Capacity-Approaching Joint Source-Channel Coding for Asymmetric Channels with Low-Density Parity-Check Codes

Cen Feng
The Department of Control Science & Engineering, Tongji University,
1239 Siping Road, Shanghai, 200092, P. R. China
E-mail: feng.cen@tongji.edu.cn

Received: 23 January 2014 / Accepted: 7 March 2014 / Published: 30 April 2014

Abstract: By only sending the parity bits, joint source-channel coding can be natively achieved with low-density parity-check codes. However, the code ensemble design of optimal low-density parity-check codes for joint source-channel coding over asymmetric communication channels is difficult. To circumvent such a difficulty, source-channel adaptors is proposed in this paper. By using the source-channel adaptors both the asymmetric communication channel and the asymmetric correlation channel can be converted to symmetric channel and the conventional design method for channel coding can be used straightforwardly. To demonstrate the effectiveness of the proposed scheme, a code ensemble for joint source-channel coding over asymmetric communication channel is designed for binary memoryless sources with a range of \(a\) priori probability. The experiment results show that the performance of the devised codes is very close to the theory limit.

Keywords: Low-density parity-check codes, Joint source-channel coding, Asymmetric channel, Channel adaptor, Asymmetric correlation.

1. Introduction

The separation theorem in information theory indicates that there is nothing to be gained from joint source-channel coding (JSCC) for stationary channels and sources given unbounded delay. However, if the communication system is heavily constrained in terms of complexity or delay, independently source and channel coding strategies can be largely suboptimal. Another disadvantage of the separate source and channel coding is that the errors, induced by channel noises in channel decoding, tend to be propagated in most fixed-to-variable length capacity approaching sources codes (e.g., Lempel-Ziv and arithmetic codes). The JSCC was therefore gained considerable attention as viable alternative for reliable communication across noisy channels. The advantage of the JSCC is due to its efficiently and flexibly exploiting the redundancy of the source to recover the signal from channel noises. Most of the developed JSCC schemes, so far, combine the source coding and channel coding by transforming the decoding process of variables length codes. A more detailed review can be found in [1]. However, the combination scheme cannot efficiently exploit the redundancy left in the source to improve the error correction capability.

The simple and efficient approach to deal with JSCC is directly employing the channel codes. The basic concept behind the application of channel codes
simultaneously for source coding and channel coding is based on Slepian-Wolf coding [2], which can be seen as a problem of channel coding with side information and is perhaps the best-known application of channel codes for source coding. The combination of Slepian-Wolf coding and channel coding is a problem of joint source-channel coding of correlated sources. If the side information is fixed, such as all-zero sequences, or known ahead, the problem reduces to JSCC. Syndrome based scheme and parity based scheme are two dominated SW coding schemes. The former follows from the scheme proposed by Wyner [3] and is straightforward to be implemented with low-density parity-check (LDPC) coset codes [4]. For the latter, systematic channel codes are generally utilized and the side information is regarded as the corrupted version of systematic bits that could be decoded upon receiving some parity bits formed by the encoder. Turbo codes are often adopted for the latter [5-7].

With channel code approach, the efficiency of JSCC is highly dependent of the performance of channel codes. Turbo codes and LDPC codes are two best-known capacity approaching channel codes. The application of turbo codes for JSCC has been extensively studied under the framework of parity based scheme. Aaron et al. are the pioneers exploiting the turbo codes for JSCC [6]. Their coding scheme has been widely applied in distributed video coding [8, 9].

But, to the best of our knowledge, there are few of literatures reporting the JSCC with LDPC codes due to the difficulty in recovering the syndrome corrupted by the channel noises with belief propagation algorithm for syndrome based scheme. Suppose \(X\) and \(Y\) are two dependent random signals and a LDPC code is defined by the parity matrix \(H\). In the syndrome based approach, \(X\) is encoded by \(C = XH\) and \(C\) is transmitted over a noisy channel. At the receiver side, the decoder observes \(\hat{C}\), representing the corrupted \(C\) by channel noises, and tries to recover \(X\) from \(\hat{C}\) with the help of the side information \(Y\). According to the belief propagation algorithm of coset codes, the most powerful decoding approach for syndrome based scheme [10], the syndrome bit is not include in the calculation of log \(a posteriori\) probability ratio and cannot be changed in message passing. Hence, the corrupted syndrome codeword \(\hat{C}\) cannot be recovered such that the syndrome based scheme is not suitable for JSCC.

