Optimizing Parameters in Soft-hard BPGD for Lossy Source Coding

Masoumeh Alinia and David G. M. Mitchell
Klipsch School of Electrical and Computer Engineering, New Mexico State University, Las Cruces, NM 88003
e-mail: {aliniama,dgmm}@nmsu.edu

2023 12th International Symposium on Topics in Coding (ISTC) | 979-8-3503-2611-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISTC57237.2023.10273545

***Abstract*—In this paper, we investigate lossy source coding based on a soft-decision based belief propagation guided decimation (BPGD) algorithm for low-density generator matrix (LDGM) codes, called soft-hard BPGD. For this algorithm, the optimal performance can only be achieved by carefully selecting the softness parameters which are conventionally found by exhaustive empirical search. To address this issue, we have introduced a framework that leverages the cavity method to predict the values of softness parameters at which phase transition occurs. This framework can significantly reduce the need for time-consuming empirical searches to determine the best parameter values for soft-hard BPGD. Our approach has been found to deliver superior rate-distortion performance compared to the hard-decimated BPGD algorithm and faster implementation than the soft-decimated algorithm.**

I. Introduction

Lossy source coding is used in various applications such as data compression [1], information hiding [2], and data embedding [3]. Linear codes have been proven to be an effective method to approach Shannon's rate-distortion (RD) bound for binary sources under Hamming distortion measure [4]–[6]. However, it is important to identify encoding schemes that are both low-complexity and capable of attaining the RD limit. Unfortunately, applying a conventional belief propagation (BP) algorithm cannot achieve good performance for lossy source coding due to unreliable marginal approximation in quantization. Consequently, BP algorithms are modified to include a decimation step proposed in terms of hard-decimation [7] or soft-decimation [8] algorithms. The idea of decimation is to sequentially fix bits and allow the BP encoder to converge. Soft-decimation based algorithms for low-density generator matrix (LDGM) codes, such as [8], [9], achieve linear complexity in code block length by substituting the hard decimation step with a soft indicator function. However, the performance of the algorithm depends heavily on the optimal parameter selection for the indicator function, which controls the distortion.

The lossy compression encoding problem in certain scenarios can be viewed as the study of the zero temperature limit of the Gibbs measure associated with a Hamiltonian [11]. The Gibbs measure can be considered as the posterior measure of the dual test-channel problem and it was shown in [6] that the inverse temperature is half the log-likelihood parameter of a binary symmetric channel (BSC) with crossover probability p, e.g., $\frac{1}{2}\ln\frac{1-p}{p}$, for a spatially coupled LDGM code ensemble. Spin glass theory in the cavity method predicts the dynamical and condensation phase transition thresholds for the Gibbs measure which, in turn, informs a number of values that can be related to the test-channel noise thresholds for channel coding.

In this paper, we apply the cavity method to soft-hard BPGD [9] and use the phase transition of the Gibbs measure at a specific temperature to tune the softness parameters of the algorithm. Determining a dynamic temperature threshold enables us to compute a corresponding threshold value for the algorithm parameters under which good RD performance can be obtained. The ability to compute this threshold value can reduce or eliminate the need for expensive numerical search techniques based on computer simulation [8], [9]. In fact, the possible solution space for finding the optimal parameter value is reduced to the point where the phases of the Gibbs measure change. We show that the soft-hard BPGD algorithm yields good RD performance using the threshold values, outperforming the (hard-decimation) BPGD algorithm.

II. LDGM code ensembles for lossy source coding

In this paper, we quantize a source sequence $\mathbf{s} = (s_1, s_2, \ldots, s_N)$ consisting of N independent and identically distributed Bernoulli $(p = 1/2)$ random variables to the nearest codeword $\mathbf{w} = (w_1, w_2, \ldots, w_M)$ from an LDGM code with a compression rate of $R = \frac{M}{N}$. The codeword $\mathbf{w}$ is used to reconstruct the source sequence as $\hat{\mathbf{s}}$, where the mapping $\mathbf{w} \rightarrow \hat{\mathbf{s}}(\mathbf{w})$ depends on the LDGM code. For the binary quantization problem, we use the Hamming metric $d(\mathbf{s}, \hat{\mathbf{s}}) = \frac{1}{N}\sum_{i=1}^{N}|\mathbf{s}_i - \hat{\mathbf{s}}_i|$ for calculating the distortion. In lossy source coding, the final goal is to minimize the average distortion $D = \mathbb{E}[d(\mathbf{s}, \hat{\mathbf{s}})]$, where $\mathbb{E}[\cdot]$ is the expectation taken over all possible source sequences $\mathbf{s}$. The rate-distortion function is in the form $R(D) = 1 - H(D)$ for $D \in [0, 0.5]$ and 0 otherwise, where H is the binary entropy function.

