

# Spatially-Coupled Codes for Side-Information Problems

Santhosh Kumar, Avinash Vem, Krishna Narayanan, Henry D. Pfister  
Department of Electrical and Computer Engineering, Texas A&M University

**Abstract**—For compound LDGM/LDPC codes with maximum a posteriori (MAP) processing, Wainwright and Martinian showed that the information-theoretic rate regions of the Wyner-Ziv (WZ) and Gelfand-Pinsker (GP) problems are achievable. For the same ensemble, these rates do not appear to be achievable with message-passing guided decimation (GD). Fortunately, spatially-coupled (SC) codes seem to provide an elegant remedy when iterative decoding falls short of MAP decoding. In particular, Aref et al. recently introduced SC LDGM codes that approach the rate-distortion region with belief-propagation guided decimation (BPGD). In this paper, we show that SC compound LDGM/LDPC codes with BPGD can approach the rate regions of the WZ and GP problems.

**Index Terms**—Belief-propagation, channel coding, convolutional LDPC codes, LDGM codes, LDPC codes, rate distortion.

## I. INTRODUCTION

In this paper, we focus on the Wyner-Ziv (WZ) and Gelfand-Pinsker (GP) coding problems. For lossy compression with a fixed distortion constraint and side information at the receiver, the WZ rate  $R_{WZ}$  is the minimum achievable rate for source encoding [1]. For channel coding with an input constraint and side information at the transmitter, the GP rate  $R_{GP}$  is the maximum achievable information rate [2]. Source and channel coding problems with side information arise naturally in network information theory. Solutions to these problems often require the use of “binning”, where a set of codewords is partitioned into separate bins such that each bin is a good source code or channel code [3].

Sparse-graph codes with iterative decoding have been very successful for channel coding problems. This good news does not extend automatically, however, to problems that require the quantization of an arbitrary sequence to a nearby codeword (e.g., lossy compression, WZ, and GP). These problems are not easily solved via iterative decoding because the iterations only converge for a vanishing fraction of the entire sequence space. In the past few years, a number of practical constructions for these problems have been considered [4], [5], [6], [7]. In all these cases, there are non-negligible performance gaps relative to what is achievable for related single-user problems.

For lossy compression, low-density generator matrix (LDGM) codes with modified versions of BP decoding can achieve practically interesting results but they appear to have a non-negligible gap to the optimal compression rate [8], [9]. Better results can be obtained by using guided decimation

techniques where an iterative algorithm (e.g., BP or survey propagation (SP)) is used to sequentially identify bits with a large bias and then fix them to match the bias [10], [11]. Building on these results, spatially-coupled (SC) LDGM codes with BP guided decimation (BPGD) were recently shown by Aref et al. to approach the optimal compression rate for a binary symmetric source (BSS) [12], [13].

In this paper, we extend the approach of Aref et al. to the WZ and GP problems defined by BSSs and binary symmetric channels (BSCs). For these problems, Wainwright and Martinian showed that compound LDGM/LDPC codes can achieve the optimal rates when maximum-likelihood processing is used [14]. In this work, we show empirically that BPGD of SC compound LDGM/LDPC codes can approach the optimal WZ and GP rates.

A new complication with compound LDGM/LDPC codes is that the BPGD encoding now has the possibility of failure. For iterative encoding with LDGM codes, encoding always succeeds but the resulting distortion depends on the code and encoder details. For compound LDGM/LDPC codes, the compressed bits are also required to satisfy some parity checks and the iterative encoder is not guaranteed to find a valid encoding regardless of the target distortion. This is similar to what happens when BPGD is used for constraint satisfaction problems [15]. In fact, without spatial-coupling the BPGD encoder failed in every experiment.

Finally, it is important to note that the WZ and GP problems can also be solved using polar codes [16]. In fact, polar codes allow the deterministic construction of codes that approach the optimal rates. The main drawback of polar codes is that their rate approaches the theoretical limit quite slowly as the block length increases. Thus, they require very large block lengths to operate at rates close to the limit. The precise trade-offs between SC codes and polar codes for these problems remains an interesting open problem.

## II. BACKGROUND

In this section, we briefly describe the WZ and GP side-information problems, and a coding scheme that achieves the rate regions of these problems. We closely follow the formulation in [14].

### A. Side-Information Problems

1) *Rate Distortion with Side Information*: Consider a linear code  $\mathcal{C} \subset \{0, 1\}^n$  with rate  $R = k/n$  and an iid sequence  $X^n = (X_1, \dots, X_n) \in \{0, 1\}^n$ , where each  $X_i$  is  $\text{Ber}(\frac{1}{2})$  and  $\text{Ber}(\delta)$  denotes a Bernoulli random variable with parameter  $\delta$ . In the rate distortion problem, the objective is to optimally

This material is based upon work supported in part by the National Science Foundation (NSF) under Grant No. 1320924. Any opinions, findings, conclusions, and recommendations expressed in this material are those of the authors and do not necessarily reflect the views of these sponsors.

