Postprint: Latency Improvement of Successive Cancellation Decoding Algorithm for Polar Codes

Authors: Zhang Yuguo, Zhou Jian

Date: 2018-05-20T00:00:00+00:00

Abstract

Polar codes, proposed by Arikan, have attracted widespread attention due to their simple encoding and decoding structures. As a high-performance channel coding scheme, polar codes exhibit excellent performance when the code length exceeds #1;. Based on the successive cancellation (SC) decoding algorithm structure, their decoding latency intensifies with increasing code length. Through analysis of the SC decoding algorithm, an improved SC decoding algorithm scheme based on frozen bits is proposed, which effectively reduces the latency of traditional SC decoding algorithms. The improved algorithm can reduce decoding latency by 50% compared to the original, and introduces the successive cancellation single-bit-flip decoding algorithm as decoding compensation to further enhance the error correction capability.

Full Text

Abstract

Polar codes, proposed by Arikan, have attracted widespread attention due to their simple encoding and decoding structure. As a high-performance channel coding scheme, polar codes exhibit excellent performance when the code length exceeds 2^{10}. However, the decoding latency of the successive cancellation (SC) algorithm intensifies with increasing code length. By analyzing the SC decoding algorithm, this paper proposes an improved SC decoding scheme based on frozen bits that effectively reduces the latency of traditional SC decoding. The improved algorithm can reduce decoding latency by 50% compared to the original, and introduces successive cancellation single-bit-flipping decoding as compensation to further enhance error correction capability.

Keywords: polar code, SC decoder algorithm, frozen bit, decoder latency

0 Introduction

In 1984, Shannon demonstrated that reliable communication over unreliable channels can be achieved through channel coding techniques that enable effective and reliable information transmission in the presence of interference and noise [1]. A core research direction in channel coding involves constructing channel codes that achieve or approach channel capacity, along with developing feasible corresponding decoding algorithms. Over the past two decades, channel coding techniques such as Turbo codes and LDPC codes have been shown to approach Shannon' s channel capacity. Although these coding techniques perform well in practical applications, their theoretical foundations remain incomplete, making it difficult to theoretically prove their optimal transmission characteristics. In 2009, Arikan proposed polar codes based on channel polarization theory, sparking a research boom in the theoretical community. Based on the successive cancellation decoding algorithm, polar codes are the first channel coding method that can be rigorously proven to achieve channel capacity [2,3]. Compared with Turbo codes and LDPC codes, polar codes can be implemented with simple encoders and decoders, and the decoding algorithm does not require complex iterative calculations. The computational complexity of encoding and decoding is only $O(N \log N)$, growing logarithmically with code length.

Although the polar code encoding scheme offers excellent anti-interference performance, the latency issue in SC decoding systems becomes increasingly prominent as the number of encoded bits increases [4,5]. To address this deficiency, we briefly elaborate on the construction principle of polar codes and focus on analyzing the SC decoding algorithm structure to propose improvements for decoding latency.

1 Polar Codes

Polar codes are constructed through the concept of channel polarization, which consists of two stages: channel combining and channel splitting. Erdal Arikan proved that through these linear operations, the symmetric capacities of the split sub-channels exhibit a polarization phenomenon. As the number of combined channels N (i.e., the code length) increases, the symmetric capacities of some sub-channels approach 1, while the symmetric capacities of the remaining sub-channels approach 0. Utilizing this channel polarization phenomenon, polar codes are constructed by transmitting K information bits over sub-channels whose symmetric capacities tend to 1, and transmitting frozen bits over the remaining sub-channels. Frozen bits are fixed bits known to both transmitter and receiver, defaulting to zero. The SC algorithm improvement proposed in this paper is precisely based on this characteristic of frozen bits. The encoding scheme proposed according to this polarized channel is the polar code, with code rate K/N.

1.1 Encoding Steps

Arikan' s paper discusses the polar code encoding scheme based on binary-input discrete memoryless channels (B-DMC) [6], denoted as $W : X \to Y$. The encoding steps based on channel polarization are as follows:

a) **Reliability estimation of polarized channels.** For B-DMC channel W, the Bhattacharyya parameter $Z(W)$ is used to measure channel reliability. The Bhattacharyya parameter value is negatively correlated with channel reliability; that is, the smaller the Bhattacharyya parameter, the greater the channel reliability, and vice versa. According to Arikan' s Bhattacharyya parameter calculation formula (for binary erasure channel (BEC)):

$$Z(W_N^{(i)}) = \sum_{y_1^N \in Y^N} \sum_{u_1^{i-1} \in X^{i-1}} \sqrt{W_N^{(i)}(y_1^N, u_1^{i-1}|0) W_N^{(i)}(y_1^N, u_1^{i-1}|1)}$$

