$$= \mathbb{E}\{(S - T)^2\} + \mathbb{E}\{(T - \hat{S})^2\}, \quad (26)$$

where  $T \triangleq \mathbb{E}\{S|\mathbf{U}\}$ . Note that (26) holds since

$$\mathbb{E}\{(S - T)(\hat{S} - T)\} = 0, \quad (27)$$

as the MMSE error is orthogonal to any function<sup>2</sup> of the observation,  $\mathbf{U}$ . The first term in (26), i.e., the estimation error  $D_{est} \triangleq \mathbb{E}\{(S - T)^2\}$  is constant with respect to  $p(\hat{s}|\mathbf{u})$ , i.e., a fixed function of  $\mathbf{U}$  and  $S$ . Hence, the minimization is over the densities that satisfy a distortion constraint of the form  $\mathbb{E}\{(T - \hat{S})^2\} \leq D_{rd}$  and  $R = \min I(\mathbf{U}; \hat{S})$ . Hence, we write (26) as

$$D = D_{rd} + D_{est}. \quad (28)$$

Note that due to their Gaussianity,  $T$  is sufficient statistics of  $\mathbf{U}$  for  $S$ , i.e.,  $S - T - \mathbf{U}$  forms a Markov chain in that order and  $T \sim \mathcal{N}(0, \sigma_T^2)$ . Hence,  $R = \min I(\mathbf{U}; \hat{S}) = \min I(T; \hat{S})$  where minimization is over  $p(\hat{s}|t)$  that satisfy  $\mathbb{E}\{(T - \hat{S})^2\} \leq D_{rd}$ , where all variables are Gaussian. This is the classical Gaussian rate distortion problem, and hence:

$$R = \frac{1}{2} \log(\sigma_T^2 / D_{rd}). \quad (29)$$

Noting that

$$T = \mathbb{E}\{S\mathbf{U}^*\}(\mathbb{E}\{\mathbf{U}\mathbf{U}^*\})^{-1}\mathbf{U}, \quad (30)$$

and using standard linear estimation principles, we obtain

$$\sigma_T^2 = \sigma_S^2 \frac{1}{1 + \frac{\sigma_{W_A}^2 \sigma_{W_T}^2}{\sigma_S^2 (K \sigma_{W_T}^2 + M \sigma_{W_A}^2)}}, \quad (31)$$

and

$$D_{est} = \sigma_S^2 \frac{\sigma_{W_T}^2 \sigma_{W_A}^2}{K \sigma_{W_T}^2 + M \sigma_{W_A}^2 + \sigma_{W_T}^2 \sigma_{W_A}^2}. \quad (32)$$

## APPENDIX B

### WITSENHAUSEN'S LEMMA

In this section, we recall Witsenhausen's lemma [18], which is used in the proof of Theorem 2.

*Lemma 1:* Consider two sequences of i.i.d. random variables  $X(i)$  and  $Y(i)$ , generated from a joint density  $P_{X,Y}$ , and two (Borel measurable) arbitrary functions  $f, g: \mathbb{R} \rightarrow \mathbb{R}$  satisfying

$$\mathbb{E}\{f(X)\} = \mathbb{E}\{g(Y)\} = 0, \quad (33)$$

$$\mathbb{E}\{f^2(X)\} = \mathbb{E}\{g^2(Y)\} = 1. \quad (34)$$

Define

$$\rho^* \triangleq \sup_{f,g} \mathbb{E}\{f(X)g(Y)\} \quad (35)$$

<sup>2</sup>Note that  $\hat{S}$  is also a deterministic function of  $\mathbf{U}$ , since optimal reconstruction can always be achieved by deterministic codes.

Then for any (Borel measurable) functions  $f_N, g_N: \mathbb{R}^N \rightarrow \mathbb{R}$  satisfying

$$\mathbb{E}\{f_N(\mathbf{X})\} = \mathbb{E}\{g_N(\mathbf{Y})\} = 0, \quad (36)$$

$$\mathbb{E}\{f_N^2(\mathbf{X})\} = \mathbb{E}\{g_N^2(\mathbf{Y})\} = 1, \quad (37)$$

for length  $N$  vectors  $\mathbf{X}$  and  $\mathbf{Y}$ , we have

$$\sup_{f_N, g_N} \mathbb{E}\{f_N(\mathbf{X})g_N(\mathbf{Y})\} \leq \rho^*. \quad (38)$$

Moreover, the supremum above is attained by linear mappings, if  $P_{X,Y}$  is a bivariate normal density.

