

# Distributed Lossy Source Coding Using Real-Number Codes

Mojtaba Vaezi and Fabrice Labeau

Department of Electrical and Computer Engineering  
McGill University, Montreal, Quebec H3A 2A7, Canada  
Email: mojtaba.vaezi@mail.mcgill.ca, fabrice.labeau@mcgill.ca

**Abstract**—We show how real-number codes can be used to compress correlated sources, and establish a new framework for distributed lossy source coding, in which we quantize compressed sources instead of compressing quantized sources. This change in the order of binning and quantization blocks makes it possible to model correlation between continuous-valued sources more realistically and correct quantization error when the sources are completely correlated. The encoding and decoding procedures are described in detail, for discrete Fourier transform (DFT) codes. Reconstructed signal, in the mean-squared error sense, is seen to be better than or close to quantization error level in the conventional approach.

**Index Terms**—Distributed source coding, real-number codes, BCH-DFT codes, channel coding, Wyner-Ziv coding.

## I. INTRODUCTION

The distributed source coding (DSC) deals with compression of correlated sources which do not communicate with each other [1]. Lossless DSC (Slepian-Wolf coding), has been realized by different binary channel codes, including LDPC [2] and turbo codes [3]. The Wyner-Ziv coding problem [4], deals with lossy data compression with side information at the decoder, under a fidelity criterion. Current approach in the DSC of a continuous-valued source is to first convert it to a discrete-valued source using quantization, and then to apply Slepian-Wolf coding in the binary field. Similarly, a practical Wyner-Ziv encoder is realized by cascading a quantizer and Slepian-Wolf encoder [5], [6]. In other words, the quantized source is compressed. There are, hence, source coding (or quantization) loss and channel coding (or binning) loss. This approach is based on the assumption that there is still correlation remaining in the quantized version of correlated sources.

In this paper, we establish a new framework for the Wyner-Ziv coding. We propose to first compress the continuous-valued source and then quantize it, as opposed to the conventional approach. The compression is thus in the real field, aiming at representing the source with fewer samples.

To do compression, we generate either syndrome or parity samples of the input sequence using a real-number channel code, similar to what is done to compress a binary sequence

of data using binary channel codes. Then, we quantize these syndrome or parity samples and transmit them. There are still coding (binning) and quantization losses; however, since coding is performed before quantization, error correction is in the real field and quantization error can be corrected when two sources are completely correlated over a block of code. A second and more important advantage of this approach is the fact that the correlation channel model can be more realistic, as it captures the correlation between continuous-valued sources rather than quantized sources. In the conventional approach, it is implicitly assumed that quantization of correlated signals results in correlated sequences in the discrete domain which is not necessarily correct due to nonlinearity of quantization operation. In addition, most of previous works assume that this correlation, in the binary field, can be modeled by a binary symmetric channel (BSC) with a known crossover probability. To avoid the loss due to inaccuracy of correlation model, we exploit correlation between continuous-valued sources before quantization.

Specifically, we use real BCH-DFT codes [7], for compression in the real field. Owing to the DFT codes, the loss due to quantization can be decreased by a factor of  $k/n$  for an  $(n, k)$  DFT code [8], [9]. Additionally, if the two sources are perfectly correlated over one codevector, reconstruction loss vanishes. This is achieved in view of modeling the correlation between the two sources in the continuous domain. Finally, the proposed scheme seems more suitable for low-delay communication because using short DFT codes a reconstruction error better than quantization error is achievable.

The rest of this paper is organized as follows. In Section II, we motivate and introduce a new framework for lossy DSC. In Section III, we briefly review encoding and decoding in real DFT codes. Then in Section IV, we present the DFT encoder and decoder for the proposed system, both in the syndrome and parity approaches. These two approaches are also compared in this section. Section V discusses the simulation results. Section VI provides our concluding remarks.

