LINEAR PRECODING AND DECODING FOR DISTRIBUTED DATA COMPRESSION

Azadeh Vosoughi and Anna Scaglione

School of Electrical and Computer Engineering
Cornell University, Ithaca, NY 14853
av78@cornell.edu; anna@ece.cornell.edu

ABSTRACT

Considering two correlated vector sources $\boldsymbol{x},\ \boldsymbol{y} \in \mathbb{R}^N$, we address the problem of lossy coding of $\boldsymbol{x}$ with uncoded side information $\boldsymbol{y}$ available at the decoder. The general non-linear mapping between $\boldsymbol{y}$ and $\boldsymbol{x}$ capturing their correlation can be approximated through a linear model $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}$ in which $\boldsymbol{n}$ is independent of $\boldsymbol{x}$. Viewing this model as a virtual communication channel with input $\boldsymbol{x}$ and output $\boldsymbol{y}$ we utilize *linear precoding and decoding* technique to convert the original vector source coding problem into a set of manageable scalar source coding problems. The scalar source coding problems can be solved using the existing distributed source coding algorithms that are primarily designed for the simple correlation model $y = x + n$ where x and y are scalar jointly Gaussian sources.

1. INTRODUCTION

Distributed source coding (DSC) refers to the compression of multiple correlated sensor outputs that do not communicate with each other [1]. Exploiting the inherent correlation structure which exists between the data sensed by the sensors would enable the sensors to compress their outputs efficiently without knowing explicitly other sensors' outputs. In fact if the joint distribution characterizing the correlation structure is known at both the encoders and the joint decoder, the sensors with separate encoders can achieve a compression rate as low as the sensors with joint encoders under certain circumstances. The theoretical foundation of this result has been established in the information theory literature under the name of Wyner and Ziv (WZ) source coding theorem [2].

Consider two arbitrary correlated scalar sources x and y. The WZ lossy source coding problem refers to encoding x with respect to a fidelity criterion, e.g., distortion measured by mean squared error (MSE), assuming that the decoder has access to y. In their landmark paper [2] Wyner and Ziv derived the rate-MSE distortion function for encoding source x when the related source y is known at the decoder. They proved that unless x and y are jointly Gaussian the WZ coding suffers rate loss comparing with lossy coding of x when y is available at both the encoder and the decoder. Nested lattice codes were first introduced in [3] as codes that can achieve the WZ limit asymptotically. Based on the partitioning ideas of [3] the authors in [1] considered trellis-based nested codes as a way of realizing nested lattice codes. We note that the work in [1] was restricted to Gaussian scalar sources x and y where x and y are related through $y = x + n$ and the noise term n is Gaussian and is independent of x.

In this paper we propose DSC algorithms which are applicable to a wider range of correlation model. In particular, we consider the scenario where the relationship between the vector sources $\boldsymbol{x}$ and $\boldsymbol{y}$

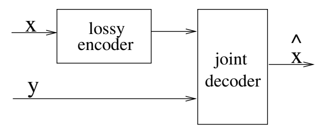

Fig. 1. WZ lossy source coding with side information at the decoder.

can be modelled as:

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n} \tag{1}$$

where $\boldsymbol{n}$ is independent of $\boldsymbol{x}$. This model can be viewed as a first order approximation of a general non-linear mapping between $\boldsymbol{y}$ and $\boldsymbol{x}$. Viewing this dependency as a fictitious communication channel with input $\boldsymbol{x}$ and output $\boldsymbol{y}$ and following the ideas in [3], one might think of searching for lattice channel codes suitable for this type of channels [4]. However, the complexity of these constellations are prohibitively high. As a simple albeit suboptimal architectural alternative communication systems utilize *linear precoding* to cope with the distortion imposed on $\boldsymbol{x}$ by $\boldsymbol{H}$ [5].The linear precoding design paradigm is based on an optimal pair of linear transformation $\boldsymbol{F}$ (precoder) and $\boldsymbol{G}$ (decoder) of blocks of transmit symbols $\boldsymbol{x}$ and receive samples $\boldsymbol{y}$ that operate jointly. Given channel state information (CSI) at both the transmitter and the receiver, the design of $(\boldsymbol{F}, \boldsymbol{G})$ appropriately takes advantage of this knowledge to use the transmit resources efficiently while maintaining a reasonable complexity [5].