LDGM codes, as duals of LDPC codes, can be represented by generator matrix $\mathbf{G} \in \{0,1\}^{N\times M}$.We define the factor graph of this code as $\mathcal{G} = (V, C, E)$ where $V = \{1, \ldots, M\}$, $C = \{1, \ldots, N\}$, and $E = \{\ldots, (a, i), \ldots\}$ denote the code bit nodes, the check nodes and the edges connecting them, respectively. The vector (a, i) denotes an edge between check node a and code bit i, which occurs iff $\mathbf{G}_{a,i} = 1$. We will use indices $a, b, c \in C$ to denote check nodes and indices $i, j, k \in V$ to denote information bits. We define the sets $C(i) = \{a \in C \mid (a, i) \in E\}$, $V(a) = \{i \in V \mid (a, i) \in E\}$.

In this paper, we follow the construction of [6], where each edge emanating from a regular check node with degree k is connected uniformly at random to one of the bit nodes. The degree of bit nodes is a random variable with Binomial distribution $Bi(kN, 1/M)$. In the asymptotic regime of large N, M, the bit node degrees have i.i.d. Poisson distribution with an average degree k/R. For an LDGM code $\mathcal{C}$, defined by the generator matrix $\mathbf{G}$, and for a codeword $\mathbf{w}$, the reconstructed source sequence is given by $\hat{\mathbf{s}} = \mathbf{G}\mathbf{w}$.

III. Soft-Hard BPGD

The algorithm used in this paper is a combination of both hard and soft decimation. Following [9], our algorithm uses soft decimation equations like [8], but after each iteration it searches for a bit node with maximum bias value. This bit node is then fixed in all future steps, and the graph is reduced with the hard decimation process. The soft-decimation BP algorithm equations are updated as follows

$$R_i^{(t+1)} = \sum_{a \in C(i)} \widehat{R}_{a\to i}^{(t)}, \quad R_{i\to a}^{(t+1)} = \sum_{b \in C(i)\backslash a} \widehat{R}_{b\to i}^{(t)} + \frac{1}{\mu} R_i^{(t)}, \tag{1}$$

$$\widehat{R}_{a\to i}^{(t+1)} = 2(-1)^{s_a+1} \tanh^{-1}\left(\beta \prod_{j \in V(a)\backslash i} B_{j\to a}^{(t)}\right), \tag{2}$$

$$B_i^{(t)} = \tanh\left(\frac{R_i^{(t)}}{2}\right), \quad B_{i\to a}^{(t)} = \tanh\left(\frac{R_{i\to a}^{(t)}}{2}\right), \tag{3}$$

where $R_{i\to a}^{(t)}, \widehat{R}_{a\to i}^{(t)}$, and $B_{i\to a}^{(t)}$ denote the message sent from code node i to check node node a, the message sent from check node a to code node i, and the bias associated with $R_{i\to a}^{(t)}$ at iteration t, respectively; $R_i^{(t)}$ and $B_i^{(t)}$ denote the likelihood ratio of code bit i and the bias associated with $R_i^{(t)}$, respectively; and $\beta = \tanh(\gamma)$ and μ are non-negative parameters. The γ parameter reflects the effort of the message-passing algorithm to find the resulting codeword $\hat{\mathbf{s}} = \mathbf{G}\mathbf{w}$ as close to $\mathbf{s}$ as possible. The larger the γ, the stronger is the effort. On the other hand, the structure of the code imposes a limit on how strong this effort can be. The term $\frac{1}{\mu} R_{i\to a}^{(t)}$ is added to the plain BP equation to make the decimation softer. The soft indicator function $I_S(B_{i\to a}^{(t)}) = \frac{2}{\mu}\tanh^{-1}(B_{i\to a}^{(t)}) = \frac{1}{\mu} R_{i\to a}^{(t)}$ approximates the hard-indicator function [9], given as

$$I_H\left(B_{i\to a}^{(t)}\right) = \begin{cases} -\infty, & B_{i\to a}^{(t)} = -1, \\ 0, & -1 < B_{i\to a}^{(t)} < 1, \\ +\infty, & B_{i\to a}^{(t)} = 1, \end{cases}$$

where μ controls the softness of the approximation and is called the *softness parameter*. Algorithm 1 describes the procedure, where t indicates the iteration number, $\mathcal{G}^{(t)}$ is the LDGM code graph at iteration t, and w_i represents the binary value assigned to code node i. The initial information to check node node messages, $R_{i\to a}^{(0)}$, are set to ± 0.1 with $\mathbb{P}\left(R_{i\to a}^{(0)} = 0.1\right) = 0.5$, and reset to 0 at iteration 1.

