axioms

MDPI

Article

Polar Code BP Decoding Optimization for Green 6G Satellite Communication: A Geometry Perspective

Chuanji Zhu [1,2,†], Yuanzhi He [1,2,*,†] and Zheng Dou [1,2]

[1] College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China; zhuchuanjie@hrbeu.edu.cn (C.Z.); douzheng@hrbeu.edu.cn (Z.D.)
[2] 32002 Troops of PLA, Beijing 100141, China
* Correspondence: he_yuanzhi@126.com
† These authors contributed equally to this work.

Abstract: The rapid evolution of mega-constellation networks and 6G satellite communication systems has ushered in an era of ubiquitous connectivity, yet their sustainability is threatened by the energy-computation dilemma inherent in high-throughput data transmission. Polar codes, as a coding scheme capable of achieving Shannon's limit, have emerged as one of the key candidate coding technologies for 6G networks. Despite the high parallelism and excellent performance of their Belief Propagation (BP) decoding algorithm, its drawbacks of numerous iterations and slow convergence can lead to higher energy consumption, impacting system energy efficiency and sustainability. Therefore, research on efficient early termination algorithms has become an important direction in polar code research. In this paper, based on information geometry theory, we propose a novel geometric framework for BP decoding of polar codes and design two early termination algorithms under this framework: an early termination algorithm based on Riemannian distance and an early termination algorithm based on divergence. These algorithms improve convergence speed by geometrically analyzing the changes in soft information during the BP decoding process. Simulation results indicate that, when E_b/N_0 is between 1.5 dB and 2.5 dB, compared to three classical early termination algorithms, the two early termination algorithms proposed in this paper reduce the number of iterations by 4.7–11% and 8.8–15.9%, respectively. Crucially, while this work is motivated by the unique demands of satellite networks, the geometric characterization of polar code BP decoding transcends specific applications. The proposed framework is inherently adaptable to any communication system requiring energy-efficient channel coding, including 6G terrestrial networks, Internet of Things (IoT) edge devices, and unmanned aerial vehicle (UAV) swarms, thereby bridging theoretical coding advances with real-world scalability challenges.

Keywords: satellite communication; 6G network; polar codes; BP decoding; early rermination algorithm; information geometry; mathematical modeling

MSC: 94-10

check for updates

Academic Editor: Behzad Djafari-Rouhani

Received: 26 January 2025
Revised: 21 February 2025
Accepted: 26 February 2025
Published: 27 February 2025

Citation: Zhu, C.; He, Y.; Dou, Z. Polar Code BP Decoding Optimization for Green 6G Satellite Communication: A Geometry Perspective. *Axioms* **2025**, *14*, 174. https://doi.org/10.3390/axioms14030174

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Low-Earth Orbit (LEO) satellite communication systems are emerging as a vital component of future 6G networks, offering the potential for high-speed communication across the globe thanks to their low latency, high coverage, and robustness. However, with the expansion of LEO satellite communication systems and the development of mega-constellations, energy consumption has become a core challenge limiting their sustainable development.

Satellites rely on limited solar power, and efficient energy utilization in communication links not only determines the lifespan and operational efficiency of satellites but also profoundly impacts the green development and long-term reliability of the entire constellation. Therefore, reducing energy consumption while ensuring communication performance has become a critical issue driving advancements in LEO satellite communication technology. This poses higher requirements for technologies such as signal processing, communication protocols, and channel coding within the 6G network architecture, while also presenting new research opportunities for achieving a sustainable 6G network architecture.

Among them, channel coding technology serves as the core of wireless communication systems, directly impacting the reliability and energy consumption performance of communication systems. In recent years, polar codes, as the first channel coding scheme theoretically proven to achieve Shannon's limit [1], have not only been incorporated into the 5G standard but are also regarded as one of the important candidate coding technologies for 6G networks. Leveraging their soft-input soft-output (SISO) characteristics and high parallel decoding capability, the Belief Propagation (BP) decoding algorithm for polar codes excels in high-throughput communication scenarios [1–3]. However, a major drawback of BP decoding is its high number of iterations and slow convergence speed, requiring numerous iterations to complete decoding in low signal-to-noise ratio (SNR) communication environments. This not only consumes considerable decoding time but also occupies excessive hardware resources and incurs high energy consumption costs, contradicting the sustainability requirements of 6G networks. Therefore, developing efficient early termination algorithms to reduce BP decoding energy consumption and enhance energy efficiency and convergence speed has become one of the key focuses in polar code research.

To address the issues of high iteration counts and slow convergence speed in BP decoding algorithms, various early termination algorithms have been proposed. In summary, current early termination algorithms can be categorized into the following three approaches:

1. The first approach involves designing a verification condition to judge the result of each iteration. If the condition is met, the iteration stops, and the current decoding result is output. For instance, the early termination criterion based on the G-matrix proposed in Ref. [4] determines whether to terminate early by comparing the estimated messages and codewords at the end of each BP decoding iteration. Ref. [5] improves upon the method in [4] by proposing the Bit Difference Ratio (BDR) criterion, which focuses on the number of differing bits between the estimated codeword and the codeword calculated based on the estimated messages. Ref. [6] introduces an early termination criterion based on the H-matrix, which has lower complexity compared to the G-matrix criterion in [4]. Ref. [7] incorporates Cyclic Redundancy Check (CRC) to more accurately perceive convergence properties. This approach draws inspiration from early termination algorithms for LDPC codes and belongs to the general idea of linear block code decoding. Its advantage lies in its simplicity and directness, while the disadvantage is that it does not incorporate the characteristics of polar codes.
2. The second approach analyzes each information bit or frozen bit based on the characteristics of polarized channels and designs early termination algorithms according to their reliability. Ref. [8] points out that observing a group of information bits, known as the Worst of Information Bits (WIBs), can detect successful decoding. WIBs are bits transmitted through channels with higher error probabilities and thus require more decoding iterations for successful decoding compared to other bits. Additionally, Ref. [9] proposes an early termination criterion based on the bit error rate (BER) for frozen bits of polar codes. In Ref. [10], researchers suggest using optimal frozen bits to terminate decoding, as these partially frozen bits have lower average error probabilities. The advantage of this approach is that it leverages channel polarization

theory and combines the characteristics of polar codes, while the disadvantage is that it does not incorporate the transmission of soft information for information nodes and does not utilize the advantages of BP decoding.

3. The third approach leverages the SISO characteristics of polar code BP decoding to design early termination algorithms based on the Log-Likelihood Ratio (LLR). Since the BP decoding algorithm outputs soft information, the LLR of each information node represents the reliability of the information. Designing early termination algorithms by calculating the LLR exploits the advantages of polar code BP decoding algorithms. Ref. [4] proposes the min Log-Likelihood Ratio (minLLR) criterion, which was later simplified in subsequent research [11] with the decision condition that all LLR values are greater than a set threshold indicating successful decoding. Furthermore, in Ref. [7], an LLR magnitude-aided (LMA) early termination scheme based on the convergence properties of LLR messages is proposed, which terminates iterations early when the LLR of an information node remains unchanged for two consecutive times. Ref. [12] introduces a partial LLR-aided (PLA) early termination scheme, which terminates iterations early when the difference in LLR values of an information node over two consecutive times is less than a threshold. This approach fully exploits the characteristics of soft information transmission in polar code BP decoding, but existing algorithms directly use Euclidean distance to calculate LLR differences, lacking a mathematical analysis of LLR.

