Analog Coding for Gaussian Source and State Interference Estimation

Ahmad Abou Saleh, Fady Alajaji, and Wai-Yip Chan
Queen's University, Kingston, ON, K7L 3N6
Email: ahmad.abou.saleh@queensu.ca, fady@mast.queensu.ca, chan@queensu.ca

Abstract—We consider zero-delay analog coding of a Gaussian source over a Gaussian channel with additive correlated Gaussian interference known to the transmitter. The receiver aims to jointly estimate the source signal and the state interference. We propose a layered parametric analog coding scheme based on linear and sawtooth mappings. We derive an upper bound on the distortion for the parametric scheme by assuming a suboptimal decoder. To optimize the system parameters, we use two suboptimal methods. The first one is partially numerical and part of the parameters are derived assuming no sawtooth mapping; the other one, however, is based on minimizing the derived upper bound. To improve the performance whenever storage and offline design complexity are not an issue, we design a nonparametric mapping through an iterative process based on joint optimization between the encoder and the decoder using the necessary conditions for optimality. Numerical results show that the nonparametric and parametric mappings outperform the linear scheme and overcome the saturation effect.

I. INTRODUCTION

In a point-to-point communication system, it is well known that for a Gaussian source transmitted over an additive white Gaussian channel (AWGN) with equal source-channel bandwidth, a simple analog coding that outputs a scaled version of the source is optimal in the mean square sense [1]. This linear analog approach, which is referred to as uncoded transmission, has low complexity and zero coding delay; this is not the case, however, for the optimal separate source-channel coding (tandem scheme) which requires infinite delay and complexity [2]. Moreover, for a Gaussian source sent over an AWGN channel with Gaussian additive interference known (non-causally) to the transmitter, a tandem scheme which comprises an optimal quantizer followed by Costa's dirty paper channel code [3] is optimal in the absence of correlation between the source and the interference. In [4], the authors consider a communication scenario with additive interference at the transmitter and side information at the receiver in which they propose a scheme based on lattice coding. For correlated source-interference, the authors in [5] propose a hybrid digital-analog (HDA) scheme for the matched source-channel bandwidth case; although this scheme is able to benefit from the correlation between the source and the interference, the optimal performance is not yet known. In [6], we study HDA coding under a similar communication scenario but with mismatch source-channel bandwidth. In some cases, the receiver is simultaneously interested in recovering the source signal and the state interference. This communication

problem was first considered in [7]; the authors show that the optimal rate-state-distortion tradeoff is achieved by using analog state transmission and message transmission via dirty paper coding. In [6], we also study this problem when the source and the interference are correlated; we propose a high delay HDA scheme that performs close to the outer bound.

Low delay analog coding is getting more relevant than ever with the increase popularity of wireless sensor networks [8]–[14]. Different from the works in [6], [7] which assume infinite complexity and delay, in this paper, we focus on zero delay and low complexity analog coding and study the reliable communication of memoryless Gaussian source and state interference over AWGN channels with interference known to the transmitter.
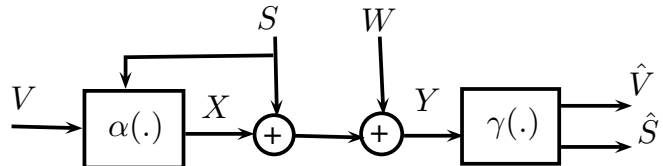

Fig. 1. System model.

II. PROBLEM FORMULATION

We consider the transmission of analog source and channel state interference over an AWGN channel with additive interference that is known to the transmitter. Our aim is to estimate the source V and the interference S which are assumed to have a Gaussian distribution. As shown in Fig. 1, the source and the interference are transformed into a channel input X using a nonlinear mapping, $\alpha(\cdot) : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. The received signal is given by $Y = \alpha(V, S) + S + W$, where $X = \alpha(V, S)$ is the channel input and W is additive Gaussian noise, independent of V and S, with variance σ_W^2 $(W \sim \mathcal{N}(0, \sigma_W^2))$. The channel operates under an input average power constraint P given by

$$\mathbb{E}[\alpha(V, S)^2] \leq P \tag{1}$$

where $\mathbb{E}[\cdot]$ denotes expectation. Moreover, we assume that (V, S) are correlated via the following covariance matrix

