ON ROBUST SYNDROME-BASED DISTRIBUTED SOURCE CODING OVER NOISY CHANNELS USING LDPC CODES

Anoosheh Heidarzadeh and Farshad Lahouti

Wireless Multimedia Communications Lab.,
School of Electrical and Computer Engineering, University of Tehran, Iran
anoosheh@ieee.org, lahouti@ut.ac.ir

ABSTRACT

In this paper, we consider distributed source coding (DSC) problem over actual noisy channel addressed as distributed joint source-channel coding (DJSCC) while targeting the important applications of real-world transmission. Our objective is to design robust syndrome-based Slepian-Wolf decoding schemes guaranteeing both efficient distributed source compression and error protection. We utilize Low-Density Parity-Check (LDPC) codes and systematically design a decoder matrix based on the parity-check matrix of the encoder which is optimized for efficient decoding in DJSCC scenario.

***Index Terms*—** Slepian-Wolf coding, LDPC, distributed joint source channel coding (DJSCC), syndrome-based decoding.

1. INTRODUCTION

The elegant Slepian-Wolf theorem [1] affirms that the lossless compression of the correlated sources that do not communicate their outputs to each other, can be done as efficient as if they communicated. Driven by a host of emerging applications such as wireless video and distributed sensor networks, recently there has been a flurry of activities on the design of practical distributed source coding schemes. The Slepian-Wolf coding problem can be solved using two general methods named as DSC using syndrome (DISCUS) and DSC using parity (DISCUP) where the optimality of DISCUS approach is proven in [2].

An important assumption in Slepian-Wolf theorem is that the channels through which compressed information of sources are sent to the decoder are perfect. On the other hand, the real-world applications usually deal with noisy channels. This impractical premise is based on exploiting high rate powerful channel codes after Slepian-Wolf encoder to protect the compressed information from channel errors. However, there are two fundamental shortcomings with the optimal separate design. First, the feasibility of the separation theorem for

This work is supported in part by the Iran Telecommunications Research Center (ITRC).

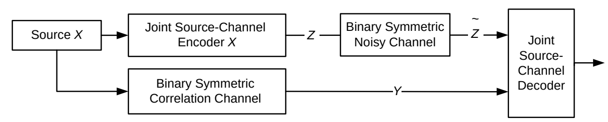

Figure 1. The joint source-channel coding of X with side information Y at the decoder.

source/channel coding with side information depends on the sufficiently large block-length codes [3]. Second, the practical channel codes are imperfect. Then, the residual errors at the input of the source decoder could therefore compromise the efficiency of the source coding system. Therefore, the practical systems are expected to perform better when employing joint source-channel coding (JSCC) [4].

Most efforts in this realm have focused on the distributed joint source-channel coding (DJSCC) problem in asymmetric case so far. As it is shown in Figure 1, this case considers that one of the sources is available at the decoder as the side information for the decoding of the other source, which its compressed version is transmitted through an actual noisy channel [4]. Most earlier DJSCC systems are based on DISCUP approaches such as [5], [6], and some references therein. Besides, Li *et al.* [7] propose error resilient compression systems for DSC over noisy channels using convolutional codes based on syndrome . An inspiring idea in [7] is that the parity-check matrix under DISCUS decoding can be changed to a new matrix for an equivalent channel decoding problem. A family of powerful codes which are designed based on their parity-check matrices are Low-Density Parity-Check (LDPC) codes [8]. Also, it is shown in [9] that the well-designed LDPC codes work well for DSC problems. On the other hand, in the asymmetric DJSCC, the used Slepian-Wolf code has to achieve the capacities of two different channels, i.e. the virtual correlation channel and the actual noisy channel. However, the design process of LDPC codes does not take both channels into account. As a result, LDPC-like codes such as Irregular Repeat Accumulate (IRA) [5] and Raptor codes [6] have been of more interest in DJSCC systems.