## II. PROPOSED SYSTEM AND MOTIVATIONS

We introduce the use of real-number codes in lossy compression of correlated signals. Specifically, we use DFT codes [7], a class of real Bose-Chaudhuri-Hocquenghem (BCH) codes, to perform compression. Similar to error correction in

This work was supported by Hydro-Québec, the Natural Sciences and Engineering Research Council of Canada and McGill University in the framework of the NSERC/Hydro-Québec/McGill Industrial Research Chair in Interactive Information Infrastructure for the Power Grid.

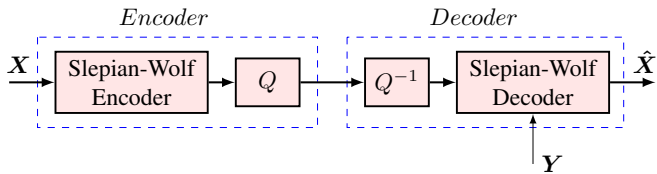


Fig. 1. The Wyner-Ziv coding using real-number codes.

finite fields, the basic idea of error correcting codes in the real field is to insert redundancy to a message vector of  $k$  samples to convert it to a codeword of  $n$  samples ( $n > k$ ) [7]. But unlike that, the insertion of redundancy in the real field is performed before quantization and entropy coding. The insertion of *soft redundancy* in the real-number codes has advantages over *hard redundancy* in the binary field. By using soft redundancy, one can go beyond quantization error, and thus reconstruct continuous-valued signals more accurately. This makes real-number codes more suitable than binary codes for lossy distributed source coding.

The proposed system is depicted in Fig. 1. Although it consists of the same blocks as existing practical Wyner-Ziv coding scheme [5], [6], the order of these blocks is changed here. That is, we perform Slepian-Wolf coding before quantization. This change in the order of the DSC and quantization blocks brings some advantages as described in the following.

- **Realistic correlation model:** In the existing framework for lossy DSC, correlation between two sources is modeled after quantization, i.e., in the binary domain. More precisely, correlation between quantized sources is usually modeled as a BSC, mostly with known crossover probability. Admittedly though, due to nonlinearity of quantization operation, correlation between the quantized signals is not known accurately even if it is known in the continuous domain. This motivates investigating a method that exploits correlation between continuous-valued sources to perform DSC.
- **Alleviating quantization error:** In lossy data compression with side information at the decoder, soft redundancy, added by DFT codes, can be used to correct both quantization errors and (correlation) channel errors. The loss due to quantization error thus can be recovered, at least partly if not wholly. More precisely, if the two sources are exactly the same over a codeword, quantization error can be corrected completely. That is, perfect reconstruction is achieved over corresponding samples. The loss due to quantization error is decreased even if correlation is not perfect, i.e., when (correlation) channel errors exist.
- **Low-delay communication:** If communication is subject to low-delay constraints, we cannot use turbo or LDPC codes, as their performance is not satisfactory for short code length. Whether low-delay requirement exists or not depends on the specific applications. However, even in the applications that low-delay transmission is not imperative, it is sometimes useful to consider low-dimensional systems for their low computational complexity.

### III. ENCODING AND DECODING WITH BCH-DFT CODES

Real BCH-DFT codes, a subset of complex BCH codes [7], are linear block codes over the real field. Any BCH-DFT code satisfies two properties. First, as a DFT code, its parity-check matrix is defined based on the DFT matrix. Second, similar to other BCH codes, the spectrum of any codeword is zero in a block of  $d - 1$  cyclically adjacent components, where  $d$  is the designed distance of that code [10]. A real BCH-DFT codes, in addition, has a generator matrix with real entries, as described below.

#### A. Encoding

An  $(n, k)$  real BCH-DFT code is defined by its generator and parity-check matrices. The generator matrix is given by

$$\mathbf{G} = \sqrt{\frac{n}{k}} \mathbf{W}_n^H \mathbf{\Sigma} \mathbf{W}_k, \quad (1)$$

