

# Power-Constrained Bandwidth-Reduction Source-Channel Mappings for Fading Channels

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 the transmission of a memoryless Gaussian source over a power-constrained Rayleigh fading channel with additive white Gaussian noise. We propose the use of low-delay joint source-channel mappings and consider optimizing the non-parametric mappings through an iterative process. A design algorithm for joint source-channel mapping is proposed and numerically evaluated for 2:1, 3:1, and 4:1 bandwidth reductions. Parametric mappings are also studied. We assume three cases of fading knowledge; in the case of presence of channel state information at both encoder and decoder, optimal power allocation is solved for the parametric mappings in terms of fading gains and average power constraint. It is shown that the proposed non-parametric and parametric mappings, which have a non-linear structure, achieve a graceful and robust performance and considerably surmount the saturation effect of linear systems.

## I. INTRODUCTION

The traditional approach for analog source transmission is to use separate source and channel coders. This separation is optimal given unlimited delay and complexity in the coders [1]. In practice, joint source-channel coding (JSCC) can lead to a better performance when delay and complexity are constrained. It is well known that for the case of Gaussian source over a Gaussian channel with matched bandwidth, uncoded, or linear, transmission is optimal [2]. However, when there is a bandwidth mismatch between source and channel and in the presence of fading [3], this result does not hold anymore. A common approach for JSCC design is to jointly optimize the components of a tandem system with respect to the channel and source characteristics. Another approach based on nonlinear analog mapping is treated in [4]–[7].

With the increasing popularity of wireless sensor networks (WSNs), reliable transmission with delay and complexity constraints is more relevant than ever. A sensor node, often conceived as having limited lifetime and processing power, communicates its sensed field information to a fusion center over a noisy wireless channel. To meet these challenges, in this paper, we investigate using a low delay and low complexity lossy source-channel mappings in WSNs.

We study the reliable transmission under bandwidth reduction of a memoryless Gaussian source over memoryless Rayleigh fading channels subject to low-delay/complexity constraints. More specifically, our scheme is based on joint optimization of the encoder and the decoder under an average power constraint. A 2:1 parametric mapping using the

Archimedes' spiral was recently considered in [8] for fading channel. In this work, we use non-parametric and parametric mappings under different bandwidth reduction ratios. The case of bandwidth reduction/expansion over additive white Gaussian noise (AWGN) channels was studied in [9]–[11]. In [9], [10], the approach used is based on mapping the output of a vector quantizer to a specific point in a channel signal set. A direct source-channel mapping approach, however, was considered in [11]. Source-channel mappings for the relay and the MAC channels were studied in [12], [13]. Our system, that uses a nonlinear direct source-channel mappings over fading channel, is shown to overcome the performance saturation which is unavoidable when using linear systems and achieves a graceful performance. The rest of the paper is organized

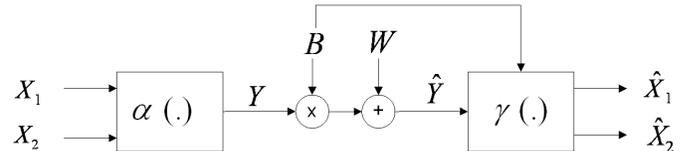


Fig. 1. System structure for 2:1 bandwidth reduction over fading channel with AWGN. CSI is only available at the decoder side, while the encoder knows only the distribution of the fading.

as follows. Section II presents the problem formulation. Section III describes the theoretical Shannon limit and the linear system over fading channels. In Section IV, we consider non-parametric mappings for which we derive necessary conditions for optimality and present the iterative design algorithm along with the implementation aspects. Section V presents parametric mappings over fading channels for different reduction ratios. Simulation results are included in Section VI. Finally, conclusions are drawn in Section VII.

## II. PROBLEM FORMULATION

We consider the transmission of a Gaussian source  $\mathbf{X} = (X_1, \dots, X_M) \in \mathbb{R}^M$  over a memoryless fading channel with AWGN. Fig. 1 shows the system structure for  $M = 2$  where the channel state information (CSI) is available at the decoder only (DCSI), and the transmitter is assumed to know the fading distribution. We also consider two other cases: 1) when CSI is available at both the encoder and the decoder (full CSI or FCSI); 2) when CSI is not available at the transmitter and the receiver. The source vector  $\mathbf{X}$ , which is composed of independent and identically distributed (i.i.d.) samples, is transformed into a one-dimensional channel input  $Y \in \mathbb{R}$

using a nonlinear mapping function  $\alpha(\cdot)$  (i.e.,  $Y = \alpha(\mathbf{X})$ ). The received symbol can be expressed as follows

$$\hat{Y} = BY + W \quad (1)$$

where  $B$  is the fading gain drawn from an i.i.d. Rayleigh distributed process, and  $W$  is the additive noise drawn from an i.i.d. Gaussian process that is independent from both  $B$  and  $\mathbf{X}$ . The mapping operates under an average power constraint  $P_c$  given by

$$\mathbb{E}[\alpha^2(\mathbf{X})] \leq P_c \quad (2)$$

where  $\mathbb{E}[\cdot]$  denotes the expectation operator. At the receiver, the transmitted source vector is estimated using minimum mean square error (MMSE) decoding (assuming DCSI/FCSI)

$$\hat{\mathbf{X}} \triangleq \gamma(\hat{Y}, B) = \mathbb{E}[\mathbf{X}|\hat{y}, b]. \quad (3)$$

We aim to find the optimal source mapping  $\alpha$  and receiver  $\gamma$  that minimize the overall mean square error (MSE) distortion  $\mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}\|^2]$  under the average power constraint in (2).