Exploiting the paradigm of linear precoding/decoding, that has never been explicitly exploited in the DSC arena, we propose to convert the original vector source coding problem into a set of parallel scalar source coding problems with the scalar side information at the decoder, where there is a one-to-one correspondence between each scalar source and scalar side information. With this conversion, we are able to use the existing constructive WZ codes (e.g., the code in [1]) to implement the general problem of vector sources. We consider a codec where the source encoder and the joint decoder have the knowledge of the correlation model parameters. In practice the encoder and the decoder learn the correlation model parameters in a training phase. These parameters are incorporated into designing the encoder and the decoder. Furthermore, utilizing the correlation model the encoder calculates the encoding rates which minimize the MSE-distortion occurred due to the reconstruction of $\boldsymbol{x}$ at the decoder. The encoder encodes using these calculated rates.

The paper is organized as follows: Section 2 describes the problem of lossy coding of $\boldsymbol{x}$ with side information $\boldsymbol{y}$ where $\boldsymbol{x}$ and $\boldsymbol{y}$ are related through (1). Section 3 overviews DISCUS, a specific DSC construction [1] which we use as the baseline of our work. Section 4 clarifies the role of linear preocder and decoder in our proposed algorithm. Section 5 elaborates the proposed coding algorithm and the codec architecture. Section 6 explains the rate allocation procedure. Numerical results and conclusion follow in Section 7.

This work is supported in part by the National Science Foundation under grant CCR-0133635.

Notation: Upper case and lower case boldface letters denote matrices and vectors, respectively. The i-th entry of vector $\boldsymbol{a}$ and the (i,j)-th entry of matrix $\boldsymbol{A}$ are presented as a_i and $[\boldsymbol{A}]_{ij}$, respectively.

2. PROBLEM STATEMENT

Consider the class of vector sources with outputs $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$ whose statistical dependency can be approximately captured through the model in (1). Regarding (1) we assume **(a1)** $\boldsymbol{n}$ and $\boldsymbol{x}$ are independent and $\boldsymbol{H}$ is a constant matrix, **(a2)** $\boldsymbol{x}$ is zero mean with a covariance matrix $\boldsymbol{R}_{xx}$ and $\boldsymbol{n} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_{nn})$. Our setup can be described as follows: the decoder has access to $\boldsymbol{y}$ (side information). The encoder compresses $\boldsymbol{x}$, considering the fact that the decoder knows $\boldsymbol{y}$. The decoder reconstructs $\boldsymbol{x}$ using the side information $\boldsymbol{y}$ as well as the compressed data transmitted by the encoder. To be able to reconstruct $\boldsymbol{x}$ with a reasonable MSE-distortion, the decoder needs to fully exploit the correlation between $\boldsymbol{x}$ and $\boldsymbol{y}$. This requires the knowledge of the correlation model parameters at the decoder, i.e., $\boldsymbol{H}$ and $\boldsymbol{R}_{nn}$[1]. Relying on the knowledge of $\boldsymbol{H}$ and $\boldsymbol{R}_{nn}$ the encoder calculates the rate at which it should compress $\boldsymbol{x}$ such that the decoder is able to reconstruct $\boldsymbol{x}$ with a certain MSE-distortion. The compression is performed in transform domain, i.e., the encoder encodes the vector $\boldsymbol{s} = \boldsymbol{F}\boldsymbol{x}$ where $\boldsymbol{F}$ is a linear transform. Our goal is to design compression algorithms of polynomial complexity that are suitable for the correlation model in (1) and provide a good approximation of $\boldsymbol{x}$.

3. OVERVIEW OF DISCUS

In an attempt to achieve the bound predicted in [2] Pradhan and Ramchandran proposed a practical DSC scheme called DISCUS (distributed source coding using syndromes) [1]. Inspired by the random binning idea used in [2] to establish the theoretical results, the authors propose a code construction for the simple correlation model $y = x + n$ where x and y are jointly Gaussian. In their scheme the encoder consists of (i) a *source space partition*, in which the encoder partitions the real line into disjoint quantization intervals and finds the index of the interval to which the quantized x belongs to[2], and (ii) a *coset code partition*, in which the encoder partitions the reconstruction points of the quantizer into bins (cosets). Encoding with side information consists of finding the index of the coset containing the quantized x and sending the index at a rate (which we refer to it as *coset rate* and restrict it to be less than or equal to the quantization rate) to the decoder over an error free channel [1]. The decoder in [1] consists of (i) *source code recovery*, in which the decoder decides the quantized x as the one which is closest to y (in the sense of Euclidean distance) in the coset whose index is sent by the encoder, and (ii) an *estimator*, in which the decoder forms the optimal estimate $\hat{x} = E\{x|y, x \in \Gamma_i\}$ where Γ_i is the decided quantization interval. The following example describes the code construction at the encoder and the reconstruction at the decoder (See figure2):