$$\Sigma_{VS} = \begin{bmatrix} \sigma_V^2 & \rho \sigma_V \sigma_S \\ \rho \sigma_V \sigma_S & \sigma_S^2 \end{bmatrix} \tag{2}$$

where σ_V^2, σ_S^2 are, respectively, the variance of the source and the interference, and ρ is the source-interference correlation coefficient. At the receiver side, the estimated source-interference pairs are reconstructed as follows $(\hat{V}, \hat{S}) = \gamma(Y) = (\gamma_v(Y), \gamma_s(Y))$, where $\gamma_v(\cdot), \gamma_s(\cdot) : \mathbb{R} \to \mathbb{R}$.

In this work, we aim to find an encoder α decoder γ that minimize the mean square error (MSE) distortion defined by

$$D \triangleq \underbrace{\mathbb{E}[(V-\hat{V})^2]}_{D_v} + \underbrace{\mathbb{E}[(S-\hat{S})^2]}_{D_s}. \qquad (3)$$

This minimization is carried out under the average power constraint in (1). In the sequel, we will focus on zero delay analog joint source-channel coding techniques. More precisely, we study parametric and nonparametric (nonlinear) mappings. The main attempts in solving this problem without imposing delay and complexity constraints have used hybrid coding techniques [6], [7] which perform close to the derived outer bound (note that the optimal limit for this system is not known). Our main contributions can be summarized as follows

- To benefit from nonlinearity whenever possible, we study a parametric analog mapping based on the sawtooth (modulo) function which has a low design complexity. We derive an upper bound on the system's distortion by assuming a suboptimal decoder at the receiver side. To optimize the system parameters, we use two suboptimal methods; the first one is partially numerical in which part of the parameters are derived by minimizing a (partial) distortion expression that assumes no sawtooth mapping is used at the encoder. The other method minimizes the derived upper bound expression.
- Whenever storage and offline design complexity are not an issue, we design a nonlinear mapping; this is done by deriving the necessary conditions for optimality and proposing an iterative algorithm based on joint optimization between the transmitter and the receiver. We refer to this scheme as a "nonparametric" analog mapping.

III. DISTORTION LOWER BOUND

In our previous work [6, Lemma 4], we derived an outer bound on the distortion region (D_v, D_s) for the same problem over fading channels. The bound on D_v was derived by assuming knowledge of S at the decoder; while the bound on D_s was obtained without assuming any additional knowledge. For the AWGN channel, the outer bound simplifies as follows

$$
\begin{aligned}
D_v &\geq D_v^{ob} \triangleq \frac{\mathrm{Var}(V|S)\sigma_W^2}{\zeta P + \sigma_W^2}, \\
D_s &\geq D_s^{ob} \triangleq \frac{\sigma_S^2(\zeta P + \sigma_W^2)}{P + \sigma_S^2 + 2\sqrt{(1-\zeta)P\sigma_S^2} + \sigma_W^2}
\end{aligned} \qquad (4)
$$

where $\mathrm{Var}(V|S) = \sigma_V^2(1-\rho^2)$ is the variance of V given S and $\zeta \in [0\ 1]$. As a result, the lower bound on the system's distortion can be expressed as follows $D \geq \inf_\zeta \{D_v^{ob} + D_s^{ob}\}$.

IV. LINEAR SCHEME

In this section, we assume that the encoder transforms the source-interference pair (V, S) into a channel input X using a linear transformation according to $X = \alpha(V, S) = \sqrt{a}(\alpha_1 V + \alpha_2 S)$, where $\alpha_1, \alpha_2 \in [-1, 1]$ and $a = P/(\alpha_1^2\sigma_V^2 + \alpha_2^2\sigma_S^2 + 2\alpha_1\alpha_2\rho\sigma_V\sigma_S)$ is a gain factor related to the power constraint P. In such case, the received signal Y is Gaussian and the minimum MSE (MMSE) decoder is a linear estimator

$$\hat{V} = \frac{\mathbb{E}[VY]}{\mathbb{E}[Y^2]}Y, \quad \hat{S} = \frac{\mathbb{E}[SY]}{\mathbb{E}[Y^2]}Y \qquad (5)$$