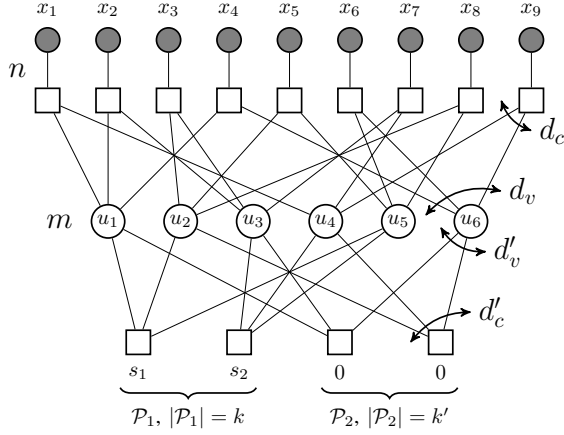


Fig. 1. The Tanner graph representation of a compound LDGM/LDPC code.

encode  $X^n$  to some codeword  $\hat{X}^n \in \mathcal{C}$  so as to minimize the average normalized Hamming distortion  $D$  between  $X^n$  and  $\hat{X}^n$ ,  $D = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i - \hat{X}_i]$ . For this problem, Shannon's rate-distortion theory [3] shows that any rate  $R > 1 - h(D)$  is achievable, where  $h(\cdot)$  is the binary entropy function. Since  $R \geq 1 - h(D)$  is also necessary, roughly  $n(1 - h(D))$  bits are required to specify  $X^n$  up to normalized distortion  $D$ .

In the WZ formulation of the rate distortion problem, there is side information  $Z^n$  at the decoder about  $X^n$ . Suppose the side information  $Z^n$  takes the form  $Z_i = X_i \oplus \text{Ber}(\delta)$ , where  $\oplus$  denotes addition modulo 2. Then, [1] shows that any rate,

$$R > R_{WZ}(D, \delta) = \text{l.c.e}\{h(D * \delta) - h(D), (\delta, 0)\}, \quad (1)$$

where  $D * \delta = D(1 - \delta) + \delta(1 - D)$ , is achievable, to describe  $X^n$  (up to normalized distortion  $D$ ) along with side information  $Z^n$ . In (1), *l.c.e* denotes the lower convex envelope.

2) *Channel Coding with Side Information*: Suppose again that  $\mathcal{C} \subset \{0, 1\}^n$  is a linear code with rate  $R = k/n$  and that  $W^n \in \{0, 1\}^n$  is distributed iid according to  $\text{Ber}(\delta)$ . In the channel coding problem, the objective is to encode a message  $m^k \in \{0, 1\}^k$  into a codeword  $X^n \in \mathcal{C}$  so that the decoder can reliably estimate the message  $m^k$  from the decoder output  $X^n \oplus W^n$ . In the GP formulation, the output at the receiver is given by  $Y^n = X^n \oplus Z^n \oplus W^n$ , where the side information  $Z^n \in \{0, 1\}^n$  is available a priori to the encoder but not the decoder. Additionally, the average weight of the codeword  $X^n$  is required to satisfy  $\frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i] \leq p$ ,  $\delta < p \leq \frac{1}{2}$ .

Under these conditions, [2] shows that the rate constraint

$$R < R_{GP}(p, \delta) = h(p) - h(\delta) \quad (2)$$

is both necessary and sufficient.

### B. Coding Scheme via Compound LDGM/LDPC Codes

We now describe coding schemes for the above side-information problems that achieve the rate regions in (1) and (2) under optimal encoding and decoding. The emphasis here is not on the practical nature of the encoding or decoding technique. In the next section, we modify the coding scheme presented here to have efficient encoding and decoding algorithms that appear to achieve the rate regions (1) and (2).