To circumvent this problem, an IRA code, which is a subset of LDPC-like codes, based approach is proposed by [11] and a low-density generator matrix (LDGM) code based approach is proposed for JSCC in [12]. But, both codes have fixed or regular degree edge connections between parity variable nodes (variable nodes associated to parity bits) and check nodes in bipartite graph. The regular degree distribution between parity variable nodes and check nodes restricts the performance optimization of the codes. Both [11] and [12] did not show the capacity approaching capability for their schemes. Furthermore, they did not discuss the transmission over asymmetric communication channels.

In this paper, we study the parity based scheme with general LDPC codes, i.e. no limitation is applied to the degree distribution between parity variable nodes and check nodes, for JSCC coding even on asymmetric communication channels. The IRA and LDGM based JSCC scheme can be considered as the particular cases of general LDPC code based scheme.

We proposed source-channel adaptors to symmetrize both the communication channel and the correlation channel. By using the source-channel adaptors, the codeword averaging can be avoid in the density evolution for code design and massive research results on conventional LDPC codes for channel coding can be applied to the proposed scheme.

The rest of this paper is organized as follows. Section 2 outlines the basic coding scheme of the parity based scheme with LDPC codes, followed by the introduction of the source-channel adaptors and density evolution formulas for the proposed scheme in Section 3. Section 4 gives simulation results to demonstrate the effectiveness of the proposed scheme. Finally, Section 5 concludes the paper.

2. Basic Coding Scheme

Consider an independent and identically distributed (i.i.d) Bernoulli source \(X\) with success probability \(p = \Pr(X = 1) \leq \frac{1}{2}\) and the output sequence of the source is divided into length \(k\) blocks. Suppose that the LDPC codes are defined by a \(m \times n\) parity check matrix \(H_{mn}\) and the corresponding systematic generator matrix is given by \(G_{mn} = [P_{mn} \ I_n]\), each source block \(X_k = \{x_1, \ldots, x_k\}\) is encoded by

\[
C_{u} = X_{k}P_{mn}
\]

at the coding rate of \(R_c = \frac{n}{m}\). \(C_{u}\) is then transmitted through a communication channel and corrupted by channel noises. The channel outcome \(C_{u}\) is used to recover \(X_k\) at the decoder. The decoding can be accomplished with any conventional LDPC decoding algorithm based on the bipartite graph of \(H_{mn}\) as shown in Fig. 1. In our simulation, belief propagation algorithm [13] is employed.

Suppose that in Fig. 1, the first \(m\) variable nodes associated to parity check bits when \(H_{mn}\) is applied for conventional channel coding. By \(m_{vi}^{(i)}\) and \(m_{vi}^{(t)}\), we denote the messages sent from the \(v\)th variable node to the \(c\)th check node and from the \(c\)th check node to the \(v\)th variable node in the \(l\)th iteration.

204
under belief propagation, respectively. At the decoder, the messages are updated in the $l$th iteration by

$$m_{w}^{(l)} = \begin{cases} m_{w0}^{(l)} & l = 0, \nu = 1, \ldots, m \\ m_{w}^{(l)} + \sum_{c \in \mathcal{C}} m_{v}^{(l-1)} & l \geq 1 \end{cases}$$

(2)

$$m_{v}^{(l)} = \gamma^{l} \left( \sum_{c \in \mathcal{C}} r^{(l-1)} \right)$$

(3)

where $m_{w0} = \ln \frac{p_{x}(y|x=0)}{p_{x}(y|x=1)}$ and $m_{w} = \ln \frac{p_{x}(y|x=0)}{p_{x}(y|x=1)}$. Here, $P_{c,c}$ denotes the conditional probability density function.

The above JSCC scheme can achieve the theory limit rate with asymptotically vanishing decoding errors probability. This can be explained as follows.

