

# Low-Latency Source-Channel Coding for Fading Channels with Correlated Interference

Ahmad Abou Saleh, Fady Alajaji, and Wai-Yip Chan

**Abstract**—We investigate the problem of sending a Gaussian source over a Rayleigh fading channel with Gaussian correlated interference known to the transmitter using low-latency codes. For the matched bandwidth case between the source and the channel, we show that among all single-letter codes, the uncoded scheme achieves the lowest mean square error distortion under full correlation between source and interference, and hence it is optimal. To benefit from nonlinear strategies for other scenarios, we derive the necessary conditions for optimality and propose an iterative algorithm based on joint optimization between the encoder and the decoder. A reduced-complexity approach for the implementation of the design algorithm is presented based on Monte-Carlo (at the encoder side) and importance sampling (at the decoder side) techniques. Furthermore, the scalability of our low-latency scheme is improved by modifying the search process at the encoder side using a targeted search method.

**Index Terms**—Shannon-Kotel’nikov mappings, joint source-channel coding, correlated interference, dirty paper coding.

## I. INTRODUCTION

A tandem Costa scheme which comprises an optimal quantizer followed by Costa’s dirty paper channel code [1] is optimal in the absence of fading and of correlation between the source and the interference. Recently, [2] proposed a high delay scheme for Gaussian sources over additive white Gaussian noise (AWGN) channels with correlated interference.

With the increasing popularity of wireless sensor networks (WSNs), reliable transmission with latency and complexity constraints has become an important practical objective. In [3], [4], the authors present joint source-channel coding schemes with low delay and complexity based on analog mappings over AWGN channels. In [5], the authors consider lossless symbol-by-symbol coding for finite alphabet source with interference. In this work, we study the reliable transmission of a memoryless Gaussian source over a Rayleigh fading channel with known correlated interference at the transmitter. Motivated by WSN applications, we focus on low-latency codes. The rest of the paper is organized as follows. Section II presents the problem formulation. Section III introduces the optimality of the uncoded scheme for the case of matched bandwidth with fully correlated source-interference. In Section IV, we design low-latency codes for a general set-up via a reduced complexity algorithm. Simulation results and comparisons to reference systems are included in Section V.

Manuscript received September 25, 2013. The associate editor coordinating the review of this letter and approving it for publication was M. Xiao.

This work was supported in part by NSERC of Canada.

A. Abou Saleh and W.-Y. Chan are with the Department of Electrical and Computer Engineering, Queen’s University, Kingston, ON, K7L 3N6, Canada (e-mail: {ahmad.abou.saleh, chan}@queensu.ca).

F. Alajaji is with the Department of Mathematics and Statistics, Queen’s University, Kingston, ON, K7L 3N6, Canada (e-mail: fady@mast.queensu.ca).  
Digital Object Identifier 10.1109/WCL.2013.120513.130680

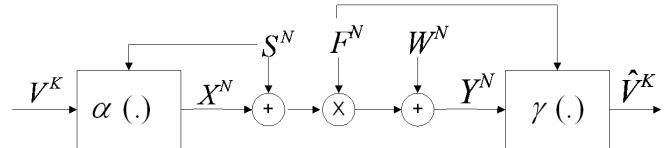


Fig. 1. A  $K : N$  system structure over a fading channel with interference known at the transmitter. The fading coefficient is known at the receiver while the transmitter knows only its distribution.