At the heart of these coding schemes is a compound LDGM/LDPC code [17], [18], [14]. We illustrate these codes through an example. The upper portion of the Tanner graph in Fig. 1 represents a length- $n$  LDGM code with information bit length  $m$ , where the codeword is  $x^n = (x_1, \dots, x_n) \in \{0, 1\}^n$  and the information bits are  $(u_1, \dots, u_m) \in \{0, 1\}^m$ . In the compound code, additionally, the information bits are required to satisfy certain parity constraints. These constraints are split into two groups  $\mathcal{P}_1$  and  $\mathcal{P}_2$  with  $|\mathcal{P}_1| = k$  and  $|\mathcal{P}_2| = k'$ . The parities in  $\mathcal{P}_1$  specify a fixed constraint given by  $s^k = (s_1, \dots, s_k) \in \{0, 1\}^k$ . For the example in Fig. 1, this means  $u_1 \oplus u_2 \oplus u_5 = s_1$  and  $u_3 \oplus u_4 \oplus u_5 = s_2$ . The parities in group  $\mathcal{P}_2$  are the usual even parities and for the example,  $u_1 \oplus u_3 \oplus u_6 = 0$ ,  $u_2 \oplus u_4 \oplus u_6 = 0$ . The bottom part therefore represents constraints akin to an LDPC code. As such, these codes are referred to as compound LDGM/LDPC codes. We distinguish between LDGM and LDPC check-nodes and we refer to the nodes representing  $x_i$ 's as LDGM bit-nodes and  $u_i$ 's as LDPC bit-nodes. The codebook is the set of all sequences  $x^n$  generated by the LDPC bit-nodes that satisfy the required constraints.

A few observations are in order. We note that the design rate of this code is  $\frac{m-k-k'}{n}$ . The codebook specified by a compound code is linear if and only if the parity constraints in  $\mathcal{P}_1$  satisfy  $s^k = 0^k$ . There is also a natural coset decomposition of these codes. Ignore for a moment the parity constraints specified by  $\mathcal{P}_1$ , and denote the resulting linear code by  $\mathcal{C}$ . Denote by  $\mathcal{C}(s^k)$ , the original codebook when the parities in  $\mathcal{P}_1$  satisfy  $s^k \in \{0, 1\}^k$ . It is easy to see that  $\mathcal{C} = \bigcup_{s^k \in \{0, 1\}^k} \mathcal{C}(s^k)$ .

Now, we describe another codebook  $\mathcal{C}'$  that is of interest for the WZ problem. The codebook  $\mathcal{C}'$  is obtained from  $\mathcal{C}$  by ignoring all the constraints in  $\mathcal{P}_2$ . In particular,  $\mathcal{C}'$  is the linear code generated from the LDGM part of  $\mathcal{C}$ . In describing  $\mathcal{C}$ , the constraints in  $\mathcal{P}_1$  are not active. When the constraints in  $\mathcal{P}_1$  are active, the codebook generated by ignoring  $\mathcal{P}_2$  is denoted by  $\mathcal{C}'(s^k)$ . Note that  $\mathcal{C}'(s^k)$  is a compound code in itself (with an empty  $\mathcal{P}_2$ ). We again have the coset decomposition of  $\mathcal{C}'$  into  $\mathcal{C}'(s^k)$ .

The main attraction of the compound codes is that they are simultaneously good for rate distortion and channel coding. In this article, we use the term *good* to mean the following.

**Remark 1.** We call a code “good for rate distortion” if the design rate of the code satisfies  $R = 1 - h(D) + \varepsilon$  for a small  $\varepsilon$  and, when the code is used to optimally encode a  $\text{Ber}(\frac{1}{2})$  sequence, the average normalized Hamming distortion is at most  $D$ . Similarly, we call a code “good for channel coding” if  $R = 1 - h(\delta) - \varepsilon$  and, when the code is used for the channel coding problem with the channel  $\text{Ber}(\delta)$ , the message can be reliably estimated under optimal decoding with probability of error at most  $\varepsilon$ .

We know that the codebooks  $\mathcal{C}$ ,  $\mathcal{C}(s^k)$ ,  $\mathcal{C}'(s^k)$  are good for both rate distortion and channel coding for any  $s^k \in \{0, 1\}^k$  [14]. But  $\mathcal{C}'$  is not a good channel code, since LDGM codes with fixed degrees exhibit non-negligible error floors.

However,  $\mathcal{C}'$  is good for rate distortion [11].

Below, we describe coding schemes for our side-information problems, using the codebooks  $\mathcal{C}$ ,  $\mathcal{C}'$ ,  $\mathcal{C}(S)$ ,  $\mathcal{C}'(S)$ . That these schemes achieve the rate regions in (1) and (2) will be an immediate consequence of the fact that compound codes are simultaneously good for rate distortion and channel coding, and LDGM codes are good for rate distortion. We only make heuristic arguments in the following, but these can easily be made rigorous.

1) *Coding Scheme for Wyner-Ziv*: For small  $\varepsilon > 0$ , choose  $n$ ,  $m$ ,  $k$  such that

$$\frac{m}{n} = 1 - h(D) + \varepsilon/2, \quad \frac{k}{n} = h(D * \delta) - h(D) + \varepsilon.$$

Consider a compound code with block length  $n$ , information bit length  $m$ ,  $|\mathcal{P}_1| = k$  and  $|\mathcal{P}_2| = 0$  such that  $\mathcal{C}'$  is a good rate distortion code with rate  $m/n$  and  $\mathcal{C}'(s^k)$  is a good channel code with rate  $1 - h(D * \delta) - \varepsilon/2$ .

