

OPTIMAL DISTRIBUTED LINEAR TRANSCEIVERS FOR SENDING INDEPENDENTLY CORRUPTED COPIES OF A COLORED SOURCE OVER THE GAUSSIAN MAC

Onkar Dabeer

School of Technology and Computer Science
Tata Institute for Fundamental Research
Mumbai, India
Email: onkar@tcs.tifr.res.in

Aline Roumy, Christine Guillemot*

IRISA-INRIA
35042 Rennes Cedex
France

ABSTRACT

We consider distributed linear transceivers for sending a second-order wide-sense stationary process observed by two noisy sensors over a Gaussian multiple-access channel (MAC). We derive the minimum mean-square error (MSE) distributed linear transceiver. The optimal linear transmitter exploits bandwidth expansion by repeating transmission and the transmitters at the two sensors are the same except for a constant factor. When the source is white, uncoded transmission is the best linear code for any SNR. But for a colored source, whitening transmit filter is sub-optimal. In high SNR regime, the magnitude response of the optimal transmission filter is inversely proportional to fourth-root of the power spectrum of the process (while that for the whitening filter is inversely proportional to the square-root of the spectrum). In the special case of a single sensor with Gaussian source, we also quantify the performance loss of linear source-channel codes with respect to the Shannon limit.

Index Terms— Colored source, linear codes, joint source-channel coding, multiaccess communications, sensor networks

1. INTRODUCTION

Motivated by applications in sensor networks, several researchers are considering the transmission of dependent sources over multiple-access channels. The problem is particularly interesting because the source-channel separation theorem does not hold in general ([1]). In fact, it is known that separation can be exponentially worse than joint source-channel coding ([2]). In this paper, we look at linear (over the real field) joint source-channel codes. Our motivation is two-fold. First, linear processing is simple to implement and by now there is vast experience in efficient hardware implementation of linear processing. This is important to keep the sensors simple and reduce their cost. Second, it has recently been shown that for transmitting memoryless, bivariate Gaussian sources over the Gaussian MAC, uncoded transmission is optimal below a certain SNR threshold ([3]). The class of linear transmitters includes uncoded transmission, and it is insightful to understand the nature of the best linear transmitters. We note that there is extensive literature on linear transceiver optimization for sending independent, memoryless sources over the MAC (see [4] and references therein). In [4], the problem is formulated as a semi-definite program, and numerical algorithms are proposed for the same. In contrast, in this paper we consider a single colored source being observed by two sensors with independent noises and

derive closed-form solution to the optimal linear transceiver. The problem of distributed source coding of multiple independently corrupted copies of a source is commonly referred to as the CEO problem ([5]). Thus in this paper we consider linear joint source-channel codes for the CEO problem over a Gaussian MAC.

Our goal is to find the performance limit of linear source-channel codes for a colored source. So we consider non-causal transceivers, which may be viewed as the limit of block transceivers as the block size goes to infinity. Our results show that bandwidth expansion is exploited by the optimal transmitter by repeating the transmission. Moreover the two sensors employ the same transmit filter (except for a scale factor), which in effect reduces the two sensor case to a single sensor case. In the high SNR regime, the optimal transmitter filter has magnitude response inversely proportional to the fourth-root of the power-spectrum (while that for the whitening filter is inversely proportional to the square-root of the spectrum). We also provide an expression for the least MSE in the high SNR regime, which shows that the MSE is proportional to the integral of the square-root of the spectrum. In the case of a single sensor with Gaussian source, we also quantify the loss with respect to the Shannon limit. For example, in the high SNR regime, for a Gaussian first-order Markov process with correlation sequence $0.7^{|k|}$, the optimal linear transceiver is about 1.4 dB from Shannon limit but it is 1.5 dB better than the whitening filter.

The remainder of this paper is organized as follows: the precise problem definition is given in Section 2, the main results are given in Section 3, and the conclusion in Section 4.

Notation: All vectors are column vectors. Superscript T denotes transpose and superscript H denotes conjugate transpose. The minimum MSE achievable using linear codes is denoted by MSE^* , while that achievable using *any* code is denoted by MSE^{**} .

