

基于编码矩阵结构特征的非删余极化码参数盲识别算法

王垚^{1,2}, 王翔¹, 杨国东³, 黄知涛¹

(1. 国防科技大学电子科学学院, 湖南 长沙 410073; 2. 陆军工程大学通信士官学校, 重庆 400036;
3. 中国人民解放军 92001 部队, 山东 青岛 266023)

摘 要: 针对目前标准非删余极化码盲识别存在的算法复杂、误码适应能力不足等问题, 首先证明了能够表征极化码码长、码率关系, 区分信息子信道和冻结子信道的定理与命题, 并基于此提出了一种高效的盲识别算法。该算法只需设定可能的最大码长, 构建对应的软判决码字矩阵与克罗内克矩阵, 利用所证明的定理与命题对两矩阵校验关系进行判决, 估计码率及冻结比特位置, 最终完成参数的识别。在校验关系判决中, 引入平均似然差作为检验量, 基于其理论概率分布与极小化极大准则确定判决门限。仿真结果表明, 推导的定理、命题与仿真结果一致, 当信噪比为 6 dB 时, 对于码长为 1 024 的极化码, 其参数识别率仍接近 100%, 识别性能和计算复杂度均优于现有软判决算法。

关键词: 信道编码; 标准极化码; 非删余; 识别

中图分类号: TN911.7

文献标志码: A

DOI: 10.11959/j.issn.1000-436x.2022033

Recognition algorithm of non punctured polarization codes based on structural characteristics of coding matrix

WANG Yao^{1,2}, WANG Xiang¹, YANG Guodong³, HUANG Zhitao¹

1. College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China
2. Communication Sergeant School, Army Engineering University, Chongqing 400036, China
3. Unit 92001 of PLA, Qingdao 266023, China

Abstract: In order to solve the problems of complexity and poor error adaptability in the blind recognition of standard non-punctured polarization codes, the theorems and propositions that could characterize the relationship between the code length and code rate, and distinguish information subchannel and frozen subchannel were proved. Based on the theorem and proposition, an efficient blind recognition algorithm was proposed. The proposed algorithm only needed to set the possible maximum code length, and the corresponding soft decision codewords matrixes and Kronecker matrixes were constructed. Based on the theorem and proposition proved, the check relationship between the two matrices was judged and the code rate and frozen bit position were estimated. The average likelihood difference was introduced as the test quantity, and the decision threshold was determined based on its theoretical probability distribution and minimax criterion. The simulation results show that the deduced theorem and propositions are consistent with the simulation results. Under the signal-to-noise ratio of 6 dB and code length of 1 024, the parameter recognition rate is still close to 100%. The recognition performance and computational complexity are better than the existing soft decision algorithms.

Keywords: channel code, standard polar code, non-punctured, recognition

收稿日期: 2021-11-16; 修回日期: 2022-01-18

基金项目: 湖南省自然科学基金创新群体资助项目 (No.2019JJ10004)

Foundation Item: The Program for Innovative Research Groups of the Natural Science Foundation of Hunan Province (No.2019JJ10004)

0 引言

为了实现稳定可靠的通信，信道编码技术被广泛应用于各类数字通信系统中，该技术的合理运用可以有效提高信息传输的质量。其中，极化码是已知的唯一一种被严格证明在二进制离散无记忆信道 (B-DMC, binary discrete memoryless channel) 下能够达到香农极限的信道编码方法^[1]，并被确定为 5G 增强型移动宽带 (eMBB, enhanced mobile broadband) 场景下的控制信道编码方案^[2]。极化码是 5G 通信中采用的新型编码方式，对其进行参数盲识别的研究，具有重要的研究价值。

在认知通信和非合作通信领域，信道编码盲识别受到了国内外越来越多的关注。从编码过程上看，极化码属于线性分组码，且当前线性分组码盲识别课题已经有了较多的研究成果与积累。文献[3]利用接收码字的秩特性，提出一种基于秩准则的编码类型识别算法。文献[4]提出一种基于游程特征的线性分组码与卷积码类型识别算法，该算法可以较好地线性分组码与卷积码进行区分。针对循环码识别问题，文献[5]从循环码定义出发，利用欧几里得算法，完成循环码生成多项式的识别。文献[6-7]利用循环码生成多项式为 $x^n + 1$ 的因子这一特性，通过因式分解实现循环码识别，但误码适应能力不足。针对 BCH (Bose-Chaudhuri-Hocquenghem) 码识别问题，文献[8]提出了一种利用码根差熵和码根统计的本原 BCH 码识别算法，但在高码率情况下识别性能变差，且算法计算量较大。文献[9]提出了一种基于伽罗华域傅里叶变换 (GFFT, Galois field Fourier transform) 的识别算法，通过算法的优化避免了错误码字对识别结果的影响，误码适应能力得到提高。文献[10]首次引入软判决序列对本原 BCH 码进行识别，通过计算码字多项式的码根 KL (Kullback-Leibler) 散度，利用码根概率分布识别出码长、本原多项式以及生成多项式，识别性能得到较大的提高。针对 RS (Reed-Solomon) 码识别问题，文献[11]提出了基于伽罗华域中高斯消元的识别算法，该算法在码长较短的情况下具有一定实用性，但是计算复杂度会随着码长增加而急剧增加，且容错性能较差。文献[12]从提高容错性能出发，提出基于 GFFT 的识别算法，虽然算法容错性能得到了提升，但运算量随着码长增加会急剧增加。文献[13]建立了基于似然判决的二元域统计判决模型，引入了能够衡