Moreover, some researchers have combined two of the above three approaches, as seen in [4,7,13]. Existing methods are typically designed based on empirical rules and lack support from a systematic theoretical framework, with the potential in the direction of early termination yet to be fully explored.

In recent years, information geometry theory has gradually emerged as a novel research tool in the field of channel coding and wireless communication [14,15]. This theory transforms complex communication problems into geometric optimization problems by mapping probability distribution functions to points on geometric manifolds, utilizing geometric tools such as metrics, curvature, and divergence for performance optimization [16]. This characteristic of information geometry provides a new perspective for optimizing the iterative process of polar code BP decoding, aiding in significantly reducing the number of iterations without sacrificing the bit error rate (BER), thereby achieving the goal of reducing energy consumption.

To better leverage the SISO characteristics of polar code BP decoding, this paper conducts a mathematical analysis of the changes in LLR values during the iterative process of polar code BP decoding and proposes a geometric framework for polar code BP decoding based on information geometry theory. This framework not only aids in intuitive understanding of iterative decoding but also facilitates the use of various geometric tools such as metrics, Riemannian distance, and divergence to address the issue of high iteration counts. Under this framework, this paper proposes two early termination algorithms for polar code BP decoding based on information geometry, which improve the convergence speed of polar code BP decoding without affecting the BER. These algorithms conduct a geometric analysis of the changes in soft information during the Belief Propagation decoding iterative process, effectively reducing energy consumption while significantly improving convergence speed. This not only underscores their broad application potential in green 6G satellite communications but also introduces a novel approach suitable for all resource-constrained scenarios in green communications, ranging from large-scale Internet of Things (IoT) deployments equipped with battery-limited sensors to unmanned aerial vehicle (UAV) communication networks.

The innovations of this paper are as follows:

1. To mathematically analyze the changes in LLR values during the BP decoding process of polar codes, this paper utilizes mathematical tools—information geometry and modern differential geometry—to analyze the BP decoding algorithms of polar codes, establish a geometric framework for polar code BP decoding algorithms, and propose a theoretical explanation for polar code BP decoding algorithms. It maps information nodes to statistical manifolds, equates the probability information contained in nodes to points on manifolds, and represents the iterative process as finding convergence points on manifolds, providing an intuitive representation of the iterative process of polar code BP decoding algorithms. Under this framework, information geometry and modern differential geometry can be utilized to improve polar code BP decoding algorithms, providing new insights for research on polar code BP decoding algorithms.
2. Addressing the issue of high iteration counts in BP decoding algorithms, this paper studies early termination algorithms for polar code BP decoding within the geometric framework. By calculating metrics, distance, and divergence in statistical manifolds, it optimizes existing algorithms and proposes early termination algorithms for polar code BP decoding based on Riemannian distance and divergence. These algorithms improve the convergence speed of polar code BP decoding without affecting the BER.

The remainder of this paper is organized as follows. Section 2 reviews the BP decoding process of polar codes. Section 3 introduces the BP decoding process of polar codes from an information geometry perspective. Section 4 proposes two early termination algorithms for polar code BP decoding based on information geometry. Section 5 presents simulation analysis. Section 6 concludes the paper and discusses future directions.The main symbols used in this paper are listed in Table 1.

Table 1. Summary of key symbols.

Symbol	Meaning
System Parameters	
N	Code length
K	Information bits length
E_b/N_0	Energy-per-bit-to-noise power spectral density ratio
$I_{\max}$	Maximum number of iterations
Algorithmic Variables	
$L_{i,j}$,$R_{i,j}$	Leftward LLR message at node (i,j)
$R_{i,j}$	Rightward LLR message at node (i,j)
ξ^t	Leftward propagation vector at iteration t
ζ^t	Rightward propagation vector at iteration t
Geometric Framework	
λ	Natural parameter of Bernoulli distribution
$G(\lambda)$	Metric matrix
η_L	Soft information at leftmost bit nodes
θ_R	Soft information at rightmost bit nodes
$D(\cdot)$	Riemannian distance or divergence measure
$I(\lambda)$	Fisher information matrix

Note: LLR = Log-Likelihood Ratio.

2. BP Decoding Algorithm for Polar Codes

In the BP decoding algorithm, for a polar code with a code length of $N = 2^n$, the factor graph of the polar code contains n stages and N levels, totaling $N(n+1)$ nodes, represented by (i,j) for the ith node at the jth stage. Each stage contains $N/2$ butterfly decoding units, and information propagates through the computing units, as shown in Figure 1. $L_{i,j}$ and $R_{i,j}$ denote the soft information for left-going and right-going messages, respectively,

at node (i, j) on the factor graph, represented by Log-Likelihood Ratios (LLRs). According to reference [3], the information update process for a computing unit can be expressed as follows:

$$\begin{cases} L_{i,j}^{t+1} = f(L_{i+1,j}^{t}, L_{i+1,j+2^{n-1}}^{t} R_{i,j+2^{n-1}}^{t}) \\ L_{i,j+2^{n-1}}^{t+1} = L_{i+1,j+2^{n-1}}^{t} f(L_{i+1,j}^{t}, R_{i,j}^{t}) \\ R_{i+1,j}^{t+1} = f(R_{i,j}^{t}, L_{i+1,j+2^{n-1}}^{t} R_{i+1,j+2^{n-1}}^{t}) \\ R_{i+1,j+2^{n-1}}^{t+1} = R_{i,j+2^{n-1}}^{t} f(R_{i,j}^{t}, L_{i+1,j}^{t}) \end{cases} \tag{1}$$

where

$$f(\alpha, \beta) = \ln\left(\frac{e^{\alpha+\beta}+1}{e^{\alpha}+e^{\beta}}\right) \approx sign(\alpha, \beta)\min(|\alpha|, |\beta|) \tag{2}$$

Equation (2) follows the min–sum approximation commonly used in BP decoding to reduce computational complexity [4].

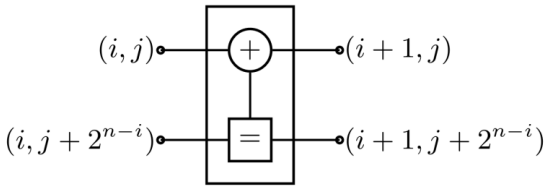


Figure 1. Minimal computational unit for BP decoding [3].

Figure 2 shows an example of a polar code factor graph with N = 8. From Equation (1), it can be seen that the iterative information of each intermediate node is determined by the iterative information of the remaining three nodes of the left and right butterfly decoders, where the node was located the previous time.