In this paper, we focus on the syndrome-based LDPC decoding in the asymmetric DJSCC scenario. Motivated by [7], we first consider a decoding scheme over an alternative parity-check matrix. As it is shown later, this decoding approach leads us to design an equivalent parity-check matrix which is specially optimized for the efficient decoding in DJSCC problems. Strictly speaking, both channels are considered in the design of the parity-check matrix used at the decoder side. Therefore, this work extends the successful LDPC-based DISCUS decoding to the case of DSC over noisy channels.

2. ASYMMETRIC CASE OF DJSCC

We consider the system of asymmetric DJSCC with the following notations which are used for the rest of this paper. In this case, first the memoryless binary source sequence $X^n = [x_1, ..., x_n]$ where $x_i \in \{0,1\}$ for $i = 1, 2, ..., n$, is encoded with a joint source-channel (JSC) encoder based on a Slepian-Wolf code with code rate $\frac{k}{n}$, resulting in syndrome bits $Z^{n-k} = [z_1, ..., z_{n-k}]$. We assume that the actual noisy channel through which the syndrome bits are sent to the decoder is a memoryless binary symmetric channel (BSC) with crossover probability $q < 0.5$. Then, the distorted syndrome sequence $\tilde{Z}^{n-k} = [\tilde{z}_1, ..., \tilde{z}_{n-k}]$ arrives at the decoder. In line with the DSC literatures, e.g., [9], the correlation between the source X and the side information $Y^n = [y_1, ..., y_n]$ where $y_i \in \{0,1\}$ is modeled by a BSC with crossover probability $p < 0.5$.

The objective of the joint source-channel decoder is to reconstruct X^n sequence, $\hat{X}^n$, utilizing both present information Y^n and $\tilde{Z}^{n-k}$ at the receiver. Different possible decoding procedures at the JSC decoder motivated us to evaluate their performances. To describe the intuitions behind our ideas we first discuss the performance of the conventional syndrome-based decoding using LDPC codes in DJSCC scenario.

3. LDPC-BASED DISCUS DECODING

A (n, k) LDPC code is determined by its sparse parity-check matrix $H_{n-k\times n}$ or, equivalently by its bipartite graph which is used in the message-passing (MP) decoding algorithm. In DSC problems, the received syndrome bits $\tilde{z}_i$'s are exactly the same as the transmitted z_i's because of the perfect channel assumption. Then, the value of the syndrome bits are considered as the values of the check nodes in LDPC-based DISCUS decoding [9]. Consequently, the MP decoder tries to find the best estimate of X^n from side information Y^n (as the distorted version of X^n) subject to the values of the check nodes in the decoder bipartite graph which are fixed and equal to the syndrome bits z_i's.

However, it is clear that this decoding rule in DJSCC setup tries to estimate wrong sequence $\hat{X}^n$ whenever even if at least one bit of $\tilde{z}_i$'s is corrupted by noise. This phenomenon shows that above LDPC-based DISCUS decoding cannot counteract with distortions caused by the actual noisy channel due to fixing $\tilde{z}_i$'s within the decoding process. Strictly speaking, we found that DISCUP decoding in DJSCC allows the received parity bits to flip if required. So, we investigate providing this capability for DISCUS decoding using special features of LDPC codes.

4. ROBUST DISCUS DECODING OVER NOISY CHANNELS

We consider a syndrome-based distributed source encoder using an LDPC code, similar to [9]. In the following we present two generalizations of the encoder's parity-check matrix to propose new matrices labeled as decoderI and decoderII when the decoding algorithm provides superior performance.

- *Decoder-I*

At the first sight, the generalization of the DISCUS decoding using LDPC codes to a robust decoding method over noisy channels seems to be hard and nontrivial. But a closer look reveals that the syndrome bits of one specific code with parity-check matrix $H_{n-k\times n}$ can be stated as the parity bits of another code with larger parity-check matrix $[H_{n-k\times n}|I_{n-k}]$ [7]. This matrix can be taken as the MP decoder matrix where I_{n-k} is an identity matrix appended to the original matrix H. The $[H|I]$ parity-check matrix demonstrates a systematic LDPC code with the input X^n and the output $[X^n|Z^{n-k}]$. Since the z_i's can be viewed as the parity bits of X^n, the decoder can then treat $[Y^n|\tilde{Z}^{n-k}]$ as a noisy version of the transmitted codeword $[X^n|Z^{n-k}]$ along two different channels. The initializations of LLR values sent from check nodes to their neighboring bit nodes can be done separately for both parts of the received bits Y^n and $\tilde{z}_i$'s by setting them to $(1 - 2y_i)\log\frac{1-p}{p}$ and $(1 - 2\tilde{z}_j)\log\frac{1-q}{q}$, respectively. Finally, it is clear that if the original LDPC code is powerful enough to do channel coding efficiently, $[H|I]$ relatively allows successful recovering of X^n, since source decoding problem, badly sensitive to the channel errors, is converted to its equivalent channel decoding setup which is naturally more powerful in sense of error protection.