Example [1]: Consider a uniform quantizer with $Q = 8$ quantization levels. Let $\mathcal{C} = \{c_1, \ldots, c_8\}$ and $\mathcal{G} = \{\Gamma_1, \ldots, \Gamma_8\}$ denote the set of reconstruction points and the quantization intervals, respectively. The quantization rate is $\log_2 8 = 3$ bits/sample. Targeting the transmission rate of 1 bit/sample we partition $\mathcal{C}$ into $2^1 = 2$ cosets $\mathcal{R}_1 = \{c_1, c_3, c_5, c_7\}$ and $\mathcal{R}_2 = \{c_2, c_4, c_6, c_8\}$. Assuming that $x \in \Gamma_5$, the encoder transmits the coset index 1 at the rate of 1 bit/sample. Knowing that the quantized x belongs to coset 1 and observing y the decoder declares Γ_5 as the interval (i.e., error free detection) and forms $\hat{x} = E\{x|y, x \in \Gamma_5\}$.

[1]Estimating $\boldsymbol{H}$ and $\boldsymbol{R}_{nn}$ at the encoder and decoder is itself an interesting problem and deserves a special treatment which is beyond the scope of this paper.

[2]For simplicity we focus on the case of scalar quantization and memoryless coset construction in [1].

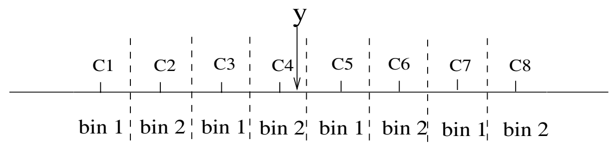

Fig. 2. Quantization intervals for a uniform quantizer with $Q = 8$ levels .

While applying the minimum distance (MD) rule to decide the quantization intervals is justified for the correlation model $y = x+n$ it is not justified for the case where the correlation model is more elaborate, e.g., the correlation model is (1).

4. LINEAR PRECODING AND DECODING IN SOURCE CODING

Using a pair of linear precoder and decoder $(\boldsymbol{F}, \boldsymbol{G})$ we convert the dependency model of (1) into a simple model:

$$\boldsymbol{z} = \boldsymbol{\Lambda}\boldsymbol{s} + \boldsymbol{w} \tag{2}$$

where $\boldsymbol{s} = \boldsymbol{F}\boldsymbol{x}$, $\boldsymbol{\Lambda}$ is a diagonal matrix, and $\boldsymbol{z} = \boldsymbol{G}\boldsymbol{y}$ and $\boldsymbol{w} = \boldsymbol{G}\boldsymbol{n}$ are the new side information and the new noise term, respectively. By establishing a one-to-one correspondence between the entries of $\boldsymbol{s}$ and $\boldsymbol{z}$ we propose to use a scheme where the decoder performs an entry by entry based MD detection as in [1]. Unless $\boldsymbol{n}$ is colored, $\boldsymbol{w}$ is white and hence the MD detection followed by entry by entry based symbol estimation is optimal. For the case where $\boldsymbol{n}$ is colored we apply a vector estimation at the decoder to improve the estimate by taking in to account the noise color. Assuming that $\boldsymbol{R}_{xx}$ is known at the encoder, we let $\boldsymbol{u} = \boldsymbol{U}_x^T\boldsymbol{x}$ where $\boldsymbol{U}_x$ is obtained from the EVD of $\boldsymbol{R}_{xx}$, i.e., $\boldsymbol{R}_{xx} = \boldsymbol{U}_x\boldsymbol{\Sigma}\boldsymbol{U}_x^T$ and the transformed coefficients $\boldsymbol{u}$ are uncorrelated. We rewrite (1) as:

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{U}_x\boldsymbol{u} + \boldsymbol{n} \tag{3}$$