2. PROBLEM DEFINITION

Consider a stationary second-order stochastic process $\{s_t\}_{t=-\infty}^{\infty}$ with zero mean and covariance function $c(t)$, $c(0) = 1$. We assume that $c(t)$ is integrable, and this implies that the process has a continuous bounded spectral density $\phi(\omega)$, $\omega \in (-\pi, \pi]$ ([6]). The process is observed by two sensors in the presence of additive noise. The observations at sensor i are

$$x_{i,t} = a_i s_t + v_{i,t},$$

where a_i are the signal amplitudes, the additive noise is i.i.d. $\mathcal{N}(0, \sigma_o^2)$ and the noise processes are independent across the sensors. These

*This research was partly funded by French National Research Agency (ANR) under the Essor project.

observations are to be transmitted over a channel. In typical scenarios, the channel has more bandwidth than the source. Let B be the bandwidth expansion factor; then for every source symbol we are allowed B channel uses. We only consider $B = 1, 2, \dots$. The transmitter at sensor i processes the observations through B linear filters $H_{i,b}(\omega)$ arranged in parallel. The resulting B data-streams are multiplexed into a single stream, which is then transmitted over the channel. We note that we have restricted ourselves to time-invariant filters. This choice is motivated by the wide-sense stationary nature of the source. Each sensor has a transmit power constraint of P and this leads to the following constraints on the transmit filters:

$$\int_{-\pi}^{\pi} |H_{i,b}(\omega)|^2 \phi(\omega) d\omega \leq P, \quad i = 1, 2; b = 1, \dots, B. \quad (1)$$

The data-streams at the two sensors are then transmitted over the Gaussian MAC. The Gaussian MAC has additive i.i.d. $\mathcal{N}(0, \sigma^2)$ noise and scales the signal transmitted by sensor i by the channel gain g_i . At the receiver, the samples are passed through a demultiplexer that converts the single data-stream into B parallel streams. If the transmitter filters have absolutely summable impulse responses $\{h_{i,b,t}\}$, then the signal on the b^{th} output branch is given by

$$y_{b,t} = \sum_{i=1}^2 g_i a_i (h_{i,b,t} * s_t) + \sum_{i=1}^2 g_i (h_{i,b,t} * v_{i,t}) + w_{b,t}$$

where $\{w_{b,t}\}$ are i.i.d. $\mathcal{N}(0, \sigma^2)$ and $*$ denotes convolution. In general, (1) does not guarantee the existence of absolutely summable filter impulse response. But under (1), the filter operations are well-defined in the Fourier domain ([6]) and we work with these more general filters.

For fixed transmitter filters, the least MSE is achieved by the non-causal Wiener filter at the receiver. Let $Y_t = [y_{1,t}, \dots, y_{B,t}]^T$ and $H_i = [H_{i,1}, \dots, H_{i,B}]^T$. Then using the orthogonality condition ([6]), the resulting MSE is found to be

$$\text{MSE} = \int_{-\pi}^{\pi} \left[\phi(\omega) - \phi_{sY}^H(\omega) \phi_{YY}^{-1}(\omega) \phi_{sY}(\omega) \right] d\omega \quad (2)$$

where

$$\begin{aligned} \phi_{YY}(\omega) &= \left(\sum_{i=1}^2 g_i a_i H_i(\omega) \right) \left(\sum_{i=1}^2 g_i a_i H_i(\omega) \right)^H \phi(\omega) \\ &+ \frac{\sigma_o^2}{2\pi} \sum_{i=1}^2 g_i^2 H_i(\omega) H_i^H(\omega) + \frac{\sigma^2}{2\pi} I \end{aligned} \quad (3)$$

and

$$\phi_{sY}(\omega) = \left(\sum_{i=1}^2 g_i a_i H_i^*(\omega) \right) \phi(\omega). \quad (4)$$

Problem Statement: Our goal is to choose the transmit filters such that the MSE (2) is minimized subject to the power constraints (1). We solve this problem analytically in the next section. We denote the minimum MSE achievable using linear codes by MSE_* .

3. MAIN RESULTS

3.1. Derivation of Main Results

Our first step to solve the problem posed in Section 2 is to find a lower bound on the MSE (2). We need the following matrix property

for this purpose. For a positive definite matrix R , let $\lambda_{\max}(R)$ denote its maximum eigenvalue. Then from Weyl's theorem [7, (9), pp. 75], we know that

$$\lambda_{\max}(R + xx^H) \leq \lambda_{\max}(R) + \lambda_{\max}(xx^H) = \lambda_{\max}(R) + \|x\|^2$$

with equality if x is an eigenvector of R corresponding to eigenvalue $\lambda_{\max}(R)$. Applying this property repeatedly to (3), we get

$$\lambda_{\max}(\phi_{YY}) \leq \frac{\sigma^2}{2\pi} + \frac{\sigma_o^2}{2\pi} \sum_{i=1}^2 g_i^2 \|H_i\|^2 + \left\| \sum_{i=1}^2 g_i a_i H_i \right\|^2 \phi$$

where for simplicity we have dropped the argument ω from the various quantities. We note that equality holds in this bound if

$$H_1(\omega) = \alpha H_2(\omega), \quad \text{for some constant } \alpha. \quad (5)$$