## II. PROBLEM FORMULATION AND MAIN CONTRIBUTIONS

We consider (Fig. 1) the transmission of a memoryless Gaussian source  $V^K = (V(1), \dots, V(K))^T \in \mathbb{R}^K$  with each component having variance  $\sigma_V^2$  over a Rayleigh fading channel in the presence of Gaussian interference  $S^N \in \mathbb{R}^N$  known to the transmitter, where  $(\cdot)^T$  denotes the transpose operator. The source vector  $V^K$  is transformed into an  $N$ -dimensional channel input  $X^N \in \mathbb{R}^N$  using a nonlinear mapping function, in general,  $\alpha(\cdot) : \mathbb{R}^K \times \mathbb{R}^N \rightarrow \mathbb{R}^N$ . The received symbol is  $Y^N = F^N(X^N + S^N) + W^N$ , where addition and multiplication are component-wise,  $F^N$  represents an  $N$ -dimensional Rayleigh fast fading random vector that is independent of  $(V^K; S^N; W^N)$  and known only to the receiver,  $X^N = \alpha(V^K, S^N)$ ,  $S^N$  is a Gaussian interference vector with independent and identically distributed (i.i.d.) components each of variance  $\sigma_S^2$ , and each sample in the additive noise  $W^N$  is i.i.d. drawn from a Gaussian distribution  $(W(i) \sim \mathcal{N}(0, \sigma_W^2), i = 1, \dots, N)$  independently from all samples in  $S^N$  and  $V^K$ . Unlike the typical dirty paper problem which assumes an AWGN channel with interference [1], we consider a fading channel and assume that  $V^K$  and  $S^N$  are jointly Gaussian and correlated. Since the fading realization is known only at the receiver, we have partial knowledge of the total actual interference  $F^N S^N$  at the transmitter. The channel operates under an input average power constraint  $P$  given by

$$\mathbb{E}[|\alpha(V^K, S^N)|^2]/N \leq P \quad (1)$$

where  $\mathbb{E}[\cdot]$  denotes the expectation operator. The reconstructed signal is given by  $\hat{V}^K = \gamma(Y^N, F^N)$ , where the decoder is a mapping from  $\mathbb{R}^N \times \mathbb{R}^N \rightarrow \mathbb{R}^K$ . The rate of the system is given by  $r = \frac{N}{K}$  channel use/source symbol.

In this work, we aim to find an encoder  $\alpha$  and decoder  $\gamma$  that minimize the mean square error (MSE) distortion  $D = \mathbb{E}[|V^K - \hat{V}^K|^2]/K$  under the average power constraint in (1) and a low-latency constraint ( $K$  and  $N$  take on small values). To the best of our knowledge, this problem with no delay constraint (infinite  $K$  and  $N$ ) is not yet solved (even for the uncorrelated source-interference setting). The main attempt in solving this problem without latency constraint is by using a generalized Costa scheme [6]. This, however, is not proved to be optimal for finite noise levels. Our main contributions are:

- We show that for matched bandwidth between the source

and the channel (i.e.,  $r = 1$ ), the uncoded scheme is optimal among all single-letter codes when the source and the interference are fully correlated.

- To benefit from nonlinearity whenever possible, we derive the necessary conditions for optimality and propose an iterative algorithm based on joint optimization between the encoder and the decoder.
- To lower the complexity of our algorithm, we use Monte Carlo (at the encoder) and importance sampling (at the decoder) techniques. We also use a targeted (focused) search method in order to make our algorithm more scalable to larger dimensions ( $K$  and  $N$ ).

### III. OPTIMALITY OF THE UNCODED SCHEME

In this section, we assume that the encoder transforms the  $K$ -dimensional signal  $V^K$  into an  $N$ -dimensional channel input  $X^N$  using a linear transformation according to

$$X^N = \alpha(V^K, S^N) = \mathbf{T}V^K + \mathbf{M}S^N \quad (2)$$

where  $\mathbf{T}$  and  $\mathbf{M}$  are  $\mathbb{R}^{N \times K}$  and  $\mathbb{R}^{N \times N}$  matrices, respectively. In such case,  $Y^N$  is Gaussian and the minimum MSE decoder is a linear estimator, viz,  $\hat{V}^K = \Sigma_{VY} \Sigma_Y^{-1} Y^N$ , where  $\Sigma_{VY} = \mathbb{E}[(V^K)(Y^N)^T]$  and  $\Sigma_Y = \mathbb{E}[(Y^N)(Y^N)^T]$ . The matrices  $\mathbf{T}$  and  $\mathbf{M}$  can be found (numerically) by minimizing the MSE distortion  $D_{linear} = \mathbb{E}_F \left[ \frac{1}{K} \text{tr} \left\{ \sigma_V^2 \mathbf{I}_{k \times k} - \Sigma_{VY} \Sigma_Y^{-1} \Sigma_{VY}^T \right\} \right]$  under the power constraint in (1), where  $\text{tr}(\cdot)$  is the trace operator and  $\mathbf{I}_{k \times k}$  is a  $k \times k$  identity matrix. For the correlation model used in our simulations (Sec. V),  $\mathbf{T}$  and  $\mathbf{M}$  are found to be diagonal matrices. Note that by setting  $\mathbf{M}$  to be the zero matrix and  $\mathbf{T} = \sqrt{P/\sigma_V^2} \mathbf{I}_{N \times k}$ , the system reduces to the uncoded scheme. In the rest of this section we assume that  $K = N$ . In such case, the source-interference model can be described by the correlation coefficient  $\rho$  between  $V$  and  $S$ .

