Design of Degree Distributions for LDPCA Codes

Feng Cen, Member, IEEE

Abstract—A degree distribution design method of Low-Density Parity-Check Accumulate (LDPCA) codes [6] is proposed for rate-compatible asymmetric Slepian-Wolf coding (SWC) of correlated binary memoryless sources. The degree distribution is designed by seeking the simultaneous optimum at the lowest and the highest compression rates of the specified range with density evolution algorithm. The experimental results show that the coding efficiency of codes constructed with the designed degree distribution is close to the Slepian-Wolf limit over the entire target range and better than all other reported results of the rate-compatible asymmetric SWC.

Index Terms—Rate-compatible codes, distributed source coding, Slepian-Wolf coding, LDPCA codes, LDPC codes.

I. INTRODUCTION

Rate-compatible SWC [1] has received a great deal of attention by researchers recently, since it can dramatically reduce the complexity of implementation and facilitate type-II hybrid automatic-repeat-request (ARQ) transmission. Two classes of powerful channel codes: turbo codes and LDPC codes, have been employed for rate-compatible SWC.

In spite of the superior coding efficiency of LDPC codes to the turbo codes in conventional channel coding and fixed rate SWC [4], most of practical rate-compatible SWC schemes are designed with turbo codes [2] [3], since the performance of LDPC codes in rate-compatible SWC is still inferior to turbo codes. Sartipi et al [5] and Varodayan et al [6] are those who have addressed the practical rate-compatible LDPC codes for SWC in the literature up to now, as far as we know. In [5], they addressed the symmetric rate-compatible SWC and the code design method. But, in their coding scheme, the codes have optimal coding efficiency only for the rate-compatible coding of the correlated sources with fixed crossover probability p. When the crossover probability between two correlated sources is changed, the codes designed in their paper have to be redesigned so as to achieve the optimal coding efficiency. The LDPCA coding scheme proposed in [6] is suitable for more general rate-compatible SWC. It can handle the variable p. Furthermore, the asymmetric coding scheme is capable to be extended to cover the same Slepian-Wolf rate region as the symmetric SWC with the time sharing method.

However, the underlying LDPC codes employed in [6] are optimally designed for conventional channel coding with fixed coding rate rather than for rate-compatible SWC. So, only a small range of achievable compression rates without rate-adaption are close to the Slepian-Wolf limit. Design of LDPC codes optimal for rate-compatible SWC is still an open issue.

In this letter, we study the problem of the degree distribution design of the LDPCA codes for rate-compatible asymmetric SWC. Finding the truly optimized degree distribution for a given range of compression rates is practically very difficult. Instead, assuming the degree distribution of the check nodes is in a concentrated fashion, we propose to optimize the degree distribution of the variable nodes simultaneously only at the lowest and the highest compression rates of the specified range. The experimental results show that the coding efficiency of the codes constructed with the degree distribution optimized with the proposed method has small gap with respect to the Slepian-Wolf limit over the entire target range of compression rates and is better than all other reported results of the rate-compatible asymmetric SWC.

II. LDPCA CODING SCHEME

Let \( X = \{x_1, x_2, \ldots \} \) and \( Y = \{y_1, y_2, \ldots \} \), \( x_j \in \{0, 1\} \) and \( y_j \in \{0, 1\} \) be any two i.i.d binary memoryless sources. They are correlated to each other with \( p = Pr[x_j \neq y_j] < 0.5 \). In the letter, \( Y \) is assumed to be taken as the side information in the asymmetric SWC.

The LDPCA encoder is the concatenation of a LDPC syndrome-former and an accumulator. The LDPC syndrome-former is defined by the degree distribution pair \( (\lambda(x), \rho(x)) \). Here, \( \lambda(x) = \sum_{j=2}^{\infty} \lambda_j x^{j-1} \) and \( \rho(x) = \sum_{j=2}^{\infty} \rho_j x^{j-1} \) specify the degree distributions of variable nodes and check nodes, respectively. More precisely, \( \lambda_j (\rho_j) \) is the fraction of edges emanated from variable nodes (check nodes) with degree \( j \). When encoding, the source bits \( \{x_1, x_2, \ldots, x_N\} \) of length \( N \) are first summed modulo 2 at the check nodes according to the bipartite graph of the LDPC syndrome-former, yielding syndrome bits \( \{s_1, s_2, \ldots, s_N\} \). These syndrome bits are in turn accumulated modulo 2, producing the accumulated syndrome \( \{a_1, a_2, \ldots, a_N\} \). Assuming \( M(\leq N) \) accumulated syndromes are transmitted to the decoder, the compression rate of \( X \) can be computed by

\[
R_x = M/N.
\]