In [8], there is no analytical way to tune the value of μ or β in order to give the best distortion performance. There, and in subsequent papers, e.g., [9], the best value of μ is found numerically by exhaustive and expensive code simulation. In the next section, we apply spin glass theory in the cavity method to carefully tune the softness parameters in the soft-hard BPGD algorithm to ensure good RD performance.

Algorithm 1 Soft-Hard Decimation Algorithm

Require: At iteration $t = 0$, initialize graph instance $\mathcal{G}^{(t=0)}$;
Generate a Bernoulli symmetric source word s;
while $V \neq \emptyset$ **do**
  Update $R_{i\to a}^{(t+1)}$ according to (1) for all $(a, i) \in E$;
  Update $\widehat{R}_{a\to i}^{(t)}$ according to (2) for all $(i, a) \in E$;
  Compute bias $B_{i\to a}^{(t)}$ and $B_i^{(t)}$ according to (3);
  Find $B^{(t)} = \max_i \left\{ \left| B_i^{(t)} \right| \mid \text{i not fixed} \right\}$;
  if $B^{(t)} > 0$ **then**
    $w_i \leftarrow '0'$;
  else
    $w_i \leftarrow '1'$;
  end if
  $\forall a \in C(i), s_a \leftarrow s_a \oplus w_i$ (update source);
  Reduce the graph $\mathcal{G} \leftarrow \mathcal{G}\backslash\{i\}$;
  $\mathcal{G}^{(t+1)} = \mathcal{G}^{(t)}\backslash\{i\}$ (remove code node i and all its edges);
end while

IV. Tuning parameters for soft-hard BPGD

In order to find the softness parameters, the solution space $\{0,1\}^{NR}$ of the equation $\hat{\mathbf{s}} = \mathbf{G}\mathbf{w}$ can be equipped with a conditional probability which is linked to spin glass theory in statistical mechanics. We consider the general class of Gibbs distribution of the form

$$P_\gamma(\mathbf{w} \mid \mathbf{s}) = \frac{1}{Z_\gamma} \prod_{a \in C} e^{-2\gamma N d(\mathbf{s}, \hat{\mathbf{s}}(\mathbf{w}))}. \tag{4}$$

The random Hamiltonian $2Nd\left(\mathbf{s}, \hat{\mathbf{s}}(\mathbf{w})\right)$ is a cost-function for assignments of variables $w_i \in \{0,1\}$; for which the different source bits and different graph ensembles of LDGM codes give different cost functions. The parameter γ is the inverse temperature, and the normalizing factor Z_γ is the partition function. From a statistical mechanics point of view, finding the most reliable code word $\mathbf{w}^* = \arg\max_{\mathbf{w}} P_\gamma\left(\mathbf{w} \mid \mathbf{s}\right)$ is to find the minimum energy configuration [6]. The minimum energy per node is equal to $2d_{N,\min}$, in which $d_{N,\min} = \min_{\mathbf{w}} d\left(\mathbf{s}, \hat{\mathbf{s}}(\mathbf{w})\right)$. The one-step replica symmetry breaking (1RSB) [10] in the cavity method gives the exact value for the internal energy, which allows the computation of the optimal distortion numerically by the population dynamics method.

Instead of estimating the optimal code word $\mathbf{w}^*$, the cavity method assumes that (4) can be decomposed into a convex superposition of measures

$$P_\gamma(\mathbf{w} \mid \mathbf{s}) = \sum_{\lambda=1}^{\mathcal{N}} \eta_\lambda P_{\gamma,\lambda}(\mathbf{w} \mid \mathbf{s}). \tag{5}$$

The summation of weights $\eta_\lambda = e^{-\gamma N(f_\lambda - f)}$ should be one, where f_λ is the free energy. Therefore

$$e^{-\gamma N f} \approx \sum_{\lambda=1}^{\mathcal{N}} e^{-\gamma N f_\lambda} \approx e^{-\gamma N \min_\varphi\left(\varphi - \gamma^{-1}\Sigma(\varphi;\gamma)\right)}, \tag{6}$$

where $e^{N\Sigma(\varphi;\gamma)}$ counts the number of extremal states P_γ with free energy $f_\lambda \approx \varphi$. The cavity method seeks two thresholds,