The Bhattacharyya parameter values for each channel are calculated, where $W_N^{(i)}$ represents the channel transition probability of the i-th polarized channel.

b) **Bit mixing.** Based on the first step, according to the relationship between Bhattacharyya parameters and channel reliability, the top K most reliable split sub-channels are selected to transmit message bits, while the remaining $N-K$ split sub-channels transmit frozen bits. At this point, the original bit sequence u_1^N and the received signal y_1^N can be constructed.

c) **Generator matrix construction.** Arikan proposed the generator matrix expression:

$$G_N = B_N F^{\otimes n}$$

where $F^{\otimes n}$ denotes the n-th Kronecker product of matrix F, and B_N is a permutation matrix used to complete bit-reversal reordering operations. The final encoding formula is:

$$x_1^N = u_1^N G_N$$

where x_1^N is the encoded bit sequence with code length $N = 2^n$.

1.2 Successive Cancellation Decoding

To obtain the decoding output of information bits, the key is to calculate the log-likelihood ratio (LLR) $L_N^{(i)}(y_1^N, \hat{u}_1^{i-1})$, which can be computed recursively [7~9]. The receiver obtains y_1^N and recursively calculates from the receiver side to obtain the LLR for subsequent decoding decisions.

The above SC decoding process shows that the decoding decision for the current i-th bit depends on the previous $i-1$ polarized channel inputs $\hat{u}_1^{i-1}$. This linear dependency characteristic determines the sequential output of decoding. Assuming the inputs $\hat{u}_1^{i-1}$ of the first $i-1$ polarized channels are all frozen bits, then according to the previous estimated bits, the decoding of $\hat{u}_i$ can directly utilize the default values of frozen bits without waiting for the decoding of $\hat{u}_{i-1}$. At this point, the decoding of $\hat{u}_i$ causes unnecessary decoding delay and overhead for the information bits we need.

The output of polarized channel $W_N^{(i)}$ includes not only the channel received signal y_1^N but also the partial output sequence $\hat{u}_1^{i-1}$ of the first $i-1$ polarized channel inputs. The channel transition probability is $W_N^{(i)}(y_1^N, \hat{u}_1^{i-1}|u_i)$. The bit estimation function for estimating bit $\hat{u}_i$ based on received signal y_1^N and partial output sequence $\hat{u}_1^{i-1}$ is:

$$\hat{u}_i = h_i(y_1^N, \hat{u}_1^{i-1}) = \begin{cases} 0, & i \in A^c \\ \chi(L_N^{(i)}(y_1^N, \hat{u}_1^{i-1})), & i \in A \end{cases}$$

where A denotes the set of information-carrying channels, and A^c denotes its complement. When $i \in A^c$, it indicates that the bit is a frozen bit and can be directly decided as the default value according to the prior agreement between transmitter and receiver to avoid decoding errors at fixed positions. When $i \in A$, it indicates that the bit carries information, and the decision function is $\chi(L_N^{(i)})$.

The log-likelihood ratio calculation formula is:

$$L_N^{(i)}(y_1^N, \hat{u}_1^{i-1}) = \ln \frac{W_N^{(i)}(y_1^N, \hat{u}_1^{i-1}|0)}{W_N^{(i)}(y_1^N, \hat{u}_1^{i-1}|1)}$$

2 Improved SC Decoding Algorithm

Regarding research on SC decoding latency, various schemes have been proposed [10~12]. In [10], merged processing elements (MPE) for outputs F and G were proposed, along with related definitions. [Figure 1: see original paper] shows the MPE calculation principle diagram proposed in [10]. Although these decoding schemes improve decoding latency, they present certain difficulties in hardware implementation.

[Figure 1: see original paper] MPE module unit calculation

[Figure 2: see original paper] shows the traditional SC decoding algorithm structure for $N = 8$ [10], and shows the specific clock cycles for decoding calculation units. Analyzing the data in , for N-bit polar codes, the traditional SC decoding algorithm requires $N-1$ clock cycles for decoding. shows the specific decoding output for $N = 8$ polar codes. The decoding of N-bit polar codes can be divided

into $N/2$ decoding output steps, with each step outputting two decoded bits. For stage i, 2^{i-1} MPEs are needed, hence the MPE calculation is:

$$\mathrm{MPE}_{\mathrm{calc}} = \sum_{i=1}^{\log N} 2^{i-1} \times \frac{N}{2^i} = N \log N - N + 1$$

Generalizing to N-bit polar codes, according to the above decoding scheme, decoding latency is effectively reduced.