where $\mathbb{E}[VY] = \sqrt{a}\alpha_1\sigma_V^2 + (1 + \sqrt{a}\alpha_2)\rho\sigma_V\sigma_S$, $\mathbb{E}[SY] = \sqrt{a}\alpha_1\rho\sigma_V\sigma_S + (1 + \sqrt{a}\alpha_2)\sigma_S^2$ and $\mathbb{E}[Y^2] = P + \sigma_S^2 + \sigma_W^2 + 2\sqrt{a}(\alpha_1\rho\sigma_V\sigma_S + \alpha_2\sigma_S^2)$. The resulting distortion from reconstructing V and S based on $\hat{V}$ and $\hat{S}$, respectively, can be expressed as follows

$$D_v^{Lin} = \sigma_V^2 - \frac{\mathbb{E}[VY]^2}{\mathbb{E}[Y^2]}, \quad D_s^{Lin} = \sigma_S^2 - \frac{\mathbb{E}[SY]^2}{\mathbb{E}[Y^2]}. \qquad (6)$$

The overall system distortion is then given by $D_{Linear} = \inf_{\alpha_1, \alpha_2}\{D_v^{Lin} + D_s^{Lin}\}$.

Remark 1 *For $\rho = 1$, the problem reduces to just estimating the interference S. The linear scheme, which simplifies to the uncoded scheme (i.e., $X = \sqrt{P/\sigma_S^2}S$), is optimal. The tightest lower bound on D_s can be obtained by setting $\zeta = 0$ in (4); in such case the distortion in (4) reduces to $D_s \geq \sigma_S^2\sigma_W^2/((\sqrt{P} + \sigma_S)^2 + \sigma_W^2)$. It is easy to prove that this distortion coincides with the one of the uncoded scheme (use (6) with $\alpha_1 = 0$ and $\alpha_2 = 1$). The optimality of the uncoded scheme for $\rho = 1$ was also noticed in [7].*

V. PARAMETRIC MAPPING

Recall that since V and S are correlated, we can write the source as $V = \frac{\rho\sigma_V}{\sigma_S}S + N_\rho$, where $N_\rho \sim \mathcal{N}(0, (1 - \rho^2)\sigma_V^2)$ is independent of S. Motivated by the high delay scheme that uses analog and hybrid layers [6], we herein propose a layered scheme based on linear coding and sawtooth mapping (sawtooth coding is used in [15] for the relay channel). The sawtooth mapping can be seen as one dimensional lattice coding; high delay lattice coding is widely studied for AWGN channel with side information [4], [16].

A. System Structure

The proposed scheme is composed of two superposed layers and outputs

$$X = c(X_1 + X_2) \qquad (7)$$

where c is a gain factor related to the power constraint (defined later). The first layer, which outputs $X_1 = \sqrt{P_s/\sigma_S^2}S$, simply scales the interference S, where $P_s \leq P$ represents the power consumed by this layer. The second layer, starts by forming a linear combination of the partial information of the source N_ρ and the interference S; this is given by $X_a = \alpha_1 N_\rho + \alpha_2 S$, where α_1, α_2 are real parameters. We then use a sawtooth mapping $\mathcal{S}(\cdot)$ on X_a to output X_2 as follows

$$\mathcal{S}(X_a) = (X_a - 2\Lambda m) \text{ for } X_a \in \big[\Lambda(2m-1), \Lambda(2m+1)\big) \quad (8)$$

where m is an integer and Λ is a nonnegative parameter dependent on the channel condition. The gain factor c in (7) is given by $c = \sqrt{P/(P_s + \mathbb{E}[X_2^2] + 2\sqrt{P_s/\sigma_S^2}\mathbb{E}[SX_2])}$, where $\mathbb{E}[X_2^2]$ can be written as follows

$$\mathbb{E}[X_a^2] + \sum_m \underbrace{-4\Lambda m \int_{D_m} x_a p(x_a)dx_a}_{I_1} + 4\Lambda^2 m^2 \underbrace{\int_{D_m} p(x_a)dx_a}_{I_2} \qquad (9)$$

and $\mathbb{E}[SX_2]$ is given by

$$\alpha_2\sigma_S^2 - \sum_m 2\Lambda m \int\int_{D_m} sf(x_a - \alpha_2 s)p(s)dx_a ds \qquad (10)$$