The simple way of explanation is to construct a $\mathbf{G}_{cv}$ such that $\gamma^{l} \left( \sum_{c \in \mathcal{C}} r^{(l-1)} \right)$ approaches 0 as $l \to \infty$. Then, $S_{c} = X_{v} \mathbf{H}_{cv}$ at the rate approaching $H(X)$ with asymptotically vanishing decoding error probability [3]. Then, $S_{c}$ can be encoded with linear codes at the rate approaching the capacity of the communication channel with asymptotically vanishing decoding errors probability, i.e., $Z_{c} = S_{c} \mathbf{G}_{cv}$, where $\mathbf{G}_{cv}$ denotes the generator matrix of the linear block codes. If $\mathbf{P}_{cv}$ is defined by $\mathbf{P}_{cv} = \mathbf{H}_{cv} \mathbf{G}_{cv}$, the encoders of JSCC with the parity based scheme and the optimal separate coding approach become identical. According to the separation theorem in information theory, $C_{w}^{*}$ can achieve the theory limit rate with asymptotically vanishing decoding error probability.

Even the above JSCC scheme is a capacity achievable scheme, how to design the capacity approaching codes is still a problem, especially for the asymmetric communication channel.

It is easily observed that the above JSCC coding scheme is equivalent to conventional channel coding for a virtual channel consisting of two subchannels: a communication subchannel and a correlation subchannel. The input and the outcome of the correlation subchannel are the source sequence and the all-zeros sequence, respectively. Hence, the correlation subchannel can be viewed as a binary asymmetric channel with transition probability of $p(0|1) = 1$ and $p(1|0) = 0$. The communication subchannel is the channel actually transmitting $C_{w}$.

The difficulty for the design of capacity achievable codes rises from the asymmetric property of subchannels. It is well known that for asymmetric channel, the error probability of the density evolution depends on the codewords [13]. In the above JSCC scheme, the correlation channel is obviously asymmetric. When we transmit codewords on an asymmetric communication channel, two asymmetric subchannels will make the problem more complicated and conventional channel coding design method inapplicable. To simplify the code design process, we propose source-channel adaptors to symmetrize the virtual channel.

### 3. Source-Channel Adaptors

Density evolution is a powerful analysis tool for LDPC code design. In conventional channel coding, if the channel is symmetric, all-zeros codes can be assumed for density evolution to reduce the computational complexity of density evolution. While for the asymmetric channel, the density evolution is codeword dependent. Generally, either the codeword-averaged density evolution[14] or channel adaptor approach [15] has to be involved in code design for asymmetric channels. The codeword-averaged density evolution evaluates the average performance of all possible codewords such that the performance prediction of individual codeword might be imprecise. The channel adaptor method symmetrizes the channel by inserting an adaptor into the communication link.

The codeword dependent performance is analogously encountered in the design of LDPC codes for JSCC over asymmetric communication channels as well. To avoid the complicatedly deriving of the codeword-averaged density evolution, we introduce source-channel adaptors to symmetrize the communication channel and the equivalent correlation channel.

As mentioned in section 2, the proposed scheme can be viewed as a channel coding for a virtual channel consisting of a communication subchannel $Ch$ and a correlation subchannel $Ch$. For such a
parallel concatenation channel, the symmetrizing adaptors applied for the communication subchannel and the correlation subchannel should be different. In Fig. 2, we illustrate coding scheme including the source-channel adaptors.

![Fig. 2. Coding scheme with source-channel adaptors.](image)

As shown in Fig. 2, before encoding, a random sequence $T_{in} = \{t_0, \ldots, t_m\}$ generated by an i.i.d source, denoted as i.i.d source 0, is added to $X_i$ in mod-2 manner, i.e., $u_i = x_i \oplus t_i$. The result sequence $U_i = \{u_0, \ldots, u_i\}$ is sent to the encoder as the systematic bits of a LDPC codeword and the encoder output the parity bits of the LDPC codeword, a length $m$ sequence $C_n$, as the encoded data. Then, the encoded data is adjusted by a channel adaptor. The channel adapter, as described in [15], consists of three modules: an i.i.d source (i.i.d source 1 in Fig. 2), a mod-2 adder and a sign adjuster. The i.i.d source generates binary sequence $mT$ according to i.i.d equiprobable distribution. The mod-2 adder perform following operation: $d_i \equiv c_i \oplus u_i$. The result sequence $D_n$ is passed through the communication channel.