量校验关系成立大小的平均校验符合度概念，通过遍历的方式对参数进行匹配，该算法具有较好的低信噪比适应能力。此外，针对卷积码^[14-15]、低密度奇偶校验 (LDPC, low density parity check) 码^[16-17]、Turbo 码^[18-19]等不同编码方式的识别问题，也有大量的研究。

由于极化码特有的编码理论与方法，传统的线性分组码识别算法并不适用于极化码，且目前针对极化码识别的研究较少。其中，文献[20]通过对信道删除概率和极化码码长进行遍历，求得对应的生成矩阵及其零空间，并利用零空间对硬判决码字矩阵依次进行校验，确定编码参数。文献[21]通过对码长和冻结比特进行遍历，依次选取冻结比特求得所对应的对偶向量，利用对偶向量对软判决码字矩阵进行校验，判断码字是否选取该比特作为其冻结比特，但在对统计量计算的过程中进行了近似处理。现有算法由于未有效利用极化码的相关性质，均需要采用参数遍历的方式对参数空间进行搜索，得到参数的估计值，导致算法过程较复杂、计算量较大，同时在计算过程中采用硬判决序列或对统计判决量进行近似，算法的误码适应能力不足。

针对上述问题，本文同样对标准非删余极化码的编码过程展开了研究。通过对极化码相关性质的研究，证明了描述极化码生成矩阵、码长、信息比特和冻结比特之间特有关系的定理与命题；利用信息比特的分布规律，实现码长的识别；引入似然差 (LD, likelihood difference) 作为检验关系的判决统计量，并从理论上分析了似然差的统计特性，基于最大最小准则对最优判决门限进行了求解，实现信息比特与冻结比特的识别，对码率进行估计。所提算法充分利用了极化码生成矩阵的结构特征，在保证识别性能的前提下，显著降低了运算量。仿真证明了所提算法的有效性。若无特殊说明，下文极化码均指标准极化码。

1 信道极化

极化码的产生来源于信道极化理论。对于一个二进制离散无记忆信道 $W: \mathcal{X} \rightarrow \mathcal{Y}$ ，其转移概率为 $W(y|x)$ ， $x \in \mathcal{X}$ ， $y \in \mathcal{Y}$ 。信道极化的过程包含 2 个阶段：信道合并和信道分裂。信道合并是指 N 个独立信道经过线性变换合并为信道 $W_N: \mathcal{X}^N \rightarrow \mathcal{Y}^N$ ，信道分裂则是将信道 W^N 分解成 N 个子信道 $W_N^{(i)}$ 。图 1 为标准极化码一级信道结构，将 2 个比

特进行编码, 可得

$$(x_0, x_1) = (u_0 \oplus u_1, u_1) \quad (1)$$

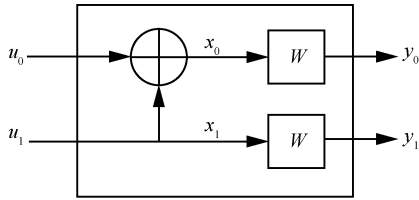


图 1 标准极化码一级信道结构

将 x_0 和 x_1 分别通过信道 W 进行传输, 得到 y_0 和 y_1 , 则合并后的信道 W_2 的转移概率为

$$W_2(y_0, y_1 | u_0, u_1) = W(y_0 | u_0 \oplus u_1)W(y_1 | u_1) \quad (2)$$

多级级联进行上述操作, 最终可以实现多级极化。经过 $n = \text{lb}N$ 次极化后, 得到的合并信道 W_N 的转移概率^[1]为

$$W_N(y_0^{N-1} | u_0^{N-1}) = W^N(y_0^{N-1} | u_0^{N-1} \mathbf{B}_N \mathbf{F}_N) \quad (3)$$