### III. PRELIMINARIES: THEORETICAL BOUND AND LINEAR SYSTEM

#### A. Theoretical Limit

Shannon's separation theorem states that in a point-to-point communication system, optimal performance can be achieved by optimizing separately (in tandem) the source and the channel coders given unlimited complexity and delay in the coders [1]. As a result, the optimal performance theoretically attainable (OPTA) can be derived by equating the source rate-distortion function times the system's rate to the channel capacity.

For a memoryless Gaussian source  $X$  with variance  $\sigma_X^2$  and MSE distortion measure, the rate-distortion function is [14]

$$R(D) = \max \left\{ 0, \frac{1}{2} \log \left( \frac{\sigma_X^2}{D} \right) \right\} \quad (4)$$

where  $D$  is the distortion. The capacity of a fading channel when considering DCSI/FCSI and with power input  $P$  can be expressed as [15]

$$C(P) = \mathbb{E}_B \left[ \frac{1}{2} \log \left( 1 + \frac{P(b)b^2}{\sigma_W^2} \right) \right] \quad (5)$$

where  $\sigma_W^2$  is the variance of the channel noise, and  $P(b) = P_c$  for DCSI. When CSI is available at both the encoder and the decoder (i.e., FCSI), the transmitted power  $P(b)$  in (5) waterfills over the fading states as follows

$$P(b) = \max \left\{ 0, \left( \frac{1}{\delta} - \frac{\sigma_W^2}{b^2} \right) \right\} \quad (6)$$

where  $\delta$  satisfies  $\mathbb{E}[P(b)] = P_c$ . To find OPTA we set  $R(D) = rC(P)$ , where  $r$  is the system's rate measured in channel uses per source sample. Solving this will lead to OPTA

$$D_{\text{OPTA}} = \sigma_X^2 \exp \left( r \mathbb{E}_B \left[ \log \left( \frac{\sigma_W^2}{P(b)b^2 + \sigma_W^2} \right) \right] \right). \quad (7)$$

Note that this bound is achievable asymptotically using long block codes.

#### B. Linear Transmission

In [3], it was proven that linear coding is optimal among all single-letter codes (i.e., 1:1 matched-bandwidth systems). For an  $M:1$  bandwidth reduction, we use a simple coder that removes samples to perform bandwidth reduction. In this case the encoder  $\alpha(\cdot)$  is a multiplication by a constant matrix and the transmitted symbol can be expressed as follows

$$Y = \sqrt{\frac{P(b)}{\sigma_X^2}} \mathbf{I}_{1 \times M} \mathbf{X} \quad (8)$$

where  $\mathbf{I}_{1 \times M}$  is a  $1 \times M$  matrix with ones on the main diagonal and zeros elsewhere, and

$$P(b) = \begin{cases} P_c, & \text{for DCSI} \\ \max \left\{ 0, \frac{1}{b} \left( \frac{1}{\delta'} - \frac{\sigma_W^2}{b} \right) \right\}, & \text{for FCSI} \end{cases} \quad (9)$$

where  $\delta'$  is calculated by solving  $\mathbb{E}[P(b)] = P_c$ . The optimal decoder (assuming DCSI/FCSI)  $\gamma(\hat{Y}, B)$  is given by

$$\hat{\mathbf{X}} = \mathbb{E}[\mathbf{X}|\hat{y}, b] = \frac{b\sqrt{P(b)\sigma_X^2}}{P(b)b^2 + \sigma_W^2} \mathbf{I}_{M \times 1} \hat{y}. \quad (10)$$

When no CSI is available at the receiver, the MMSE decoder can be expressed as follows

$$\hat{\mathbf{X}} = \mathbb{E}[\mathbf{X}|\hat{y}] = \frac{\int \int \mathbf{x} p(\hat{y}|\alpha(\mathbf{x}), b) p(b) p(\mathbf{x}) db d\mathbf{x}}{\int \int p(\hat{y}|\alpha(\mathbf{x}), b) p(b) p(\mathbf{x}) db d\mathbf{x}} \quad (11)$$