γ_d and γ_c, in which the nature of the decomposition (5) changes. For $\gamma < \gamma_d$ this measure is extremal and $\mathcal{N} = 1$. For $\gamma_d < \gamma < \gamma_c$, the measure is a convex superposition of an exponentially large number of extremal states. The exponent $\varphi - \gamma^{-1}\Sigma(\varphi;\gamma)$ in (6) is minimized at a value $\varphi_{\text{int}}(\gamma)$ such that $\Sigma(\varphi_{\text{int}}(\gamma);\gamma) > 0$. Then the complexity function

$$\Sigma(\gamma) \equiv \Sigma(\varphi_{\text{int}}(\gamma);\gamma) = \gamma(\varphi_{\text{int}}(\gamma) - f(\gamma)) \tag{7}$$

describes the growth rate of extremal states (5). The complexity function decreases as γ increases, becoming negative at γ_c, which is the point where it loses its meaning. At γ_d, there are no singularities in the free and internal energies, and their analytical expressions do not change in the range $0 < \gamma < \gamma_c$. The transition at γ_d is dynamic, and Markov chain Monte Carlo algorithms have an equilibration time that diverges when γ approaches γ_d. The cavity method uses the population dynamic approach to predict phase transitions by solving fixed point integral equations [10].

The goal of this paper is to apply the cavity method to soft-hard BPGD and thus the general framework and derivation will follow that of [6], which applied the method for hard-decision BPGD of spatially coupled LDGM codes; however, the BP formulation with a soft indicator function that we use will result in certain changes to the computation. We first derive the cavity equations needed to compute the complexity function (7) for our soft-hard BPGD algorithm. For BP equations (1) and (2), we need to replace $(-1)^{s_a} = J$ as well as express the bias in terms of variable to check node updates using (3), since the average field h_i in the cavity method must be defined as $R_{i\to a}^{(t)} = \gamma h_i$. To derive the fixed point equations, the cavity equations [10] are written for an average ensemble of the graph and source word, which involve messages on the edges of the graph. For the integral fixed point equations and their relations, the interested reader may refer to appendices B and C of [6].

We define two functions $\widehat{F}$ and F based on BP equations (1) and (2) as

$$\begin{aligned}&\widehat{F}(h_1,\ldots,h_{k-1} \mid J) = \frac{2J}{\gamma}\tanh^{-1}\Big(\frac{1-\beta}{1+\beta}\prod_{i=1}^{k-1}\tanh\frac{\gamma h_i}{2}\Big),\\ &F\left(\widehat{h}_1,\ldots,\widehat{h}_z\right) = \sum_{i=1}^{z}\Big(\frac{1}{\mu}+1\Big)\widehat{h}_i - \widehat{h}_z,\end{aligned} \tag{8}$$

where h_i, $\widehat{h}_i$ are average fields in the cavity messages $Q_{i\to a}$ and $\widehat{Q}_{a\to i}$, which are distributions on $\widehat{R}_{a\to i}$ and $R_{i\to a}$.The $Q_{i\to a}(R_{i\to a})$ and $\widehat{Q}_{a\to i}(\widehat{R}_{a\to i})$ can be considered as i.i.d. realization of the random variables $Q(r)$ and $Q(\hat{r})$. We note that the primary difference in these equations compared to the $g(\cdot)$ and $\hat{g}(\cdot)$ equations in [6] is the addition of the soft parameters β and μ. The probability distributions for the average fields $h_{i\to a}$ and $\widehat{h}_{a\to i}$ satisfy equations

$$q(h) = \sum_{z=0}^{\infty}P(z)\int\prod_{a=1}^{z}d\widehat{h}_a\widehat{q}(\widehat{h}_a)\times\delta(h - F(\widehat{h}_1,\ldots,\widehat{h}_z)), \tag{9}$$

$$\begin{aligned}&\widehat{q}(\widehat{h}) =\\ &\int\prod_{i=1}^{k-1}dh_i q(h_i)\times\frac{1}{2}\sum_{J=\pm1}\delta\left(\widehat{h} - \widehat{F}(h_1,\ldots,h_{k-1} \mid J)\right),\end{aligned} \tag{10}$$

where $P(z) = \frac{(k/R)^z}{z!}e^{-k/R}$. The conditional measure over spin variables $\sigma_1,\ldots,\sigma_{k-1}$ is

$$\begin{aligned}&\nu_1(\sigma_1,\ldots,\sigma_{k-1} \mid J\sigma, h_1,\ldots,h_{k-1})\\ &= \frac{1+\beta+J\sigma(1-\beta)\prod_{i=1}^{k-1}\sigma_i}{1+\beta+J\sigma(1-\beta)\prod_{i=1}^{k-1}\tanh\frac{\gamma h_i}{2}}\prod_{i=1}^{k-1}\frac{1+\sigma_i\tanh\frac{\gamma h_i}{2}}{2}.\end{aligned} \tag{11}$$