in which  $\mathbf{W}_k$  and  $\mathbf{W}_n^H$  respectively are the DFT and IDFT matrices of size  $k$  and  $n$ , and  $\mathbf{\Sigma}$  is an  $n \times k$  matrix with  $n - k$  zero rows [11]–[14]. Particularly, for odd  $k$ ,  $\mathbf{\Sigma}$  has exactly  $k$  nonzero elements given as  $\Sigma_{00} = 1$ ,  $\Sigma_{i,i} = \Sigma_{n-i,k-i} = 1$ ,  $i = 1 : \frac{k-1}{2}$  [11], [12]. This guarantees the spectrum of any codeword to have  $n - k$  consecutive zeros, which is required for any BCH code [10]. The parity-check matrix  $\mathbf{H}$ , on the other hand, is constructed by using the  $n - k$  columns of  $\mathbf{W}_n^H$  corresponding to the  $n - k$  zero rows of  $\mathbf{\Sigma}$ . Therefore, due to unitary property of  $\mathbf{W}_n^H$ ,  $\mathbf{H}\mathbf{G} = \mathbf{0}$ .

In the rest of this paper, we use the term DFT code in lieu of real BCH-DFT code. Besides, we only consider odd numbers for  $k$  and  $n$ ; thus, the error correction capability of the code is  $t = \lfloor \frac{n-k}{2} \rfloor = \frac{n-k}{2}$ .

#### B. Decoding

For decoding, we use the extension of the well-known Peterson-Gorenstein-Zierler (PGZ) algorithm to the real field [10]. This algorithm, aimed at detecting, localizing, and estimating errors, works based on the syndrome of error. We summarize the main steps of this algorithm, adapted for a DFT code of length  $n$ , in the following.

- 1) Compute vector of syndrome samples
- 2) Determine the number of errors  $\nu$  by constructing a syndrome matrix and finding its rank
- 3) Find coefficients  $\Lambda_1, \dots, \Lambda_\nu$  of error-locating polynomial  $\Lambda(x) = \prod_{i=1}^{\nu} (1 - xX_i)$  whose roots are the inverse of error locations
- 4) Find the zeros  $X_1^{-1}, \dots, X_\nu^{-1}$  of  $\Lambda(x)$ ; the errors are then in locations  $i_1, \dots, i_\nu$  where  $X_1 = \alpha^{i_1}, \dots, X_\nu = \alpha^{i_\nu}$  and  $\alpha = e^{-j\frac{2\pi}{n}}$
- 5) Finally, determine error magnitudes by solving a set of linear equations whose constants coefficients are powers of  $X_i$ .

As mentioned, the PGZ algorithm works based on the syndrome of error, which is the syndrome of the received codeword, neglecting quantization. Let  $\mathbf{r} = \mathbf{y} + \mathbf{e}$  be the received vector, then

$$\mathbf{s} = \mathbf{H}\mathbf{r} = \mathbf{H}(\mathbf{y} + \mathbf{e}) = \mathbf{H}\mathbf{e}, \quad (2)$$

where  $\mathbf{s} = [s_1, s_2, \dots, s_{2t}]^T$  is a complex vector of length  $n - k$ . In practice however, the received vector is distorted by quantization ( $\mathbf{r} = \hat{\mathbf{y}} + \mathbf{e}$ ,  $\hat{\mathbf{y}} = \mathbf{y} + \mathbf{q}$ ) and its syndrome is no longer equal to the syndrome of error because

$$\tilde{\mathbf{s}} = \mathbf{H}\mathbf{r} = \mathbf{H}(\mathbf{y} + \mathbf{q} + \mathbf{e}) = \mathbf{s}_q + \mathbf{s}_e, \quad (3)$$

where  $\mathbf{s}_q \equiv \mathbf{H}\mathbf{q}$  and  $\mathbf{s}_e \equiv \mathbf{H}\mathbf{e}$ . While the “exact” value of errors is determined neglecting quantization, the decoding becomes an *estimation* problem in the presence of quantization. Then, it is imperative to modify the PGZ algorithm to detect errors reliably [10]–[13]. Error detection, localization, and also estimation can be largely improved using least squares methods [14].