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

### IV. NON-PARAMETRIC MAPPINGS

Using the Lagrange multiplier method [16], the constrained optimization problem of minimizing the overall MSE  $\mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}\|^2]$  distortion subject to (2) can be recast into an unconstrained problem via the Lagrange cost function

$$\min_{\alpha, \gamma} J(\alpha, \gamma) \quad (12)$$

where the cost function  $J(\alpha, \gamma)$  is given by

$$J(\alpha, \gamma) = \mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}\|^2] + \lambda \mathbb{E}[\alpha^2(\mathbf{X})] \quad (13)$$

and the Lagrange multiplier  $\lambda$  is used to control the average power. For a given  $\lambda$ , if the solution of the unconstrained problem in (12) fulfills the average power constraint in (2), the obtained solution is also proven to solve the constrained problem [16]. However, the unconstrained minimization in (12) is still hard to solve due to the interdependencies between the optimized components, and since the encoder/decoder mappings are, in general, nonlinear functions. To overcome these challenges, we proceed in a similar way to classical design problems (e.g., vector quantizer design [17]) by formulating the necessary conditions for optimality. This is done by finding the optimal encoder  $\alpha$  given the decoder  $\gamma$ , and vice versa. In what follows, we assume DCSI to derive the necessary conditions for optimality.

### A. Necessary Conditions for Optimality

The problem of finding the optimal source mapping  $\alpha$  (assuming  $\gamma$  is fixed) is given by

$$\alpha = \arg \min_{\alpha} \left\{ \mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}\|^2] + \lambda \mathbb{E}[\alpha^2(\mathbf{X})] \right\}. \quad (14)$$

Using Bayes' rule, the overall MSE  $\mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}\|^2]$  distortion can be expressed as follows

$$\mathbb{E}[\|\mathbf{X} - \hat{\mathbf{X}}\|^2] = \int \int \int p(\mathbf{x})p(b)p(\hat{y}|\alpha(\mathbf{x}), b)\|\mathbf{x} - \hat{\mathbf{x}}\|^2 d\mathbf{x}dbd\hat{y}. \quad (15)$$

Note that this factorization follows from the fact that channel noise, source, and fading are assumed to be independent of each other. The mapping average power is given by

$$P = \int p(\mathbf{x})\alpha^2(\mathbf{x})d\mathbf{x}. \quad (16)$$

Since  $p(\mathbf{x})$  in (15)–(16) is nonnegative, the optimal source mapping  $\alpha$  can be optimized for each  $\mathbf{x}$  individually according to

$$\alpha(\mathbf{x}) = \arg \min_{y \in \mathbb{R}} \left\{ \int \int p(b)p(\hat{y}|y, b)\|\mathbf{x} - \hat{\mathbf{x}}\|^2 dbd\hat{y} + \lambda y^2 \right\}. \quad (17)$$

Hence, (17) is a necessary condition for  $\alpha$  to be the optimal mapping.

On the receiver side, the optimal decoder in the MSE sense (assuming  $\alpha$  is fixed) is found using the conditional expectation of the source given the received symbol and the fading gain

$$\hat{\mathbf{X}} \triangleq \gamma(\hat{Y}, B) = \mathbb{E}[\mathbf{X}|\hat{y}, b] = \frac{\int \mathbf{x}p(\hat{y}|\mathbf{x}, b)p(\mathbf{x})d\mathbf{x}}{\int p(\hat{y}|\mathbf{x}, b)p(\mathbf{x})d\mathbf{x}}. \quad (18)$$

### B. Design Algorithm

Based on the above necessary conditions for optimality, it is possible to optimize the mapping at the sensor nodes and the receiver using an iterative process. This is done by fixing one part while optimizing the other. One common problem with such iterative technique is that the final solution will depend on the initialization of the algorithm and does not guarantee convergence to the global optimum solution. To get around these challenges, we use noisy channel relaxation [18]. This method suggests to design the system for a noisy channel, and uses the solution obtained as an initialization when designing the system for a less noisy channel. For a given channel signal-to-noise ratio ( $\text{CSNR} \triangleq P_c \mathbb{E}[B^2]/\sigma_w^2$ ), the design algorithm is stated as follows

- 1) Choose some initial mapping for the encoder  $\alpha$ .
- 2) Find the optimal receiver  $\gamma$  according to (18).
- 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 (17).
- 6) Find the optimal receiver  $\gamma$  according to (18).
- 7) Evaluate the cost function  $J^{(i)}$ . If the relative improvement of  $J^{(i)}$  compared to  $J^{(i-1)}$  is less than some positive threshold  $\varepsilon$ , stop iterating. Otherwise go to step 4.

In our simulations, we used  $\varepsilon = 10^{-4}$  and a linear mapping for initializing the encoder  $\alpha$  at low CSNR.