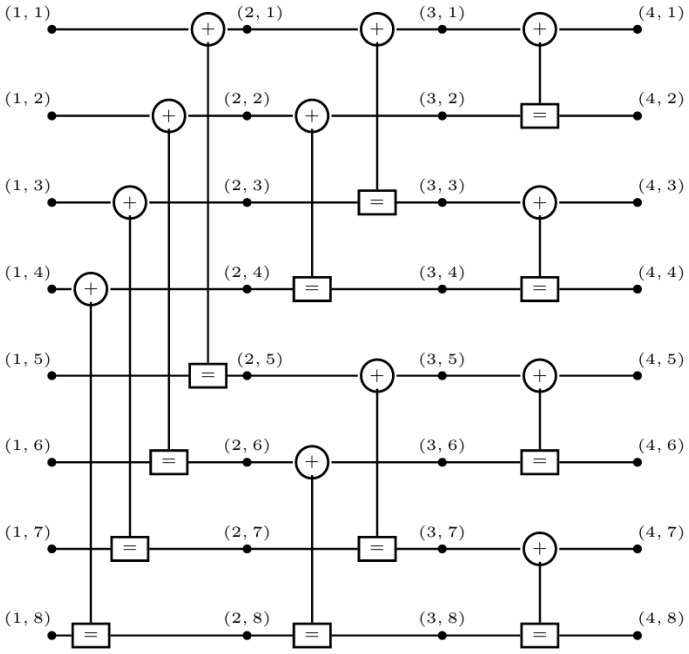


Figure 2. An example of a factor graph for a polar code with N = 8 [3].

As shown in Figure 3, u represents the original codeword, G represents the generator matrix, x represents the codeword after encoding and transmission through the channel, and y represents the codeword sequence received at the receiver. η_L denotes the soft information of the bit nodes at the leftmost side of the factor graph, which, after hard

decision, becomes the decoded result. θ_R represents the soft information of the bits at the rightmost side of the factor graph, which, after hard decision, becomes $\hat{x}$ (the predicted value of x).

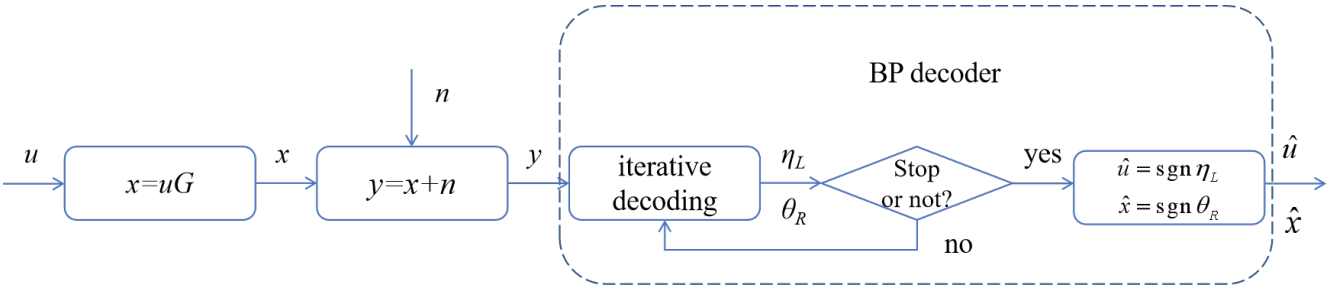


Figure 3. A schematic diagram of the BP decoding process for polar codes *.

* Note:

- When Erdal Arıkan introduced the BP decoding for polar codes, he did not specify a concrete stopping criterion. In practical applications, a maximum number of iterations is often set or one of the three approaches mentioned in the introduction is adopted to prematurely terminate the iterations.
- In the literature and practical applications, attention is often solely focused on $\hat{u}$, while $\hat{x}$ is overlooked. Ref. [4] points out that when the iterations are stopped, $\hat{x} = \hat{u}G$.

The process of polar code BP decoding is represented as Algorithm 1:

Algorithm 1 Polar Code BP Decoding

Require: Received signal vector y (length N), maximum iterations $I_{\max}$, frozen bit set $\mathcal{A}$
Ensure: Decoded result

for $i = 1$ to N **do**
  Initialize $L_{i,n+1} = y_i$
  for all j **do**
    Set $L_{i,j} = 0$ and $R_{i,j} = 0$
  end for
  if $i \in \mathcal{A}$ **then**
    Set $R_{i,1} = +\infty$ {Indicate frozen bit position}
  end if
end for
for iter $= 1$ to $I_{\max}$ **do**
  for $j = 1$ to N **do**
    Compute right-to-left messages based on Equation (1)
  end for
  for $j = 1$ to N **do**
    Compute left-to-right messages based on Equation (1)
  end for
  if convergence criterion met (e.g., $L_{i,1}^{\text{iter}} = L_{i,1}^{\text{iter}+1}$ for all i) **then**
    Break
  end if
end for
Decision: Output decoded result based on final $L_{i,1}$ values

3. Information Geometry of BP Polar Decoding

3.1. Geometric Representation of Soft Information

The probability distribution of each codeword is a Bernoulli distribution, which we usually use as follows:

$$\begin{pmatrix} 0 & 1 \\ p & 1-p \end{pmatrix}. \tag{3}$$

The exponential deformation of Bernoulli distribution is defined as follows:

$$p(x) = p^x(1-p)^{1-x} = \exp[x \ln x + (1-x)\ln(1-p)]. \tag{4}$$

Let $\lambda = LLR = \ln \frac{p}{1-p}$, and then $p = \frac{e^\lambda}{1+e^\lambda}$; it follows that

$$\begin{aligned} p(x;\lambda) &= \exp[x \ln\left(\tfrac{e^\lambda}{1+e^\lambda}\right) + (1-x)\ln\left(1-\tfrac{e^\lambda}{1+e^\lambda}\right)] \\ &= \exp[\lambda x - \ln(1+e^\lambda)]. \end{aligned} \tag{5}$$

The statistical manifold on which the LLR (Log-Likelihood Ratio) lies is

$$M = \{p(x;\lambda) = \exp[\lambda x - \ln(1+e^\lambda)] | x \in \{0,1\}, \lambda \in R\}. \tag{6}$$

This is a one-dimensional manifold.

For a polar code with code length (N, K), the output contains soft information for K information bits. As mentioned earlier, the soft information for one information bit can be represented as a point on a statistical manifold consisting of one-dimensional Bernoulli distributions. Therefore, the soft information for K information bits can be represented as a point on a K-dimensional statistical manifold, which is composed of the joint probability distribution of K one-dimensional Bernoulli distributions. The joint probability distribution can be expressed as

$$p(x;\lambda) = \prod_i \exp[\lambda_i x - \ln(1+e^{\lambda_i})] = \exp[\sum_i \lambda_i x - \sum_i \ln(1+e^{\lambda_i})] \ . \tag{7}$$

In this formula, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \cdots, \lambda_K), i \in [1, K]$.

This K-dimensional statistical manifold can be represented as

$$M = \{p(x;\boldsymbol{\lambda}) = \exp[\sum_i \lambda_i x - \sum_i \ln(1+e^{\lambda_i})]\}. \tag{8}$$

Soft information is mapped to a point on a statistical manifold, whose value can be represented by the coordinates on the manifold, while the iterative process of soft information can be regarded as the movement of points on the statistical manifold. This geometric treatment brings the following three advantages:

- Intuitiveness: The geometric representation dynamically demonstrates the process of polar code BP decoding, providing a clear perspective for a deep understanding of the nature of decoding.
- Flexibility: The construction of the manifold can be flexibly adjusted according to research objectives. For example, when studying a particular information node, a two-dimensional manifold can be constructed through the Cartesian product, with left information and right information corresponding to different manifold coordinates, respectively. When studying the left information of N nodes, an N-dimensional manifold can be constructed, with each coordinate component representing the left information of one node.