**Lemma 1** *Among all single-letter (symbol-by-symbol) codes with  $K = N = 1$ , the uncoded scheme is optimal for  $\rho = 1$ .*

*Proof:* Due to the correlation between the source and the interference, we can write  $S = S_I + S_D$ , where  $S_I$  and  $S_D$  are i.i.d. Gaussian vectors that are independent of each other with  $S_D = \frac{\rho\sigma_S}{\sigma_V} V$  and  $S_I \sim \mathcal{N}(\mathbf{0}, (1 - \rho^2)\sigma_S^2)$ . Assuming knowledge of  $\tilde{S} = \beta_1 S_I + \beta_2 S_D$  at the decoder, where  $\beta_1, \beta_2 \in \mathbb{R}$  (this is motivated by [7]), we have the following using the rate distortion theorem

$$\frac{1}{2} \log \frac{\text{Var}(V|\tilde{S})}{\mathbb{E}[(V - \hat{V})^2]} \leq I(V; \hat{V}|\tilde{S}) \quad (3)$$

where  $\text{Var}(V|\tilde{S}) = \sigma_V^2 \left( 1 - \frac{\beta_1^2 \rho^2}{\beta_1^2 (1 - \rho^2) + \beta_2^2 \rho^2} \right)$ . Moreover, using the data processing inequality for the mutual information, the facts that conditioning reduces differential entropy and that the Gaussian distribution maximizes differential entropy,  $I(V; \hat{V}|\tilde{S})$  can be upper bounded for any memoryless single-letter codes  $(\alpha, \gamma)$  with  $X = \alpha(V, S)$  and  $\hat{V} = \gamma(Y, f)$  by

$$I(V; Y|\tilde{S}) = h(Y|\tilde{S}) - h(W) \leq \frac{1}{2} \log \frac{\text{MSE}(Y; \tilde{S})}{\sigma_W^2} \quad (4)$$

where  $\text{MSE}(Y; \tilde{S})$  is the distortion from estimating  $Y$  based on  $\tilde{S}$  using a linear MSE estimator. In (4), we have used the fact that  $h(Y|\tilde{S}) \leq h(Y - \gamma_{lmse}(\tilde{S}))$ , where  $\gamma_{lmse}(\tilde{S})$  is the linear MSE estimator of  $Y$  based on  $\tilde{S}$ .  $\text{MSE}(Y; \tilde{S})$  is a function of  $\beta_1, \beta_2, \mathbb{E}[XS_I]$  and  $\mathbb{E}[XS_D]$ . By Cauchy-Swartz, we have  $|\mathbb{E}[XS_D]| \leq \sqrt{\mathbb{E}[X^2]\mathbb{E}[S_D^2]}$  and  $|\mathbb{E}[XS_I]| \leq \sqrt{\mathbb{E}[X^2]\mathbb{E}[S_I^2]}$ . For a given  $\beta_1$  and  $\beta_2$ , the maximum value of

$\text{MSE}(Y; \tilde{S})$  has to be used in (4). Combining (3) and (4), we have  $\mathbb{E}[(V - \hat{V})^2|f] \geq \frac{\text{Var}(V|\tilde{S})\sigma_W^2}{\text{MSE}(Y; \tilde{S})}$ . As a result,

$$\mathbb{E}[(V - \hat{V})^2] \geq \mathbb{E}_F \left[ \frac{\text{Var}(V|\tilde{S})\sigma_W^2}{\text{MSE}(Y; \tilde{S})} \right]. \quad (5)$$