### C. Implementation Aspects

For the actual implementation of (17) and (18), some modifications are required. By the fact that it is impossible to evaluate the formulas for all vector  $\mathbf{X}$  in  $\mathbb{R}^M$ , we form as in [13] a set  $\mathcal{X}$  composed of Monte-Carlo samples drawn from the distribution of  $\mathbf{X}$ . In our simulations, we use  $10^4$  samples to define  $\mathcal{X}$ . Since the channel input and output spaces are real valued, we discretize them using a pulse amplitude modulation (PAM) alphabet  $\mathcal{Y}$

$$\mathcal{Y} = \left\{ -d\frac{L-1}{2}, -d\frac{L-3}{2}, \dots, d\frac{L-3}{2}, d\frac{L-1}{2} \right\} \quad (19)$$

where  $d$  and  $L$  determine the resolution and the cardinality of the discrete set  $\mathcal{Y}$ , respectively. This set becomes close to analog by taking  $d$  to be small in relation to the standard deviation of the noise and by choosing a sufficiently large  $L$ . In our simulations, we use an  $L$  in the range [300 600], and set  $d = 12/(L-1)$ .

Since complexity is one of our main concerns, it is important to note that the decoder side can be approximated with a table-lookup, thereby avoiding having to compute a numerical integration for each received symbol. This is done by first discretizing the fading gain  $B$  using a discrete set  $\mathcal{B}$  of cardinality  $|\mathcal{B}| = N$  (e.g., using a  $N$ -level quantizer) and mapping the discretized receiver input  $\hat{Y}$  and the fading  $B$  onto a decoded symbol  $\hat{\mathbf{X}} = \gamma(\hat{Y}, B)$  for  $(\hat{Y}, B) = (\hat{y}, b)$ . Note that for a given  $N$ , one can design an optimal quantizer for  $B$ . However, in our simulation results we used a uniform quantizer for simplicity and set  $N$  to 256. In a similar way to the decoder side, the encoder can be also implemented via a table-lookup by quantizing the source input. Using this approach, the system complexity is reduced to that of a table lookups.

The discretized versions of (17) and (18), which are used in the implementation of the design algorithm, are, respectively, expressed as follows

$$\alpha(\mathbf{x}) = \arg \min_{y \in \mathcal{Y}} \left\{ \sum_{b \in \mathcal{B}} \sum_{\hat{y} \in \mathcal{Y}} Q(b)Q(\hat{y}|y, b)\|\mathbf{x} - \hat{\mathbf{x}}\|^2 + \lambda y^2 \right\} \quad (20)$$

and

$$\gamma(\hat{y}, b) = \frac{\sum_{\mathbf{x} \in \mathcal{X}} \mathbf{x}Q(\hat{y}|\alpha(\mathbf{x}), b)}{\sum_{\mathbf{x} \in \mathcal{X}} Q(\hat{y}|\alpha(\mathbf{x}), b)} \quad (21)$$

where  $Q(\cdot)$  and  $Q(\cdot|\cdot)$  denote a probability mass function (pmf) and a conditional pmf, respectively. Note that in [11], the necessary condition for the encoder mapping over the AWGN channel is solved using variational calculus.

### D. Non-Parametric Mapping with Full CSI

In this section, we assume that the CSI is available at both encoder and decoder (FCSI). The main change in the necessary conditions for optimality is in finding the optimal encoder mapping  $\alpha$ . The encoder mapping is now optimized for each  $(\mathbf{x}, b)$  according to

$$\alpha(\mathbf{x}, b) = \arg \min_{y \in \mathcal{Y}} \left\{ \sum_{\hat{y} \in \mathcal{Y}} Q(\hat{y}|y, b)\|\mathbf{x} - \hat{\mathbf{x}}\|^2 + \lambda y^2 \right\}. \quad (22)$$

Note that the optimal decoder is given by (21).

### E. Non-Parametric Mapping with No CSI

In this section, we assume that the CSI is not available at both encoder and decoder. The main change in the necessary conditions for optimality from the one with CSI at the decoder, is in finding the optimal decoder mapping. The decoder averages over the fading gain and can be expressed as follows

$$\gamma(\hat{y}) = \frac{\sum_{\mathbf{x} \in \mathcal{X}} \sum_{b \in \mathcal{B}} \mathbf{x} Q(\hat{y} | \alpha(\mathbf{x}), b)}{\sum_{\mathbf{x} \in \mathcal{X}} \sum_{b \in \mathcal{B}} Q(\hat{y} | \alpha(\mathbf{x}), b)}. \quad (23)$$

Note that the optimal encoder is given by (20).

## V. PARAMETRIC MAPPINGS

### A. 2:1 Spiral Mapping

1) *System Structure*: The Archimedes' spiral was shown to perform well for AWGN channels [7]. In this section, we extend the work of [7] to optimize the spiral mapping over fading channels. Recently, we discovered that spiral mapping was also considered over fading channels in [8]. However, our system differs from the one in [8] in parts by the structure, optimization process and in applying power allocation when we have full CSI.