### C. Performance Compared to Binary Codes

DFT codes by construction are capable of decreasing quantization error. When there is no error, an  $(n, k)$  DFT code brings down the mean-squared error (MSE), below the level of quantization error, with a factor of  $R_c = k/n$  [8], [9]. This is also shown to be valid for channel errors, as long as channel can be modeled as by additive noise. To appreciate this, one can consider the generator matrix of a DFT code as a tight frame [9]; it is known that frames are resilient to any additive noise, and tight frames reduce the MSE  $k/n$  times [15]. Hence, DFT codes can result in a MSE even better than quantization error level whereas the best possible MSE in a binary code is obviously lower-bounded by quantization error level.

## IV. WYNER-ZIV CODING USING DFT CODES

The concept of lossy DSC and Wyner-Ziv coding in the real field was described in Section II. In this section, we use DFT codes, as a specific means, to do Wyner-Ziv coding in the real field. This is accomplished by using DFT codes for binning, and transmitting compressed signal, in the form of either syndrome or parity samples.

Let  $\mathbf{x}$  be a sequence of i.i.d random variables  $x_1 x_2 \dots x_n$ , and  $\mathbf{y}$  be a noisy version of  $\mathbf{x}$  such that  $y_i = x_i + e_i$ , where  $e_i$  is continuous, i.i.d., and independent of  $x_i$ . Since  $e$  is continuous, this model precisely captures any variation of  $\mathbf{x}$ , so it can model correlation between  $\mathbf{x}$  and  $\mathbf{y}$  accurately. For example, the Gaussian, Gaussian Bernoulli-Gaussian, and Gaussian-Erasure correlation channels can be modeled using this model [16]. These correlation models are practically important in video coders that exploit Wyner-Ziv concepts, e.g., when the decoder builds side information via extrapolation of previously decoded frames or interpolation of key frames [16]. In this paper, the virtual correlation channel is assumed to be a Bernoulli-Gaussian channel, inserting at most  $t$  random errors in each codeword; thus,  $\mathbf{e}$  is a sparse vector.

### A. Syndrome Approach

1) *Encoding*: Given  $\mathbf{H}$ , to compress an arbitrary sequence of data samples, we multiply it with  $\mathbf{H}$  to find the corresponding syndrome samples  $\mathbf{s}_x = \mathbf{H}\mathbf{x}$ . The syndrome is then quantized ( $\hat{\mathbf{s}}_x = \mathbf{s}_x + \mathbf{q}$ ), and transmitted over a noiseless digital communication system, as shown in Fig. 2. Note that  $\mathbf{s}_x$ ,  $\hat{\mathbf{s}}_x$  are both complex vectors of length  $n - k$ .

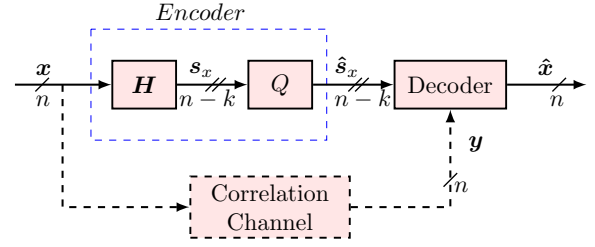


Fig. 2. The Wyner-Ziv coding using DFT codes: Syndrome approach.

2) *Decoding*: The decoder estimates the input sequence from the received syndrome and side information  $\mathbf{y}$ . To this end, it needs to evaluate the syndrome of channel (correlation) errors. This can be simply done by subtracting the received syndrome from syndrome of side information. Then, neglecting quantization, we obtain,

$$\mathbf{s}_e = \mathbf{s}_y - \mathbf{s}_x, \quad (4)$$

and  $\mathbf{s}_e$  can be used to precisely estimate the error vector, as described in Section III-B. In practice, however, the decoder knows  $\hat{\mathbf{s}}_x = \mathbf{s}_x + \mathbf{q}$  rather than  $\mathbf{s}_x$ . Therefore, only a distorted syndrome of error is available, i.e.,

$$\tilde{\mathbf{s}}_e = \mathbf{s}_y - \hat{\mathbf{s}}_x = \mathbf{s}_e - \mathbf{q}. \quad (5)$$

Hence, using the PGZ algorithm, error correction is accomplished based on (5). Note that, having computed the syndrome of error, decoding algorithm in DSC using DFT codes is exactly the same as that in the channel coding problem. This is different from DSC techniques in the binary field which usually require a slight modification in the corresponding channel coding algorithm to customize for DSC.