The bound in (5) can be tightened by maximizing the right hand side of (5) over  $\beta_1$  and  $\beta_2$ . If we choose  $\beta_1 = 1$  and  $\beta_2 = 0$ , (5) reduces to

$$\mathbb{E}[(V - \hat{V})^2] \geq \mathbb{E}_F \left[ \frac{\sigma_V^2}{\left( 1 + \frac{|f|^2(\sqrt{P} + \rho\sigma_S)^2}{\sigma_W^2} \right)} \right]. \quad (6)$$

The MSE of the uncoded scheme (using  $D_{linear}$ ) is

$$D_{uncoded} = \mathbb{E}_F \left[ \frac{|f|^2 \sigma_V^2 \sigma_S^2 (1 - \rho^2)}{|f|^2 (P + \sigma_S^2 + 2\sqrt{P}\rho\sigma_S) + \sigma_W^2} \right] + \mathbb{E}_F \left[ \frac{\sigma_V^2}{1 + \frac{|f|^2 (P + \sigma_S^2 + 2\sqrt{P}\rho\sigma_S)}{\sigma_W^2}} \right]. \quad (7)$$

It is easy to see that equality occurs in (6) using the uncoded scheme for  $\rho = 1$  (see (7)). ■

**Remark 1** *To obtain an analytical bound on the system's distortion for the case of bandwidth reduction ( $r \leq 1$ ), we follow our previous model in [8] and assume that only the first  $N$  pairs  $\{(V(i), S(i))\}_{i=1}^N$  are correlated with correlation coefficient  $\rho$ . In such case, for a  $K : N$  system, we can derive the following distortion lower bound by assuming knowledge of  $\tilde{S}^N = \beta_1 S_I^N + \beta_2 S_D^N$  at the decoder:<sup>1</sup>*

$$D \geq \min_{\eta} \left( r \frac{\text{Var}(V|\tilde{S})}{\exp \left\{ \mathbb{E}_F \left[ \log \left( \frac{\text{MSE}(Y; \tilde{S})}{\eta P f^2 + \sigma_W^2} \right) \right] \right\}} + (1 - r) \frac{\sigma_V^2}{\exp \left\{ \mathbb{E}_F \left[ \frac{N}{K - N} \log \left( \frac{\eta P f^2 + \sigma_W^2}{\sigma_W^2} \right) \right] \right\}} \right) \quad (8)$$

where  $\eta \in [0, 1]$ . This bound generalizes the bounds in [2] as it allows fading and bandwidth reduction and is tighter than the bounds in [2] for some values of  $\rho$ . For  $r = 1$ , setting  $(\beta_1 = 1, \beta_2 = 1, \eta = 0)$  and  $(\beta_1 = 1, \beta_2 = 0, \eta = 0)$  in (8) lead to the generalization of the two bounds in [2] under matched bandwidth (for the fading channels). Note that the low-latency coding design algorithm (of Sec. IV) is for a general (correlation model) set-up.

**Remark 2** *For an AWGN channel or when the fading  $F$  is constant (almost surely), the uncoded scheme is optimal for the matched bandwidth case ( $r = 1$ ) among all  $\alpha : \mathbb{R}^K \times \mathbb{R}^K \rightarrow \mathbb{R}^K$  if  $\rho = 1$ . This can be shown by choosing  $(\beta_1 = 1, \beta_2 = 0, \eta = 0)$  in (8) and cancelling the expectation in (7) and (8).*

### IV. DESIGN ALGORITHM FOR LOW-LATENCY CODING

In this section, we present a scheme based on joint optimization between the encoder and the decoder through an iterative algorithm. It is difficult to optimize the encoder and the decoder in order to minimize the MSE distortion, and closed form expressions for  $\alpha(\cdot)$  and  $\gamma(\cdot)$  may not exist. Using the Lagrange multiplier method, the constrained minimization

<sup>1</sup>The derivation of the second term of (8) requires knowledge of  $V^N$  in addition to  $\tilde{S}^N$  at the decoder.

of the MSE subject to (1) can be recast into an unconstrained minimization via the Lagrange cost function

$$J(\alpha, \gamma) = \frac{\mathbb{E}[\|V^K - \gamma(Y^N, F^N)\|^2]}{K} + \lambda \frac{\mathbb{E}[\|\alpha(V^K, S^N)\|^2]}{N} \quad (9)$$

where  $\lambda$  is used to control the average power. The unconstrained minimization is still hard to solve due to interdependencies between the optimized components. To overcome this, we proceed in a way similar to classical design problems [9] by deriving the necessary conditions for optimality.