Suppose we want to encode  $X^n \in \{0, 1\}^n$  up to normalized distortion  $D$ , with side information  $Z^n$  in the form  $Z_i = X_i \oplus \text{Ber}(\delta)$ . The sequence  $X^n$  can be encoded to  $\hat{X}^n \in \mathcal{C}'$  with distortion at most  $D$ , since  $\mathcal{C}'$  is a good rate distortion code. The sequence  $\hat{X}^n$  belongs to a unique coset  $\mathcal{C}'(\hat{s}^k)$  for some  $\hat{s}^k \in \{0, 1\}^k$ . The encoder transmits  $\hat{s}^k$  to the decoder, which requires a rate of  $R = \frac{k}{n} = h(D * \delta) - h(D) + \varepsilon$ .

The decoder, together with side information  $Z^n$  and  $\hat{s}^k$ , tries to recover  $\hat{X}^n$ . Note that

$$Z_i = X_i \oplus \text{Ber}(\delta), \quad \text{and} \quad \frac{1}{n} \sum_{i=1}^n \mathbb{E}|X_i - \hat{X}_i| \approx D.$$

This however implies  $Z_i \approx \hat{X}_i \oplus \text{Ber}(D * \delta)$ . Since  $\mathcal{C}'(\hat{s}^k)$  is a good channel code with rate  $1 - h(D * \delta) - \varepsilon/2$ , the decoder can recover  $\hat{X}^n$  reliably. As such, any rate  $R > h(D * \delta) - h(D)$  is sufficient to describe  $X$  up to distortion  $D$ . Thus, the rate region in (1) is achievable with this scheme.

**Remark 2.** The coding scheme we described above is slightly different from [14]. The difference is in using  $\mathcal{C}'$  instead of  $\mathcal{C}$  when compressing  $X^n$  to  $\hat{X}^n$ . The reason for our choice will be apparent later.

2) *Coding Scheme for Gelfand-Pinsker*: First, choose  $n$ ,  $m$ ,  $k$ ,  $k'$  such that

$$\frac{m - k'}{n} = 1 - h(\delta) - \varepsilon/2, \quad \frac{m - k - k'}{n} = 1 - h(p) + \varepsilon/2.$$

Consider a compound code with parameters  $n$ ,  $m$ ,  $k$ ,  $k'$  such that  $\mathcal{C}$  is a good channel code with rate  $\frac{m - k'}{n}$  and  $\mathcal{C}(s^k)$  is a good rate distortion code with rate  $\frac{m - k - k'}{n}$ .

We want to transmit a message  $m^{k'} \in \{0, 1\}^{k'}$ , when the encoder has side information  $Z^n$ . First, at the encoder, the side information  $Z^n$  is compressed to  $\hat{Z}^n \in \mathcal{C}(m^k)$ . Since  $\mathcal{C}(m^k)$  is a good rate distortion code with rate  $1 - h(p) + \varepsilon/2$ , we have

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E}|Z_i - \hat{Z}_i| \approx p, \quad \text{or} \quad \hat{Z}_i \approx Z_i \oplus \text{Ber}(p).$$

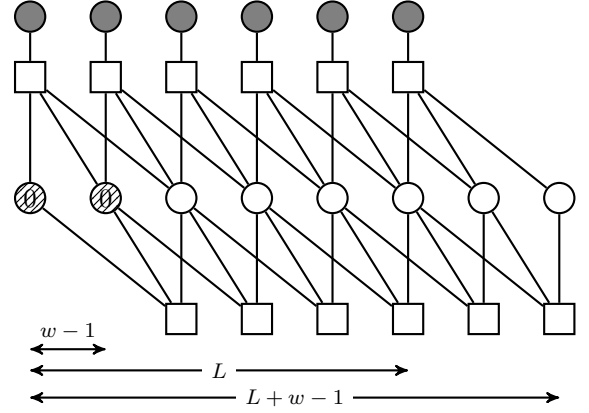


Fig. 2. Illustration of edge connections in a spatially-coupled compound LDGM/LDPC code. LDPC bit-nodes in the first  $w - 1$  groups are set to 0.