### B. Parity Approach

Syndrome-based Wyner-Ziv coding is straightforward but not very efficient because, in a real DFT code, syndrome samples are complex numbers. This means that to transmit each sample we need to send two real numbers, one for the real part and one for the imaginary part. Thus, the compression ratio, using an  $(n, k)$  DFT code, is  $\frac{n}{2(n-k)}$  whereas it is  $\frac{n}{n-k}$  for a similar binary code. This also imposes a constraint on the rate of code, i.e.,  $n < 2k$  or  $R_c > \frac{1}{2}$ , since otherwise there is no compression. In the sequel, we explore parity-based approach to the Wyner-Ziv coding.

1) *Encoding*: To compress  $\mathbf{x}$ , the encoder generates the corresponding parity sequence  $\mathbf{p}$  with  $n - k$  samples. The parity is then quantized and transmitted, as shown in Fig. 3, instead of transmitting the input data. The first step in parity-based system is to find the systematic generator matrix, as  $\mathbf{G}$  in (1) is not in the systematic form. Let  $\mathbf{H}$  be partitioned as  $\mathbf{H} = [\mathbf{H}_1 | \mathbf{H}_2]$ , where  $\mathbf{H}_1$  is a matrix of size  $n - k \times k$ , and  $\mathbf{H}_2$  is a square matrix of size  $n - k$ . Since  $\mathbf{H}_2$  is a Vandermonde matrix,  $\mathbf{H}_2^{-1}$  exist and we can write

$$\mathbf{H}_{\text{sys}} = \mathbf{H}_2^{-1} \mathbf{H} = [\mathbf{P} | \mathbf{I}_{2t}], \quad (6)$$

in which  $\mathbf{P} = \mathbf{H}_2^{-1} \mathbf{H}_1$  is an  $(n - k) \times k$  matrix, and  $\mathbf{I}_{2t}$  is an identity matrix of size  $2t$ .

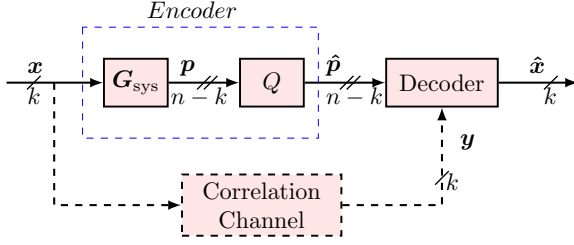


Fig. 3. The Wyner-Ziv coding using DFT codes: Parity approach.

The systematic generator matrix corresponding to  $\mathbf{H}_{\text{sys}}$  is given by

$$\mathbf{G}_{\text{sys}} = \begin{bmatrix} \mathbf{I}_k \\ -\mathbf{P} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_k \\ -\mathbf{H}_2^{-1}\mathbf{H}_1 \end{bmatrix}. \quad (7)$$

Clearly,  $\mathbf{H}_{\text{sys}}\mathbf{G}_{\text{sys}} = \mathbf{0}$ . It is also easy to check that

$$\mathbf{H}\mathbf{G}_{\text{sys}} = \mathbf{0}. \quad (8)$$

Therefore, we do not need to calculate  $\mathbf{H}_{\text{sys}}$  and the same parity-check matrix  $\mathbf{H}$  can be used for decoding in the parity approach.