We introduce the SVD of $\boldsymbol{H}\boldsymbol{U}_x = \boldsymbol{U}_{hu}\boldsymbol{\Lambda}\boldsymbol{V}_{hu}^T$. Defining $\boldsymbol{s} = \boldsymbol{V}_{hu}^T\boldsymbol{u}$ and post multiplying (3) with $\boldsymbol{U}_{hu}^T$ we obtain $\boldsymbol{z} = \boldsymbol{U}_{hu}^T\boldsymbol{y} = \boldsymbol{\Lambda}\boldsymbol{s} + \boldsymbol{w}$ where $\boldsymbol{w} = \boldsymbol{U}_{hu}^T\boldsymbol{n}$ and has the covariance matrix $\boldsymbol{R}_{ww} = \boldsymbol{U}_{hu}^T\boldsymbol{R}_{nn}\boldsymbol{U}_{hu}$. In a nutshell using the pair of linear precoder and decoder:

$$\boldsymbol{F} = \boldsymbol{V}_{hu}^T\boldsymbol{U}_x^T \qquad \boldsymbol{G} = \boldsymbol{U}_{hu}^T$$

we convert the dependency model of (1) into the simple model in (2). In fact the output of the decoder $\boldsymbol{z}$ provides us with an *initial estimate* of $\boldsymbol{s}$. Incorporating the data sent by the encoder (i.e., the coset indices) the decoder *refines* this initial estimate and obtains an improved estimate that results in less MSE-distortion.

5. ALGORITHM DESCRIPTION

The encoder of our proposed system consists of (i) a linear precoder $\boldsymbol{F}$ to transform $\boldsymbol{x}$ to $\boldsymbol{s}$, (ii) N fixed rate scalar uniform quantizers, and (iii) N coset partitioners that operate in parallel. The encoder quantizes the entries of $\boldsymbol{s}$ and finds the indices of the cosets to which the quantized entries belong to. Let r_i bits/sample denote the rate at which the coset index corresponding to encoding s_i is sent to the decoder, i.e., r_i is the *coset rate*. The encoder sends these indices at the total rate of $R = \sum_{i=1}^N r_i$ bits/vector. The decoder has three parts (i) a linear decoder $\boldsymbol{G}$ to form $\boldsymbol{z}$ using $\boldsymbol{y}$, (ii) N MD rule detectors that operate in parallel, and (iii) a vector estimator.

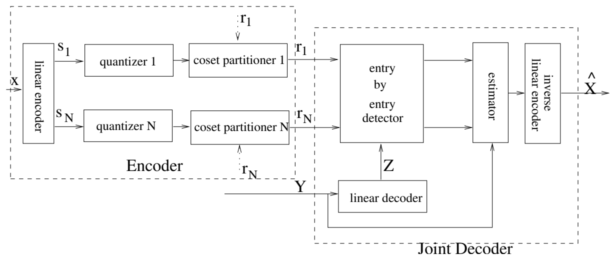

Fig. 3. The architecture of the proposed codec

5.1. Encoder Side: Quantization and Coset Partitioning

The entry s_i is quantized with the i-th quantizer which has Q_i levels and a step size Δ_i. We define the *quantization rate* as $r_{s_i} = \log_2 Q_i$. The i-th quantizer partitions the real line into Q_i disjoint intervals. Each interval corresponds to an input signal amplitude in that interval and is associated with a reconstruction point c_q^i $q = 1, \ldots, Q_i$. We call the set $\{c_q^i\}_{q=1}^{Q_i}$ the *source codebook* $\mathcal{C}_i$ associated with the i-th quantizer. We construct the cosets as follows: let r_i bits/sample be the coset rate and define the total number of cosets as $P_i = 2^{r_i}$. We partition $\mathcal{C}_i$ into P_i cosets, namely $\mathcal{R}_p^i$ $p = 1, \ldots, P_i$, such that each coset $\mathcal{R}_p^i$ includes $M_i = Q_i/P_i$ reconstruction points. Clearly we have $r_i = r_{s_i} - r_{c_i}$. Let $\mathcal{R}_p^i = \{r_{p,1}^i, \ldots, r_{p,M_i}^i\}$ for $p = 1, \ldots, P_i$. The coset partitioning is such that the elements of $\mathcal{R}_p^i$ are related to the elements of $\mathcal{C}_i$ through the relationship $r_{p,m}^i = c_{P_i(m-1)+p}^i$ $m = 1, \ldots, M_i$ $p = 1, \ldots, P_i$. In words, considering $\mathcal{C}_i$ we let the first P_i elements of $\mathcal{C}_i$ to be the first elements of the P_i cosets, the second P_i elements of $\mathcal{C}_i$ to be the second elements of the P_i cosets and so on, while we preserve the order during assigning (See Fig. 2). Given the index of the quantization interval q the index of the coset p can be found as $p = q \; mod \; P_i$. Let $N = 1$ and consider Example 1. Given $s_1 \in \Gamma_5^1$ the quantizer quantizes s_1 to c_5^1 and finds the index of the coset as $1 = 5 \; mod \; 2$.