Simulated results of the decoderI confirm its improvement in comparison with the original MP decoder. On the other hand, there are three major necessary features for one specific sparse matrix when MP decoding over this matrix works well. (i) the number of 1's in each column should be large enough. (ii) the number of 1's in each row must be relatively small. (iii) the factor graph of the code must include no cycles of length four [10]. As can be seen the decoderI essentially suffers the large number $n-k$ of weight-1 columns in $[H|I]$. Therefore, these columns compromise the expected decoding performance. In [11] it is pointed out that attempting to convert $[H|I]$ to a better matrix is technically challenging. So, they propose to start with a new, good sparse matrix $H'_{n-k\times 2n-k}$, diagonalize it using Gaussian Elimination method to $[P_{n-k\times n}|I_{n-k}]$. Then, they use P (which may be dense) to compute the syndromes and also utilize H' as the

MP decoder parity-check matrix. However, our simulations show that this approach fails especially when the correlation between sources decreases. Since, the sub-matrix P is not designed to generate optimal syndromes for Slepian-Wolf coding, it really can not achieve the capacity of the correlation channel.

- *Decoder-II*

This method is related to MP decoding over an equivalent larger parity-check matrix H^* than $[H|I]$ which satisfies all conditions (i) through (iii) above. Generally, the problem of producing equivalent parity-check matrices of one specific code defined by its parity-check matrix, are presented from different points of view in [10], [12] and [13]. Yedidia *et al.* propose a systematic approach to design a generalized parity-check matrix in [10] while their purpose of design is to satisfy all of the above conditions. However, their method does not guarantee the resulting matrix to be without any cycles of length four and more. As generalizations of this work, [12] and [13] present different methods to generate the equivalent matrices aiming at eliminating the loops in the original matrix. These methods are typically of interest in the channel coding scenarios of some codes such as Reed-Solomon codes which contain a large number of short cycles. But, at the outputs of these algorithms with input matrix $[H|I]$, the columns of weight-1 still remain. We present a novel algorithm which specially increase the number of 1's in the problematic columns of weight-1 besides satisfying other conditions (ii) and (iii). The fact is that each check constraint (row) of the parity-check matrix can be rewritten in terms of some new secondary bit nodes as pointed out in [10].

Our design procedure is general and it can be applied to any LDPC matrix within the context under consideration where there is no cycle initially in the parity-check matrix H. The design process of the decoder-II is detailed below.

- **Step 1:** Initialize $w = 1$.
- **Step 2:** Let ρ_w be w'th row of $[H|I]$. The set B_w shows the positions of 1's in ρ_w and $|B_w|$ denotes the number of 1's in the set B_w.
- **Step 3:** Choose randomly $|B_w| - 1$ elements of B_w which are between 1 and n, and the element $w + n$ in B_w as the set $G_{1,w}$. Introduce a new secondary bit node (column), indexed by $S_{1,w}$. Set the entries in the ρ_w, indexed by $G_{1,w}$, to zero and add an element 1 to ρ_w corresponding to $S_{1,w}$. Add a new row of length $S_{1,w}$ and set the entries in this row, indexed by $G_{1,w}$ and $S_{1,w}$, to one.
- **Step 4:** Choose the element $w + n$ in $G_{1,w}$ and an element which is in B_w and it is not a member of $G_{1,w}$ as the set $G_{2,w}$. Introduce another new secondary bit node, indexed by $S_{2,w}$. Add a new row of length $S_{2,w}$ to the two created rows in Step 3. Set the entries in this row, indexed by $G_{2,w}$ and $S_{2,w}$, to one. Add a new row of length $S_{2,w}$ and set the entries in this row, indexed by $S_{2,w}$ and $G_{1,w}$ except $w + n$ to one.