Finally, the equations for distributions $q^{\sigma}(r \mid h)$ and $\widehat{q}^{\sigma}(\widehat{r} \mid \widehat{h})$, for $\sigma = \pm1$, are

$$\begin{aligned}&q^{\sigma}(r \mid h)q(h) =\\ &\sum_{z=0}^{\infty}P(z)\int\prod_{a=1}^{z}d\widehat{h}_a\widehat{q}(\widehat{h}_a)\times\delta(h - F(\widehat{h}_1,\ldots,\widehat{h}_z))\\ &\times\int\prod_{a=1}^{z}d\widehat{r}_a\widehat{q}^{\sigma}\left(\widehat{r}_a \mid \widehat{h}_a\right)\times\delta\left(r - F(\widehat{r}_1,\ldots,\widehat{r}_z)\right),\end{aligned} \tag{12}$$

$$\begin{aligned}&\widehat{q}^{\sigma}(\widehat{r} \mid \widehat{h})\widehat{q}(\widehat{h}) = \int\prod_{i=1}^{k-1}dh_i q(h_i)\\ &\times\frac{1}{2}\sum_{J=\pm1}\sum_{\sigma_1,\ldots,\sigma_{k-1}=\pm1}\nu_1(\sigma_1,\ldots,\sigma_{k-1} \mid J\sigma, h_1,\ldots,h_{k-1})\\ &\times\delta\left(\widehat{h} - \widehat{F}(h_1,\ldots,h_{k-1} \mid J)\right)\\ &\times\int\prod_{i=1}^{k-1}dr_i q^{\sigma_i}(r_i \mid h_i)\,\delta\left(\widehat{r} - \widehat{F}(r_1,\ldots,r_{k-1} \mid J)\right).\end{aligned} \tag{13}$$

The equations (9), (10), (12), and (13) constitute the fixed point equations for six probability distributions that comprise the cavity method for our formulation of soft-hard BPGD. We are now ready to give the expression for the complexity (7) in terms of the above densities which will allow us to determine the threshold values for our softness parameter μ. This next formulation of the complexity follows the standard framework derived for hard-decimation BPGD in [6], with only minor differences according to the code parameters.

The Bethe free energy functional which approximates $\phi(\gamma)$ is given by

$$\begin{aligned}&-\gamma\phi_{\text{int}} =\\ &\ln(1+e^{-2\gamma}) + (R-1)\ln 2 - (k-1)\int\prod_{i=1}^{k}dh_i q(h_i)\\ &\times\frac{1}{2}\sum_{J=\pm1}\sum_{\sigma_1,\ldots,\sigma_k=\pm1}\nu_1(\sigma_1,\ldots,\sigma_k \mid J, h_1,\ldots,h_k)\\ &\times\int\prod_{i=1}^{k}dr_i q^{\sigma_i}(r_i \mid h_i)\ln Z_1(r_1,\ldots,r_k \mid J)\end{aligned}$$

$$+ R\sum_{z=0}^{\infty} P(z) \times \int \prod_{a=1}^{z} d\widehat{h}_a \widehat{q}(\widehat{h}_a) \sum_{\sigma} \nu_2(\sigma \mid \widehat{h}_1, \ldots, \widehat{h}_z)$$

$$\times \int \prod_{a=1}^{z} d\widehat{r}_a \widehat{q}^{\sigma}(\widehat{r}_a \mid \widehat{h}_a) \ln Z_2(\widehat{r}_1, \ldots, \widehat{r}_z), \tag{14}$$

where

$$Z_1(h_1, \ldots, h_k) = 1 + J\frac{1-\beta}{1+\beta}\prod_{i=1}^{k}\tanh\frac{\gamma h_i}{2},$$

$$Z_2(\widehat{h}_1, \ldots, \widehat{h}_k) = \frac{1}{2}\sum_{\sigma=\pm 1}\prod_{i=1}^{z}(1+\sigma\tanh\frac{\gamma h_i}{2}).$$

and

$$\nu_2(\sigma \mid \widehat{h}_1, \ldots, \widehat{h}_z) = \frac{\prod_{a=1}^{z}(1+\sigma\tanh\frac{\gamma\widehat{h}_a}{2})}{\prod_{a=1}^{z}(1+\tanh\frac{\gamma\widehat{h}_a}{2}) + \prod_{a=1}^{z}(1-\tanh\frac{\gamma\widehat{h}_a}{2})}. \tag{15}$$