An even easier way to come up with systematic generator matrix is to partition  $\mathbf{G}$  as  $\begin{bmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \end{bmatrix}$  where  $\mathbf{G}_1$  is a square matrix of size  $k$ . Then, from  $\mathbf{H}\mathbf{G} = \mathbf{0}$  and the fact that  $\mathbf{H}_2$  is invertible one can see  $\mathbf{G}_2 = -\mathbf{H}_2^{-1}\mathbf{H}_1\mathbf{G}_1$ ; thus, we have

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{I}_k \\ -\mathbf{H}_2^{-1}\mathbf{H}_1 \end{bmatrix} \mathbf{G}_1. \quad (9)$$

Note that  $\mathbf{G}_1$  is invertible because using (1) any  $k \times k$  submatrix of  $\mathbf{G}$  can be represented as product of a Vandermonde matrix and the DFT matrix  $\mathbf{W}_k$ . This is also proven using a different approach in [9], where it is shown that any subframe of  $\mathbf{G}$  is a frame and its rank is equal to  $k$ . Hence, since  $\mathbf{G}_1$  is invertible, the systematic generator matrix is given by

$$\mathbf{G}_{\text{sys}} = \mathbf{G}\mathbf{G}_1^{-1}. \quad (10)$$

Again  $\mathbf{H}\mathbf{G}_{\text{sys}} = \mathbf{0}$  because  $\mathbf{H}\mathbf{G} = \mathbf{0}$ . Therefore, the same parity-check matrix  $\mathbf{H}$  can be used for decoding in the parity approach. It is also easy to see that  $\mathbf{G}_{\text{sys}}$  is a real matrix. The question that remains to be answered is whether  $\mathbf{G}_{\text{sys}}$  corresponds to a BCH code? To generate a BCH code,  $\mathbf{G}_{\text{sys}}$  must have  $n - k$  consecutive zeros in the transform domain.  $\mathbf{W}_n\mathbf{G}_{\text{sys}} = (\mathbf{W}_n\mathbf{G})\mathbf{G}_1^{-1}$ , the Fourier transform of this matrix satisfies this condition because  $\mathbf{W}_n\mathbf{G}$ , the Fourier transform of original matrix, satisfies that.

Note that, since parity samples, unlike syndrome samples, are real numbers, using an  $(n, k)$  DFT code a compression ratio of  $\frac{k}{n-k}$  is achieved. Obviously, a compression ratio of  $\frac{n}{n-k}$  is achievable if we use a  $(2n - k, n)$  DFT code.

2) *Decoding*: A parity decoder estimates the input sequence from the received parity and side information  $\mathbf{y}$ . Similar to the syndrome approach, at the decoder, we need to find the syndrome of channel (correlation) errors. To do this, we append the parity to the side information and form a

vector of length  $n$  whose syndrome, neglecting quantization, is equal to the syndrome of error. That is,

$$\mathbf{z} = \begin{bmatrix} \mathbf{y} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{x} \\ \mathbf{p} \end{bmatrix} + \begin{bmatrix} \mathbf{e}_k \\ \mathbf{0} \end{bmatrix} = \mathbf{G}_{\text{sys}}\mathbf{x} + \mathbf{e}_n, \quad (11)$$

hence,

$$\mathbf{s}_z = \mathbf{s}_e. \quad (12)$$

Similarly, when quantization is involved ( $\hat{\mathbf{p}} = \mathbf{p} + \mathbf{q}$ ), we get

$$\tilde{\mathbf{z}} = \begin{bmatrix} \mathbf{y} \\ \hat{\mathbf{p}} \end{bmatrix} = \mathbf{z} + \begin{bmatrix} \mathbf{0} \\ \mathbf{q} \end{bmatrix} = \mathbf{G}_{\text{sys}}\mathbf{x} + \mathbf{e}_n + \mathbf{q}_n, \quad (13)$$

and

$$\mathbf{s}_{\tilde{\mathbf{z}}} = \mathbf{s}_e + \mathbf{s}_q, \quad (14)$$

in which,  $\mathbf{s}_q \equiv \mathbf{H}\mathbf{q}_n$ . Therefore, we obtain a distorted version of error syndrome. In both cases, the rest of the algorithm, which is based on the syndrome of error, is similar to that in the channel coding problem using DFT codes.