5.2. Decoder Side: MD rule Detecting and Vector Estimation

In (2) we establish a one-to-one correspondence between the entries of $\boldsymbol{z}$, $\boldsymbol{s}$, and $\boldsymbol{w}$. Equivalently, we have $z_i = [\boldsymbol{\Lambda}]_{ii} s_i + w_i$ $i = 1, \ldots, N$ where w_i has the variance $\sigma_{w_i}^2 = [\boldsymbol{R}_{ww}]_{ii}$. Given the coset index p for each entry s_i, the decoder applies the MD rule to decide the quantized value of s_i in p-th coset that is closet in distance to z_i, i.e., $r_{p,m^*}^i = \arg\min_{r_{p,m}^i \in \mathcal{R}_p^i} ||z_i - [\boldsymbol{\Lambda}]_{ii} r_{p,m}^i||$ and declares $P_i(m^* - 1) + p$ as the index of the interval to which s_i belongs.

The entry by entry MD detector provides the indices of the quantization intervals. Knowing these intervals we form the vector estimate $\hat{\boldsymbol{s}}$ and consequently $\hat{\boldsymbol{x}} = \boldsymbol{F}^{-1}\hat{\boldsymbol{s}}$. To find a closed form expression for $\hat{\boldsymbol{x}}$ we assume that $\boldsymbol{s}$ and $\boldsymbol{y}$ are *jointly Gaussian*. For jointly Gaussian $\boldsymbol{s}$ and $\boldsymbol{y}$ the conditional statistics are [6]:

$$\begin{aligned} \boldsymbol{\mu} &= (\boldsymbol{R}_{ss}^{-1} + \boldsymbol{\mathcal{H}}^T \boldsymbol{R}_{nn}^{-1} \boldsymbol{\mathcal{H}})^{-1} \boldsymbol{\mathcal{H}}^T \boldsymbol{R}_{nn}^{-1} \boldsymbol{y} \\ \boldsymbol{R}_{s|y} &= (\boldsymbol{R}_{ss}^{-1} + \boldsymbol{\mathcal{H}}^T \boldsymbol{R}_{nn}^{-1} \boldsymbol{\mathcal{H}})^{-1} \end{aligned}$$

where $\boldsymbol{\mathcal{H}} = \boldsymbol{H}\boldsymbol{F}^{-1}$ and $\boldsymbol{R}_{ss} = \boldsymbol{V}_{hu}^T \boldsymbol{\Sigma} \boldsymbol{V}_{hu}$. Lemma 1 provides the expression for $\hat{\boldsymbol{s}}$. The proofs are omitted due to lack of space [6]:

lemma 1: *Let $\mathcal{V}_s$ denote the N dimensional volume whose edges are determined by the decided quantization interval. Define $\boldsymbol{v} = \boldsymbol{U}^T(\boldsymbol{s} - \boldsymbol{\mu})$ and let $\mathcal{V}_v$ denote the new volume we wish to integrate over which is obtained from transforming $\mathcal{V}_s$ in the new coordinate system. The optimal vector estimate $\hat{\boldsymbol{s}} = E\{\boldsymbol{s} | \boldsymbol{s} \in \mathcal{V}_s, \boldsymbol{y}\}$ is:*