where $D_m = [\Lambda(2m-1), \Lambda(2m+1))$ is the mth domain region of $\mathcal{S}(\cdot)$, $p(\cdot)$ denotes a probability density function (pdf) and $f(\cdot)$ is the pdf of $\alpha_1 N_\rho \sim \mathcal{N}(0, \alpha_1^2 \sigma_V^2 (1-\rho^2))$. Note that the integrals in (9) can be simplified as follows

$$
\begin{aligned}
I_1 &= \frac{\sqrt{\mathbb{E}[X_a^2]}}{\sqrt{2\pi}} \left[-\exp\left(\frac{-(\Lambda(2m+1))^2}{2\mathbb{E}[X_a^2]} \right) \right. \\
&\quad \left. +\exp\left(\frac{-(\Lambda(2m-1))^2}{2\mathbb{E}[X_a^2]} \right) \right], \\
I_2 &= \frac{1}{2} \left[\mathrm{erf}\left(\frac{\Lambda(2m+1)}{\sqrt{2\mathbb{E}[X_a^2]}} \right) - \mathrm{erf}\left(\frac{\Lambda(2m-1)}{\sqrt{2\mathbb{E}[X_a^2]}} \right) \right] \quad (11)
\end{aligned}
$$

where $\mathrm{erf}(\cdot)$ is the Gaussian error function. At the decoder side, to obtain an estimate of the source and the interference, we use the optimal MMSE estimator ($\hat{V} = \mathbb{E}[V|Y], \hat{S} = \mathbb{E}[S|Y]$). The use of an optimal decoder comes at the expense of computational and design complexity. To lower the design complexity, we resort to two suboptimal methods for choosing the system parameters as described next.

B. System Optimization

1) Method 1: In this method, the optimized parameters α_1 and α_2 are found by assuming that no sawtooth mapping is used. In such case, the parameters α_1 and α_2 are found by minimizing the MSE distortion D_{N_ρ} from reconstructing N_ρ using a linear MMSE (LMMSE) estimator. This distortion is given by (assuming no sawtooth mapping) $D_{N_\rho} = \sigma_V^2(1 - \rho^2) - \frac{\mathbb{E}[N_\rho Y]^2}{\mathbb{E}[Y^2]}$. The sawtooth parameter Λ and P_s, however, are found numerically to minimize the overall MSE distortion D by performing a grid search. This is done by generating a large set of (V, S, W) triplets and computing D empirically for each possible (P_s, Λ) in the search space.

2) Method 2: In this method, we optimize the system parameters by minimizing an upper bound on the system's distortion. To get a closed form expression on the upper bound, we propose the use of a suboptimal decoder. Let us first note that the sawtooth mapping, which uses the symmetric modulo function (8) over the interval $[-\Lambda \ \Lambda]$, can be written as $\mathcal{S}(X_a) = X_a \ mod \ \Lambda$. To reconstruct the interference, we simply use an LMMSE estimator based on the received signal Y. The distortion from reconstructing S is given by

$$
\begin{aligned}
(D_s)_{parametric} &= \sigma_S^2 - \frac{\mathbb{E}[SY]^2}{\mathbb{E}[Y^2]} \\
&= \sigma_S^2 - \frac{(c(\mathbb{E}[S(X_a \ mod \ \Lambda)] + \sqrt{P_s}\sigma_S) + \sigma_S^2)^2}{P + \sigma_S^2 + \sigma_W^2 + 2c(\sqrt{P_s}\sigma_S + \mathbb{E}[S(X_a \ mod \ \Lambda)])}
\end{aligned}
$$

where $\mathbb{E}[S(X_a \ mod \ \Lambda)]$ can be written as in (10).