### C. Comparison Between the Two Approaches

As we saw earlier, using an  $(n, k)$  code the compression ratio in the syndrome and parity approaches, respectively, is  $\frac{n}{2(n-k)}$  and  $\frac{k}{n-k}$ . Hence, the parity approach is  $2k/n = 2R_c > 1$  times more efficient than the syndrome approach. Conversely, we can find two different codes that result in same compression ratio, say  $\frac{n}{n-k}$ . We know that in the parity approach, a  $(2n - k, n)$  code can be used for this matter. It is also easy to verify that, in the syndrome approach, a code with rate  $R_c = \frac{n+k}{2n}$  results in the same compression. For odd  $n$  and  $k$ , the  $(n, \frac{n+k}{2})$  DFT code gives the desired compression ratio. Thus, for a given compression ratio the parity approach implies a code with smaller rate compared to the code required in the syndrome approach.

## V. SIMULATION RESULTS

We evaluate the performance of the proposed systems using a Gauss-Markov source with zero mean, unit variance, and correlation coefficient 0.9; the effective range of the input sequences is thus  $[-4, 4]$ . The sources sequences are binned using a  $(7, 5)$  DFT code. The compressed vector, either syndrome or parity, is then quantized with a 6-bit uniform quantizer, and transmitted over a noiseless communication media. The correlation channel randomly inserts one error ( $t = 1$ ), generated by a Gaussian distribution. The decoder localizes and decodes errors. We compare the MSE between transmitted and reconstructed codevectors, to measure end to end distortion. In all simulations, we use 20,000 input frames for each channel-error-to-quantization-noise ratio (CEQNR). We vary the CEQNR and plot the resulting MSE. The result are presented in Fig. 4, and compared against the quantization error level in the existing lossy DSC methods.

It can be observed that the MSE in the syndrome approach is lower than quantization error except for a small range of CEQNR. Similarly, in the parity approach, the MSE is less than quantization error for a wide range of CEQNR. Note that in lossy DSC using binary codes, the MSE can be equal



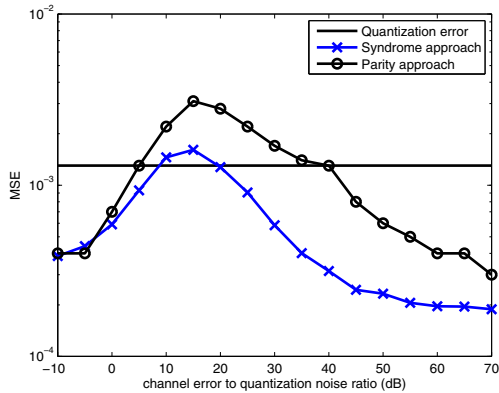


Fig. 4. Reconstruction error in the syndrome and parity approaches, using a (7, 5) DFT code in Fig. 2, 3. For both schemes, the virtual correlation channel inserts one error at each channel error to quantization noise ratio.

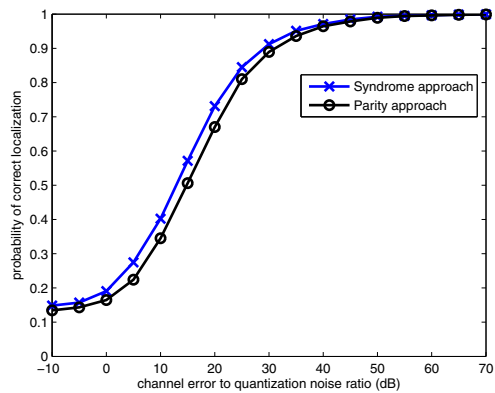


Fig. 5. Relative frequency of correct localization of correlation channel error in the syndrome and parity approaches, using a (7, 5) DFT code.

to quantization error only if the probability of error is zero. The performance of both algorithms improves as CEQNR is very high. This improvement is due to better error localization, since the higher the CEQNR the better the error localization, as shown in Fig. 5 and [11]. At very low CEQNRs, although error localization is poor, the MSE is still very low because, compared to quantization error, the errors are so small that the algorithm may localize and correct some of quantization errors instead. Additionally, reconstruction error is always reduced with a factor of  $R_c = k/n$ , in an  $(n, k)$  DFT code.