其中, \mathbf{B}_N 表示比特逆序重排矩阵, $\mathbf{F}_N = \mathbf{F}^{\otimes n}$, $\mathbf{F} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, \otimes 表示 Kronecker 积。对于非标准极化码, 也可采用 3×3 或更高阶次的核 \mathbf{F} 。分裂后的子信道 $W_N^{(i)}$ 的转移概率^[1]为

$$W_N^{(i)}(y_0^{N-1}, \hat{u}_0^{i-1} | u_i) = \sum_{u_0^{N-1} \in X^{N-1}} \frac{1}{2^{n-1}} W_N(y_0^{N-1} | u_0^{N-1}) \quad (4)$$

其中, $(y_0^{N-1}, \hat{u}_0^{i-1})$ 与 u_i 为 $W_N^{(i)}$ 的输入与输出。

1.1 极化信道可靠性估计

对于二进制删除信道 (BEC, binary erasure channel), 为描述极化信道的可靠性, 可以引入巴氏参数^[1]作为表征信道 W 质量的参数, 其定义为

$$Z(W) = \sum_{y \in \mathcal{Y}} \sqrt{W(y|0)W(y|1)} \quad (5)$$

当信道 W 只用来发射单个比特时, 巴氏参数是使用最大后验概率判决时错误概率的上界。当信道为 BEC 时, $Z(W)$ 可以通过式(6)计算得到^[1]。

$$\begin{cases} Z(W_N^{(2i)}) = 2Z(W_{N/2}^{(i)}) - Z(W_{N/2}^{(i)})^2 \\ Z(W_N^{(2i+1)}) = Z(W_{N/2}^{(i)})^2 \end{cases} \quad (6)$$

其中, $Z(W_1^{(0)}) = \varepsilon$, ε 为信道删除概率。 $Z(W_N^{(i)})$ 越大, 信道就越不可靠^[1]。因此, 编码过程中需要选择巴氏参数较小的 k 个子信道为信息比特, 其余子信道则为冻结比特。随着码长的增加, 子信道的容量将趋近于 2 个极端, 一部分子信道容量趋近于 1,

剩余的子信道容量则会趋近于 0。图 2 给出了 BEC 在码长为 1 024、删除概率为 0.5 时, 各子信道巴氏参数的分布情况。从图 2 中可以明显看出, 子信道的巴氏参数分布趋近于 0 和 1 这 2 个极端。

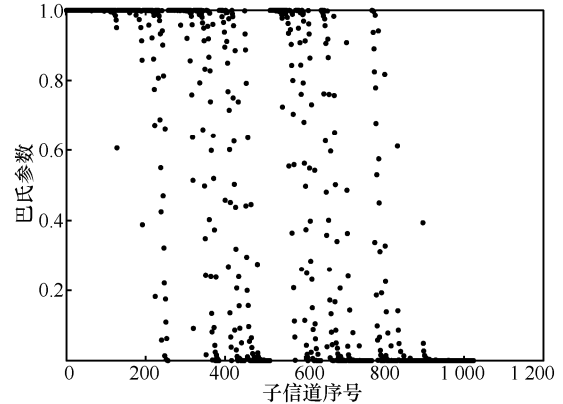


图 2 BEC 子信道巴氏参数分布

1.2 极化码编码

极化码利用可靠性较高的子信道来传输有用的信息比特, 而利用可靠性较差的子信道传输冻结比特。极化码的编码过程用生成矩阵的形式可以表示为

$$\mathbf{c}_{1 \times N} = \mathbf{u}_{1 \times N} \mathbf{G}_N \quad (7)$$

其中, $\mathbf{u}_{1 \times N}$ 为原始比特序列, $\mathbf{c}_{1 \times N}$ 为编码后的比特序列, \mathbf{G}_N 为生成矩阵, $N = 2^n$ 为码长, n 为正整数。生成矩阵表示为

$$\mathbf{G}_N = \mathbf{B}_N \mathbf{F}^{\otimes n} \quad (8)$$

容易证明, 生成矩阵 \mathbf{G}_N 是一个 $N \times N$ 维的满秩矩阵, 其行向量与子信道一一对应。假设 A 是一个元素个数为 k 的正整数的集合, A 中每个元素为信息比特所对应子信道的行标号, 则其补集 A^c 中的元素为冻结比特所对应子信道的行标号, 码率为 $\frac{\text{card}(A)}{N}$ 。 \mathbf{u}_A 为要传输的信息比特, \mathbf{u}_{A^c} 为冻结比特, 于是编码过程可表示为

$$\mathbf{c}_{1 \times N} = \mathbf{u}_A \mathbf{G}_N(A) + \mathbf{u}_{A^c} \mathbf{G}_N(A^c) \quad (9)$$