$$\hat{\boldsymbol{s}} = \boldsymbol{\mu} + \sum_{j=1}^{N} \boldsymbol{U}\boldsymbol{e}_j \frac{\int_{\boldsymbol{v} \in \mathcal{V}_v} v_j e^{-\frac{1}{2}\sum_{i=1}^{N} v_i^2/[\boldsymbol{\Lambda}_{s|y}]_{ii}} d\boldsymbol{v}}{\int_{\boldsymbol{v} \in \mathcal{V}_v} e^{-\frac{1}{2}\sum_{i=1}^{N} v_i^2/[\boldsymbol{\Lambda}_{s|y}]_{ii}} d\boldsymbol{v}} \quad (4)$$

in which $\boldsymbol{e}_j$ is an all-zero vector whose j-th entry is one, $\boldsymbol{U}$ and $\boldsymbol{\Lambda}_{s|y}$ are obtained from the EVD of $\boldsymbol{R}_{s|y} = \boldsymbol{U}\boldsymbol{\Lambda}_{s|y}\boldsymbol{U}^T$. The integral can be numerically evaluated using Gaussian quadrature technique [6].

Remark: The estimator $\hat{\boldsymbol{s}}$ is decomposed into two terms: $\boldsymbol{\mu}$ is the optimal estimate of $\boldsymbol{s}$ when the encoder does not send any data to the decoder and the decoder forms an estimate using the side information $\boldsymbol{y}$. The second term is the refinement in the estimation resulting from the extra bits (coset indices) that the encoder sends.

6. RATE ALLOCATION POLICY

We let the quantization rates r_{s_i} $i = 1, \ldots, N$ to be:

$$r_{s_i} = \max\{0, \lceil \frac{1}{2} \log_2 \frac{\sigma_{s_i}^2}{\theta} \rceil\} \quad i = 1, \ldots, N \quad (5)$$

where $\sigma_{s_i}^2 = [\boldsymbol{R}_{ss}]_{ii}$ and θ is a parameter which for simplicity we choose to be identical for $\forall i$. Consider reconstruction of $\{s_i\}_{i=1}^N$ at the decoder given a set of coset rates $\{r_i\}_{i=1}^N$ and a value of θ. Let p_{e_i} represent the average probability of error occurred during the detection of the quantization interval to which s_i belongs. Clearly p_{e_i} depends on $\{r_{s_i}\}_{i=1}^N$ (which is determined by the choice of θ), the target total coset rate $R = \sum_{i=1}^N r_i$ (which for compression purpose we wish $R < \sum_{i=1}^N r_{s_i}$), and $\{r_i\}_{i=1}^N$. The MSE-distortion occurred during the reconstruction of $\{s_i\}_{i=1}^N$ can be bounded if the MD rule detectors at the decoder make least error in deciding the quantization intervals. Given θ and the target rate R bits/vector we define the optimal coset rates $\{r_i^*\}_{i=1}^N$ as:

$$\{r_i^*\}_{i=1}^N = \arg\min \sum_{i=1}^{N} p_{e_i} \quad (6)$$

subject to N inequality constraints $0 \leq r_i \leq r_{s_i}$ and an equality constraint $\sum_{i=1}^N r_i = R$. In the following we provide an upper bound on p_{e_i}: recall $z_i = [\boldsymbol{\Lambda}]_{ii} s_i + w_i$ $i = 1, \ldots, N$. We rewrite s_i as $s_i = c_q^i + \varepsilon_i$ where c_q^i and ε_i denote the quantized value of s_i and the quantization error, respectively. Combining these two relations we obtain $z_i = [\boldsymbol{\Lambda}]_{ii} c_q^i + \nu_i$ where $\nu_i = [\boldsymbol{\Lambda}]_{ii}\varepsilon_i + w_i$ is the equivalent noise term. Assuming that ε_i and w_i are uncorrelated we have $\sigma_{\nu_i}^2 = ([\boldsymbol{\Lambda}]_{ii})^2 E\{\varepsilon_i^2\} + \sigma_{w_i}^2$ where $E\{\varepsilon_i^2\}$ is the MSE-distortion induced by the i-th quantizer.

Lemma 2: *Suppose $0 \leq r_i < r_{s_i}$. Assuming that $\nu_i \sim \mathcal{N}(0, \sigma_{\nu_i}^2)$ and is independent of c_q^i, an upper bound on p_{e_i} is:*

$$p_{e_i} < 0.5 exp\{-\frac{\Delta_i^2 2^{2r_i}}{8\sigma_{\nu_i}^2}\} \quad (7)$$

for $r_i = r_{s_i}$ by the definition of p_{e_i} we let $p_{e_i} = 0$.