- Practicality: Analyzing and optimizing characteristics such as metrics, curvature, and divergence on the manifold with geometric tools can effectively improve the convergence performance of decoding algorithms, providing powerful tool support for subsequent performance optimization.

3.2. Geometric Representation of Node Information

From a microscopic perspective, at time t, the left information of any node can be represented by the coordinate λ_L^t, and the manifold it resides on can be denoted as M_1; similarly, the right information can be represented by the coordinate λ_R^t, and the manifold it resides on can be denoted as M_2. Therefore, at time t, the coordinates of the node's information on the manifold can be expressed as $\lambda^t(\lambda_L^t, \lambda_R^t)$. By taking the Cartesian product of the manifold where the left information coordinate λ_L^t resides and the manifold where the right information coordinate λ_R^t resides, we obtain

$$M = M_1 \times M_2, \tag{9}$$

which is a two-dimensional manifold with coordinates representing left information and right information, respectively.

For any node, its iterative information at any time can be represented by coordinates on M. In this way, the iterative process of the node's information can be represented as the trajectory of a point moving on the manifold.

As shown in Figure 4, if four information nodes are represented as (A, B, C, D), then Equation (1) can be expressed as

$$\begin{aligned} \lambda_{AL}^{t+1} &= f(\lambda_{BL}^t, \lambda_{DR}^t \lambda_{CL}^t) \\ \lambda_{DL}^{t+1} &= \lambda_{CL}^t f(\lambda_{BL}^t, \lambda_{AR}^t) \\ \lambda_{BR}^{t+1} &= f(\lambda_{AR}^t, \lambda_{CL}^t \lambda_{DR}^t) \\ \lambda_{CR}^{t+1} &= \lambda_{DR}^t f(\lambda_{AR}^t, \lambda_{BL}^t) \end{aligned} \tag{10}$$

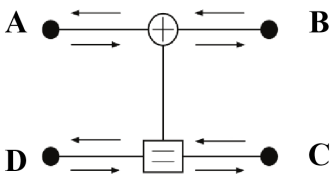
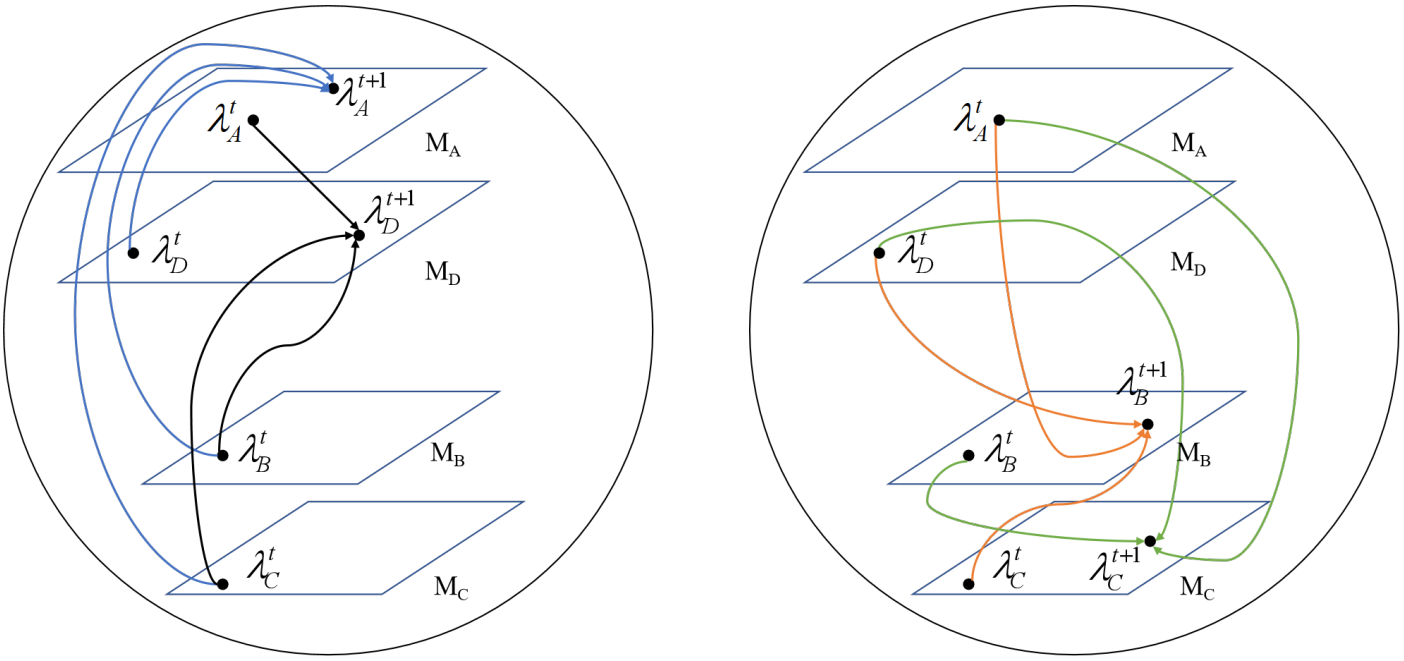


Figure 4. Another representation of the decoding unit.

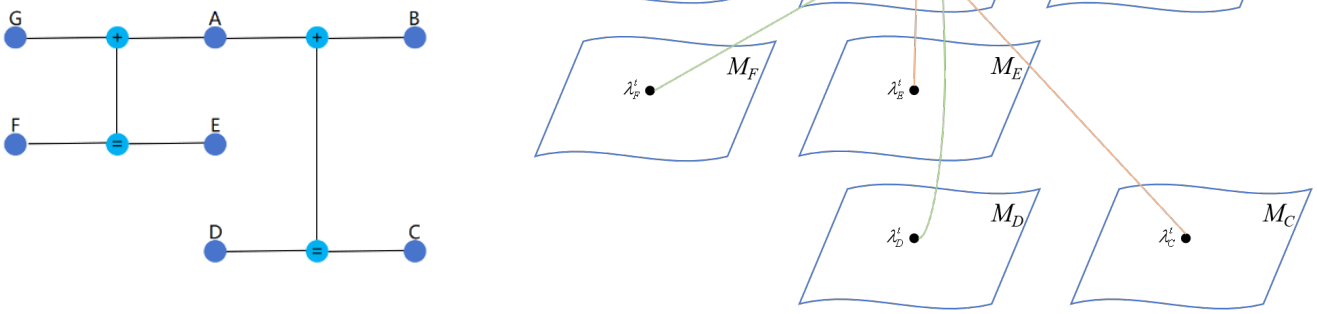
We correlate the four information nodes (A, B, C, D) with the four statistical manifolds (M_A, M_B, M_C, M_D). The probabilistic information embodied within these information nodes is equivalent to a point on the statistical manifold where the probability density function resides, thereby yielding four information points: λ_A^t, λ_B^t, λ_C^t, λ_D^t. The process of information transmission on the butterfly decoding unit can be represented by Figure 5.