一般情况下, 极化码的冻结比特均取 0, 此时编码过程可进一步简化为

$$\mathbf{c}_{1 \times N} = \mathbf{u}_A \mathbf{G}_N(A) \quad (10)$$

由极化码编码原理可知, 当给定编码长度, 极化码的编译码结构将被唯一确定, 且可以通过生成矩阵的形式表示编码过程。由此可知, 标准非删余

极化码识别最终目的是完成对极化码码长、信息比特及冻结比特的识别。通常在实际通信系统中，采用线性分组码的比特流序列通常有固定的帧结构和相应的同步码^[21]，且已有相应方法能够实现对帧同步码的盲识别^[22-23]，故本文假设已完成帧同步，集中对极化码参数识别问题进行讨论。除特别说明外，本文矩阵运算均在二元域 $F(2)$ 上进行。

2 极化码识别原理与算法

2.1 识别原理

根据极化码编码原理可知，标准非删余极化码码长通常为 2^n ($n > 0$)，不同码长的极化码对应着不同的生成矩阵 G_N 。在给出具体的识别算法之前，首先对矩阵 G_N 的相关性质进行研究，给出相关的定理及命题，作为极化码参数识别算法的理论基础。

定理 1 对于矩阵 $F^{\otimes n}$ ，具有如下性质

$$F^{\otimes n} F^{\otimes n} = I \quad (11)$$

定理 1 利用数学归纳法易证得。根据定理 1 可知，对于矩阵 $F^{\otimes n}$ ，当剔除其第 i 行后，得到的新矩阵为 $F_{-i}^{\otimes n}$ ，矩阵 $F_{-i}^{\otimes n}$ 的对偶向量即 $F^{\otimes n}$ 的第 i 列。根据极化码编码性质可知，生成矩阵 G_N 与矩阵 $F^{\otimes n}$ 由相同的行向量组成，只是行向量的排列顺序不同，其对应关系由矩阵 B_N 所决定，易证得

$$G_N F^{\otimes n} = B_N \quad (12)$$

显然，矩阵 G_N 的行向量与矩阵 $F^{\otimes n}$ 的列向量之间仍有一一对应的关系，即对于矩阵 G_N 中任意一行向量，矩阵 $F^{\otimes n}$ 中有且仅有一个列向量不与其正交。该性质是本文算法的基础，对于非标准极化码，如采用 3×3 或更高阶次的核 \hat{F} ，当 \hat{F} 满足 $\hat{F}^{\otimes n} \hat{F}^{\otimes n} = I$ 时，下文所证明的性质及所提策略仍然适用，只需对具体算法做出适当修改；但当 \hat{F} 不满足 $\hat{F}^{\otimes n} \hat{F}^{\otimes n} = I$ 时，本文算法将不再适用。

考虑码长为 $N = 2^n$ 、码率为 k/N 的极化码，当码字 $(c_1, c_2, \dots, c_N), \dots, (c_{(L-1)N+1}, c_{(L-1)N+2}, \dots, c_{LN})$ 被拓展为 $(c_1, c_2, \dots, c_N, \dots, c_{(L-1)N+1}, c_{(L-1)N+2}, \dots, c_{LN})$ 时，拓展后的码字等价于码长为 LN ($L = 2^m$) 的新极化码，拓展后的极化码生成矩阵可记为 $G_N^L = I_L \otimes G_N$ ，其中 I_L 为维数为 $L \times L$ 的单位矩阵。

命题 1 对于拓展后码长为 LN ($L = 2^m$) 的新极化码，记其对应的 $(n+m)$ 次克罗内克积为

$F^{\otimes(n+m)}$ ，生成矩阵为 G_N^L ，则有

$$G_N^L F^{\otimes(n+m)} = F^{\otimes m} \otimes B_N \quad (13)$$

证明

$$\begin{aligned} G_N^L F^{\otimes(n+m)} &= (I_L \otimes G_N)(F^{\otimes(n+m)}) = \\ &= (I_L \otimes G_N)(F^{\otimes m} \otimes F^{\otimes n}) = \\ &= F^{\otimes m} \otimes (G_N F^{\otimes n}) = F^{\otimes m} \otimes B_N \end{aligned} \quad (14)$$