The encoder transmits the vector  $X^n = \hat{Z}^n \oplus Z^n$ . It is important to note that this choice of  $X^n$  has an average weight of  $p$ . The output at the decoder is given by  $Y^n = X^n \oplus Z^n \oplus W^n = \hat{Z}^n \oplus W^n$ , where the channel  $W^n$  is distributed iid according to  $\text{Ber}(\delta)$ . Since  $\mathcal{C}$  is a good channel code with rate  $1 - h(\delta) - \varepsilon/2$ ,  $\hat{Z}^n$  can be reliably decoded at the receiver. Now, since  $\hat{Z}^n$  belongs to a unique coset  $\mathcal{C}(m^k)$ , the message  $m^k$  can be recovered reliably. Thus, any rate  $R = k/n < h(p) - h(\delta)$  is achievable.

### III. SPATIALLY-COUPLED CODES AND MESSAGE-PASSING

In this section, we describe coding schemes for the side-information problems based on spatially-coupled compound LDGM/LDPC codes. The encoding and decoding are based on practically implementable, polynomial time message-passing algorithms. In essence, we describe new codebooks  $\mathcal{SC}$ ,  $\mathcal{SC}(s^k)$ ,  $\mathcal{SC}'$ ,  $\mathcal{SC}'(s^k)$ , analogous to  $\mathcal{C}$ ,  $\mathcal{C}(s^k)$ ,  $\mathcal{C}'$ ,  $\mathcal{C}'(s^k)$ , that appear to be good for rate distortion and/or channel coding when encoding and decoding is done with message-passing algorithms [13], [12], [19], [20]. Given that these are good for rate distortion and/or channel coding, the coding scheme for the side-information problems is same as in the previous section. We begin with the construction of spatially-coupled compound codes.

#### A. Spatially-Coupled Compound LDGM/LDPC Codes

For convenience, we assume regular degrees in both LDGM and LDPC components of the compound code. The construction can be easily generalized to irregular degrees. Suppose  $d_c$  and  $d_v$  denote the degrees of the check- and bit-nodes, respectively, in the LDGM component. Similarly, define  $d'_v$  and  $d'_c$  for the LDPC component. See Fig. 1 for an illustration. Consider  $L + w - 1$  groups of LDPC bit-nodes at positions  $\mathcal{N}_v = \{1, 2, \dots, L + w - 1\}$ ,  $L$  groups of LDGM check-nodes at positions  $\mathcal{N}_c = \{1, \dots, L\}$ ,  $L$  groups of LDPC check-nodes at positions  $\mathcal{N}'_c = \{w, 2, \dots, L + w - 1\}$ . See Fig. 2.

Choose  $N$  large enough so that  $Nd_v/d_c, Nd_v/w, Nd'_v/d'_c, Nd'_v/w \in \mathbb{N}$ . We first show the coupling structure in the LDGM part. At each position  $i \in \mathcal{N}_v$ , place  $N$  LDPC bit-nodes each with  $d_v$  edge sockets. Similarly, at each position  $j \in \mathcal{N}_c$ , place  $Nd_v/d_c$  LDGM check-nodes each with  $d_c$

---

**Algorithm 1** Belief-Propagation Guided Decimation

---

**Require:** Sequence  $x^n \in \{0, 1\}^n$  to encode, parameters  $(T, \beta)$ , graph  $G(V, U, C)$ .

Set  $m_{i \rightarrow a} = \hat{m}_{a \rightarrow i} = 0$  for  $i \in V \cup U$ ,  $a \in C$  and  $(i, a) \in G$ .

Initialize  $V_{\text{dec}}$  to be LDPC bit-nodes in first  $w - 1$  sections.

**while**  $V_{\text{dec}} \neq V$  **do**

**for**  $t = 1$  to  $T$  **do**

$m_{i \rightarrow a} = (-1)^{x_i} \tanh(\beta)$  for  $i \in U$  and  $a \in C$ .

$m_{i \rightarrow a} = (-1)^{u_i} \cdot \infty$  for  $i \in V_{\text{dec}}$  and  $a \in C$ .

$m_{i \rightarrow a} = \sum_{b \in \partial i \setminus \{a\}} \hat{m}_{b \rightarrow i}$  for  $i \in V \setminus V_{\text{dec}}$  and  $a \in C$ .

$\hat{m}_{a \rightarrow i} = \tanh^{-1} \prod_{j \in \partial a \setminus \{i\}} \tanh m_{j \rightarrow a}$  for  $i \in V \setminus V_{\text{dec}}$

    and  $a \in C$ .

**end for**

Evaluate  $m_i = \sum_{a \in \partial i} \hat{m}_{a \rightarrow i}$  for all  $i \in V \setminus V_{\text{dec}}$ .

Set  $B$  to be max. of  $|m_i|$  when  $i$  varies over left-most  $w$  sections of  $V \setminus V_{\text{dec}}$ ; denote the resulting bit-node by  $i^*$ .

