修士副テーマレポート／修士副テーマ論文

副テーマ研究題目：Study of Slepian Wolf Code using LDPC Codes

氏　　名：YUAN Lishuai

学生番号：s2110402

副テーマ指導教員：KURKOSKI, Brian Michael

令和 5 年 5 月

北陸先端科学技術大学院大学

Study of Slepian-Wolf Code using LDPC Codes

Abstract: This paper provides an overview of Slepian-Wolf (SW) codes, establishing a foundational understanding of this coding scheme. We introduce the concept of general Slepian-Wolf codes, which can be either symmetric or asymmetric, achieved by partitioning a single systematic channel code [1]. Next, Low-Density Parity-Check (LDPC) code is introduced and used as an advanced channel code in the Slepian-Wolf code. The simulations are conducted using MATLAB, allowing us to process the data and analyze the results. The objective of this research is to explore the Slepian-Wolf codes and LDPC codes. The findings demonstrate that Slepian-Wolf codes with LDPC codes can approach the theoretical limits.

1. Introduction

The Slepian-Wolf code, named after David Slepian and Jack Wolf, is a coding scheme that allows for the efficient compression of correlated sources without the need for explicit communication between encoders [2].

The key idea behind the Slepian-Wolf code is to design encoders that generate codewords based on the joint statistics of the correlated sources, without exchanging information with each other. These codewords are then decoded jointly by a decoder using the statistical properties of the correlated sources. The Slepian-Wolf code finds applications in various fields, including multimedia communication, sensor networks, distributed sensing, and distributed storage systems.

2. Objectives

There are two objectives of this research:

1. To learn the basic knowledge of Slepian-Wolf code, and principles and encoding and decoding of LDPC code.
2. To learn how to encode and decode messages using Slepian-Wolf code with LDPC code, to conduct the simulation and result analysis.

3. Slepian-Wolf Code

For the relevant sources X, Y, even if they are separately encoded and jointly decoded at the receiver, it is guaranteed to recover X, Y with arbitrarily small error probability, provided that the joint encoding rate, $R = H(X, Y)$, satisfied:

$$R_1 \geq H(X|Y)$$
$$R_2 \geq H(Y|X)$$
$$R = R_1 + R_2 \geq H(X, Y)$$

The achievable rate region of the Slepian-Wolf code theorem is shown in Figure 1. In the figure, R_1 and R_2 denote the coding rates of sources X, and Y, respectively. Therefore, any point on the two-dimensional plane in the figure represents a rate pair (R_1, R_2) [2]. That achievable region denotes a point with rate pair (R_1, R_2) in this region that could achieve error-free transmission.

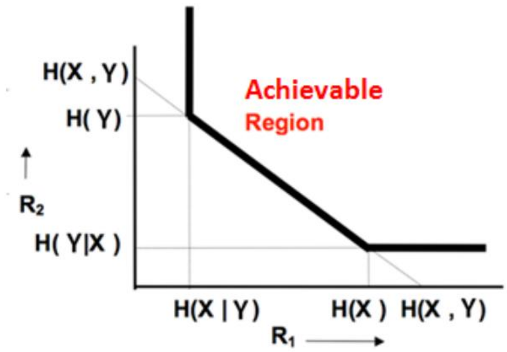

Figure 1 SW bound.

One advantage of the Slepian-Wolf code technique is its reliance on a single good channel code. Specifically, when considering $L = 2$, if the binary code C approaches the capacity of a binary symmetric channel (BSC), the general Slepian-Wolf code (C, M) will approach the Slepian-Wolf limit if the joint correlation between X_1 and X_2 can be modeled with the same BSC. However, when $L > 2$, finding a channel that accurately models the correlation among sources becomes more complex. Efficient design of a single channel code C becomes possible if the correlation satisfies the condition that $X_1 \oplus \cdots \oplus X_L$ follows a Bernoulli-p process [1]. This condition can occur in scenarios such as the remote multiterminal setting where the encoder observes only a noisy version of the source. Nevertheless, the constraint on the correlation model limits the general applicability of this approach compared to other cases.

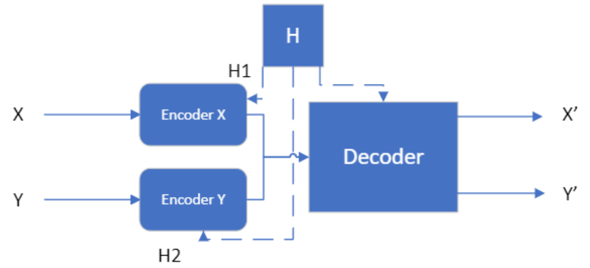

Figure 2 Block diagram for Slepian-Wolf coding: independent encoding of two correlated data streams X and Y [1].