In terms of compression, the parity approach is  $2R_c = \frac{10}{7}$  times more efficient than the syndrome approach, as discussed earlier in Section IV-C. Not surprisingly though, the performance of the parity approach is not as good as that of the syndrome approach, because it contains fewer redundant samples. On top of that, in this simulation, 1/5 of samples are corrupted in the parity approach while this figure is 1/7 for the syndrome approach. The parity approach, however, suffers from the fact that dynamic range of parity samples, generated by (10), could be much higher than that of syndrome samples as  $t$  increases. This implies more precision bits to achieve the same accuracy. Finally, it is worth mentioning that when data and side information are the same over a block of code, reconstruction error becomes zero in both approaches.

## VI. CONCLUSIONS

We have introduced a new framework for distributed lossy source coding in general, and Wyner-Ziv coding specifically. The idea is to do binning before quantizing the continuous-valued signal, as opposed to the conventional approach where binning is done after quantization. By doing binning in the real field, the virtual correlation channel can be modeled more accurately, and quantization error can be corrected when there is no error. In the new paradigm, Wyner-Ziv coding is realized by cascading a Slepian-Wolf encoder with a quantizer. We employ real BCH-DFT codes to do the Slepian-Wolf in the real field. At the decoder, by introducing both syndrome-based and parity-based systems, we adapt the PGZ decoding algorithm accordingly. From simulation results, we conclude

that our systems, specifically with short codes, can improve the reconstruction error, so that they may become viable in real-world scenarios, where low-delay communication is required.

## REFERENCES

- [1] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. IT-19, pp. 471–480, July 1973.
- [2] A. D. Liveris, Z. Xiong, and C. N. Georghiades, "Compression of binary sources with side information at the decoder using LDPC codes," *IEEE Communications Letters*, vol. 6, pp. 440–442, Oct. 2002.
- [3] A. Aaron and B. Girod, "Compression with side information using turbo codes," in *Proc. IEEE Data Compression Conference*, pp. 252–261, 2002.
- [4] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Transactions on Information Theory*, vol. 22, pp. 1–10, Jan. 1976.
- [5] B. Girod, A. M. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," *Proceedings of the IEEE*, vol. 93, pp. 71–83, Jan. 2005.
- [6] Z. Xiong, A. D. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Processing Magazine*, vol. 21, pp. 80–94, Sept. 2004.
- [7] T. Marshall Jr., "Coding of real-number sequences for error correction: A digital signal processing problem," *IEEE Journal on Selected Areas in Communications*, vol. 2, pp. 381–392, March 1984.
- [8] V. K. Goyal, J. Kovačević, and J. A. Kelner, "Quantized frame expansions with erasures," *Applied and Computational Harmonic Analysis*, vol. 10, no. 3, pp. 203–233, 2001.
- [9] G. Rath and C. Guillemot, "Frame-theoretic analysis of DFT codes with erasures," *IEEE Transactions on Signal Processing*, vol. 52, pp. 447–460, Feb. 2004.
- [10] R. E. Blahut, *Algebraic Codes for Data Transmission*. New York: Cambridge Univ. Press, 2003.
- [11] G. Rath and C. Guillemot, "Subspace-based error and erasure correction with DFT codes for wireless channels," *IEEE Transactions on Signal Processing*, vol. 52, pp. 3241–3252, Nov. 2004.
- [12] G. Takos and C. N. Hadjicostis, "Determination of the number of errors in DFT codes subject to low-level quantization noise," *IEEE Transactions on Signal Processing*, vol. 56, pp. 1043–1054, March 2008.
- [13] A. Gabay, M. Kieffer, and P. Duhamel, "Joint source-channel coding using real BCH codes for robust image transmission," *IEEE Transactions on Image Processing*, vol. 16, pp. 1568–1583, June 2007.
- [14] M. Vaezi and F. Labeau, "Least squares solution for error correction on the real field using quantized DFT codes," to appear in EUSIPCO 2012.
- [15] J. Kovačević and A. Chebira, *An introduction to frames*. Now Publishers, 2008.
- [16] F. Bassi, M. Kieffer, and C. Weidmann, "Source coding with intermittent and degraded side information at the decoder," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2941–2944, 2008.