The LDPCA decoder modifies the decoding graph to handle the received accumulated syndromes in a rate-compatible way. Assuming \( a_{i-1} \) and \( a_{i+1} \) are transmitted to the decoder and \( a_i \) is not sent to the decoder, the decoder will merge the \( i \)-th check node with the \((i + 1)\)-th check node of the original decoding graph to form the \( k \)-th check node of the new decoding graph. In other words, the edges connected to the \( k \)-th check node of the new decoding graph consist of all edges connected to the \( i \)-th check node and \((i + 1)\)-th check node of the original decoding graph. The syndrome at the \( k \)-th check node is obtained by

\[
s_k = a_{i+1} \oplus a_{i-1} = s_i \oplus s_{i+1}.
\]
Thus, the degree distributions of the LDPCA decoding graph are defined by the pair \((\lambda(x), \rho_d(x))\), where \(\rho_d(x)\) is the degree distribution of the check nodes and constrained by
\[
\int_0^1 \rho_d(x) \, dx = R_c \int_0^1 \rho(x) \, dx.
\]

III. Degree Distribution Design

To simplify the degree distribution design problem, we assume that the degree distribution of the check nodes is in a concentrated fashion, since it does not noticeably degrade the performance of LDPC codes in conventional channel coding [7]. According to the construction method of the bipartite graph proposed in [6], the degree distribution of the check nodes of the decoding graph can be written in concentrated fashion at any compression rates of a specified range, i.e.
\(
\rho_d(x) = \rho_d x^2 + (1 - \rho_d)x^{3+1},
\)
provided that it is in the concentrated fashion at the lowest compression rate of the range. For the concentrated degree distribution, \(\rho_d(x)\) can be determined by \(\lambda(x)\) and the compression rate \(R_c\), i.e.
\[
j = \left\lfloor \frac{1}{R_c} \sum_{i \geq 2} \frac{\lambda_i}{i} \right\rfloor, \quad \rho_d = (j^2 + j)R_c \sum_{i \geq 2} \frac{\lambda_i}{i} - j,
\]
where \(\lfloor \cdot \rfloor\) denotes the floor function. Based on the above assumption, the objective of the design of degree distribution pair becomes optimizing \(\lambda(x)\) such that for a specified range of compression rates \([R_{c,\text{min}}, R_{c,\text{max}}]\), the coding efficiencies are close to Slepian-Wolf limit. Obviously, it is quite hard to find the truly optimized \(\lambda(x)\), which should have optimal coding efficiencies at all the compression rates within the range \([R_{c,\text{min}}, R_{c,\text{max}}]\). Instead of finding such a true optimum, we propose to search for \(\lambda(x)\) having the maximum thresholds of admissible correlations between the source and the side information simultaneously for \(R_{c,\text{min}}\) and \(R_{c,\text{max}}\).

The density evolution (DE) algorithm [9] is utilized to estimate the decoding error probability. It is well known that in asymmetric SWC the correlation between \(X\) and \(Y\) can be modelled as a virtual binary symmetric channel (BSC) with the error probability equal to \(p\). Hence, the initial message density of the variable nodes for DE algorithm is determined by the observation of the virtual BSC. Provided that the maximum number of DE iterations \(m\) is fixed, the decoding error probability actually depending on \(\lambda(x)\), \(R_c\) and \(p\) can be written as \(P_e(\lambda(x), R_c, p)\). Assuming that \(p_{\text{min}}\) and \(p_{\text{max}}\) are the channel parameters of the virtual BSCs corresponding to the Slepian-Wolf limits when compression is performed at \(R_{c,\text{min}}\) and \(R_{c,\text{max}}\), respectively, we define the maximum of decoding error probabilities at \(R_{c,\text{min}}\) and \(R_{c,\text{max}}\) as
\[
P_{e,\text{max}} = \max(P_e(\lambda(x), R_{c,\text{min}}, p_{\text{min}} - \delta_1), P_e(\lambda(x), R_{c,\text{max}}, p_{\text{max}} - \delta_2)),
\]
where \(\delta_1\) and \(\delta_2\) are positive real values and indicate the gaps of the virtual BSCs with respect to Slepian-Wolf limits in terms of the channel parameter.

In practice, the optimization scheme adopted by us consists of two successive steps: global optimization and local optimization. In the global optimization, first, define small enough target \(\delta_1\) and \(\delta_2\), and then make use of means of differential evolution [8] to search for \(\lambda(x)\) having the minimum \(P_{e,\text{max}}\) after \(m\) iterations of DE. In order to guarantee that the coding efficiencies at \(R_{c,\text{min}}\) and \(R_{c,\text{max}}\) have similar small gaps with respect to Slepian-Wolf limits, the relationship between \(\delta_1\) and \(\delta_2\) should satisfy
\[
H(p_{\text{min}}) - H(p_{\text{min}} - \delta_1) = H(p_{\text{max}}) - H(p_{\text{max}} - \delta_2). \tag{7}
\]
In the local optimization, the hill-climbing approach with linear programming described in [9] is employed and the result of the global optimization is used as the initial degree distribution. For each basic step of hill-climbing, \(\lambda(x)\) is optimized at \(R_{c,\text{min}}\) and \(R_{c,\text{max}}\) by turns. When optimization is performed at one rate, the probability of decoding errors at the other rate should be guaranteed against increase. This is taken as an additional constraint in linear programming. After each basic step of the local optimization, the adopted new degree distribution should have either larger admissible channel parameters, i.e., one or both of \(\delta_1\) and \(\delta_2\) can take on smaller values, or at least smaller decoding error probability for one or both of the virtual BSCs after \(m\) iterations of DE.