4. Low-Density Parity-Check Code

LDPC (Low-Density Parity-Check) codes are a class of error-correcting codes used in digital communication systems to detect and correct transmission errors. LDPC codes are characterized by a sparse parity-check matrix with a low density of non-zero elements.

The basic idea behind LDPC codes is to encode the information bits into a longer codeword by adding parity-check bits. These parity-check bits are generated based on the relationships between the information bits and the parity-check matrix. The codeword is then transmitted over a noisy channel, and at the receiver, the original information bits are reconstructed by applying an iterative decoding algorithm [3].

4.1 Approximate lower triangular (ALT) encoding.

In approximate lower triangular encoding, the parity-check matrix H of the LDPC code is transformed into an approximate lower triangular form. This transformation simplifies the encoding process by reducing the complexity of encoding operations.

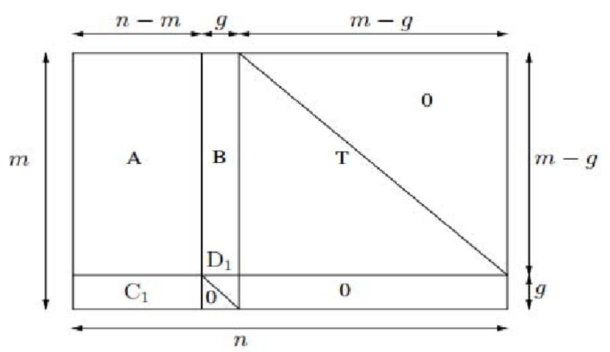

Figure 3 Approximate lower triangular H matrix

4.2 Sum-Product Decoding

Sum-product is a message-passing algorithm that operates on the factor graph representation of the code. Sum-product decoding aims to iteratively estimate the most likely values of the transmitted bits given the received codeword.

The Sum-product decoding algorithm follows these general steps:

Initialization: Initialize the messages between the variable nodes and the check nodes to some reasonable values by the log-likelihood ratio:

$$L_i = \log_2 \frac{\Pr\left[y_i|c_i=0\right]}{\Pr\left[y_i|c_i=1\right]} = \log_2 \frac{\Pr\left[c_i=0|y_i\right]}{\Pr\left[c_i=1|y_i\right]}$$

To initialize the decoder messages, the channel message at node i is passed to all the connected check nodes:

$$Q_{i\to j} = L_i \; for \; all \; j \in M_i, i \in [1,n]$$

Check node function: Check node j is connected to d variable nodes:

$$R_{i\to j} = 2\tan^{-1}\Big(\prod_{e\in N_{j\backslash i}} \tanh\frac{Q_{i\to j}}{2}\Big) \; for \; all \; i \in N_j$$

Variable node function: Variable node i is connected to d check nodes.

$$Q_{i\to j} = L_i + \sum_{e\in M_{i\backslash j}} R_{e\to j} \; for \; all \; j \in M_i$$

Hard decision and syndrome check: At variable node i, a hard decision is made using all available inputs.

$$t_i = L_i + \sum_{e\in M_i} R_{e\to i}$$

$$\widehat{c_i} = \begin{cases} 0, & t_i \geq 0 \\ 1, & t_i < 0 \end{cases}$$

5. Slepian-Wolf Code with LDPC

LDPC H matrix generation: Input the number of H matrix rows k and columns n ($\frac{k}{n} = 0.5$), and make sure each column with the only constraint being that the 1's should be placed in distinct rows (full-rank matrix). After obtaining a full-rank H matrix, eliminate cycles of length 4 in the factor graph and reorder to ALT form.

SW with LDPC Encoding: Splitting H matrix to generate new matrices $[H_1 H_2]$. Then two binary sources

X and Y (n length) are split into three parts in the form:

$$x = [a_1, v_1, q_1], y = [u_2, a_2, q_2]$$

Where $a_1\ v_1\ u_2\ a_2$ are row-vectors of length: $m = k/2$, q_1, q_2 are a row-vector of length: $n-k = n-2m$.