At the decoder side, the channel output sequence $V_i$ is a different notation from $V_n$, due to probable corruptions induced by channel noises, is considered as the parity bits of the received LDPC codeword in the decoder.

For the bits in $D_n$, their corresponding initial log-likelihood ratios (LLR) $F_n$ should be post-processed by the sign adjuster of the channel adapter which functions as follows: $g_i = f_i \cdot (1 - 2 \cdot t_i)$. It is easy to check that the sign adjuster undoes the effect of the mod-2 adder at the encoder side.

Considering the channel adaptor and communication channel together as a new augmented channel, we can observe that it satisfies the symmetry condition and is a binary input output symmetric (BIOS) channel according to [15]. The transition probability of the new augmented communication channel becomes

$$p(g_i = g \mid c_i = 0) = p(g_i = -g \mid c_i = 1)$$

$$= p(g_i = g \mid x_i = 0) p(x_i = 0)$$

$$+ p(g_i = -g \mid x_i = 1) p(x_i = 1)$$

(4)

Let $V_n$ denote the sequence of the systematic bits of the received LDPC codewords at the decoder. $V_n$ is produced by doing mod-2 addition of the side information sequence $Y_i$ to the random sequence $T_{in}$, i.e., $v_i = y_i \oplus t_i$.

After calculating the initial LLRs of the received codeword, the decoder proceeds to do iterative message-passing to restore the codeword and eventually output the reconstructed $X_i$, denoted by $\hat{X}_i$, due to probable errors, even very few, occurred in decoding.

In such a coding scheme, $X_i$ is encoded at rate

$$R_e = \frac{m}{k}$$

(5)

Comparing the above coding scheme to the basic coding scheme, we can easily observe that there are two extra adaptors involved: a channel adaptor and a source adaptor.

The channel adapter as described above performs the channel dithering [15]. The source adapter, being composed of the i.i.d source 0 and two mod-2 adders, is intended for source dithering according to following theorem, i.e., uniforming the source and symmetrizing the correlation channel.

**Theorem 1.** For any binary correlated memoryless sources $X$ and $Y$, the correlation between $kU$ and $kV$ as defined above satisfies the symmetry condition, i.e.,

$$p(v_i = 0 \mid u_i = 0) = p(v_i = 1 \mid u_i = 1), \quad \forall i = 0, \ldots, k-1,$$

and $u_i$ is uniform distributed, i.e.,

$$p(u_i = 0) = p(u_i = 1) = 0.5.$$

**Proof:** Since for $\forall i = 0, \ldots, k-1$ both $t_i$ and $u_i$ are i.i.d equiprobable random variables, we have

$$p(v_i = 0 \mid u_i = 0)$$

$$= p(v_i = 0 \mid t_i = 0, u_i = 0) p(t_i = 0 \mid t_i = 0)$$

$$+ p(v_i = 0 \mid t_i = 1, u_i = 0) p(t_i = 1 \mid t_i = 0)$$

$$= p(v_i = 0 \mid t_i = 0, u_i = 0) p(x_i = 0)$$

$$+ p(v_i = 0 \mid t_i = 1, u_i = 0) p(x_i = 1)$$

$$= p(v_i = 0 \mid x_i = 0) p(x_i = 0)$$

$$+ p(v_i = 1 \mid x_i = 1) p(x_i = 1)$$

(6)

Similarly, we have

$$p(v_i = 1 \mid u_i = 1)$$

$$= p(v_i = 0 \mid x_i = 0) p(x_i = 0)$$

$$+ p(v_i = 1 \mid x_i = 1) p(x_i = 1)$$

(7)
Then symmetry condition is fulfilled. The distribution of \( u_i \) is given by

\[
p(u_i = 0) = p(t_{ui} = 0) p(x_i = 0) + p(t_{ui} = 1) p(x_i = 1) = 0.5 p(x_i = 0) + 0.5 p(x_i = 1)
\]  

(8)

Similarly, we have

\[
p(u_i = 1) = 0.5 p(x_i = 0) + 0.5 p(x_i = 1)
\]  

(9)

Then the uniform condition follows directly. It is easy to check that the correlation channel between \( U_s \) and \( V_s \) has the transition probability of

\[
p(v_i = 0 | u_i = 1) = p(v_i = 1 | u_i = 0) \\
p(v_i = 0 | x_i = 1) p(x_i = 1) \\
+ p(v_i = 1 | x_i = 0) p(x_i = 0)
\]  

(10)

By using the channel and source dithering, both the new augmented communication channel and the new augmented correlation channel fulfill symmetry condition. Thus, all previously developed code design and coding method for symmetric channels can be applied directly and the difficulty in the coding for asymmetric channel and nonuniform source cases is dramatically reduced.