**if**  $B = 0$  **then**

  Pick a bit-node  $i^*$  uniformly in left-most  $w$  sections of  $V \setminus V_{\text{dec}}$  and set  $u_{i^*}$  to be 0 or 1 uniformly randomly.

**else**

  Set  $u_{i^*}$  to 0 or 1 with prob.  $\frac{1 + \tanh m_{i^*}}{2}$  or  $\frac{1 - \tanh m_{i^*}}{2}$ .

**end if**

Set  $V_{\text{dec}} = V_{\text{dec}} \cup \{i^*\}$ .

**end while**

If  $\{u_i\}$  fail to satisfy LDPC checks, then **re-encode**.

---

edge sockets. At each bit- and check-node group, partition the  $Nd_v$  edge sockets into  $w$  groups using a uniform random permutation, and denote these partitions, respectively, by  $\mathcal{N}_{i,k}^v$ ,  $\mathcal{N}_{j,k}^c$  where  $1 \leq i \leq L + w - 1$ ,  $1 \leq j \leq L$ ,  $1 \leq k \leq w$ . The coupled LDGM component is constructed by connecting the sockets in  $\mathcal{N}_{j,k}^c$  to sockets in  $\mathcal{N}_{j+k-1,k}^v$ . The coupling structure in the LDPC part can be constructed similarly. These connections are depicted in Fig. 2 for  $L = 6$  and  $w = 3$ .

We note that such a construction leaves some edge sockets of the LDPC bit-nodes at the boundary unconnected. We also shorten the LDPC bit-nodes in the first  $w - 1$  groups to 0. This shortening and the unconnected edge sockets at the boundary are necessary for the spatial-coupling phenomenon to take effect [21]. It is implicit that the LDPC check-nodes at each position in  $\mathcal{N}_c^c$  are partitioned into two groups  $\mathcal{P}_1$  and  $\mathcal{P}_2$  so as to have the desired design rate. To avoid small error floors when decoding, each LDPC bit-node should have some connections to the checks in  $\mathcal{P}_2$ . The characterization of the codebooks  $\mathcal{SC}$ ,  $\mathcal{SC}(s^k)$ ,  $\mathcal{SC}'$ ,  $\mathcal{SC}'(s^k)$  follows implicitly.

### B. Message-Passing Algorithms for Encoding

Message-passing algorithms for channel coding are now a standard part of the coding theory literature. We refer the reader to [22] for their description. It has been recently shown [19], [20] that SC compound LDGM/LDPC codes are good for channel coding under message-passing. As such,  $\mathcal{SC}$ ,  $\mathcal{SC}(s^k)$ ,  $\mathcal{SC}'(s^k)$  are good for channel coding under message-passing. In the following, we focus exclusively on the message-passing

algorithm for rate distortion. In particular, we describe a variation of the so-called BPGD algorithm [13], [12].

Consider an instance of the spatially-coupled compound LDGM/LDPC described above. Denote its Tanner graph by  $G(V, U, C)$ , where  $V$  denotes the LDPC bit-nodes,  $U$  denotes the LDGM bit-nodes,  $C$  denotes the check-nodes (both LDGM and LDPC). Place a sequence  $x^n \in \{0, 1\}^n$  at the top of LDGM bit-nodes as in Fig. 1. The message-passing rules here are same as in the channel coding setup by assuming that  $x_i$  have come through a BSC channel (parameterized by  $\beta$ ). However, every  $T$  iterations, an LDPC bit-node is decimated (shortened based on the current LLR). This encoding procedure for the codebook  $\mathcal{C}(s^k)$  is described in Algorithm 1 assuming  $s^k = 0^k$ . For  $s^k \neq 0^k$ , the update for each check node in  $\mathcal{P}_1$  is modified to include the appropriate  $s_k$  hard-decision message. Also, the decimated sequence  $u^n$  may not satisfy all the LDPC check constraints. In this case, successive encoding attempts often result in a valid codeword due to the randomization in Algorithm 1. In practice, we found that removing double edges and 4-cycles from the code essentially eliminated this problem at moderate block lengths. For example, see the results in Table I.

One variation in Algorithm 1 from [13] is the choice of the LDPC bit-node for decimation. In [13], bit-node with maximum bias over entire graph is selected, but we restrict the search for maximum biased bit to only left-most  $w$  sections of the bits that are not already decimated. We observed that this change increases the chances of encoding to a valid codeword.