After separately encoding, a pair of syndromes s_1, s_2 and t_1, t_2 are computed and obtained.

$$s_1{}^T = H_1 x^T, \qquad s_2{}^T = H_2 y^T$$

Decoding: After transmitting in Binary Symmetric Channel (BSC), conducting the joint decoding with Sum-product decoding. $t_1 + t_2$ is decoded and $\hat{a}_1, \hat{a}_2$ is obtained as the systematic part of the recovered codeword. Finally, the $\hat{x}\ \hat{y}$ are reconstructed:

$$\hat{x} = [\hat{a}_1\ O_{1\times m}]G\ +\ t_1$$
$$\hat{y} = [O_{1\times m}\ \hat{a}_2]G\ +\ t_2$$

6. Simulation & Result

In the simulation, SW encodes two independent identically distributed binary discrete sources X and Y whose correlation is modeled as a BSC with crossover probability p.

The information block length is set as $k = 10^1, 10^2, 10^3$ bits. The code rate is 1/2, achieved by setting the number of rows and columns of the H matrix as k and $2k$, respectively. The decoding process is limited to 25 iterations, and both the error probability (p) and Bit Error Rate (BER) are constrained to 0.1. Result, shown as Figure 4, is given as BER over the error rate probability, the joint entropy could be calculated by $H(X,Y) = 1 + H(p)$.

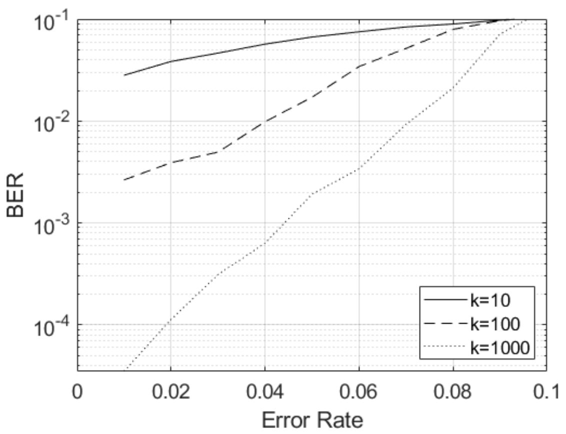

Figure 4 BER vs. Error rate

The symmetric scenario is used in the simulation, the number of rows: $m_1 = m_2 = k/2$, and the $R_1 = R_2 = \frac{n-k/2}{n}$, so $(R_1, R_2) = (\frac{3}{4}, \frac{3}{4})$. Also, $H(X|Y) = H(Y|X) = H(p) = 0.47$. Result obtained with the LDPC based scheme with the SW bound shown in Figure 5. The result is shown that the distance to the bound is 0.0219.

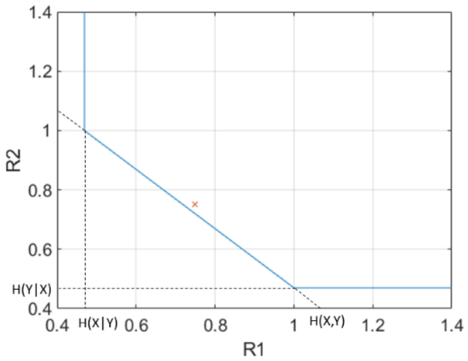

Figure 5 SW Bound

7. Conclusion

In this research, we investigate the combination of independent encoding and joint decoding using an advanced channel code. Additionally, we explore the integration of an LDPC code into the Slepian-Wolf (SW) code framework to approach the SW bound, which represents the theoretical limit of achievable.

To evaluate the performance of the proposed approach, simulations were conducted. However, it is worth noting that the simulations did not consider on the Additive White Gaussian Noise (AWGN) channel. Through this research, I understand the coding schemes that can achieve high compression efficiency and reliable transmission in real-world communication systems.

8. References

[1] V. Stankovic, A. D. Liveris, Z. Xiong, and C. N. Georghiades, “Design of Slepian-Wolf codes by channel code partitioning,” Data Compression Conference, 2004. Proceedings. DCC 2004, Snowbird, UT, USA, 2004, pp. 302-311, March 2004.

[2] D. Splepian and J.K. Wolf, “Noiseless coding of correlated information sources,” IEEE Trans. On Information Theory, vol. IT-19, pp.471-480, July 1973.

[3] A. D. Liveris, Z. Xiong, and C. N. Georghiades, "Compression of binary sources with side information at the decoder using LDPC codes," in IEEE Communications Letters, vol. 6, no. 10, pp. 440-442, Oct. 2002.