Bandwidth reduction is achieved by first approximating  $\mathbf{X} \in \mathbb{R}^2$  to the closest point on the spiral mapping. We refer to this approximation operation as  $\mathbf{q}$ . The approximated point, represented by its radial distance  $r$  from the origin, is then mapped to the channel via an invertible operator  $\ell(\cdot)$

$$y = K\tilde{y} = K\ell(r) = K \left( \pm 0.16 \left( \frac{\pi^2}{\Delta} \right) (r^2) \right) \quad (24)$$

where  $K$  is a gain factor related to channel power constraint  $P_c = K^2 \mathbb{E}[\tilde{Y}^2]$ ,  $\Delta$  is the radial distance between any two neighboring spiral arms, and  $+$ ,  $-$  represent positive and negative channel values, respectively. At the receiver (assuming DCSI), we scale the received samples using an optimal scaling factor ( $\beta = (b\sqrt{P_c\sigma_Y^2}) / (P_c b^2 + \sigma_W^2)$ ) instead of a suboptimal factor ( $1/(Kb)$ ) as used in [8]. The reconstructed source vector is then estimated using maximum likelihood (ML) decoding. This is done by mapping the scaled received sample to a point on the spiral curve. In the simulation results, we also use the MMSE decoder which was also considered in [8].

2) *System Optimization*: The radial distance  $\Delta$  is the only parameter that needs to be optimized in order to minimize the overall MSE distortion  $\mathbb{E}[|\mathbf{X} - \hat{\mathbf{X}}|^2]$  under an average power constraint. The source signal is affected by two type of distortion: 1) the approximation distortion  $\bar{\epsilon}_a^2$  which is related to the approximation operation  $\mathbf{q}$ , and 2) the channel distortion  $\bar{\epsilon}_{\text{ch}}^2$  which is due to the transmission over a noisy environment. The approximation error  $\bar{\epsilon}_a^2$  can be approximated by the quantization error in a scalar quantizer (i.e.,  $\bar{\epsilon}_a^2 = \Delta^2/12$ ) [7], and the distortion from channel noise  $\bar{\epsilon}_{\text{ch}}^2$  is well approximated by the scaled noise variance  $\beta^2 \sigma_W^2$ . Hence the overall MSE distortion can be approximated as follows

$$\text{MSE} \approx \mathbb{E}_B[(\bar{\epsilon}_a^2 + \bar{\epsilon}_{\text{ch}}^2)|b] \approx \frac{\Delta^2}{12} + \sigma_W^2 \sigma_Y^2 \mathbb{E}_B \left[ \frac{P_c b^2}{(P_c b^2 + \sigma_W^2)^2} \right]. \quad (25)$$

The optimal radial distance  $\Delta_{\text{opt}}$  is found by minimizing the MSE distortion in (25).

3) *Power Allocation*: In this section, we assume that the CSI is also available at the transmitter side (i.e., FCSI). Instead of updating the radial distance  $\Delta$  at each time index [8], we allocate the power (along the time index) according to the CSI knowledge in a similar approach to [3].

Assuming the channel state is  $b$  and the corresponding power allocated is  $P(b)$ , the overall MSE distortion is well approximated in a similar way to (25). The optimal power allocation  $P^*(b)$  that minimizes the MSE distortion can be found by solving the following optimization problem

$$\begin{aligned} \min \quad & \mathbb{E}_B \left[ \frac{P(b)b^2}{(P(b)b^2 + \sigma_W^2)^2} \right] \\ \text{s.t.} \quad & \mathbb{E}_B[P(b)] = P_c, \quad P(b) \geq 0. \end{aligned} \quad (26)$$

As a result, for low noise level, the power allocation is well approximated as follows

$$P^*(b) = \begin{cases} \frac{1}{b} \left( \frac{1}{\sqrt{\mu}} - \frac{2\sigma_W^2}{b} \right), & \text{for } b \geq 2\sigma_W^2 \sqrt{\mu} \\ 0, & \text{otherwise} \end{cases} \quad (27)$$

where  $\mu$  is a threshold for all channel states and can be found from the average power constraint  $P_c$  and the statistics of  $b$ .

### B. M : 1 Parametric Mapping

In this section, we consider both 3:1 and 4:1 bandwidth reductions. In [19], parametric mappings over AWGN channel for these reduction ratios were presented. In what follows, we extend these mappings to accommodate the fading channel. In a similar way to the 2:1 spiral mapping, the 3:1 and the 4:1 bandwidth reductions are done by first approximating  $\mathbf{X}$  to the closest point on the mapping curve, and then performing a one-dimensional representation using (24). At the receiver side, we use ML or MMSE decoding for signal recovery.

The overall distortion can be approximated as follows

$$\begin{aligned} \text{MSE} & \approx \mathbb{E}_B[(\bar{\epsilon}_a^2 + \bar{\epsilon}_{\text{ch}}^2)|b] \\ & \approx \beta_{M:1} \Delta^{2-\theta_{M:1}} + P_c \sigma_W^2 \sigma_Y^2 \mathbb{E}_B \left[ \frac{b^2}{(P_c b^2 + \sigma_W^2)^2} \right] \end{aligned} \quad (28)$$

where  $\theta_{3:1} = 1$ ,  $\beta_{3:1} = 0.6312$ ,  $\theta_{4:1} = 1.3$ , and  $\beta_{4:1} = 1.6244$  for  $\sigma_X = 1$ , are found using a nonlinear curve fitting. Note that  $\Delta$ , which has a different meaning than in the 2:1 spiral mapping, is found by minimizing (28).