For any intermediate node A, as shown in Figure 6a, its left information is determined by the information from the other three nodes (B, C, D) in the butterfly decoder at the previous time step, and similarly, its right information is determined by the information from the other three nodes (G, F, E) in another butterfly decoder at the previous time step. This information iteration is independent of the information of node A itself at the previous time step, and such information iteration can be termed as extrinsic information.



(**a**) (**b**)

Figure 5. A schematic diagram of the iterative information process of a unit decoder from a geometric perspective. (**a**) A description of the schematic diagram of the left information iterative process. (**b**) A description of the schematic diagram of the right information iterative process.



(**a**) (**b**)

Figure 6. (**a**) The information carried by typical node A is related to six adjacent information nodes. (**b**) A schematic diagram of the information iteration process on typical node A from a geometric perspective.

From the perspective of information geometry, seven information nodes (A, B, C, D, E, F, G) correspond to seven statistical manifolds $(M_A, M_B, M_C, M_D, M_E, M_F, M_G)$. The iterative process can be expressed as

$$\lambda_A^{t+1} = g(\lambda_{BL}^t, \lambda_{DR}^t, \lambda_{CL}^t, \lambda_{EL}^t, \lambda_{FR}^t, \lambda_{GR}^t), \tag{11}$$

or as

$$\lambda_A^{t+1} = h(\lambda_B^t, \lambda_D^t, \lambda_C^t, \lambda_E^t, \lambda_F^t, \lambda_G^t). \tag{12}$$

The information iteration process on node A is shown in Figure 6b, where the green arrows represent right information and the orange arrows represent left information.

3.3. Geometric Representation of Factor Graph Information

From the perspective of the factor graph, for a polar code with code length (N, K), at time t, the information propagated leftwards contained in all nodes of the factor graph can be represented as an $N(n+1)$ vector ξ_t as follows:

$$\xi^t = (\lambda_L^t)_{N\times(n+1)}. \tag{13}$$

All possible probability distributions of the information transmitted to the left can form an $N(n+1)$-dimensional manifold M_L, where ξ^t can be seen as a point on M_L. Among them, η_L represents the soft information transmitted to the left of the leftmost bit node in the factor graph, and the manifold where η_L is located is a K-dimensional submanifold of M_L, represented by Equation (8). η_L is the projection of ξ on this submanifold.

Similarly, at time t, the information propagated rightwards contained in all nodes of the factor graph can be represented as an $N(n+1)$ vector ζ_t as follows:

$$\zeta^t = (\lambda_R^t)_{N\times(n+1)}. \tag{14}$$

All possible probability distributions of the information transmitted to the right can form an $N(n+1)$-dimensional manifold M_R, where ζ^t can be seen as a point on M_R. Among them, θ_R represents the soft information transmitted to the left of the leftmost bit node in the factor graph, and the manifold where θ_R is located is a N-dimensional submanifold of M_R. η_L is the projection of ζ on this submanifold.

The parameter space of the codeword after hard judgment is K-dimensional Hamming space. The iterative process from a geometric perspective is shown in Figure 7:

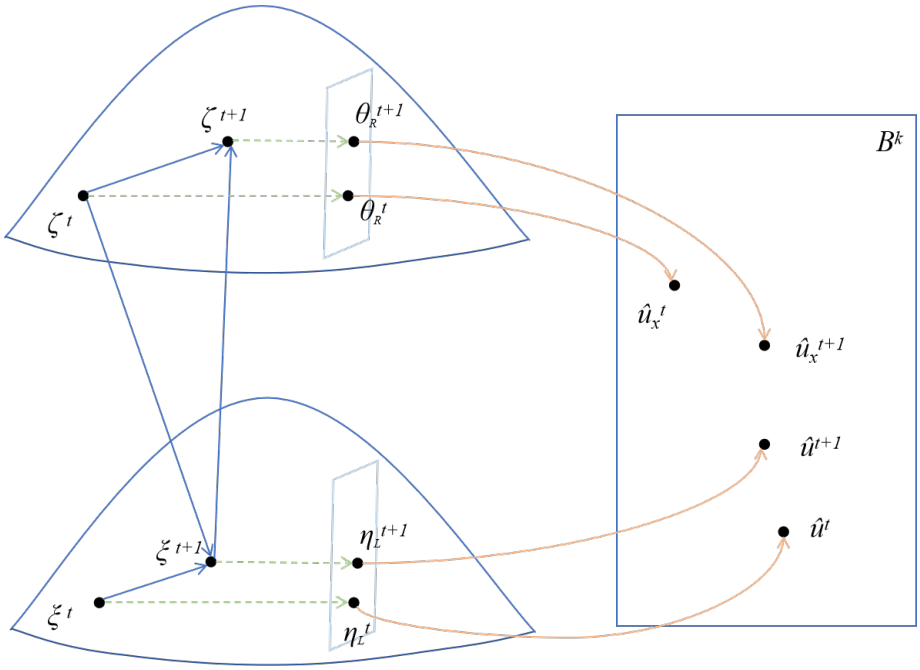


Figure 7. Polarization code BP decoding process from a geometric perspective.

Among them

$$\begin{matrix} \xi^{t+1} = \Phi_L(\xi^t, \zeta^t) \\ \zeta^{t+1} = \Phi_R(\zeta^t, \xi^{t+1}) \end{matrix}. \tag{15}$$

The process of hard judgment is

$$\begin{aligned} \hat{u}^t &= \mathrm{sgn}\eta_L^t \\ \hat{u}_x^t &= \hat{x}^t G^{-1} \\ \hat{x}^t &= \mathrm{sgn}\theta_R^t \end{aligned} . \tag{16}$$

The iterative process of polar code BP decoding can be seen as the movement of ξ and ζ on their respective manifolds. When decoding is complete, there is $\xi^t = \xi^{t+1}$, $\zeta^t = \zeta^{t+1}, \theta_R^t = \theta_R^{t+1}, \eta_L^t = \eta_L^{t+1}, \hat{u}^t = \hat{u}^{t+1}, \hat{u}_x^t = \hat{u}_x^{t+1}, \hat{u}^t = \hat{u}_x^t$.

The process of polar code BP decoding from a geometric perspective is represented as Algorithm 2:

Algorithm 2 Polar Code BP Decoding

Require: y: received signal vector (length N), maximum iteration number $I_{\max}$, frozen bits
Ensure: decoded result $\hat{u}$

```
for i = 1 to N do
    Initialize the coordinates of each node based on the received signal vector y and frozen bit positions
end for
for iter = 1 to I_max do
    Left-to-right iteration: Update the coordinates of ξ based on Equation (15)
    Right-to-left iteration: Update the coordinates of ζ based on Equation (15)
    if ξ^t = ξ^{t+1} and ζ^t = ζ^{t+1} then
        break
    end if
end for
Output: decoded result û
```

4. Early Stopping Based on Information Geometry

4.1. Algorithm Flow

In the process of polar code Belief Propagation (BP) decoding, the high number of iterations and significant computational load are crucial factors limiting decoding efficiency. Based on the geometric model for polar code BP decoding established in the previous section, this paper proposes an early convergence algorithm utilizing the geometric model. The objective is to prematurely terminate the decoding process by calculating the distance of soft information in differential geometric space, thereby enhancing decoding efficiency.