To get an estimate of the source V, we first use a modulo function on the received signal and then apply an LMMSE estimator. More precisely, we first obtain

$$
\begin{aligned}
\tilde{Y} &= (Y/c) \ mod \ \Lambda \\
&= \underbrace{\left(\alpha_1 N_\rho + (\alpha_2 + \sqrt{\frac{P_s}{\sigma_S^2}} + \frac{1}{c})S + \frac{W}{c} \right)}_{Z} \ mod \ \Lambda \quad (12)
\end{aligned}
$$

where the last equality follows from the fact that the modulo operation satisfies the "distributive law" (i.e., $[x \ mod \ \Lambda +$

$y] \ mod \ \Lambda = [x + y] \ mod \ \Lambda$). We then decode V using an LMMSE estimator based on $\tilde{Y}$. The resulting distortion is

$$
(D_v)_{parametric} = \sigma_V^2 - \frac{\mathbb{E}[V(Z \ mod \ \Lambda)]^2}{\mathbb{E}[(Z \ mod \ \Lambda)^2]} \quad (13)
$$

where Z is defined in (12), $\mathbb{E}[(Z \ mod \ \Lambda)^2]$ is given by

$$
\mathbb{E}[Z^2] - 4\Lambda \sum_m m \int_{\tilde{D}_m} zp(z)dz + 4\Lambda^2 \sum_m m^2 \int_{D_m} p(z)dz \quad (14)
$$

and $\mathbb{E}[V(Z \ mod \ \Lambda)]$ can be expressed as follows

$$
\mathbb{E}[VZ] - \sum_m 2\Lambda m \int_{\tilde{D}_m} \iint vp(v|s)p(z|v,s)p(s)dvdsdz \quad (15)
$$

where $\tilde{D}_m = [\Lambda(2m-1), \Lambda(2m+1))$ is the mth domain region of $\mathcal{S}(\cdot)$. Note that integrals in (14) can be simplified in a similar way as in (11) and distributions in (15) are Gaussian. The upper bound on the system's distortion D_{upper} for parametric mapping is then given by

$$
D_{upper} = (D_v)_{parametric} + (D_s)_{parametric}. \quad (16)
$$

Note that $(\alpha_1, \alpha_2, P_s, \Lambda)$ are found by minimizing D_{upper}.

Remark 2 *For low correlation values between the source and the interference, we propose to use a slightly different decoder (for method 2) for estimating V that gives a better performance. The only modification is that $\tilde{Y} = \frac{-\alpha_2 Y}{c\kappa} \ mod \ \Lambda$, where $\kappa = \left(\sqrt{\frac{P_u}{\sigma_S^2}} + \frac{1}{c} \right)$. After some manipulations, we can write $\tilde{Y} = (\alpha_1 V + W_{eq}) \ mod \ \Lambda$, where $Z_{eq} = \frac{-\alpha_2 W}{c\kappa} - (1 + \frac{\alpha_2}{\kappa})X_a \ mod \ \Lambda$. Note that W_{eq} can be regarded as an equivalent noise term.*

VI. Nonparametric Mapping

We next present a scheme based on joint optimization between the encoder and the decoder through an iterative algorithm. Closed form expressions for $\alpha(\cdot)$ and $\gamma(\cdot) = (\gamma_v(\cdot), \gamma_s(\cdot))$ that minimize the distortion D may not exist; this makes the optimization of the encoder and decoder difficult. The rest of this section is dedicated to the design of the source-channel mapping $\alpha(V, S)$ and the decoder $\gamma(Y)$. Using the Lagrange multiplier method, the constrained minimization of the MSE distortion D subject to the power constraint in (1) can be recast into an unconstrained minimization via the Lagrange cost function $J(\alpha, \gamma)$ given by

$$
\mathbb{E}[(V - \gamma_v(Y))^2] + \mathbb{E}[(S - \gamma_s(Y))^2] + \lambda \mathbb{E}[\alpha(V, S)^2] \quad (17)
$$

where λ is used to control the average power. The above unconstrained minimization is still hard to solve due to interdependencies between the optimized components. To solve this, we proceed in a way similar to classical design problems [17] by deriving necessary conditions for optimality.

A. Necessary Conditions for Optimality

The optimal encoder mapping α^* (assuming γ is fixed) is

$$
\arg\min_\alpha \left\{ \mathbb{E}[(V - \hat{V})^2] + \mathbb{E}[(S - \hat{S})^2] + \lambda \mathbb{E}[\alpha(V, S)^2] \right\}. \quad (18)
$$