The MPE operation modules in the first stage shown in [Figure 2: see original paper] are replaced by adders in [Figure 3: see original paper]. Analyzing the three outputs of MPE in [Figure 1: see original paper], the computational load of the adder output is 1/3 of the MPE computational load. The MPE computation amount of the improved decoding scheme is defined as follows:

[Figure 3: see original paper] Improved SC decoding algorithm structure

Through the improved encoding and decoding scheme, the latency of the SC decoding algorithm is effectively improved. The improved polar code encoding and decoding scheme is now presented:

a) **Channel reliability estimation.** The Bhattacharyya parameters proposed above target BEC channel reliability estimation. For binary symmetric channel (BSC) and additive white Gaussian noise channel (AWGN), different estimation parameters are selected. The parameters for the three channel types are shown in .

b) **Modify the original bit mixing scheme.** Change the information bits in the first $N/2$ positions to frozen bits.

c) **Perform decoding compensation for the previous step.** Change the frozen bit positions in the latter $N/2$ of the original bit mixing scheme to information bit positions.

3 Decoding Bit Positions

For polar codes (N, K, A, u_{A^c}), the expression is interpreted as calculating the reliability of each sub-channel according to Bhattacharyya parameters, selecting A as information bits, and constructing $N = 8$ mixed bits with u_{A^c} as frozen bits. According to the characteristics of traditional SC decoding, decoding the i-th bit depends not only on the transition probability of the split polarized channel but also on the previous $i-1$ decoding outputs. From , to obtain the information bit position u_4, the first three decoding outputs are required, demanding four clock cycles of work. However, for the first four decoding outputs, according to the mixed bit information, only the 4th bit is an information bit while the first three are frozen bits. The 8-bit decoding output requires computing 12 MPEs $(8/2 \times \log_2 8 = 12)$. From , when the 4th bit is decoded, 8 MPEs need to be computed, incurring approximately half the decoding delay cost for this single

bit. In summary, traditional SC decoding introduces increased decoding latency and consequent complexity.

To overcome this latency characteristic, consider the following encoding scheme: change the original 4th information bit position to a frozen bit, and make the 5th position an information bit. The previous mixed bit scheme is modified to $(8, 4, \{5, 6, 7, 8\}, (0, 0, 0, 0))$. Under this selection scheme, the first four decoding outputs are frozen bits. Referring to , the original 1-4 decoding clock cycles are redundant. Therefore, decoding latency is reduced from the original 7 clock cycles to 3 clock cycles.

According to the above scheme, sub-processes 1 and 2 in are redundant. The MPE calculation amount decreases from 12 to 4. In the SC decoding algorithm, the first-stage MPE output out_F is used for subsequent calculations of the first half of decoded output bits, and out_G is used for subsequent calculations of the second half. With the improved mixed bit scheme, $\hat{u}_1$ to $\hat{u}_4$ are frozen bits, making the out_F operations redundant. The four MPE operation modules in the first part of [Figure 2: see original paper] can be replaced by four adders, with the improved decoding scheme shown in [Figure 3: see original paper].

Channel estimation parameters for different binary discrete memoryless channels

Therefore, for 8-bit polar codes, if the first four bits are transmitted as frozen bits, the decoding scheme shown in [Figure 3: see original paper] can be adopted, reducing decoding latency by 57% and decreasing MPE computation by 60%.

Generalizing to N-bit polar codes, the decoding latency is:

$$\text{Latency} = \frac{N}{2} + \frac{N}{4} + \cdots + \frac{N}{2^{\log N}} = N - 1 - \frac{N}{2} = \frac{N}{2} - 1$$

The MPE computation amount is:

$$\text{MPE}_{\text{Num}} = \frac{N}{2} \times \frac{1}{3} + \frac{N}{4} \times \frac{2}{3} + \cdots + \frac{N}{2^{\log N}} \times \frac{\log N}{3} = \frac{2}{3}N - 1$$

compares the improved SC decoding algorithm with traditional SC decoding. Under the same code length, the improved algorithm reduces decoding latency by 50% and decreases MPE computation by 33% compared to the original.

4 Decoding Compensation and Simulation

In the improved SC algorithm, channel reliability estimation is particularly important as it directly affects the selection of bit position conversions. Converting frozen bits to information bits in the latter half will impact decoding performance. As decoding compensation, introducing a cyclic redundancy check (CRC) [13] module represents a direction for further algorithm improvement.