1) *Necessary Conditions for Optimality*: The optimal encoder mapping  $\alpha^*$  (assuming  $\gamma$  is fixed) is given by

$$\arg \min_{\alpha} \left\{ \frac{\mathbb{E}[\|V^K - \gamma(Y^N, F^N)\|^2]}{K} + \lambda \frac{\mathbb{E}[\|\alpha(V^K, S^N)\|^2]}{N} \right\} \quad (10)$$

Using Bayes' rule, the distortion  $\mathbb{E}[\|V^K - \hat{v}^K\|^2]$  is given by

$$\iiint p(v^k, s^n) p(y^n | \alpha(v^k, s^n), s^n, f^n) p(f^n) \|v^k - \hat{v}^k\|^2 dv^k ds^n dy^n df^n \quad (11)$$

where  $p(\cdot)$  and  $p(\cdot|\cdot)$  denote, respectively, a probability density function (pdf) and a conditional pdf. The average power is

$$P = \frac{1}{N} \iint p(v^k, s^n) \|\alpha(v^k, s^n)\|^2 dv^k ds^n. \quad (12)$$

Since  $p(v^k, s^n)$  in (11)–(12) is nonnegative,  $\alpha$  can be optimized “pointwise” for each  $v^k$  and  $s^n$  according to

$$\alpha^*(v^k, s^n) = \arg \min_{x^n \in \mathbb{R}^n} \left\{ \frac{1}{K} \iint p(y^n | x^n, s^n, f^n) \|v^k - \hat{v}^k\|^2 p(f^n) dy^n df^n + \frac{\lambda}{N} \|x^n\|^2 \right\}. \quad (13)$$

Thus, (13) is a necessary condition for an optimal encoder.

On the receiver side, the optimal decoder in the MSE sense (assuming  $\alpha$  is fixed) is given by  $\mathbb{E}[V^K | y^n, f^n]$  as follows

$$\gamma^*(y^n, f^n) = \frac{\iint v^k p(y^n | v^k, s^n, f^n) p(v^k, s^n) dv^k ds^n}{\iint p(y^n | v^k, s^n, f^n) p(v^k, s^n) dv^k ds^n}. \quad (14)$$

2) *Design Algorithm 1*: Based on the above necessary conditions for optimality, we optimize  $\alpha$  and  $\gamma$  using an iterative process based on (13) and (14). The update equations (13) and (14) yield a lower distortion at each iteration step; hence, with a finite amount of training data, convergence is ensured. One problem with such iterative techniques is that the final solution depends on the initialization of the algorithm and convergence to the global optimum is not guaranteed. The design Algorithm 1 is as follows

- 1) Choose some initial mapping for the encoder  $\alpha$ .
- 2) Find the optimal decoder  $\gamma$  according to (14).
- 3) Set the iteration index  $i = 0$  and the cost  $J^{(0)} = \infty$ .
- 4) Set  $i = i + 1$ .
- 5) Find the optimal mapping  $\alpha$  according to (13).
- 6) Find the optimal decoder  $\gamma$  according to (14).
- 7) Evaluate the cost function  $J^{(i)}$ . If the relative improvement of  $\frac{J^{(i-1)} - J^{(i)}}{J^{(i-1)}}$  is less than some positive threshold  $\varepsilon$  or  $i > I_{max}$ , stop iterating. Else go to step 4.

Algorithm 1 is nested inside a “bracketing” Lagrange multiplier search. We first set  $\lambda = \lambda_0$ . If the designed  $\alpha$  produces  $\frac{\mathbb{E}[\|\alpha(V^K, S^N)\|^2]}{N} > P$ ,  $\lambda_0$  is increased; else  $\lambda_0$  is decreased. The search ends if  $\frac{\mathbb{E}[\|\alpha(V^K, S^N)\|^2]}{N}$  is close enough to but  $< P$ .