Using this bound we get that

$$\begin{aligned} \phi_{sY}^H \phi_{YY}^{-1} \phi_{sY} &\geq \lambda_{\min}(\phi_{YY}^{-1}) \|\phi_{sY}\|^2 = \frac{\|\phi_{sY}\|^2}{\lambda_{\max}(\phi_{YY})} \\ &\geq \frac{\|\sum_{i=1}^2 g_i a_i H_i\|^2 \phi^2}{D(\omega)} \end{aligned} \quad (6)$$

where

$$D(\omega) := \frac{\sigma^2}{2\pi} + \frac{\sigma_o^2}{2\pi} \sum_{i=1}^2 g_i^2 \|H_i\|^2 + \left\| \sum_{i=1}^2 g_i a_i H_i \right\|^2 \phi.$$

We see that equality holds in (6) if (5) is true. Substituting (5) in (2), we get

$$\text{MSE} \geq \int_{-\pi}^{\pi} \frac{N(\omega)}{D(\omega)} \phi(\omega) d\omega$$

where

$$N(\omega) := \frac{\sigma^2}{2\pi} + \frac{\sigma_o^2}{2\pi} \sum_{i=1}^2 g_i^2 \|H_i(\omega)\|^2. \quad (7)$$

Expanding the term $\|\sum_{i=1}^2 g_i a_i H_i(\omega)\|^2$ in the denominator $D(\omega)$ and using

$$g_1 g_2 a_1 a_2 \text{Re} \left(H_1^H(\omega) H_2(\omega) \right) \leq |g_1 g_2 a_1 a_2| \|H_1(\omega)\| \|H_2(\omega)\| \quad (8)$$

for the cross-term, we get,

$$\begin{aligned} \text{MSE} &\geq \int_{-\pi}^{\pi} \frac{N(\omega) \phi(\omega) d\omega}{N(\omega) + \left(\sum_{i=1}^2 |g_i a_i| \|H_i(\omega)\| \right)^2 \phi(\omega)} \\ &=: U(H_1, H_2) \end{aligned}$$

and equality holds if (5) is true and equality holds in (8). We note that $U(H_1, H_2)$ is a lower bound on the MSE and minimizing it under the power constraints (1) yields a lower bound on MSE_* . If the filters that minimize $U(H_1, H_2)$ satisfy (5) and (8) with equality, then the lower bound on MSE_* is tight. We show below that this is indeed the case. For convenience let

$$A_{i,b}(\omega) := |H_{i,b}(\omega)|^2$$

and $A_i(\omega) = [A_{i,1}(\omega), \dots, A_{i,B}(\omega)]^T$.

Lemma 1 *There exist filters (H_1, H_2) such that they minimize the lower bound $U(H_1, H_2)$ under (1), satisfy condition (5), satisfy (8) with equality, and therefore satisfy $A_1(\omega) = A_2(\omega)$.*

Proof: Writing $U(H_1, H_2)$ as a function of (A_1, A_2) :

$$U(A_1, A_2) = \int_{-\pi}^{\pi} \frac{N(\omega)\phi(\omega)d\omega}{N(\omega) + \left(\sum_{i=1}^2 |g_i a_i| \sqrt{\sum_{b=1}^B A_{i,b}(\omega)}\right)^2} \phi(\omega)$$

The power constraints are

$$\int_{-\pi}^{\pi} A_{i,b}(\omega)\phi(\omega)d\omega \leq P, \quad i = 1, 2; b = 1, \dots, B.$$