As mentioned in the previous section, when decoding concludes, we have $\eta_L^t = \eta_L^{t+1}$. In fact, since $\hat{u}^t = \mathrm{sgn}(\eta_L^t)$, when the difference between η_L^t and η_L^{t+1} is sufficiently small, a correct decoding result can still be obtained. Therefore, the core of the algorithm lies in utilizing differential geometry theory to design a convergence threshold. This threshold, based on the non-Euclidean properties of soft information in geometric space, calculates the Riemannian distance or divergence to obtain a geometric distance, which is used to measure the minute differences in information updates during the iteration process. When the geometric distance of the soft information for the same information bit in two consecutive iterations is less than this threshold, the algorithm determines that the iteration has converged and prematurely terminates the decoding.

For Algorithm 3, during the implementation of the algorithm, the maximum number of iterations $I_{\max}$ is first set, and a convergence threshold eps is also established based on differential geometry principles. In the initialization stage, the Log-Likelihood Ratio (LLR) values for left-propagated soft information of the information nodes on the leftmost side of the factor graph required for decoding are prepared, and a geometric model is constructed to calculate the geometric distance between soft information.

During the iteration process, the algorithm updates the LLR values for each information bit one by one and calculates the geometric distance of the soft information between the current and previous iterations. If the geometric distance is less than the threshold eps, the iteration is deemed to have converged, and the iteration is terminated; if the number of iterations reaches $I_{\max}$ without convergence, the iteration is forcibly terminated.

Ultimately, the algorithm decides and outputs the decoding result based on the results of the final iteration. Since the threshold design is based on differential geometry principles, the algorithm achieves an improvement in decoding efficiency while ensuring decoding error correction performance.

Algorithm 3 Early Convergence Algorithm Based on Information Geometry

Require: Maximum number of iterations $I_{\max}$, Threshold eps
Ensure: Decoding result
Initialize
for $t = 1$ to $I_{\max}$ **do**
  for $j = 1$ to N **do**
    Update LLR
  end for
  covt $= \max_i D[(\eta_L^t)_i, (\eta_L^{t+1})_i]$
  if covt $<$ eps **then**
    break
  end if
end for
Decide and output decoding result

4.2. Calculation of Geometric Distance

4.2.1. Riemann Distance

In this context, the Fisher Information Matrix degrades into the Fisher Information. According to reference [17], the Fisher Information is expressed as follows:

$$I(\lambda) = -E[\frac{\partial^2}{\partial \lambda^2} \ln p(x;\lambda)]. \tag{17}$$

E represents the expectation with respect to x.

Firstly, we calculate the first-order derivative:

$$\frac{\partial}{\partial \lambda} \ln p(x;\lambda) = \frac{\partial}{\partial \lambda}[\lambda x - \ln(1+e^{\lambda})] = x - \frac{e^{\lambda}}{(1+e^{\lambda})}. \tag{18}$$

The second-order derivative is

$$\frac{\partial^2}{\partial \lambda^2} \ln p(x;\lambda) = \frac{\partial^2}{\partial \lambda^2}[\lambda x - \ln(1+e^{\lambda})] = -\frac{e^{\lambda}}{(1+e^{\lambda})^2}. \tag{19}$$

The Fisher information matrix of a statistical manifold composed of Bernoulli distributions is

$$I(\lambda) = -E[\frac{e^{\lambda}}{(1+e^{\lambda})^2}] = \frac{e^{\lambda}}{(1+e^{\lambda})^2}. \tag{20}$$

In this article, the inverse matrix of the Fisher information matrix is used as a metric for statistical manifolds, and the metric for statistical manifolds composed of Bernoulli distributions is defined as

$$G(\lambda) = I(\lambda)^{-1} = \frac{(1+e^{\lambda})^2}{e^{\lambda}}. \tag{21}$$

In a one-dimensional Bernoulli distribution manifold, the Riemann distance can be approximated as follows:

$$D(\lambda_1, \lambda_2) = \int_{\lambda_1}^{\lambda_2} \sqrt{G(\lambda)} d\lambda = \int_{\lambda_1}^{\lambda_2} \sqrt{\frac{\left(1+e^{\lambda}\right)^2}{e^{\lambda}}} d\lambda. \tag{22}$$

First, calculate the indefinite integral as

$$\int \sqrt{\frac{\left(1+e^{\lambda}\right)^2}{e^{\lambda}}} d\lambda = -2e^{-\frac{1}{2}\lambda} + 2e^{\frac{1}{2}\lambda} \ . \tag{23}$$

So, the Riemann distance of a one-dimensional Bernoulli distribution manifold can be expressed as

$$D(\lambda_1, \lambda_2) = 2\left| e^{\frac{\lambda_2}{2}} - e^{-\frac{\lambda_2}{2}} - e^{\frac{\lambda_1}{2}} + e^{-\frac{\lambda_1}{2}} \right|. \tag{24}$$

4.2.2. Divergence

In applications, divergence is also an important concept for describing the difference between two points on a manifold. The most widely used in statistical manifolds is Kullback–Leibler divergence (KL divergence). Using Riemann distance or divergence to calculate the difference between two probability distributions on a statistical manifold is a common method in information geometry.

The divergence formula on the manifold mentioned above is (see Appendix A for details) as follows:

$$\begin{aligned} & D_{KL}(\lambda_1 \| \lambda_2) = \sum_x p(x;\lambda_1) \log \frac{p(x;\lambda_1)}{p(x;\lambda_2)} \\ & = \sum_x \exp[\lambda_1 x - \ln(1+e^{\lambda_1})] \log \frac{\exp[\lambda_1 x - \ln(1+e^{\lambda_1})]}{\exp[\lambda_2 x - \ln(1+e^{\lambda_2})]} \\ & = (\lambda_1 - \lambda_2)\frac{e^{\lambda_1}}{1+e^{\lambda_1}} - \ln(\frac{1+e^{\lambda_1}}{1+e^{\lambda_2}}) \end{aligned} \ . \tag{25}$$

The above two distance formulas are incorporated into the decoding algorithm framework, resulting in two polar code BP decoding algorithms based on information geometry: the BP polar decoding algorithm with early stopping based on Riemann distance (ESBP-R) and the BP polar decoding algorithm with early stopping based on divergence (ESBP-D).

5. Simulation Results and Comparisons

In this section, we compare the performances of the proposed early convergence algorithms, namely ESBP-R and ESBP-D, with several early convergence algorithms designed using LLR (Log-Likelihood Ratio) through simulations. These include the min Log-Likelihood Ratio (minLLR) algorithm proposed in reference [4], an LLR magnitude-aided (LMA) early convergence algorithm based on the convergence properties of LLR messages proposed in reference [7], and a partial LLR-aided (PLA) early convergence algorithm proposed in reference [12]. The performance metrics include the block error rate (BER) and the average number of iterations. The polar codes studied in the simulations have a code length of $N = 1024$, an information bit length of $K = 512$, a code rate of $r = 0.5$, and a maximum number of iterations of 30. The channel is an Additive White Gaussian Noise (AWGN) channel, the polar code construction follows the Bhattacharyya parameter method as described in reference [1], and the modulation scheme is BPSK.