## VI. NUMERICAL RESULTS

In this section, we assume a Gaussian source vector  $\mathbf{X}$  with i.i.d. samples with unit variance  $\sigma_X^2 = 1$  and a Rayleigh fading gain  $B$  with  $\mathbb{E}[B^2] = 1$ . The results are plotted as signal-to-distortion ratio (SDR  $\triangleq E[|\mathbf{X}|^2]/E[|\mathbf{X} - \hat{\mathbf{X}}|^2]$ ) versus CSNR. The proposed system is compared to a linear system and to the theoretical performance limit. Fig. 2 shows the result for parametric spiral and non-parametric 2:1 mapping when considering DCSI and FCSI. We can notice that for most CSNR values, the spiral and the non-parametric mappings outperform the linear system and give a strong performance; there is around 2 dB gap from OPTA which is achievable asymptotically using a highly complex long block codes. For

low to moderate CSNRs, using MMSE decoding with spiral mapping gives a substantial gain over ML decoding. However as the CSNR gets large, the gap between ML and MMSE decoder diminishes. Moreover, using the optimal scaling factor with ML decoder gives a few dBs SDR gain over the one with suboptimal scaling at low CSNRs [8]. Fig. 2 shows that the non-parametric mapping gives around 0.5 dB gain over the spiral mapping with MMSE decoder. Moreover, at low CSNRs, the non-parametric mapping does not underperform the linear system, unlike the parametric spiral mapping.

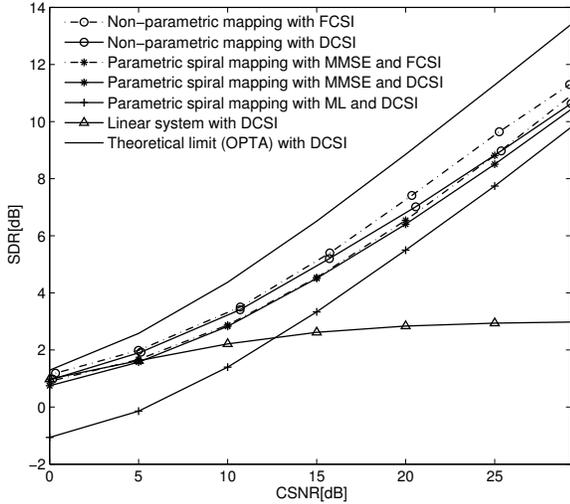


Fig. 2. The performance of the parametric and the non-parametric mappings with 2:1 bandwidth reduction over fading channel with AWGN. The performance of the linear system and the theoretical limit (OPTA) with CSI at the decoder are also included. Note that for FCSI, the performance of the linear system and the theoretical limit improves over the DCSI case by at most 0.2 dB in SDR.

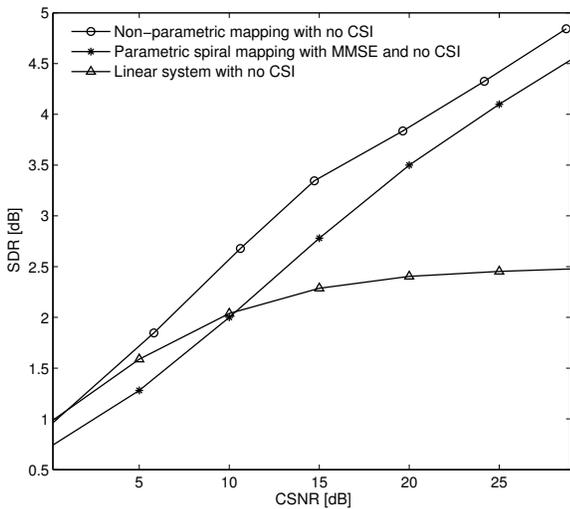


Fig. 3. The performance of the parametric and the non-parametric mappings with 2:1 bandwidth reduction over fading channel with AWGN. The performance of the linear system is also included. This figure is for no CSI at both encoder and decoder.

Assuming FCSI, we can notice around 0.5 dB gain in SDR over the DCSI case. This gain is numerically observed to reach 1 dB for CSNR = 40 dB.

Fig. 3 shows the performance of the spiral and the non-parametric 2:1 mappings when no CSI is available at both

transmitter and receiver. It is clear that the non-parametric mapping outperforms the spiral mapping and overcomes the saturation effect which is unavoidable with the linear system.