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

$$\iiint p(v,s)p(y|\alpha(v,s),s)(v-\hat{v})^2 dv ds dy \quad (19)$$

where $p(\cdot|\cdot)$ denotes a conditional pdf. Similarly, the distortion $\mathbb{E}[(S - \hat{S})^2]$ can be expressed as follows

$$\iiint p(v,s)p(y|\alpha(v,s),s)(s-\hat{s})^2 dv ds dy. \quad (20)$$

The average consumed power is given by

$$P = \iint p(v,s)\alpha(v,s)^2 dv ds. \quad (21)$$

Since $p(v,s)$ in (19), (20) and (21) is nonnegative, the encoder α^* can be optimized "pointwise" for each (v,s) according to

$$\arg\min_{x\in\mathbb{R}} \left\{ \int p(y|x,s)[(v-\hat{v})^2 + (s-\hat{s})^2]dy + \lambda x^2 \right\}. \quad (22)$$

Thus, (22) is a necessary condition for an optimal encoder.

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

$$\gamma_v^*(y) = \frac{\iint v p(y|v,s)p(v,s)dv ds}{\iint p(y|v,s)p(v,s)dv ds}. \quad (23)$$

Similarly, the optimal decoder γ_s in the MSE sense (assuming α is fixed) is given by $\mathbb{E}[S|y]$ as follows

$$\gamma_s^*(y) = \frac{\iint s p(y|v,s)p(v,s)dv ds}{\iint p(y|v,s)p(v,s)dv ds}. \quad (24)$$

B. Design Algorithm

Using the above necessary conditions for optimality, we optimize α and γ via an iterative process based on (22), (23) and (24). The update equations (22), (23) and (24) yield a lower distortion at each iteration step; Thus, with a finite amount of training data, convergence is guaranteed. The main problem with such iterative process is that the final solution depends on the choice of the initial mapping in the algorithm and convergence to the global optimum is not ensured. The design Algorithm 1 is as follows:

1) Choose some initial mapping for the encoder α.
2) Find the optimal decoder $\gamma = (\gamma_v, \gamma_s)$ according to (23) and (24).
3) Set the iteration index $i = 0$ and the cost $J^{(0)} = \infty$.
4) Set $i = i + 1$.
5) Find the optimal mapping α according to (22).
6) Find the optimal decoder $\gamma = (\gamma_v, \gamma_s)$ according to (23) and (24).
7) Evaluate the cost function $J^{(i)}$. If the relative improvement of $\frac{J^{(i-1)} - J^{(i)}}{J^{(i-1)}} < \varepsilon$ 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 α produces $\mathbb{E}[\alpha(V,S)^2] > P$, λ_0 is increased; else λ_0 is decreased. The search ends if $\mathbb{E}[\alpha(V,S)^2]$ is close enough to but less than P. For initialization of the algorithm, we notice that using the proposed parametric mapping results in better performance than the linear scheme.

C. Implementation Aspects

For the implementation of (22), (23) and (24), some modifications are required. Since it is intractable to evaluate the formulas for all real-valued (V,S), we form as in [11] a set of pairs $(\mathcal{V}, \mathcal{S})$ composed of samples drawn from $p(v,s)$. Since the channel input and output spaces are real valued, we discretize them using a pulse amplitude modulation alphabets $\mathcal{X}$ and $\mathcal{Y}$, respectively, in each direction. We use

$$\mathcal{X} = \mathcal{Y} = \left\{ -e\frac{L-1}{2}, -e\frac{L-3}{2}, ..., e\frac{L-3}{2}, e\frac{L-1}{2} \right\} \quad (25)$$

where e and L are the resolution and the cardinality of the set, respectively. In our simulations, we use 10^5 pairs $(\mathcal{V}, \mathcal{S})$, $\epsilon = 10^{-3}$, $I_{max} = 15$, $L = 700$ and $e = 12/(L-1)$.

VII. NUMERICAL RESULTS

In this section, we consider a source-interference pairs that are transmitted over an AWGN channel with Gaussian interference and power constraint $P = 1$. Fig. 2 shows the performance, defined as $10\log\frac{1}{D}$, versus the correlation ρ for channel signal-to-noise ratio $\text{CSNR} \triangleq P/\sigma_W^2 = 25$ dB.