3) *Implementation Aspects*: For the implementation of (13) and (14), some modifications are required. Since it is intractable to evaluate the formulas for all real-valued  $(V^K, S^N)$ , we form as in [10] a set of pairs  $(\mathcal{V}, \mathcal{S})$  composed of samples drawn from  $p(v^k, s^n)$ . Since the channel input and output spaces are real valued, we discretize them using a pulse amplitude modulation (PAM) alphabets  $\mathcal{X}$  and  $\mathcal{Y}$ , respectively, in each direction. We use  $\mathcal{X} = \mathcal{Y} = \{-e\frac{L-1}{2}, -e\frac{L-3}{2}, \dots, e\frac{L-3}{2}, e\frac{L-1}{2}\}$ , where  $e$  and  $L$  are the resolution and the cardinality of the set, respectively.

Even after discretizing the channel input and output using  $\mathcal{X}$  and  $\mathcal{Y}$ , the (offline) design algorithm is still computationally expensive. This is due to the fact that our problem have both fading and interference on top of AWGN. To lower the complexity, we resort to Monte-Carlo and importance sampling techniques. Using Monte-Carlo, (13) is given by

$$\alpha^* = \arg \min_{x^n \in \mathcal{X}^n} \sum_{i=1}^{\ell} \frac{\|v^k - \gamma(y_i^n, f_i^n)\|^2}{K} + \frac{\lambda}{N} \|x^n\|^2 \quad (15)$$

and (14) can be expressed as follows

$$\gamma^* = \frac{\sum_{i=1}^{\ell} v_i^k p(y_i^n | \alpha(v_i^k, s_i^n), s_i^n, f_i^n)}{\sum_{i=1}^{\ell} p(y_i^n | \alpha(v_i^k, s_i^n), s_i^n, f_i^n)} \quad (16)$$

where  $(v_i^k, s_i^n)$  and  $f_i^n$  are Monte Carlo (MC) samples that follow the distributions  $p(v^k, s^n)$  and  $p(f^n)$ , respectively, and  $y_i^n$  represents the discretized version of the corresponding MC samples using  $\mathcal{Y}$ ; equality in (15) and (16) are valid by the strong law of large numbers when  $\ell \rightarrow \infty$ . Note that the discretization of  $y_i^n$  is necessarily to couple (15) and (16); this is how we implement the “fixing” of decoder mapping when updating the encoder (15). To increase the convergence of (16) (i.e., lower  $\ell$  and the complexity), an alternative to sampling from  $p(v^k)$  is to use importance sampling. This is conducted by using samples from another distribution  $q(v^k)$ . After some manipulation, (16) is given by

$$\gamma^*(y^n, f^n) = \frac{\sum_{i=1}^{\ell} \frac{v_i^k p(y_i^n | \alpha(v_i^k, s_i^n), s_i^n, f_i^n) p(v_i^k, s_i^n)}{q(v_i^k)}}{\sum_{i=1}^{\ell} \frac{p(y_i^n | \alpha(v_i^k, s_i^n), s_i^n, f_i^n) p(v_i^k, s_i^n)}{q(v_i^k)}}. \quad (17)$$

$q(v^k)$  has to be chosen to improve the convergence speed. Since suboptimal linear decoding gives some information about the estimate, we choose  $q(v^k) \sim \mathcal{N}(v_{subopt}^k, D_{linear})$ , where  $v_{subopt}^k$  and  $D_{linear}$  are the source estimate and the MSE distortion from applying the linear decoder, respectively.

One major issue that stands against the scalability of Algorithm 1 is that for each pair of  $(v^k, s^n)$  one has to search over a set of cardinality  $|\mathcal{X}|^n$  which scales exponentially with the number of channel dimensions  $n$ . To reduce the search complexity, we use a heuristic targeted search approach. Instead of fixing the  $N$ -dimensional PAM alphabet  $\mathcal{X}^n$  and blindly searching over the whole region (fixed-search method), this targeted method starts by mapping each pair  $(v^k, s^n)$  to a channel input  $(x^n)^{(0)}$  using a parametric mapping and then searching over a region that varies for each pair and iteration  $i$ . For each source-interference pair, we choose the region to have a span of  $\mathcal{X}^{(i)} = (x)^{(i-1)} \pm 4(\sigma_{span})^{(i)}$  in each dimension, where  $(x^n)^{(i-1)}$  is the channel input found for that pair in iteration  $i - 1$ . We have chosen  $(\sigma_{span}^2)^{(i)}$  to decrease with  $i$

$$(\sigma_{span}^2)^{(i)} = c \frac{(L-1)e}{4} \left(1 - \frac{i}{I_{max}}\right)^3 \quad i = 1, \dots, I_{max}$$

TABLE I  
SDR IN DB VERSUS  $\rho$  FOR  $P = 0.6$ ,  $\sigma_S^2 = 1$ ,  $\sigma_W^2 = 0.01$  AND  $r = 1$ .