IV. Experimental Results

The edge perspective degree distribution of the variable nodes optimally designed for \(R_c \in [0.175, 0.7]\) to assess the proposed method is as follows,
\[
\lambda(x) = 0.071112 + 0.238143x^2 + 0.182737x^3 + 0.073795x^9 + 0.079317x^{14} + 0.354896x^{32}. \tag{8}
\]
The moderate length of source blocks, \(N = 16434\), is utilized to evaluate the performance so as to compare with the best known results of punctured turbo codes [3]. The online available software [10] is employed for the construction, encoding and decoding of LDPCA codes. The degree distribution employed in [6] for irregular LDPCA codes (deg 2 to 21) is re-evaluated in our experiments for comparison as well. In all simulations, more than 2000 source blocks were simulated with the maximum decoding iterations 100. The reconstructed bit error rate (BER) is measured only within the information bits. Furthermore, in order to measure the BER, we assume that the decoder does not request the increment data to recovery the source bits perfectly in all simulations. The parameter \(p\) is assumed to be known at the decoder in our simulations.

Fig. 1 shows the bit-error rates (BER) of the LDPCA codes constructed with different \(\lambda(x)\)'s at different compression rates. The LDPCA codes constructed with the \(\lambda(x)\) defined by equation (8) have steep water-fall regions and very low error-floor levels within the given range of the compression rate. The steep water-fall region of the BER curve can improve the accuracy of the prediction of the smallest achievable compression rate of the source blocks at the encoder encoder side. The more accurate the prediction is, the fewer times the decoder asks for increment data in ARQ transmission, and
then the shorter decoding delay the receiver has. In contrast, the LDPCA codes constructed with the λ(x) used in [6] for irregular LDPCA codes (deg 2 to 21) have high error-floor levels at low compression rates, since it is optimized for conventional fixed-rate channel coding. Moreover, the high error-floor level often means that some decoding errors can not be detected with the syndrome.

The results shown in Fig. 2 are obtained under the pseudo error free condition, $BER < 10^{-6}$. In typical application model of the SWC, separately encoding and jointly decoding, the transmitter needs the help of the feedback from the receiver to decide whether or not to send the increment data. When the increment data are requested by the decoder, additional decoding time delay will be induced. The high decoding BER often implies that many times of requests will be made by the decoder so as to compress the data at the lowest achievable compression rate. Thus, the additional decoding time delay may be quite long and eventually make the system fail to work in real time. So, in many situations, especially for real time applications, it is better that the system works under some type of pseudo error free condition such that the additional decoding time delay will have unnoticed impact on the system performance. The $BER < 10^{-6}$ is a commonly used pseudo error free condition in many literatures such as [3] [4]. Thus, utilizing this pseudo error free condition can facilitate the performance comparison with other codes and satisfy the requirements of most of applications.

In Fig. 2, the performance of the LDPCA codes constructed with $\lambda(x)$ defined by equation (8) is measured for $R_x \in [9/66, 47/66]$. It can be observed that although $\lambda(x)$ is only optimized for the lowest compression rate and highest compression rate of the specified range, the gaps with respect to the SW limit are small over the entire measured range and smaller than the punctured turbo codes [3]. The measured gaps for the designed codes with respect to the Slepian-Wolf limit fluctuate between 0.05 and 0.07 bits. It should be noted that the results of LDPCA codes (deg 2 to 21) in Fig. 2 and in [6] are different. This is because the increment data are assumed to be transmitted to the decoder until the data can be decoded perfectly and very few source blocks are simulated in [6]. The best known results of fixed rate asymmetric SWC reported by Liveris et al [4] under the pseudo error free condition with syndrome based scheme is also presented in Fig. 2. Although, the block length used in [4] is much longer than our designed codes, the coding efficiency of theirs are very close.

Fig. 1. Comparison of the bit-error rates achieved by LDPCA codes constructed with the $\lambda(x)$ used as LDPCA codes (deg 2 to 21) in [6] (marked hollow star) and LDPCA codes defined by (8) (marked hollow triangles) for the compression rates $R_x = 17/66$, 0.5, 47/66.

Fig. 2. Performance comparison of LDPCA codes constructed with the $\lambda(x)$s used as LDPCA codes (deg 2 to 21) in [6], LDPCA codes defined by equation (8), fixed rate LDPC codes [4], punctured turbo codes [3] and Slepian-Wolf limit.

REFERENCES