Although, the number of 1's in weight-1 column $w + n$ is increased to two, the second and the fourth newly created rows have similar columns resulting in cycles of length four which must be eliminated. To this end, we utilize the Maximum Cycle Strategy (MCS) algorithm of [13].

- **Step 5:** Define a set $\mathcal{C}_w$ which consists of the output rows of Step 4. Run MCS algorithm over the set $\mathcal{C}_w$ as the input matrix of MCS with output matrix H_w which consists of five rows equivalent to ρ_w of $[H|I]$ without any loop of length four while satisfying the conditions (i) and (ii) of good MP decoding.
- **Step 6:** Set $w = w + 1$. If $w \leq n - k$ go to Step 2, else go to Step 7.
- **Step 7:** Set $H^* = [H_1; ...; H_{n-k}]$ where the size of H^* and H_w's for $w = 1, 2, ..., n - k$, are $5(n - k) \times (5n - 4k)$ and $5 \times (5n - 4k)$, respectively.

5. PERFORMANCE ANALYSIS

In this section we compare the methods presented in sections 3 and 4. We designed a short regular $(256, 128)$ LDPC code $H_{128\times256}$ of rate $\frac{1}{2}$ which has three 1's in each column and six 1's in each row without any cycle of length four. This code can achieve $2 : 1$ compression in the DISCUS scheme. Also, the number of iterations of joint MP decoding for all of the methods are fixed to 30. Also, the transmissions of the syndromes Z^{128} are continued until 200 sequences $\hat{X}^{256}$ are erroneously reconstructed.

5.1. Simulation Results

In Figure 2, the performance of the mentioned methods in sections 3 and 4 are presented for different actual noisy channels (different q's) where the correlation of the sources is fixed to $p = 0.02$. The BER values at $q = 0$ authenticate that the used LDPC code is powerful enough to solve the ideal Slepian-Wolf coding. As we expected, the simulation results of the decoderI demonstrate noticeable improvement in BER in comparison with LDPC-based DISCUS decoding. Strictly speaking, we can see that the BER is almost reduced 91.1% in average. In the decoderII, we transformed $[H_{128\times256}|I_{128}]$ to a new MP decoder parity-check matrix $H^*_{640\times768}$. The overall BER improvement is around 95%.

Figure 3 shows the BER performance comparison from a different viewpoint where the actual noisy channel is a BSC with $q = 0.01$ for different correlations between sources. Indeed, the simulated outcomes of the decoderII are still surprising and also outperform the proposed algorithm in [11] when the correlation between sources decreases.

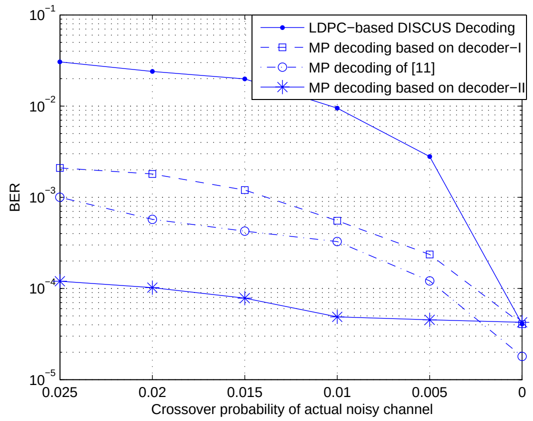

Figure 2. BER comparison for DJSCC vs. different q's where the correlation channel is modeled by a BSC with $p = 0.02$.