Figure 8 compares the block error rates of several decoding algorithms under three channel conditions. Figure 8 confirms that the two information geometry-based BP decoding algorithms for polar codes do not suffer from any loss in decoding performance.

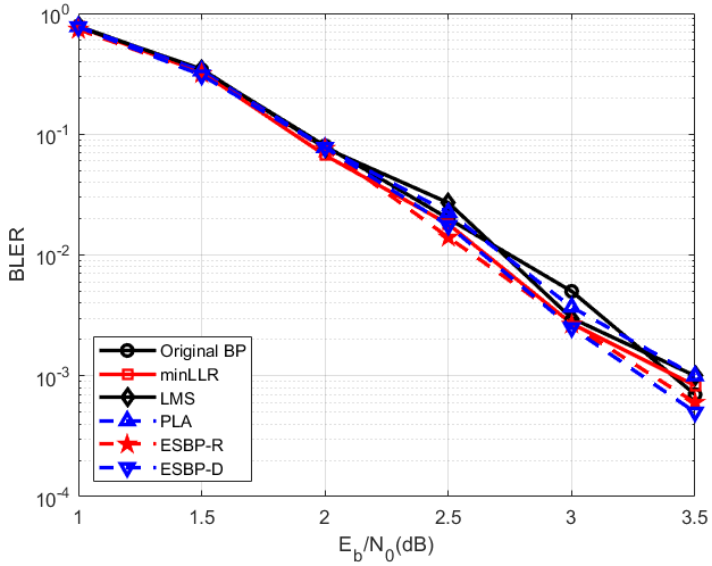


Figure 8. BLER performance for five polar codes with N = 1024 and K = 512. E_b/N_0 (energy-per-bit-to noise power spectral density ratio) measures the normalized signal quality.

The average number of iterations is shown in Figure 9. Table 2 presents the ratios of iteration reductions achieved by several early convergence algorithms. Table 3 demonstrates that compared to the average number of iterations of the three existing early convergence algorithms designed based on LLR, ESBP-R reduces the number of iterations by 2–11%, and ESBP-D reduces it by 4–15%. Notably, at E_b/N_0 values ranging from 1.5 dB to 2.5 dB, ESBP-R and ESBP-D exhibit significant improvements in convergence speed.

Simulation results indicate that under various channel conditions, the two information geometry-based BP decoding algorithms for polar codes proposed in this paper, namely ESBP-R and ESBP-D, can terminate the decoding algorithm earlier than other methods, leading to a lower average number of iterations.

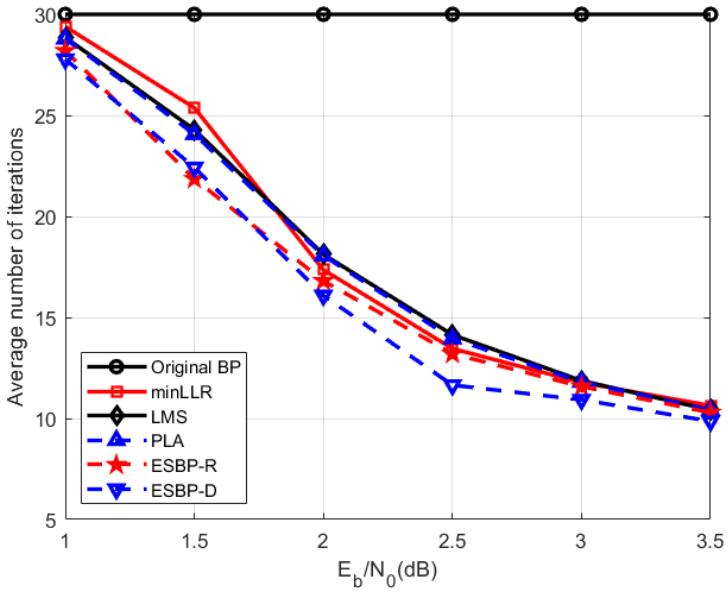


Figure 9. The average number of iterations corresponding to the simulation results in Figure 8.

Table 2. The ratio of the number of iterations reduced by several early convergence algorithms under different channel conditions.

E_b/N_0	minLLR	LMS	PLA	ESBP-R	ESBP-D
1	2.01%	3.81%	4.01%	6.01%	7.39%
1.5	15.42%	19.03%	19.78%	27.15%	25.27%
2	42.14%	39.59%	39.81%	43.97%	46.38%
2.5	55.14%	52.87%	53.51%	56.01%	61.17%
3	60.74%	60.43%	60.63%	61.40%	63.59%
3.5	64.55%	65.28%	65.31%	65.71%	67.12%

Table 3. The ratio of reduced iteration times compared to three existing early convergence algorithms.

E_b/N_0	ESBP-R	ESBP-D
1	2.8%	4.3%
1.5	11.0%	8.8%
2	5.8%	9.9%
2.5	4.7%	15.9%
3	2.0%	7.6%
3.5	2.0%	5.9%

6. Conclusions

This paper addresses the issue of high energy consumption in the Belief Propagation (BP) decoding algorithm for polar codes in Low-Earth Orbit (LEO) satellite communication systems. A geometric framework for polar code BP decoding is established, and two early convergence algorithms for polar code BP decoding based on geometric distance are proposed. Simulation results demonstrate that both algorithms improve the convergence speed of polar code BP decoding without compromising the bit error rate. Our research provides a new theoretical interpretation and method for the BP decoding of polar codes, offering valuable insights for further enhancing the decoding performance of polar codes. It also presents fresh perspectives for achieving sustainable 6G network architectures and communications in all resource-constrained node scenarios.

Previous scholars have attempted to utilize geometry to study channel coding problems in Turbo codes and LDPC codes [14,18–21]. The geometric framework for polar code BP decoding proposed in this paper can be seen as a continuation of these studies. Through research on early convergence algorithms for polar code BP decoding, this paper validates the effectiveness of the geometric framework. Future work can leverage this geometric framework to improve optimization algorithms for polar code BP decoding and optimize the error correction performance of polar code BP decoding.

Author Contributions: This research was conducted by all the authors: Y.H. and C.Z. conceived the idea, performed the analysis, and designed the scheme; C.Z. and Z.D. conducted the numerical simulations; Y.H., C.Z. and Z.D. co-wrote the manuscript. All authors have read and agreed to the published version of this manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data used to support the findings of the current study are available from the authors upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Deduction of Equation (25)

For a binary random variable $X \in \{0, 1\}$, we can assume its probability distribution is controlled by the logit parameter λ as follows:

$$p(x;\lambda) = \frac{e^{\lambda x}}{1+e^{\lambda}}.$$

That is

- When $x = 1$, $p(1;\lambda) = \frac{e^{\lambda}}{1+e^{\lambda}}$.
- When $x = 0$, $p(0;\lambda) = \frac{1}{1+e^{\lambda}}$.