$\rho$	0	0.2	0.4	0.6	0.8	1
Uncoded	1.9	2.9	4.1	5.7	8.4	17.8
Linear	3.5	5.6	8.7	12.4	15.8	17.8
Optimized mapping	6.4	6.8	9.1	12.5	15.9	17.8
SDR upper bound (5)	12.2	12.4	12.9	14.2	16.9	17.8

where  $c$  is a constant. This method limits the search (at iteration  $i$ ) to some small neighbourhood of the current state  $(x^n)^{(i-1)}$ , with neighbourhoods diminishing in size as the algorithm progresses. This method is not used at the channel output and we still need to use a fixed  $\mathcal{Y}$ . We run Algorithm 1 twice with the second round initialized with the result of the first; this helps “escaping” from a local minimum at the beginning of the second round when  $(\sigma_{span}^2)^{(1)}$  is large. In our simulation, we use  $10^5$  pairs to define  $(\mathcal{V}, \mathcal{S})$ ,  $\epsilon = 10^{-3}$ ,  $I_{max} = 14$ ,  $c = 1/8$  and  $\ell = 250$  for  $K, N \leq 3$ . Note that after the design process of the encoder and decoder ends, both elements can be implemented using a table-lookup; for the encoder, this is done by quantizing the source-interference pairs and mapping each pair to a channel input.

## V. NUMERICAL RESULTS AND DISCUSSION

In this section, we consider a source with  $\sigma_V^2 = 1$  that is correlated to the interference using the correlation model in [8] (only the first  $\min(K, N)$  pairs  $\{V(i), S(i)\}_{i=1}^{\min(K, N)}$  are correlated) and Rayleigh fading with  $\mathbb{E}[F^2] = 1$ .

### A. Matched Bandwidth Case

Table I shows the SDR  $\triangleq \frac{\sigma_V^2}{D}$  of the optimized mapping (Section IV) along with the linear (2) and the uncoded schemes for  $K = N = 1$ . We first design our system for  $\rho = 0$  (linear scheme is used as initialization); the solution obtained is then used as initialization to design the system for higher values of  $\rho$ . We perform a numerical search over the coding matrices for the linear scheme so that  $D_{linear}$  (in Sec. III) is minimized under the power constraint  $P$ . We notice that as  $\rho$  increases, the SDR gap between the optimized and the linear scheme decreases (also the upper bound on all single-letter codes (5) with optimized values for  $\beta_1$  and  $\beta_2$ ). For  $\rho = 1$ , the numerical result confirms Lemma 1; all schemes reduce to the uncoded one and achieve optimality.

### B. Mismatched Bandwidth Case

Fig. 2 shows the SDR of our proposed mapping for 2:1 and 3:2 bandwidth reductions versus  $\text{CSNR} \triangleq P/\sigma_W^2$ . We notice that the optimized mapping outperforms the other reference schemes for all CSNRs (for the same  $K : N$ ). For the 2:1 system, we initialize Algorithm 1 with a parametric mapping based on a spiral curve [3] and the search range  $\mathcal{X}^{(i)}$  is set to 100 points. As a reference, we also plot the performance of an (optimized) parametric mapping based on the combination of spiral and sawtooth (modulo technique used for interference cancelation [5]). For the 3:2 case, a linear mapping is used for initialization and  $\mathcal{X}^{(i)}$  is set to 15 points in each direction. For reference, a parametric mapping that uses a linear transmission on the first symbol and 2:1 spiral mapping on the last two symbols to achieve a 3:2 reduction is combined linearly with the interference and simulated. For the alphabet  $\mathcal{Y}$ , we use  $e = 8/(L - 1)$ ,  $L = 500$  (for the 2:1 system) and  $L = 20$  (for the 3:2 system). Fig. 2 shows the SDR upper bound