Now let

$$\bar{A} = \frac{|g_1|A_1 + |g_2|A_2}{|g_1| + |g_2|}.$$

Since the power constraints are linear in (A_1, A_2) , the pair (\bar{A}, \bar{A}) satisfies the power constraints. Moreover, since $N(\omega)$ given by (7) is a linear function of (A_1, A_2) and since the square-root function is concave, we get $U(\bar{A}, \bar{A}) \leq U(A_1, A_2)$. Thus it follows that $U(A_1, A_2)$ is minimized for some $A_1(\omega) = A_2(\omega)$, and we denote this common value by

$$C(\omega) = [C_1(\omega), \dots, C_B(\omega)]^T = A_1(\omega).$$

For choosing the filters, we note that we are completely free to choose the phase responses since the power constraints do not depend on them. So we choose $H_1(\omega) = \text{sign}(g_1 g_2 a_1 a_2) H_2(\omega)$ such that $|H_{2,b}(\omega)|^2 = C_b(\omega)$, $b = 1, \dots, B$. It is easy to check that this ensures equality in (8) and (5) also holds. ■

Let $\tilde{C}(\omega) = \sum_{b=1}^B C_b(\omega)$. The problem of finding the best linear transmitter now reduces to finding a $C(\omega) = [C_1(\omega), \dots, C_B(\omega)]$ that minimizes

$$\text{MSE} = \int_{-\pi}^{\pi} \frac{N(\omega)\phi(\omega)d\omega}{N(\omega) + (|g_1 a_1| + |g_2 a_2|)^2 \tilde{C}(\omega)\phi(\omega)} \quad (9)$$

$$\text{with } N(\omega) = \frac{\sigma^2}{2\pi} + \frac{\sigma_o^2}{2\pi} (g_1^2 + g_2^2) \tilde{C}(\omega)$$

subject to

$$\int_{-\pi}^{\pi} C_b(\omega)\phi(\omega)d\omega \leq P, \quad b = 1, \dots, B. \quad (10)$$

We note that the MSE (9) depends only on $\tilde{C}(\omega)$. The power constraints (10) imply that

$$\int_{-\pi}^{\pi} \tilde{C}(\omega)\phi(\omega)d\omega \leq BP. \quad (11)$$

If we minimize the MSE (9) w.r.t. $\tilde{C}(\omega)$ subject to constraint (11), then we get a lower bound on the MSE (since constraint (11) is weaker than conditions (10)). However, the resulting lower bound is tight since we can choose $C_{b,*}(\omega) = \tilde{C}_*(\omega)/B$, which attains the lower bound with equality and satisfies (10). Thus to find MSE_* we have to minimize the convex function (9) of $\tilde{C}(\omega)$ subject to the linear constraint (11). Let $[x]^+ = \max\{0, x\}$. Using [8, Theorem 4.4.1, pp. 87] it can be verified that the solution is given by

$$\begin{aligned} \tilde{C}_*(\omega) &= \frac{\sigma(|g_1 a_1| + |g_2 a_2|)}{\sqrt{2\pi}} \left[\lambda_* - \sqrt{\frac{\sigma^2/2\pi}{(|g_1 a_1| + |g_2 a_2|)\phi(\omega)}} \right]^+ \\ &\quad \times \frac{\sqrt{\phi(\omega)}}{\frac{\sigma_o^2}{2\pi}(g_1^2 + g_2^2) + (|g_1 a_1| + |g_2 a_2|)^2 \phi(\omega)} \end{aligned} \quad (12)$$

where λ_* is chosen to ensure equality in (11). Due to space constraints we do not show the tedious (but straightforward) calculations involved in deriving the above expression. To summarize, we have shown the following.

Proposition 1 *There exists an optimal linear transmitter that satisfies:*

- $H_{i,b}(\omega) = H_{i,1}(\omega)$, $i = 1, 2$, $b = 1, \dots, B$.
- $H_{1,b}(\omega) = \text{sign}(g_1 g_2 a_1 a_2) H_{2,b}(\omega)$, $b = 1, \dots, B$.
- $|H_{1,1}(\omega)|^2 = \tilde{C}_*(\omega)/B$, where $\tilde{C}_*(\omega)$ is given by (12).
- MSE_* is obtained by substituting (12) in (9).