However, there still exist a question, i.e., whether the symmetrized coding scheme is theory limit achievable? According to Slepian-Wolf theorem, the minimum rate required by the communication subchannel is \( H(X | Y) \) and the minimum rate based on the new augmented correlation subchannel is \( H(U | V) \). It is easily to verify that \( H(X | Y) = H(U | V) \). As for the capacity of the new binary-input communication channel, it equals the mutual information of the original binary communication channel with i.i.d equiprobable input distribution [13]. Furthermore, it was shown that the ration between the i.i.d mutual information rate and the channel capacity is lower bounded by 0.942 in [16] and proved that the absolute difference is upper-bounded by 0.011 bit/symbol in [17]. Hence, the proposed source-channel adaptor scheme will have very small (can be neglected in practical application) lost in coding efficiency, but it will dramatically reduce the design difficulty due to the BIOS characteristic of the virtual channel.

Now, we proceed to develop the codeword independent density evolution formulas for the coding scheme according to the belief-propagation algorithm. By taking two subchannel into account, we divide the variable nodes into two sets: the information variable nodes (associated to the systematic bits of LDPC codes) and the parity variable nodes. Hereafter, let the subscripts \( s \) and \( p \) describe the notations associated to the information variable nodes and the parity variable nodes, respectively. Accordingly, the code ensemble definition are modified into

\[
\hat{\lambda}_s(x) = \sum_k \lambda_{sk} x^{k-1}
\]  

(11)

\[
\hat{\lambda}_p(x) = \sum_k \lambda_{pk} x^{k-1}
\]  

(12)

\[
\rho_s(x) = \sum_k \rho_{sk} x^{k-1}
\]  

(13)

\[
\rho_p(x) = \sum_k \rho_{pk} x^{k-1}
\]  

(14)

where \( \rho_{sk} \) or \( \rho_{pk} \) denotes the fraction of edges emanating from the check node with \( k \) edges connecting to information variable nodes or parity variable nodes, respectively, and \( \lambda_{sk} \) or \( \lambda_{pk} \) is the fraction of edges emanating from a degree \( k \) information variable node or parity variable node, respectively. The corresponding node perspective degree distributions of check nodes are given by

\[
\hat{\rho}_s(x) = \sum_k \hat{\rho}_{sk} x^{k-1}
\]  

(15)

\[
\hat{\rho}_p(x) = \sum_k \hat{\rho}_{pk} x^{k-1}
\]  

(16)

where \( \hat{\rho}_{sk} = \frac{\rho_{sk}}{k} \sum_i \rho_{sk} i \) and \( \hat{\rho}_{pk} = \frac{\rho_{pk}}{k} \sum_i \rho_{pk} i \) denote the fraction of the check nodes with \( k \) edges connecting to the information variable nodes and parity-check variable node, respectively.

On the basis of the new code ensemble definition, we can then easily obtain the following density evolution update equations for irregular codes,

\[
P^{(l)}_s = P^{(0)}_s \otimes \hat{\lambda}_s \left( Q^{(l-1)}_s \right)
\]  

(17)

\[
P^{(l)}_p = P^{(0)}_p \otimes \hat{\lambda}_p \left( Q^{(l-1)}_p \right)
\]  

(18)

\[
Q^{(l-1)}_s = \Gamma^{-1} \left( \rho_s \left( \Gamma \left( P^{(l-1)}_s \right) \right) \otimes \hat{\rho}_p \left( \Gamma \left( P^{(l-1)}_p \right) \right) \right)
\]  

(19)

\[
Q^{(l-1)}_p = \Gamma^{-1} \left( \rho_p \left( \Gamma \left( P^{(l-1)}_p \right) \right) \otimes \hat{\rho}_s \left( \Gamma \left( P^{(l-1)}_s \right) \right) \right)
\]  

(20)

where the superscript \( (l) \) denotes the \( l \) th iteration and \( \Gamma() \) follows the definition in [18].