**Remark 3.** *The BPGD algorithm, when applied to the uncoupled compound LDGM/LDPC code, always failed to satisfy the LDPC check constraints and the spatial-coupling structure is required to overcome this problem. Thus, spatial coupling in compound codes not only helps to reduce the distortion, but allows the BPGD algorithm to encode to a valid codeword.*

It is also possible to create a coupling structure in a circular fashion and decimate the bit-nodes in a fixed section as done in [12]. However, this leads to a *wave* of decimations from both ends and will result in a failure to encode to a codeword, as the constraints imposed by the LDPC check-nodes are unlikely to match when the two waves meet. From a physics point of view, this is akin to growing a crystal on a torus from a single seed. When the two growth interfaces meet, they are very unlikely to mesh nicely and form a pure crystal.

## IV. NUMERICAL RESULTS

Table I shows the number of attempts required to encode 50 source sequences in  $\mathcal{C}(s^k)$ , with  $s^k = 0^k$ , over different block lengths. For example, at a block length of 9000, 21 sequences encoded in the first attempt and 9 sequences did not encode in four attempts. Without removing 4-cycles, only 5 sequences encoded at first and 35 failed after 4 attempts. At a block length of 81000, all 50 sequences encoded in the first attempt.

Tables II and III provide the simulation results with SC compound codes with message-passing algorithms. In the LDGM part of the code, 1% of the check-nodes have degree

1. This is to ensure that the decoding gets started for the channel coding problem. Encoding for the WZ problem is relatively easy, since this is performed using the codebook  $\mathcal{SC}'$ , which does not have LDPC check constraints. The optimal thresholds are calculated based on the rate of the code. For example, if a code of rate  $R$  is used to encode a  $\text{Ber}(\frac{1}{2})$ , then  $D_* = h^{-1}(1 - R)$ . The reported distortion and the optimal thresholds correspond to the saturated section of the SC system unaffected by the boundary condition.

For the GP problem, half of all the LDPC check-nodes are chosen randomly to carry message bits, except in the  $w - 1$  sections at the boundary where no message bits are placed. Since the bit-nodes in the left-most  $w - 1$  sections are initialized to 0, the LDPC check nodes cannot carry any messages in the left-most  $w - 1$  sections. Also, since the channel code for the GP problem uses fewer parity checks for the right boundary LDPC bit-nodes, there is a small error floor if we place message bits in the right-most  $w - 1$  sections. To avoid this floor, we do not place message bits in these sections and instead incur a small rate loss.

TABLE I

NO. OF ATTEMPTS FOR SUCC. ENCODING FOR 50 CODEWORDS.  
 $(d_v, d_c, d'_v, d'_c) = (6, 3, 3, 6)$ ,  $(L, w) = (15, 3)$ ,  $(\beta, T) = (0.65, 10)$ .

Block length ( $n$ )	4-cycles	$1/2/3/4/\geq 5$
9000	yes	5/3/5/2/35
9000	no	21/12/5/3/9
27000	no	35/15/0/0/0
45000	no	40/9/0/0/1
63000	no	44/6/0/0/0
81000	no	50/0/0/0/0

TABLE II

THRESHOLDS FOR WYNER-ZIV PROBLEM WITH  
 $(n \approx 140000, \beta = 1.04, T = 10)$ .

LDGM ( $d_v, d_c$ )	LDPC ( $d'_v, d'_c$ )	( $L, w$ )	( $D_*, \delta_*$ )	( $D, \delta$ )
(6, 3)	(3, 6)	(20, 4)	(0.111, 0.134)	(0.1174, 0.122)
(8, 4)	(3, 6)	(20, 4)	(0.111, 0.134)	(0.1149, 0.120)
(10, 5)	(3, 6)	(20, 4)	(0.111, 0.134)	(0.1139, 0.122)

TABLE III

THRESHOLDS FOR GELFAND-PINSKER PROBLEM WITH  
 $(n \approx 140000, \beta = 0.65, T = 10)$ .

LDGM ( $d_v, d_c$ )	LDPC ( $d'_v, d'_c$ )	( $L, w$ )	( $p_*, \delta_*$ )	( $p, \delta$ )
(6, 3)	(3, 6)	(20, 4)	(0.215, 0.157)	(0.220, 0.152)
(8, 4)	(3, 6)	(20, 4)	(0.215, 0.157)	(0.223, 0.151)
(10, 5)	(3, 6)	(20, 4)	(0.215, 0.157)	(0.220, 0.151)

## V. CONCLUSION

We have considered SC compound LDGM/LDPC codes that achieve the rate regions of the WZ and GP side-information problems. The decoding and encoding is based on message-passing algorithms. The encoding is performed using BPGD algorithm. For the WZ problem, LDGM codes were sufficient for the encoding to achieve the entire rate region. For the GP problem, we require compound codes and the encoding here is more complicated since the LDPC bit-nodes need to satisfy additional constraints. The structure enforced by

spatial-coupling seems to be crucial for the BPGD algorithm to encode to a codeword in compound codes. Numerical evidence is provided that shows that the SC compound codes can approach the desired rate regions.