Thus the two sensors use the same transmit filters except for the constant factor of ± 1 . The bandwidth expansion is exploited merely by repeating the transmission B times. For simplicity of discussion, we assume that $g_1 \geq 0$, $g_2 \geq 0$, $a_1 = a_2 = 1$, $B = 1$ below. If the source is white, then $\phi(\omega) = 1$ and from (12) we get that uncoded transmission is the optimal linear transmitter. For a whitening filter of a colored source $\tilde{C}(\omega) = \text{constant}/\phi(\omega)$, which is different from that (12) for the optimal linear filter. Hence for a colored source, the whitening filter is sub-optimal. The nature of the optimal filter simplifies significantly in the limit as $\sigma \rightarrow 0$ with σ_o/σ fixed to some value γ . From (9) we get that

$$\lim_{\substack{\sigma \rightarrow 0 \\ \sigma_o/\sigma \rightarrow \gamma}} \frac{\text{MSE}}{\sigma^2} = \frac{\gamma^2(g_1^2 + g_2^2)}{(g_1 + g_2)^2} + \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{d\omega}{(g_1 + g_2)^2 \tilde{C}(\omega)}.$$

Using a Lagrange multiplier we can easily minimize the above limit subject to (11). This gives us that as $\sigma \rightarrow 0$ with $\sigma_o/\sigma \rightarrow \gamma$,

$$\tilde{C}_*(\omega) \rightarrow \frac{BP}{\sqrt{\phi(\omega)} \int_{-\pi}^{\pi} \sqrt{\phi(\omega)} d\omega} \quad (13)$$

and

$$\lim_{\substack{\sigma \rightarrow 0 \\ \sigma_o/\sigma \rightarrow \gamma}} \frac{\text{MSE}_*}{\sigma^2} = \frac{\gamma^2(g_1^2 + g_2^2)}{(g_1 + g_2)^2} + \frac{\left\{ \int_{-\pi}^{\pi} \sqrt{\phi(\omega)} d\omega \right\}^2}{2\pi P (g_1 + g_2)^2}. \quad (14)$$

In contrast, for a whitening filter $\tilde{C}(\omega)$ is inversely proportional to $\phi(\omega)$. The loss of the whitening filter w.r.t. the optimal linear transmitter is plotted for an example in the next section.

3.2. Single Sensor Case

Consider the special case when $g_1 = 1$, $g_2 = 0$, $\sigma_o^2 = 0$ and the source is Gaussian. The main reason for considering this case is that the Shannon limit can be found. For this special case, we now compare the optimal linear transmitter, the whitening filter transmitter, and the Shannon limit for achievable MSE using *any* code. The MSE for the whitening filter is given by:

$$\text{MSE}_W = \frac{\sigma^2}{\sigma^2 + P} \int_{-\pi}^{\pi} \phi(\omega) d\omega = \frac{\sigma^2}{\sigma^2 + P}.$$

For the optimal linear transmitter, from (14) with $\gamma = 0$ we get that

$$\lim_{\sigma \rightarrow 0} \frac{\text{MSE}_*}{\sigma^2} = \frac{1}{2\pi P} \left\{ \int_{-\pi}^{\pi} \sqrt{\phi(\omega)} d\omega \right\}^2.$$

Thus we get

$$\lim_{\sigma \rightarrow 0} \frac{\text{MSE}_*}{\text{MSE}_W} = \frac{\left\{ \int_{-\pi}^{\pi} \sqrt{\phi(\omega)} d\omega \right\}^2}{2\pi \int_{-\pi}^{\pi} \phi(\omega) d\omega} =: r_1.$$

The Cauchy-Schwarz inequality easily verifies that the above limit is upper bounded by 1 with equality iff $\phi(\omega) = 1$, that is, when the source is white.

Now using Shannon theory we can also find the least MSE that can be achieved using *any* code. For this purpose assume that $\phi(\omega) \geq \delta > 0$. Then for $D < \delta$ the rate-distortion function is given by ([9, Theorem 4.6.2, pp. 133])

$$R(D) = \frac{1}{2} \log \left(\frac{Q_1}{D} \right)$$

where

$$Q_1 = 2\pi \exp \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \ln(\phi(\omega)) d\omega \right)$$

is the one-step prediction error of the source ([6]) and \ln is the logarithm w.r.t. the natural base. Since for $g_1 = 1, g_2 = 0$, the channel is a single-user AWGN channel, we can achieve $D < \delta$ if

$$R(D) < \frac{1}{2} \log \left(1 + \frac{P}{\sigma^2} \right) \implies D > \frac{Q_1}{1 + \frac{P}{\sigma^2}}.$$

(We assume that P/σ^2 is large enough so that the lower bound above is less than δ , which is necessary for the validity of the expression for $R(D$.) We know from the source-channel separation theorem [9] that there exist codes that come arbitrarily close to the above lower bound. Thus (for sufficiently large P/σ^2) the smallest achievable MSE using any code is

$$\text{MSE}_{**} = \frac{Q_1}{1 + \frac{P}{\sigma^2}}.$$

Thus we get that

$$\lim_{\sigma \rightarrow 0} \frac{\text{MSE}_{**}}{\text{MSE}_*} = \frac{2\pi Q_1}{\left\{ \int_{-\pi}^{\pi} \sqrt{\phi(\omega)} d\omega \right\}^2} =: r_2.$$

For transmitting a white Gaussian source over a Gaussian channel with $B = 1$, it is well known that uncoded transmission is optimal ([10]) and the above limit is unity in this case. However for colored sources, linear codes lead to a loss, as shown in the example below.