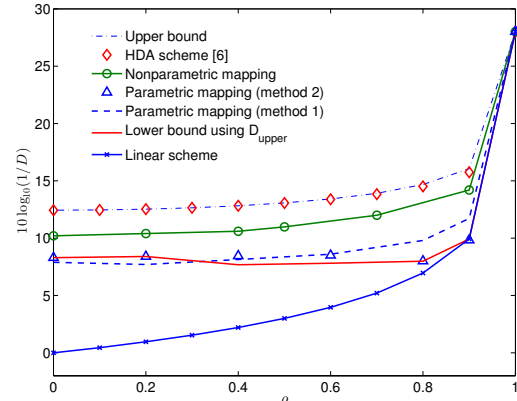Fig. 2. Performance versus source-interference correlation ρ, for $\sigma_V^2 = \sigma_S^2 = 1$ and CSNR = 25 dB.

The nonparametric mapping outperforms other zero delay coding schemes; using parametric mapping as initialization for Algorithm 1 gives 0.5 to 1 dB gain in performance (for $\rho \leq 0.8$) over the case where we use a linear mapping to initialize the algorithm. The parametric mapping (with MMSE decoding) which is easier to design, outperforms the linear scheme and performs relatively close to the nonparametric mapping. Note that using the optimized parameters resulting from method 1, gives the 'best' performance for high correlation values; for low correlation values, however, it is better to use the optimized parameters resulting from minimizing D_{upper} given by (16) (method 2). This behaviour comes from the fact that in method 2 we minimize a different objective function than the one used in method 1. As shown from Fig. 2, the lower bound obtained from the derived D_{upper} is close to the performance of the parametric mapping (optimized via method 1) with MMSE decoding for low to moderate correlation values; for high correlation values, we can notice some gap that is also manifested in the performance of method 2. Moreover, for the case of $\rho = 1$, all schemes revert to the uncoded scheme which is optimal. It is worth

mentioning that the upper bound and the HDA scheme in [6] are asymptotic in the sense of requiring infinite source and coding block lengths, hence the gap to the proposed zero delay schemes is not surprising.

Fig. 3 shows the performance versus CSNR levels. We can notice that the nonparametric mapping outperforms other zero delay schemes and that the lower bound found using D_{upper} is very close to the performance of the parametric mapping (optimized via method 1) with MMSE decoder. Moreover, the use of nonlinear mappings defeats the saturation effect which is inevitable with the use of the linear scheme.
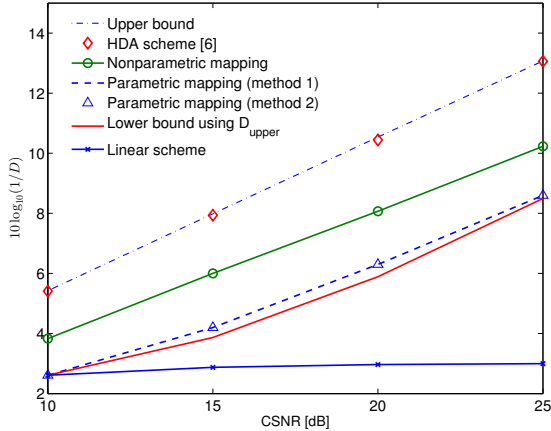

Fig. 3. Performance versus CSNR levels, for $\rho = 0.5$ and $\sigma_V^2 = \sigma_S^2 = 1$.

Figs. 4 and 5 show the encoder-decoder structure of the nonparametric mapping for two different correlation values. It is clear that the encoder and decoder mappings comprise a piecewise nonlinear function that combines hard and soft decision signalling. The proposed parametric mapping uses such combination; this explains the good performance achieved using parametric mapping. There is always a gain from using the nonparametric mapping; this is due to the fact that the nonparametric mapping has a higher degree of freedom in placing points in space without being restrained to a specific structure. Such gain comes at the expense of higher storage and offline design complexity.
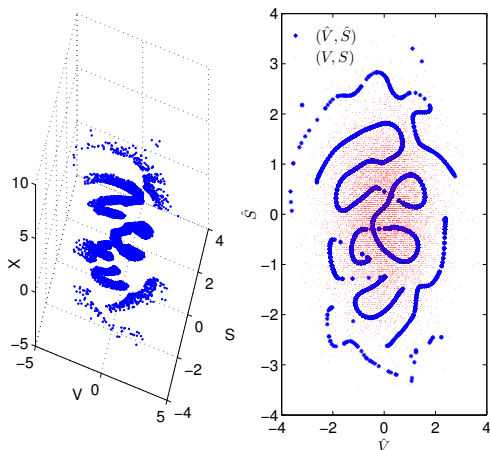