## REFERENCES

- [1] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, vol. 22, no. 1, pp. 1–10, 1976.
- [2] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Problems of Inform. Transm.*, vol. 9, pp. 19–31, 1980.
- [3] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley Series in Telecommunications, Wiley, 2nd. ed., 2006.
- [4] Y. Yang, S. Cheng, Z. Xiong, and W. Zhao, "Wyner-Ziv coding based on TCQ and LDPC codes," in *Proc. Asilomar Conf. on Signals, Systems & Computers*, vol. 1, pp. 825–829, 2003.
- [5] A. Liveris, Z. Xiong, and C. Georgiades, "Nested convolutional/turbo codes for the binary Wyner-Ziv problem," in *Proc. Intl. Conf. on Image Proc.*, vol. 1, pp. 1–601, 2003.
- [6] U. Erez and S. ten Brink, "A close-to-capacity dirty paper coding scheme," *IEEE Trans. Inform. Theory*, vol. 51, no. 10, pp. 3417–3432, 2005.
- [7] Y. Sun, Y. Yang, A. Liveris, V. Stankovic, and Z. Xiong, "Near-capacity dirty-paper code design: A source-channel coding approach," *IEEE Trans. Inform. Theory*, vol. 55, no. 7, pp. 3013–3031, 2009.
- [8] T. Murayama, "Thouless-Anderson-Palmer approach for lossy compression," *Physical Review E*, vol. 69, no. 3, p. 035105, 2004.
- [9] P. Regalia, "A modified belief propagation algorithm for code word quantization," *IEEE Trans. Commun.*, vol. 57, no. 12, pp. 3513–3517, 2009.
- [10] S. Ciliberti, M. Mézard, and R. Zecchina, "Lossy data compression with random gates," *Physical review letters*, vol. 95, no. 3, p. 38701, 2005.
- [11] M. Wainwright, E. Maneva, and E. Martinian, "Lossy source compression using low-density generator matrix codes: Analysis and algorithms," *IEEE Trans. Inform. Theory*, vol. 56, no. 3, pp. 1351–1368, 2010.
- [12] V. Aref, N. Macris, R. Urbanke, and M. Vuffray, "Lossy source coding via spatially coupled LDGM ensembles," in *Proc. IEEE Int. Symp. Inform. Theory*, (Cambridge, MA, USA), pp. 373–377, July 2012.
- [13] V. Aref, N. Macris, and M. Vuffray, "Approaching the rate-distortion limit with spatial coupling, belief propagation and decimation." [Online]. Available: <http://arxiv.org/abs/1307.5210>, July 2013.
- [14] M. J. Wainwright and E. Martinian, "Low-density graph codes that are optimal for binning and coding with side information," *IEEE Trans. Inform. Theory*, vol. 55, pp. 1061–1079, March 2009.
- [15] A. Montanari, F. Ricci-Tersenghi, and G. Semerjian, "Solving constraint satisfaction problems through belief propagation-guided decimation," *arXiv preprint arXiv:0709.1667*, 2007.
- [16] S. B. Korada and R. L. Urbanke, "Polar codes are optimal for lossy source coding," *IEEE Trans. Inform. Theory*, vol. 56, no. 4, pp. 1751–1768, 2010.
- [17] C. H. Hsu and A. Anastasopoulos, "Capacity-achieving codes with bounded graphical complexity and maximum likelihood decoding," *IEEE Trans. Inform. Theory*, vol. 56, no. 3, pp. 992–1006, 2010.
- [18] D. J. C. MacKay and R. M. Neal, "Near Shannon limit performance of low density parity check codes," *Electronic Letters*, vol. 32, pp. 1645–1646, Aug. 1996.
- [19] N. Obata, Y.-Y. Jian, K. Kasai, and H. D. Pfister, "Spatially-coupled multi-edge type LDPC codes with bounded degrees that achieve capacity on the BEC under BP decoding," in *Proc. IEEE Int. Symp. Inform. Theory*, pp. 2433–2437, July 2013.
- [20] K. Kasai and K. Sakaniwa, "Spatially-coupled MacKay-Neal codes and Hsu-Anastasopoulos codes." submitted to *IEICE* [Online]. Available: <http://arxiv.org/abs/1102.4612>, 2011.
- [21] S. Kudekar, T. J. Richardson, and R. L. Urbanke, "Threshold saturation via spatial coupling: Why convolutional LDPC ensembles perform so well over the BEC," *IEEE Trans. Inform. Theory*, vol. 57, pp. 803–834, Feb. 2011.
- [22] T. J. Richardson and R. L. Urbanke, *Modern Coding Theory*. New York, NY: Cambridge University Press, 2008.