Using the upper bound in (7) as an expression for p_{e_i} we present an *iterative rate allocation algorithm* which solves (6) subject to the constraints. A similar algorithm has been proposed in [8] for assigning bits to the subcarriers of a single user in a multiuser OFDM system in which the authors minimize the total transmit power under the single user rate constraint. Given θ and R, the bound on p_{e_i} is a convex and decreasing function of r_i over the range $0 \leq r_i < r_{s_i}$, i.e., $p_{e_i}(r_i + 1) < p_{e_i}(r_i)$ where $p_{e_i}(r)$ is the expression evaluated at $r_i = r$. We propose a greedy algorithm which assigns bits to the streams $i = 1, \ldots, N$ one bit at a time, and in each assignment the stream that results in the largest error probability reduction is selected. The bit allocation process is completed when all R bits are assigned. The pseudo code of the algorithm follows:

Let $r_i = 0$ and $\Delta p_{e_i} = p_{e_i}(0) - p_{e_i}(1)$ for $i = 1, \ldots, N$
for r=1:R
$i^* = \arg\max_i \Delta p_{e_i}$ and $r_{i^*} = r_{i^*} + 1$
if $r_{i^*} == r_{s_{i^*}}$ then $\Delta p_{e_{i^*}} = -K$
else $\Delta p_{e_{i^*}} = p_{e_{i^*}}(r_{i^*}) - p_{e_{i^*}}(r_{i^*} + 1)$
end
end

where we choose $K > 1$. The final bit allocation solution is $\{r_i^*\}_{i=1}^N$. Relying on the correlation model parameters the encoder forms the expression p_{e_i}, finds $\{r_i^*\}_{i=1}^N$, and utilizes these rates to encode s_i.

7. NUMERICAL RESULTS

In practice the source encoder has hardware limitation and thus uses a fixed set of quantizers. Also the value for R can be roughly estimated assuming that there is a constraint on the amount of energy that can be spent to transmit a bit of information and that there is a limited power budget available at the source encoder. In our simulation we further investigate the effect of the particular choices for θ and R on the performance of the codec which employs the rate allocation algorithm described in Section 6.

The experiment parameters are the following: $N = 4$, $\boldsymbol{H}$ and $\boldsymbol{R}_{xx}$ are Toeplitz matrices with first row $[1\ \rho\ \rho^2\ \rho^3]$ where $\rho = 0.6,\ 0.8$, respectively. We generate $\boldsymbol{x}$ and $\boldsymbol{n}$ independently according to $\boldsymbol{n} \sim \mathcal{N}(\boldsymbol{0}, \sigma_n^2 \boldsymbol{I})$ with $10log_{10}\sigma_n^2 = -12$ dB and $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_{xx})$. The horizontal and vertical axes correspond to the total coset rate R bits/vector and the MSE-distortion (in dB) due to the reconstruction of $\boldsymbol{x}$ at the decoder, respectively. We choose the set of θ values as $\theta = [0.03\ \ 0.05\ 0.08\ 0.09\ 0.1\ 0.15\ 0.2\ 0.25]$.

Figure 4 illustrates the performance of the proposed codec. We compare the achievable performance against $trace(\boldsymbol{R}_{x|y})$ where $\boldsymbol{R}_{x|y}$ is the error covariance of $\hat{\boldsymbol{x}} = E\{\boldsymbol{x}|\boldsymbol{y}\}$. The fact that the MSE-distortion is below $trace(\boldsymbol{R}_{x|y})$ indicates that incorporating the extra bits sent by the encoder reduces the distortion occurred during the reconstruction of $\boldsymbol{x}$. As a result of optimization over θ we have observed that as R increases less distortion can be achieved by employing finer quantizers, i.e., quantizers with higher rates. As a benchmark to evaluate the distortion-rate performance of our proposed algorithm, we use the theoretical bounds established in [7] for jointly Gaussian vectors $\boldsymbol{x}$ and $\boldsymbol{y}$. The rate-distortion function is [7]:

$$R(D) = \min_{D_i} \sum_{i=1}^{N} \max\{0.5 \log_2 \frac{\lambda_i^2}{D_i}, 0\}$$

where λ_i^2 are the N eigenvalues of the matrix $\boldsymbol{R}_{x|y}$ and the minimum is over all sets $\{D_i\}_{i=1}^N$ satisfying $\sum_{i=1}^N D_i \leq D$. Figure 4 shows a wide gap between the codecs performance and the theoretical $R(D)$ bound. In [7] it is shown that the $R(D)$ bound is achievable if the encoder applies conditional Karhünen-Loève transform (obtained from the EVD of $\boldsymbol{R}_{x|y}$) to $\boldsymbol{x}$ and encodes the transformed coefficients. Furthermore, attaining the bound requires a vector quantization over an infinite length sequence of uncorrelated transformed coefficients.