We assume that for $\gamma < \gamma_c$, equations (9) and (10) have unique solutions $q(h) = \delta(h)$ and $\widehat{q}(\widehat{h}) = \delta(h)$. Note that the initial condition $q_i^{\sigma}(r \mid h) = \delta_{+\infty}(\sigma_i r_i)$ satisfies a symmetry $q^{\sigma}(r \mid 0) = q^{-\sigma}(-r \mid 0) = \delta_{+\infty}(r\sigma)$ (even for $h \neq 0$). Now, for $h = \widehat{h} = 0$ the iterations of (12) and (13) preserve this symmetry. In other words, the solutions of these equations (for $h = \widehat{h} = 0$) found from a symmetric initial condition satisfy $q^{\sigma=+1}(r \mid 0) = q^{\sigma=-1}(-r \mid 0)$ and $\widehat{q}^{\sigma=+1}(\widehat{r} \mid 0) = \widehat{q}^{\sigma=-1}(-\widehat{r} \mid 0)$. Therefore, we look only for symmetrical solutions and set $q^{+}(r) = q^{\sigma=+1}(r \mid 0)$ and $\widehat{q}^{+}(\widehat{r}) = \widehat{q}^{\sigma=+1}(\widehat{r} \mid 0)$. Then we have

$$q^{+}(r) = \sum_{r=0}^{\infty} P(z) \times \int \prod_{a=1}^{z} d\widehat{r}_a \hat{q}^{+}(\widehat{r}_a)\, \delta\left(r - F\left(\widehat{r}_1, \ldots, \widehat{r}_z\right)\right), \tag{16}$$

$$\widehat{q}^{+}(\widehat{r}) = \int \prod_{i=1}^{k-1} dr_i q^{+}(r_i) \times \sum_{J=\pm 1} \frac{1+\beta+J(1-\beta)}{2(1+\beta)} \delta\left(\widehat{r} - \widehat{F}\left(r_1, \ldots r_{k-1} \mid J\right)\right). \tag{17}$$

This simplifies the free energy (14) as

$$-\gamma f = \ln(1+e^{-2\gamma}) + (R-1)\ln 2. \tag{18}$$

To compute the complexity, the first line in ϕ_{int} in (14) omits expression (18) and we have

$$\Sigma(\gamma) = (k-1)\Sigma_e\left[q^{+}, \widehat{q}^{+}\right] - k\Sigma_v\left[\hat{q}^{+}\right] + R\Sigma_v\left[q^{+}\right],$$

where

$$\Sigma_v\left[q^{+}\right] = \int dr q^{+}(r)\ln(1+\tanh\frac{\gamma r}{2}),$$

$$\Sigma_e\left[q^{+}, \widehat{q}^{+}\right] = \int dr d\widehat{r} q^{+}(r)\widehat{q}^{+}(\widehat{r})\ln(1+\tanh\frac{\gamma r}{2}\tanh\frac{\gamma\widehat{r}}{2}).$$

The average distortion or internal energy at temperature γ is obtained by differentiating (18). In order to compute the solutions of the cavity equations (9), (10), (12), and (13) we will use the population dynamics algorithm [6] with some modification. We omit the algorithm here, due to limited space, but describe briefly the formulation. We consider two populations for code bits and check nodes; the total size of the populations for the check nodes and the bit nodes is n, and we use a maximum iteration $t_{\max}$. First, the equations (9) and (10) are solved by the algorithm, and then equations (12) and (13) are found. From the final populations obtained after $t_{\max}$ iterations, it is straightforward to compute the complexity and the thresholds γ_d.

In order to obtain parameter thresholds for soft-hard BPGD, we now derive the corresponding values $\beta_d = (1-e^{-2\gamma_d})/(1+e^{-2\gamma_d})$ and $\mu_d = e^{4\gamma_d}$, where $\xi_d = 1/\mu_d$. Since decimated BP can correctly sample the Gibbs-Boltzmann measure up to γ_d, we can see that the optimal value for γ can be found near the dynamic phase γ_d. Therefore, the best value of μ for soft-hard BPGD can be found in the interval $\left(0, e^{4\gamma_d}\right)$. This can either be used to reduce the search space for an optimal value of μ, or the thresholds can be used directly in the BP equations. These techniques significantly reduce the complexity of exhaustive search techniques based on code simulation, such as [8], which need to be re-run for each code/code ensemble.