4. Experiments

To demonstrate the performance of the proposed scheme, we design irregular codes for JSCC over
binary asymmetric channels (BASCs). The BASC is defined by transition probabilities as follows,

\[
p_{Y|X}(y|x) = \begin{cases} (1 - \epsilon_{01})\delta(y) + \epsilon_{01}\delta(y-1) & \text{if } x = 0 \\ (1 - \epsilon_{10})\delta(y-1) + \epsilon_{10}\delta(y) & \text{if } x = 1 \end{cases}
\]  

(21)

where \( \epsilon_{01} \in [0,1] \) and \( \epsilon_{10} \in [0,1] \) are the crossover probabilities and \( \delta \) is the Dirac delta function.

The density evolution in conjunction with differential evolution [19] is employed to optimize the degree distribution of code ensemble. The code ensemble at rate \( cR \) is optimized for the binary memoryless source \( X \) with \( p \in [0.2,0.5] \) over the BASCs. To simplify the optimization, we only optimize the degree distributions simultaneously for two extreme \( p \)'s, i.e., 0.2 and 0.5.

The degree distributions of the devised codes are listed as follows,

\[
\lambda_\epsilon(x) = 0.233689x + 0.401220x^2 + 0.017851x^3 + 0.246028x^4 + 0.011678x^5 + 0.089535x^6, 
\]

(22)

\[
\lambda_p(x) = 0.235345x + 0.153562x^2 + 0.044587x^3 + 0.220126x^4 + 0.272369x^5 + 0.074011x^6, 
\]

(23)

\[
\rho_\epsilon(x) = 0.06156 + 0.93844x 
\]

(24)

\[
\rho_p(x) = 0.167658x^2 + 0.832342x^3 
\]

(25)

The code graphs are randomly selected from the code ensemble with the information block length \( m = 50000 \). Then, with these graphs (codes) fixed, we find the corresponding parity matrix \( H \) and use Gaussian elimination to find the generator matrix \( G \).

Belief propagation with maximum 200 iterations for each codeword is employed for decoding. More than 400 codewords were transmitted for each simulation. In the simulation, we examined the bit error rate of sources with different \( p \)'s over different communication channels. The simulation results are shown in Fig. 3. In Fig. 3, we plot out the bit error rate of each simulation against the new augmented communication channel capacity. The augmented communication channel, as mentioned in Section 3, consists of the channel adaptor and the communication channel. The capacity of the worst allowable new augmented communication channel with asymptotically vanishing errors is also presented in Fig. 3 as theory limits. From Fig. 3, we can observe that the designed codes have considerably good performance for the source with \( p \in [0.2,0.5] \), although the optimization is only performed at two extreme points. Measuring the gaps between the channel capacities at BER \( \leq 10^{-5} \) and the theory limits, we can observe that the maximum gap is quite small and around 0.044 bits at \( p = 0.2 \).

Since the proposed scheme is a generalization of the LDGM JSCC coding scheme and IRA JSCC coding scheme, it must be more efficient.

5. Conclusions

We present a novel simple JSCC coding scheme based on LDPC codes and source-channel adaptors to facilitate the codes design for asymmetric communication channels. The simulation results show that the proposed coding scheme with the source-channel adaptors has very good performance close to the theory limit and promising application in future. Although we employ the fixed rate codes in our experiments, the puncturing scheme can be applied to the proposed scheme to improve the feasibility of the proposed scheme. Since many sources encountered in practical communications are not well compressed, the proposed scheme can directly replace the channel coding approach in the communication system to improve the error correction ability by efficiently exploiting the redundancy left in the sources. Furthermore, if we incorporate the complicated source model, such as the hidden Markov model, the application scenarios of the proposed scheme can be extended and coding efficiency of the proposed scheme can be further improved.

It remains future work to investigate the short length codes satisfying practical application requirements and the codes for more complicated source models.

Acknowledgements

The authors would like to thank the support of the National Natural Science Foundation of China No. 60972035.
References


2014 Copyright ©, International Frequency Sensor Association (IFSA) Publishing, S. L. All rights reserved. (http://www.sensorsportal.com)