To quantify the gain we achieve by coding with side information we consider a simple codec whose encoder quantizes the entries of $\boldsymbol{s}$ using the uniform quantizers and sends the index of the quantization intervals to the decoder. The decoder receives these indices error free and uses them to form $\hat{\boldsymbol{s}}$, without knowing explicitly $\boldsymbol{y}$. Figure 4 shows that at low rates (e.g. for rates $\leq$ 8 bits/vector) there is a significant difference between the attainable distortion by the simple

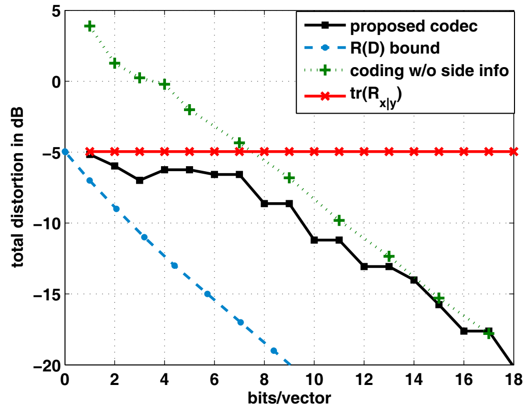

Fig. 4. codec without LT at the encoder: performance comparison with $R(D)$ lower bound, $tr(\boldsymbol{R}_{x|y})$, and the simple codec.

codec and the proposed codecs, where the difference decreases as the rates increases.

In a nutshell in this paper we proposed WZ codecs which are suitable for a wider class of vector sources $\boldsymbol{x}$ and $\boldsymbol{y}$ where $\boldsymbol{x}$ and $\boldsymbol{y}$ are related as $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}$. Viewing this dependency model as a virtual fading channel with input $\boldsymbol{x}$ and output $\boldsymbol{y}$, we utilize linear precoder and decoder to convert the problem of vector source coding into a set of parallel scalar source coding problems. Using this technique, we can use the WZ codes designed for the simple model $y = x + n$ for the more general model $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}$. We also proposed a simple rate allocation policy for assigning encoding rates to each scalar sources.

8. REFERENCES

[1] S.Sandeep and K. Ramchandran, "Distributed source coding using syndroms (DISCUS): design and construction", *IEEE Trans. Inform. Theory*, Vol. 49, Issue 3, pp. 626-643, March 2003.

[2] A. D. Wyner and J. Ziv ,"The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, Vol. IT-22, pp. 1-10, January 1976.

[3] R. Zamir, S. Shamai and U. Erez, "Nested linear lattice codes for structured multiterminal binning," *IEEE Trans. Inform. Theory*, Vol. 48, No.6, pp. 1250-1276, June 2002.

[4] X. Giraud and J. C. Belfiore, "Constellations matched to the Rayleigh fading channel" *IEEE Trans. Inform. Theory*, Vol. 42, No. 1, pp. 106-115, January 1996.

[5] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis and H. Sampath, "Optimal designs for space-time linear precoders and decoders", *IEEE Trans. Sig. Proc.*, Vol. 50, No. 5, pp. 1051-1064, May 2002.

[6] A. Vosoughi and A. Scaglione, "Precoding and Decoding Paradigms for Distributed Data Compression, " *Submitted to IEEE Tran. Sig. Proc.*, available at *http://crisp.ece.cornell.edu/*.

[7] M. Gastpar, P. L. Dragotti and M. Vetterli, "The distributed Karhunen-Loeve transform, " *EPFL technical report IC/2003/12*, also *IEEE Multimedia Sig. Proc.*, pp. 57-60, December 2002.

[8] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE JSAC*, Vol. 17, No. 10, pp. 1747-1758, October 1999.