Figs. 4 and 5 show the performance of the non-parametric and the parametric mappings for 3:1 and 4:1 bandwidth reduction, respectively. In these figures we consider only the DCSI case. Similar to the 2:1 system, we can notice that the non-parametric mappings outperform the other systems and give a graceful performance. More precisely, the non-parametric mappings are shown to give around 1 dB gain in SDR over the parametric mappings. This gain is due to the fact that the non-parametric mappings have a higher degree of freedom in placing points in space without being restrained to a specific structure.

Motivated by the broadcast scenario, we next optimize the encoder for a fixed-design CSNR level and assume that the true CSNR is known by the decoder. It is noticed that both non-parametric and parametric mappings exhibit various degree of robustness against mismatch in noise level; the non-parametric mappings still outperform the parametric mappings for most design and true CSNR levels. However, for a low design CSNR level, we have noticed that the parametric mappings give better performance when the true CSNR is very high. This can be explained from the observation that the non-parametric mappings have a different structure at low CSNR than at high CSNR and look more like a linear mapping.

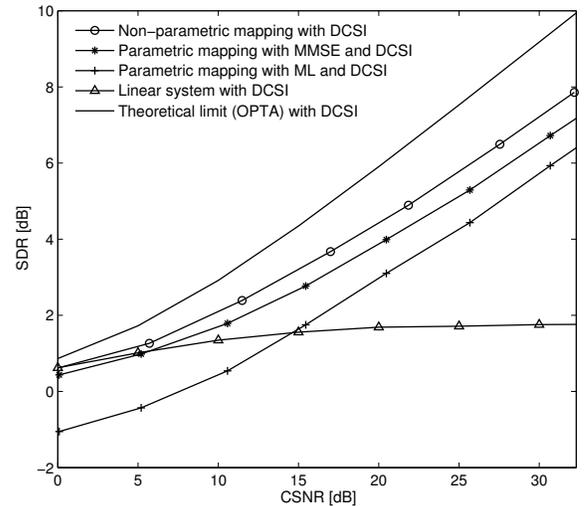


Fig. 4. The performance of the parametric and the non-parametric mappings with 3:1 bandwidth reduction over fading channel with AWGN. The performance of the theoretical limit and the linear system are also plotted.

Fig. 6 shows the optimized decoder mappings for a 2:1 bandwidth reduction ratio. We can notice that the structure of the non-parametric mapping is similar to the Archimedes' spiral. Moreover, the length of the mapping curves increase as the fading gain increase (i.e., less noisy channel). This makes the mapping better fill the power-delimited space, in order to lower the approximation error due to projecting a higher dimension source into a lower dimension, while not increasing much the distortion due to channel noise. Note also that the radial distance between the mapping arms is not uniform as in the case when using a spiral mapping with ML decoder.

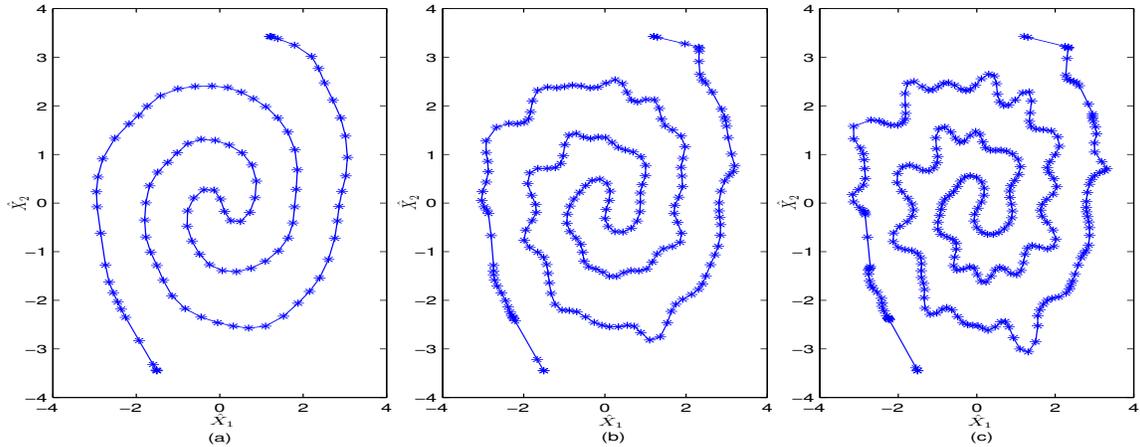


Fig. 6. Decoded pairs for 2:1 bandwidth reduction, and fading gain  $b$ : (a) 0.3, (b) 0.6, and (c) 0.9. This graph is made for CSNR=30 dB and DCSI. The lines that are drawn between asterisk points correspond to neighbours in the one-dimensional mapped source (i.e., channel space).