The KL divergence is defined as

$$D_{KL}(p(x;\lambda_1)\|p(x;\lambda_2)) = \sum_x p(x;\lambda_1)\log\frac{p(x;\lambda_1)}{p(x;\lambda_2)}.$$

Substituting the probability mass function of the Bernoulli distribution

$$D_{KL} = p(1;\lambda_1)\log\frac{p(1;\lambda_1)}{p(1;\lambda_2)} + p(0;\lambda_1)\log\frac{p(0;\lambda_1)}{p(0;\lambda_2)}.$$

Expanding the calculation

$$= \frac{e^{\lambda_1}}{1+e^{\lambda_1}}\log\left(\frac{\frac{e^{\lambda_1}}{1+e^{\lambda_1}}}{\frac{e^{\lambda_2}}{1+e^{\lambda_2}}}\right) + \frac{1}{1+e^{\lambda_1}}\log\left(\frac{\frac{1}{1+e^{\lambda_1}}}{\frac{1}{1+e^{\lambda_2}}}\right).$$

Simplifying

$$= \frac{e^{\lambda_1}}{1+e^{\lambda_1}}\log\left(e^{(\lambda_1-\lambda_2)}\frac{1+e^{\lambda_2}}{1+e^{\lambda_1}}\right) + \frac{1}{1+e^{\lambda_1}}\log\left(\frac{1+e^{\lambda_2}}{1+e^{\lambda_1}}\right).$$

Decomposing the logarithmic terms

$$= \frac{e^{\lambda_1}}{1+e^{\lambda_1}}\left[(\lambda_1-\lambda_2) + \log\frac{1+e^{\lambda_2}}{1+e^{\lambda_1}}\right] + \frac{1}{1+e^{\lambda_1}}\log\frac{1+e^{\lambda_2}}{1+e^{\lambda_1}}.$$

Extracting the common logarithmic term

$$= (\lambda_1-\lambda_2)\frac{e^{\lambda_1}}{1+e^{\lambda_1}} + \log\frac{1+e^{\lambda_2}}{1+e^{\lambda_1}}\left(\frac{e^{\lambda_1}}{1+e^{\lambda_1}} + \frac{1}{1+e^{\lambda_1}}\right).$$

Since the sum of the denominator and numerator equals 1, taking the base as e,we obtain the final expression:

$$D_{KL} = (\lambda_1-\lambda_2)\frac{e^{\lambda_1}}{1+e^{\lambda_1}} - \ln\left(\frac{1+e^{\lambda_1}}{1+e^{\lambda_2}}\right).$$

References

1. Arikan, E. Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels. *IEEE Trans. Inf. Theory* **2009**, *55*, 3051–3073. [CrossRef]
2. Arikan, E. A performance comparison of polar codes and Reed-Muller codes. *IEEE Commun. Lett.* **2008**, *12*, 447–449. [CrossRef]
3. Arikan, E. Polar codes: A pipelined implementation. In Proceedings of the 4th International Symposium on Broadband Communication (ISBC 2010), Melaka, Malaysia, 11–14 July 2010; Volume 2010, pp. 11–14.
4. Yuan, B.; Parhi, K.K. Early stopping criteria for energy-efficient low-latency belief-propagation polar code decoders. *IEEE Trans. Signal Process.* **2014**, *62*, 6496–6506. [CrossRef]

5. Zhang, S.y.; Shahrrava, B. Stopping criterion for belief propagation polar code decoders based on bits difference ratio. In Proceedings of the 2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 17–19 October 2019; pp. 285–289. [CrossRef]
6. Hasani, A.; Lopacinski, L.; Grass, E. Early stopping of BP polar decoding based on parity-check sums. In Proceedings of the 2022 IEEE 95th Vehicular Technology Conference:(VTC2022-Spring), Helsinki, Finland, 19–22 June 2022; pp. 1–5. [CrossRef]
7. Ren, Y.; Zhang, C.; Liu, X.; You, X. Efficient early termination schemes for belief-propagation decoding of polar codes. In Proceedings of the 2015 IEEE 11th International Conference on ASIC (ASICON), Chengdu, China, 3–6 November 2015; pp. 1–4. [CrossRef]
8. Simsek, C.; Turk, K. Simplified early stopping criterion for belief-propagation polar code decoders. *IEEE Commun. Lett.* **2016**, *20*, 1515–1518. [CrossRef]
9. Zhang, Q.; Liu, A.; Tong, X. Early stopping criterion for belief propagation polar decoder based on frozen bits. *Electron. Lett.* **2017**, *53*, 1576–1578. [CrossRef]
10. Yan, Y.; Zhang, X.; Wu, B. Simplified early stopping criterion for belief-propagation polar code decoder based on frozen bits. *IEEE Access* **2019**, *7*, 134691–134696. [CrossRef]
11. Choi, S.; Yoo, H. Area-efficient early-termination technique for belief-propagation polar decoders. *Electronics* **2019**, *8*, 1001. [CrossRef]
12. Yang, C.; Yang, J.; Liang, X.; Zhang, Z.; You, X.; Zhang, C. BP polar decoders with early termination and belief enhancing. In Proceedings of the 2018 IEEE 23rd International Conference on Digital Signal Processing (DSP), Shanghai, China, 19–21 November 2018; pp. 1–5. [CrossRef]
13. Cui, J.; Li, J.; Dong, Y.; Zhang, X.; Zeng, Q. Simplified early stopping criterion for polar codes. In Proceedings of the 2018 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), Qingdao, China, 14–16 September 2018; pp. 1–4. [CrossRef]
14. Ikeda, S.; Tanaka, T.; Amari, S.i. Information geometry of turbo and low-density parity-check codes. *IEEE Trans. Inf. Theory* **2004**, *50*, 1097–1114. [CrossRef]
15. Zhu, C.; He, Y.; Dou, Z. Research on Enhanced Belief Propagation List Decoding Algorithm for Polar Codes in UAV Communications for 6G. *Drones* **2024**, *8*, 737. [CrossRef]
16. Amari, S.i. *Information Geometry and Its Applications*; Applied Mathematical Sciences; Springer: Tokyo, Japan, 2016; Volume 194. [CrossRef]
17. Rao, C.R. Information and the accuracy attainable in the estimation of statistical parameters. In *Breakthroughs in Statistics: Foundations and Basic Theory*; Springer: New York, NY, USA, 1992; pp. 235–247. [CrossRef]
18. Richardson, T. The geometry of turbo-decoding dynamics. *IEEE Trans. Inf. Theory* **2000**, *46*, 9–23. [CrossRef]
19. Ikeda, S.; Tanaka, T.; Amari, S.i. Information Geometrical Framework for Analyzing Belief Propagation Decoder. In *Proceedings of the Advances in Neural Information Processing Systems*; Dietterich, T., Becker, S., Ghahramani, Z., Eds.; MIT Press: Cambridge, MA, USA, 2001; Volume 14, pp. 407–414. [CrossRef]
20. Ikeda, S.; Tanaka, T.; Amari, S.i. Information geometry of turbo codes. In Proceedings of the IEEE International Symposium on Information Theory, Lausanne, Switzerland, 30 June–5 July 2002; p. 114. [CrossRef]
21. Ikeda, S. Information geometry of turbo and LDPC codes. In Proceedings of the Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, Philadelphia, PA, USA, 23 March 2005; Volume 5, pp. v/1029–v/1032. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.