V. Numerical Results

In this section, the results of various experiments of a C++ based implementation of Algorithm 1 are reported for different LDGM code ensembles. Codes were generated randomly with regular check nodes and irregular bit nodes degrees with a binomial distribution as described in Section II.

Fig. 1 demonstrates the impact of choosing the parameter ξ. The algorithm was run for 1000 iterations of a randomly drawn LDGM code with a constant check node degree 5, $R = 1/2$, and three different codeword lengths $N = 1000$, 10000, and 100000. The results were obtained by averaging over 100 trials. We observe that, as stronger constraints are enforced (increasing ξ), the distortion value initially decreases but then becomes flat beyond the optimal value of ξ, indicating the presence of a threshold. The proposed dynamic phase transition in the cavity method described in Section IV can be applied to determine $\gamma_d = 0.851$, which results in a corresponding value of $\xi_d = 0.0331$. We note that the computed threshold ξ_d lies in the flat region of the distortion curve and thus yields good performance, since the algorithm is not sensitive to the choice of ξ in this region. This choice therefore has the advantage of avoiding exhaustive searches over ξ, or could be used to reduce the search space. Moreover, we observe that the computed ξ_d is robust over multiple different source length realizations from the ensemble.[1] Hard BPGD [6] has a corresponding algorithmic parameter that can be optimized by the cavity method, which we denote β_A. In Fig. 2, we compare the average distortion obtained by varying ξ in the soft-hard BPGD algorithm and β_A in BPGD, where

[1]The optimal value of ξ (for minimum distortion) in the BPGD algorithm for fixed check node degrees is typically observed to be slightly lower than the dynamical temperature ξ_d. While the value of ξ_d can be calculated using cavity theory, it is not clear what theoretical principles affect the optimum value of ξ. The calculation of the optimal value for ξ is left to future work.

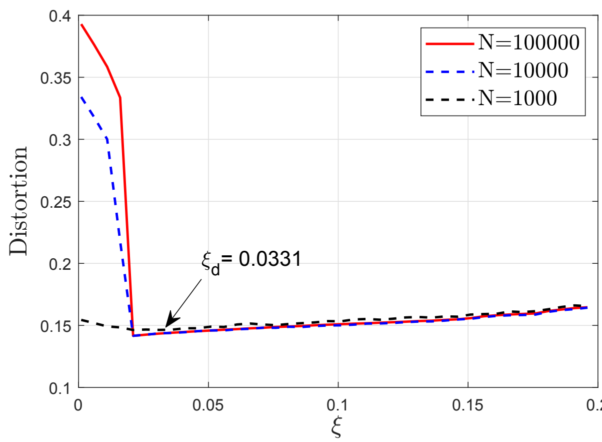

Fig. 1. BPGD distortion as a function of ξ for LDGM codes with $R = 1/2$.

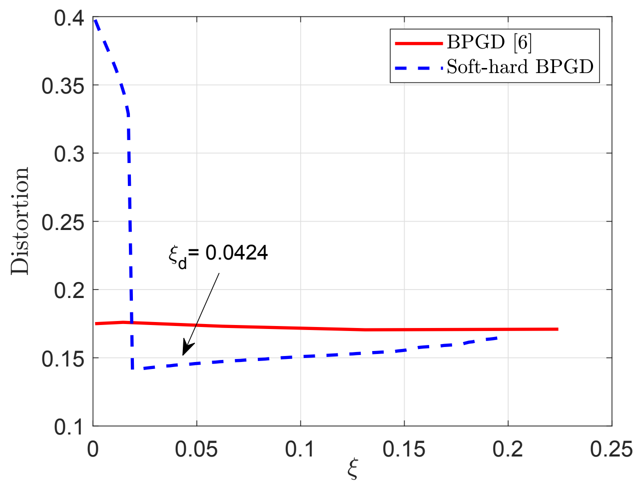

Fig. 2. Distortion comparison between BPGD and soft-hard BPGD

we set $\xi = e^{-4\beta_A}$. Both constructed LDGM codes have check node degree 5, rate $R = 1/2$, and length 128000. For this example, we compute the threshold $\xi_d = 0.0424$. We observe that, providing the softness parameter is seleceted carefully, the soft-hard BPGD outperforms BPGD; moreover, this happens at the threshold value. The soft-hard BPGD has an advantage in that it constrains the belief of the code bits in the direction of the current bit belief at each iteration rather than only constraining the beliefs at the decimation step in the hard BPGD. This allows for a refinement of the information bit beliefs at each iteration and improves the distortion results.