Example: Consider a first-order autoregressive Gaussian process with $c(k) = \rho^{|k|}$, $\rho \in (-1, 1)$. For this process

$$\phi(\omega) = \frac{(1 - \rho^2)}{2\pi(1 - 2\rho \cos(\omega) + \rho^2)}$$

and $Q_1 = 1 - \rho^2$. In Figure 1 we have plotted the loss of the whitening filter and the optimal filter w.r.t. the Shannon limit as $\sigma \rightarrow 0$ (which are respectively given by $r_1 r_2$ and r_2) for various values of ρ . We see that in the higher SNR regime, substantial gain can be achieved by using the optimal filter; for $\rho = 0.7$, the optimal filter is about 1.4 dB from Shannon limit but it is 1.5 dB better than the whitening filter.

4. CONCLUSIONS

In this paper, we derived distributed linear transceivers that minimize the MSE for transmitting independently corrupted versions of a colored source over a Gaussian MAC. The key conclusion is that the optimal transceiver can significantly outperform the whitening transmitter. In the single sensor case, the loss of linear processing

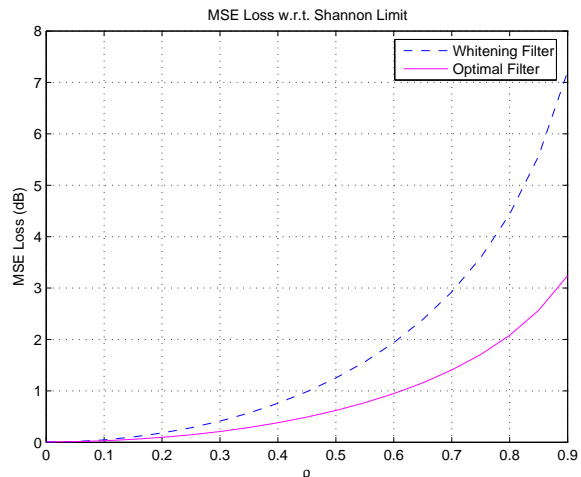


Fig. 1. The optimal filter has significantly less loss compared to the whitening filter for $\rho > 0.5$.

w.r.t. the Shannon limit may not be too large, and due to its simplicity, linear processing is an attractive choice. Our focus was on identifying the limits of linear processing. Future investigation into implementation aspects in the absence of channel information and with finite-precision arithmetic is needed to understand the true merits of linear processing in sensor communication.

5. REFERENCES

- [1] T. Cover, A. E. Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Tran. on Inform. Theory*, vol. 26, no. 6, pp. 648–657, November 1980.
- [2] M. Gastpar and M. Vetterli, "Power, spatio-temporal bandwidth, and distortion in large sensor networks," *IEEE Journal on Select. Areas in Commun.*, vol. 23, no. 4, pp. 745–754, April 2005.
- [3] A. Lapidoth and S. Tinguely, "Sending a bivariate Gaussian source over a Gaussian MAC," arXiv:cs.IT/061029, May 2006.
- [4] Z-Q. Luo et al., "Transceiver Optimization for Block-Based Multiple Access Through ISI Channels," *IEEE Tran. on Sig. Proc.*, vol. 52, no. 4, pp. 1037–1052, April 2004.
- [5] H. Vishwanathan and T. Berger, "The quadratic gaussian ceo problem," *IEEE Tran. on Inform. Theory*, vol. 43, no. 5, pp. 1549–1559, September 1997.
- [6] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*, Springer Series in Statistics. Springer-Verlag, New York, 2 edition, 1991.
- [7] H. Lütkepohl, *Handbook of Matrices*, John Wiley and Sons, New York, 1996.
- [8] R. G. Gallager, *Information Theory and Reliable Communication*, John Wiley and Sons, 1968.
- [9] T. Berger, *Rate Distortion Theory*, Information and System Sciences Series. Prentice-Hall, New Jersey, 1971.
- [10] T. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley and Sons, New York, 1991.