## REFERENCES

- [1] Z. Xiong, A.D. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Processing Magazine*, vol. 21, no. 5, pp. 80–94, 2004.
- [2] M. Gastpar and M. Vetterli, "Power, spatio-temporal bandwidth, and distortion in large sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 4, pp. 745–754, 2005.
- [3] T. Başar, "The Gaussian test channel with an intelligent jammer," *IEEE Trans. on Inf. Th.*, vol. 29, no. 1, pp. 152–157, 1983.
- [4] T. Başar and Y.W. Wu, "A complete characterization of minimax and maximin encoder-decoder policies for communication channels with incomplete statistical description," *IEEE Trans. on Inf. Th.*, vol. 31, no. 4, pp. 482–489, 1985.
- [5] T. Başar and Y.W. Wu, "Solutions to a class of minimax decision problems arising in communication systems," *Journal of Optimization Theory and Applications*, vol. 51, no. 3, pp. 375–404, 1986.
- [6] R. Bansal and T. Başar, "Communication games with partially soft power constraints," *Journal of Optimization Theory and Applications*, vol. 61, no. 3, pp. 329–346, 1989.
- [7] M. Gastpar, "Uncoded transmission is exactly optimal for a simple Gaussian sensor network," *IEEE Trans. on Inf. Th.*, vol. 54, no. 11, pp. 5247–5251, 2008.
- [8] H. Behroozi, F. Alajaji, and T. Linder, "On the optimal power-distortion region for asymmetric Gaussian sensor networks with fading," in *IEEE International Symposium on Inf. Th.* IEEE, 2008, pp. 1538–1542.
- [9] H. Behroozi, F. Alajaji, and T. Linder, "On the optimal performance in asymmetric Gaussian wireless sensor networks with fading," *IEEE Trans. on Signal Processing*, vol. 58, no. 4, pp. 2436–2441, 2010.
- [10] A. Kashyap, T. Başar, and R. Srikant, "Correlated jamming on MIMO Gaussian fading channels," *IEEE Trans. on Inf. Th.*, vol. 50, no. 9, pp. 2119–2123, 2004.
- [11] S. Shafiee and S. Ulukuş, "Mutual information games in multiuser channels with correlated jamming," *IEEE Trans. on Inf. Th.*, vol. 55, no. 10, pp. 4598–4607, 2009.
- [12] S.N. Diggavi and T.M. Cover, "The worst additive noise under a covariance constraint," *IEEE Trans. on Inf. Th.*, vol. 47, no. 7, pp. 3072–3081, 2001.
- [13] H. Viswanathan and T. Berger, "The quadratic Gaussian CEO problem," *IEEE Trans. on Inf. Th.*, vol. 43, no. 5, pp. 1549–1559, 1997.
- [14] Y. Oohama, "Rate-distortion theory for Gaussian multiterminal source coding systems with several side informations at the decoder," *IEEE Trans. on Inf. Th.*, vol. 51, no. 7, pp. 2577–2593, 2005.
- [15] Y. Oohama, "The rate-distortion function for the quadratic Gaussian CEO problem," *IEEE Trans. on Inf. Th.*, vol. 44, no. 3, pp. 1057–1070, 1998.
- [16] V. Prabhakaran, D. Tse, and K. Ramachandran, "Rate region of the quadratic Gaussian CEO problem," in *Proceedings of the International Symposium on Inf. Th.* IEEE, 2004, p. 119.
- [17] M. Gastpar, "A lower bound to the AWGN remote rate-distortion function," in *IEEE 13th Workshop on Statistical Signal Processing*. IEEE, 2005, pp. 1176–1181.
- [18] H.S. Witsenhausen, "On sequences of pairs of dependent random variables," *SIAM Journal on Applied Mathematics*, pp. 100–113, 1975.