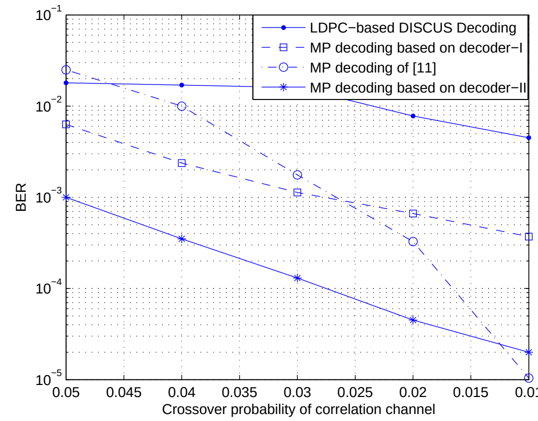

Figure 3. BER comparison in DJSCC scenario where the actual noisy channel is a BSC with $q = 0.01$ vs. different p's.

5.2. Complexity Considerations

The outstanding results of decoderII naturally come at the price of increased complexity at the decoding process which as we will demonstrate below is a reasonable computational complexity. We know that the LDPC decoding complexity corresponds to the average number of 1's per rows. Actually, in our simulated $[H|I]$ the number of the rows is increased five times in H^*. Moreover, three specific secondary bit nodes are used for each row of $[H|I]$. Hence, the number of the columns in $[H|I]$ is increased to $3(n-k)+(2n-k)$ in H^*. Each secondary column has just two ones and also the weights of the original columns have almost remained constant. Therefore, though the size of H^* is obviously larger than $[H|I]$, but the number of 1's in H^* remains small in average per rows.

6. REFERENCES

[1] J. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. 19, pp. 471–480, Jul. 1973.

[2] S.S. Pradhan, and K. Ramchandran, "Distrubuted source coding using syndromes (DISCUS): Design and construction," *Proc. IEEE DCC*, pp. 158–167, Mar. 1999.

[3] S. Shamai, S. Verdu, and R. Zamir, "Systematic lossy source/channel coding," *IEEE Trans. Inform. Theory*, vol. 44, pp. 564–579, Mar. 1998.

[4] J. Garcia-Frias, and Z. Xiong, "Distributed source and joint source-channel coding: from theory to practice," *Proc. ICASSP*, vol. 5, pp. 1093–1096, Mar. 2005.

[5] A.D. Liveris, Z. Xiong, and C.N. Georghiades, "Joint source-channel coding of binary sources with side information at the decoder using IRA codes," *IEEE Multimedia Signal Proc. Workshop*, Dec. 2002.

[6] Q. Xu, V. Stanković, and Z. Xiong, "Distributed joint source-channel coding of video using Raptor codes," *IEEE Journal on Selected Areas in Communications*, vol. 25, No. 4, May. 2007.

[7] P. Tan, and J. Li, "Enhancing the robustness of distributed compression using ideas from channel coding," *IEEE GLOBCOM*, vol. 4, Nov.-Dec. 2005.

[8] T.J. Richardson, M.A. Shokrollahi, and R.L. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inform. Theory*, vol. IT-47, pp. 619–637, Feb. 2001.

[9] A.D. Liveris, Z. Xiong, and C.N. Georghiades, "Compression of binary sources with side information at the decoder using LDPC codes," *IEEE Commun. Lett.* , vol. 6, pp. 440–442, Oct. 2002.

[10] J. Yedidia, J. Chen, and M. Fossorier, "Generating code representations suitable for belief propagation decoding," *Proc. 40th Allerton Conference on Communication, Control and Computing*, pp. 447–456, Oct. 2002.

[11] J. Li, and R. Hu, "Slepian-Wolf cooperation: A practical and efficient compress-and-forward relay scheme," *Proc. 43rd Annual Allerton Conf. on Communications, Control and Computing*, St. Louis, MO, Nov. 2005.

[12] S. Sankaranarayanan, and B. Vasic, "Iterative Decoding of Linear Block Codes: A Parity-Check Orthogonalization Approach," *IEEE Trans. Inform. Theory*, vol. 51, Issue 9, pp. 3347–3353, Sep. 2005.

[13] V. Kumar, and O. Milenkovic, "On graphical representations of algebraic codes suitable for iterative decoding," *IEEE Commun. Lett.*, vol. 9, Issue 8, pp. 729–731, Aug. 2005.