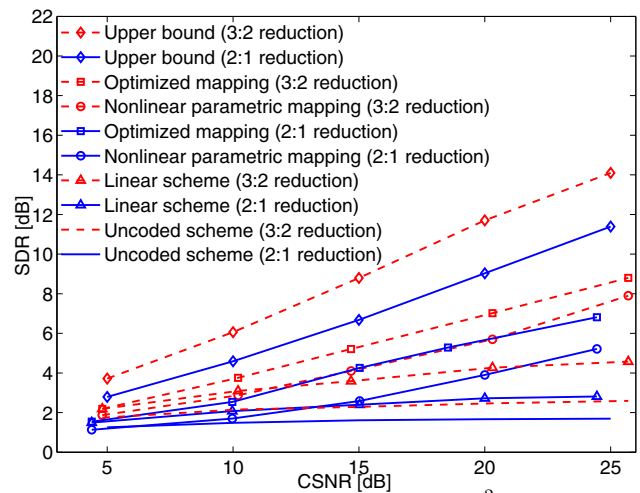


Fig. 2. SDR Performance versus CSNR for  $P = 1$ ,  $\sigma_S^2 = 1$ ,  $\rho = 0.3$ .

(using (8)); it is important to note that the bound is an asymptotic result in the sense of infinite source and coding block lengths, hence the gap to our low-delay scheme is not surprising. We also designed an optimized mapping for 1:2 expansion system; we noticed up to 1.6 dB gain in SDR over the “best” parametric system (linear in this case) for  $\text{CSNR}=20$  and  $\rho = 0.3$ . We used the optimized mapping when there is no fading as initialization to our algorithm. No gain is seen (over parametric mappings) when a linear mapping is used for initialization.

From other simulations, comparing the targeted search method to the fixed one using the same alphabets cardinality (with similar and tractable complexity), we notice that the proposed search method gives up to 1 dB gain over the fixed search method for moderate to high CSNRs (for 2:1 system). For low CSNRs, not much gain is observed; this is due to the fact that the (required) number of discrete points  $L$  depends on the noise levels (in our case,  $L$  is fixed). For high noise levels,  $L$  can be made smaller with no performance loss.

## REFERENCES

- [1] M. Costa, “Writing on dirty paper,” *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.
- [2] Y.-C. Huang and K. R. Narayanan, “Joint source-channel coding with correlated interference,” *IEEE Trans. Commun.*, vol. 60, no. 5, pp. 1315–1327, May 2012.
- [3] F. Hekland, P. A. Floor, and T. A. Ramstad, “Shannon-Kotel’nikov mappings in joint source-channel coding,” *IEEE Trans. Commun.*, vol. 57, no. 1, pp. 94–105, Jan. 2009.
- [4] E. Akyol, K. Rose, and T. Ramstad, “Optimal mappings for joint source channel coding,” in *2010 IEEE Inform. Theory Workshop*.
- [5] J. Du, E. G. Larsson, M. Xiao, and M. Skoglund, “Optimal symbol-by-symbol cost precoding for a relay-aided downlink channel,” *IEEE Trans. Commun.*, vol. 59, no. 8, pp. 2274–2284, Aug. 2011.
- [6] W. Zhang, S. Kotagiri, and J. Laneman, “Writing on dirty paper with resizing and its application to quasi-static fading broadcast channels,” in *Proc. 2007 IEEE Int. Symp. on Inform. Theory*.
- [7] Y.-K. Chia, R. Soundararajan, and T. Weissman, “Estimation with a helper who knows the interference,” in *Proc. 2012 IEEE Int. Symp. on Inform. Theory*.
- [8] A. Abou Saleh, F. Alajaji, and W.-Y. Chan, “Hybrid digital-analog coding for interference broadcast channels,” in *Proc. 2013 IEEE Int. Symp. on Inform. Theory*.
- [9] N. Farvardin and V. Vaishampayan, “On the performance and complexity of channel-optimized vector quantizers,” *IEEE Trans. Inf. Theory*, vol. 37, pp. 155–159, Jan. 1991.
- [10] J. (Karlsson) Kron and M. Skoglund, “Optimized low-delay source-channel-relay mappings,” *IEEE Trans. Commun.*, vol. 58, no. 5, pp. 1397–1404, May 2010.