After adding check codes (typically 24 bits) before encoding, although the coding redundancy increases slightly and the coding rate decreases from k/N to $(k-24)/N$, this effect becomes less noticeable as code length increases.

The successive cancellation single-bit-flipping decoding algorithm [14] can further improve decoding performance on the basis of CRC. In traditional SC decoding, channel noise and error propagation in decoding output are the main causes of decoding decision errors. Error propagation occurs between each bit of the source signal, affecting only the bit error rate. Typically, channel noise causes only 1-bit errors, and the probability of 1-bit errors increases with signal-to-noise ratio or code length. Therefore, finding this erroneous bit can greatly improve SC decoding performance. When the decoding output fails CRC verification, this algorithm flips a single unreliable information bit in the decision sequence and re-executes SC decoding. Select the T decision bits with the smallest LLR values in the output sequence, and record the corresponding information bit index set as M [14]. Each time, take one value from set M, flip the corresponding decision bit value, and re-perform SC decoding output until the new codeword passes CRC. If no valid codeword is obtained after traversing set M, decoding fails.

The single-bit-flipping algorithm can serve as decoding compensation for the improved SC decoding algorithm. The simulation results in [Figure 4: see original paper] show that decoding performance decreases linearly with the number of converted bit positions. Through the introduction of the single-bit-flipping algorithm, decoding performance compensation allows frozen bits to be converted to information bits, making decoding output more reliable.

[Figure 4: see original paper] Simulation of improved SC decoding algorithm (N=1024)

5 Conclusion

Based on analyzing frozen bit positions in polar code encoding, this paper implements a new mixed bit scheme by exchanging positions between certain information bits and frozen bits. According to the traditional SC decoding algorithm structure, redundant decoding computations are reduced, thereby decreasing decoding latency. The essence of the algorithm improvement is to sacrifice lower bit error rate performance in exchange for substantially reduced decoding latency. The introduction of successive cancellation single-bit-flipping decoding algorithm can effectively perform decoding compensation and reduce the bit error rate. Seeking better decoding compensation methods warrants further in-depth research.

References

[1] Shannon C E. A mathematical theory of communication [J]. Bell System Technical Journal, 1948, 27 (3): 379-423.

[2] Arikan E. Channel polarization: a method for constructing capacity-achieving codes for symmetric binary-input memoryless channels [J]. IEEE Trans on Information Theory, 2009, 55 (7): 3051-3073.

[3] Arikan E. On the origin of polar coding [C]// Proc of IEEE Int. Symp. Inf. Theory. 2015.

[4] 王实. 5G 移动通信发展趋势与若干关键技术 [J]. 信息通信, 2015, 12:

[5] ESLAMI, ALI, HOSSEIN P N. A practical approach to polar codes [C]// Proc of IEEE Inf. Theory Workshop. 2011.

[6] 田宝玉. 信息论基础 [M]. 2 版. 北京: 人民邮电出版社, 2016: 32-128.

[7] 陈凯. 极化编码理论与实用方案研究 [D]. 北京: 北京邮电大学, 2014.

[8] Andersson M, Rathi V, Thobaben R. Nested polar codes for wiretap and relay channels [J]. IEEE Communications Letters, 2010, 14 (8): 752-754.

[9] 李超. Polar Codes 编译码算法研究及应用 [D]. 成都: 电子科技大学,

[10] Zhang C, Parhi K K. Low-latency sequential and overlapped architectures for successive cancellation polar decoder [J]. IEEE Trans on Signal Process, 2013, 61 (10): 2429-2441.

[11] Yuan B, Parhi K K. Low-latency successive-cancellation polar decoder architecture using 2-bit decoding [J]. IEEE Trans on Circuits and Systems, 2014, 61 (4): 1241-1254.

[12] Zhang C, Parhi K K. Latency analysis and architecture design of simplified SC polar decoders [J]. IEEE Trans on Circuits Syst. II, Express Briefs, 2014, 61 (2): 115-119.

[13] 牛凯. 一种循环冗余校验辅助的极化码译码方法: 中国, 201210202279.2 [P]. 2014-03-26.

[14] Orion A, Balatsoukas-Stimming A, Andreas B. A low-complexity improved successive cancellation decoder for polar codes [C]// Proc of the 48th IEEE Asilomar Conference on Signals, Systems and Computers. 2014: 2116-2120.

Note: Figure translations are in progress. See original paper for figures.