Fig. 4. Encoder (left) and its corresponding decoder (right) mappings optimized using Algorithm 1 for CSNR = 25 dB, $\rho = 0$, $\sigma_V = \sigma_S = 1$ and $P = 1$; parametric mapping is used for the initialization of Algorithm 1. In the figure to the right, the asterisks show the reconstructed $(\hat{V}, \hat{S})$ and the small dots are samples from the distribution of (V, S).
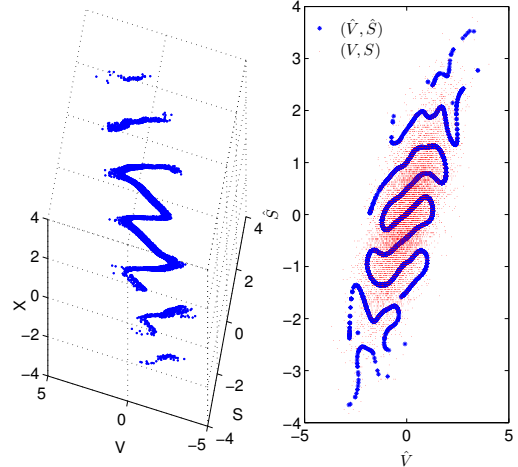

Fig. 5. Encoder (left) and its corresponding decoder (right) mappings optimized using Algorithm 1 for CSNR = 25 dB, $\rho = 0.7$, $\sigma_V = \sigma_S = 1$ and $P = 1$; parametric mapping is used for the initialization of Algorithm 1. In the figure to the right, the asterisks show the reconstructed $(\hat{V}, \hat{S})$ and the small dots are samples from the distribution of (V, S).

REFERENCES

[1] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: lossy source-channel communication revisited," *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1147–1158, May 2003.

[2] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.

[3] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[4] Y. Kochman and R. Zamir, "Joint Wyner-Ziv/dirty-paper coding by modulo-lattice modulation," *IEEE Trans. Inform. Theory*, vol. 55, no. 11, pp. 4878–4889, Nov 2009.

[5] 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.

[6] A. Abou Saleh, W.-Y. Chan, and F. Alajaji, "Source-channel coding for fading channels with correlated interference," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 3997–4011, Nov 2014.

[7] A. Sutivong, M. Chiang, T. Cover, and Y.-H. Kim, "Channel capacity and state estimation for state-dependent Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 4, pp. 1486–1495, Apr 2005.

[8] T. A. Ramstad, "Shannon mappings for robust communication," *Telektronikk*, vol. 98, no. 1, pp. 114–128, 2002.

[9] 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.

[10] Y. Hu, J. Garcia-Frias, and M. Lamarca, "Analog joint source channel coding using non-linear mappings and MMSE decoding," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3016–3026, Nov 2011.

[11] 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.

[12] P. A. Floor, A. N. Kim, N. Wernersson, T. Ramstad, M. Skoglund, and I. Balasingham, "Zero-delay joint source-channel coding for a bivariate Gaussian on a Gaussian MAC," *IEEE Trans. Comm.*, vol. 60, no. 10, pp. 3091–3102, Oct 2012.

[13] J. (Karlsson) Kron, F. Alajaji, and M. Skoglund, "Low-delay joint source-channel mappings for the Gaussian MAC," *IEEE Commun. Letters*, vol. 18, no. 2, pp. 249–252, Feb 2014.

[14] A. Abou Saleh, F. Alajaji, and W.-Y. Chan, "Low-latency source-channel coding for fading channels with correlated interference," *IEEE Wireless Commun. Letters*, vol. 3, no. 2, pp. 137–140, April 2014.

[15] S. Yao, M. Khormuji, and M. Skoglund, "Sawtooth relaying," *IEEE Commun. Letters*, vol. 12, no. 9, pp. 612–614, Sep 2008.

[16] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice strategies for canceling known interference," *IEEE Trans. Inform. Theory*, vol. 51, no. 11, pp. 3820–3833, Nov 2005.

[17] N. Farvardin and V. Vaishampayan, "On the performance and complexity of channel-optimized vector quantizers," *IEEE Trans. Inform. Theory*, vol. 37, pp. 155–159, Jan 1991.