Finally, Fig. 3 displays a comparison of the average distortion obtained from the BPGD algorithm [6] and soft-hard BPGD algorithm as a function of R for long code lengths. The constructed codes have check node degrees $k = 3$, 4, or 5 and length $N = 128000$. The distortion is computed for fixed R, different ξ_d for different check node degrees, for 50 instances, and the empirical average is taken. We compare against the RD bound $D_{sh}(R)$, shown as the lowest bold black curve. We observe that, as expected, the distortion performance of codes encoded with soft-hard BPGD outperform those with hard BPGD algorithm where, for both algorithms, the performance worsens for increasing k. This behavior follows from the fact that the BPGD threshold worsens with increasing k, while the

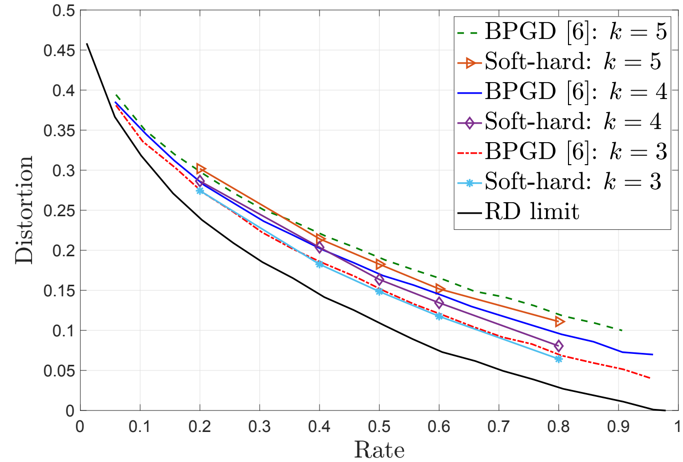

Fig. 3. Distortion performance over a range of rates for long codes with BPGD and soft-hard BPGD.

optimal encoding threshold improves to the Shannon limit [6].

VI. Conclusions

This paper proposes a framework based on the cavity method to compute a critical threshold for parameters in the soft-hard BPGD algorithm. By choosing parameters equal to, or close to the threshold value, superior rate-distortion performance is obtained when compared to the hard-decimated BPGD algorithm. Such an approach can reduce or eliminate costly computer search and simulation to optimize parameters while maintaining good rate-distortion performance.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. CCF-2145917.

References

[1] T. Berger and J. D. Gibson, "Lossy source coding," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2693-2723, 1998.

[2] P. Moulin and J. A. O'Sullivan, "Information-theoretic analysis of information hiding," *IEEE Trans. Inf. Theory*, vol. 49, no. 3, pp.563-593, 2003.

[3] R. J. Barron, B. Chen, and G. W. Wornell, "The duality between information embedding and source coding with side information and some applications," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1159-1180, 2003.

[4] M. J. Wainwright, E. Maneva and E. Martinian, "Lossy Source Compression Using Low-Density Generator Matrix Codes: Analysis and Algorithms," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1351-1368, 2010.

[5] S. Korada and R. Urbanke, "Polar Codes are Optimal for Lossy Source Coding," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1751-1768, 2010.

[6] V. Aref, N. Macris and M. Vuffray, "Approaching the Rate-Distortion Limit With Spatial Coupling, Belief Propagation, and Decimation," *IEEE Trans. Inf. Theory*, vol. 61, no. 7, pp. 3954-3979, 2015.

[7] T. Filler and J. J. Fridrich, "Binary quantization using belief propagation with decimation over factor graphs of LDGM codes," CoRR, vol.abs/0710.0192, 2007.

[8] D. Castanheira and A. Gameiro, "Lossy source coding using belief propagation and soft-decimation over LDGM codes," *IEEE Int. Symp. on Personal, Indoor and Mobile Radio Comm.*, pp. 431-436, 2010.

[9] A. Golmohammadi, D. G. M. Mitchell, J. Kliewer and D. J. Costello, "Encoding of spatially coupled LDGM codes for lossy source compression," *IEEE Trans. Comm.*, vol. 66, no. 11, pp. 5691-5703, 2018.

[10] A. Montanari, F. Ricci-Tersenghi, and G. Semerjian, "Clusters of solutions and replica symmetry breaking in random K-satisfiability," *J. Stat. Mechanics: Theory and Experiment*, vol. 2008, no. 4, p. P04004, 2008.

[11] A. G. Dimakis, M. J. Wainwright and K. Ramchandran, "Lower bounds on the rate-distortion function of LDGM codes," *2007 IEEE Information Theory Workshop*, pp. 650-655, 2007.