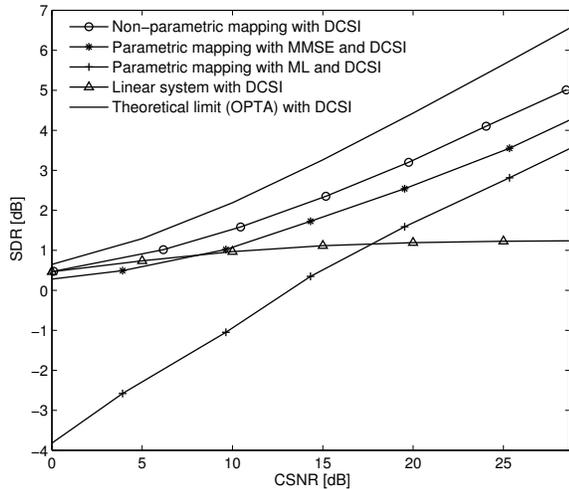


Fig. 5. The performance of the parametric and the non-parametric mappings with 4:1 bandwidth reduction over fading channel with AWGN. The performance of the theoretical limit and the linear system are also plotted.

## VII. CONCLUSIONS

In this paper, we have presented a low-delay/complexity lossy joint source-channel coding scheme for the Gaussian source over Rayleigh fading channel with AWGN. A design algorithm for optimizing the source-channel mapping has been presented based on the necessary conditions for optimality. Mappings based on parametric functions are also studied. Simulation results for 2:1, 3:1, and 4:1 bandwidth reductions have shown that both non-parametric and parametric mappings outperform the linear system and give graceful performance variation with CSNR. Moreover, these nonlinear mappings have shown to overcome the performance saturation of linear systems and achieve a robust performance against noise mismatch. The implementation aspects of the design algorithm, and some non-parametric mappings have been presented. When considering FCSI, optimal power allocation has been analytically solved for parametric mappings in terms of channel fading gains and average power constraint. In general, it is possible to design mappings for other bandwidth

expansion/reduction rates using a similar approach.

## REFERENCES

- [1] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.
- [2] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code or not to code," in *Proc. IEEE Int. Symp. on Info. Theory*, Sorrento, Italy, June 2000.
- [3] J.-J. Xiao, Z.-Q. Luo, and N. Jindal, "Linear joint source-channel coding for Gaussian sources through fading channels," in *Proc. IEEE Global Telecommunications Conference*, San Francisco, CA, Dec 2006.
- [4] C. E. Shannon, "Communication in the presence of noise," in *Proc. IRE*, 1949, pp. 10–21.
- [5] V. A. Kotel'nikov, *The Theory of Optimum Noise Immunity*. New York: McGraw-Hill, 1959.
- [6] T. A. Ramstad, "Shannon mappings for robust communication," *Teletronikk*, vol. 98, no. 1, pp. 114–128, 2002.
- [7] F. Hekland, P. A. Floor, and T. A. Ramstad, "Shannon-Kotel'nikov mappings in joint source-channel coding," *IEEE Trans. Communications*, vol. 57, no. 1, pp. 94–105, Jan 2009.
- [8] G. Brante, R. Souza, and J. Garcia-Frias, "Analog joint source-channel coding in Rayleigh fading channels," in *Proc. IEEE Int. Conf. Acous., Speech and Sig. Process. (ICASSP)*, Prague, Czech Republic, May 2011.
- [9] A. Fuldseth and T. A. Ramstad, "Bandwidth compression for continuous amplitude channels based on vector approximation to a continuous subset of the source signal space," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP)*, Munich, Germany, Apr 1997.
- [10] P. A. Floor, T. A. Ramstad, and N. Wernersson, "Power constrained channel optimized vector quantizers used for bandwidth expansion," in *Proc. IEEE ISWCS*, Trondheim, Norway, Oct 2007.
- [11] E. Akyol, K. Rose, and T. Ramstad, "Optimal mappings for joint source channel coding," in *IEEE Inform. Theory Workshop*, Cairo, Egypt, 2010.
- [12] J. (Karlsson) Kron and M. Skoglund, "Optimized low-delay source-channel-relay mappings," *IEEE Trans. Communications*, vol. 58, no. 5, pp. 1397–1404, May 2010.
- [13] J. (Karlsson) Kron, *Low-Delay Sensing and Transmission*. Ph.D. dissertation, Royal Institute of Technology, 2011.
- [14] T. M. Cover and J. A. Thomas, *Elements Information Theory*. New York: Wiley, 2006.
- [15] D. N. C. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge, U.K.: Cambridge University Press, 2005.
- [16] H. Everett III, "Generalized Lagrange multiplier method for solving problems of optimum allocation of resources," *Operations Research*, vol. 11, no. 3, pp. 399–417, 1963.
- [17] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Comm.*, vol. 28, no. 1, pp. 84–95, Jan 1980.
- [18] S. Gadkari and K. Rose, "Noisy channel relaxation for VQ design," in *Proc. IEEE International Conference Acoustics, Speech and Signal Processing (ICASSP)*, Atlanta, GA, May 1996.
- [19] P. A. Floor, *On the Theory of Shannon-Kotel'nikov Mappings in Joint Source-Channel Coding*. Ph.D. dissertation, Norwegian University of Science and Technology, 2008.