证毕。

命题 1 说明，生成矩阵 G_N^L 与矩阵 $F^{\otimes(n+m)}$ 相乘所得结果具有特定的矩阵结构，且该结构受 m 和 N 值的直接影响，通过两矩阵相乘结果能够求得 m 和 N 的值。例如，当 $m=1$, $N=4$ ，两矩阵相乘结果为

$$G_4^L F^{\otimes 3} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (15)$$

2.2 冻结比特与信息比特识别

假设接收到的码流序列共包含 M 个完整码字，记为 r ，其中 N 满足 $N = 2^n$ 。考虑命题 1 中 $m=1$ 时的情况，即码字矩阵的设定码长与真实码长相等，利用截获的码流序列 r 构建码字矩阵 R

$$R = \begin{bmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,N} \\ r_{2,1} & r_{2,2} & \cdots & r_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ r_{M,1} & r_{M,2} & \cdots & r_{M,N} \end{bmatrix} \quad (16)$$

在无误码条件下，根据式(10)，接收到的码字可表示为 $R = C = UG_N(A)$ ，其中 U 为对应的信息矩阵。记码字矩阵 R 与矩阵 $F^{\otimes n}$ 相乘所得矩阵为 Φ 。

命题 2 在无误码条件下，矩阵 Φ 中零列向量的数量为 $n-k$ ，零列向量的位置与冻结比特的位置一一对应。

证明 对于矩阵 R 与矩阵 $F^{\otimes n}$ ，可得

$$\begin{aligned} \Phi &= RF^{\otimes n} = UG_N(A)F^{\otimes n} = \\ &= UP_{k \times N} G_N F^{\otimes n} = UP_{k \times N} B_N F^{\otimes n} F^{\otimes n} = UP_{k \times N} B_N \end{aligned} \quad (17)$$

其中，矩阵 $P_{k \times N}$ 为行满秩矩阵，用于选出 G_N 中行标号属于集合 A 的行。考虑到矩阵 B_N 为行置换矩阵，每行每列有且仅有一个元素为 1，行与行之间、

列与列之间 1 的位置均不同, 可将式(17)化简为

$$\Phi = UP_{k \times N}B_N = U\Gamma \quad (18)$$

其中, $\Gamma = [\gamma_1^T \ \gamma_2^T \ \cdots \ \gamma_k^T]^T$, γ_i 为矩阵 Γ 的第 i 个行向量。根据矩阵 $P_{k \times N}$ 的性质可知, 矩阵 Γ 由矩阵 B_N 的 k 个行向量组成。矩阵 Γ 第 i 行 γ_i 的元素, 只在某一特定位置 $(\alpha(i))$ 为 1, 其他位置均为 0, 即 $\gamma_i = [0_{1 \times (\alpha(i)-1)} \ 1 \ 0_{1 \times (N-\alpha(i))}]$ 。考虑到矩阵 U 为信息矩阵, 通常该矩阵为列满秩矩阵, 则矩阵 Γ 的行向量为矩阵 $U\Gamma$ 的基。因此矩阵 $U\Gamma$ 必存在 $N-k$ 个全零列向量, 即码字矩阵 R 与矩阵 $F^{\otimes n}$ 的 $N-k$ 个列满足校验关系, 记该 $N-k$ 个列的列标号集合为 B , 即 $B = \{b_1, b_2, \dots, b_{N-k}\}$ 。

根据极化码编码性质可知, 生成矩阵 G_N 与矩阵 $F^{\otimes n}$ 中行向量所处位置的对应关系由矩阵 B_N 决定。因此根据命题 1 可知, 当剔除生成矩阵 G_N 第 i 行后得到矩阵 G_{N-i} , 存在由 B_N 决定的变换 $\pi(\cdot)$, 使矩阵 G_{N-i} 的对偶向量为 $F^{\otimes n}$ 的第 $\pi(i)$ 列。则信息比特和冻结比特所对应子信道的行标号集合分别为 $A = \{\pi^{-1}[\alpha(1)], \pi^{-1}[\alpha(2)], \dots, \pi^{-1}[\alpha(k)]\}$ 和 $A^c = \{\pi^{-1}[b_1], \pi^{-1}[b_2], \dots, \pi^{-1}[b_{N-k}]\}$ 。证毕。

因此, 根据命题 2 可知, 当码字矩阵设定的码长与真实码长相等时, 只需要考察码字矩阵 R 与矩阵 $F^{\otimes n}$ 各列之间的校验关系, 并利用映射关系 $\pi(\cdot)$ 即可求得集合 A 。当矩阵 $F^{\otimes n}$ 第 i 列与码字矩阵 R 不满足校验关系时, $\pi^{-1}(i)$ 为集合 A^c 中的一个元素, 否则 $\pi^{-1}(i)$ 为集合 A 中的一个元素。同时, 码率估计值为 $\frac{\text{card}(A)}{N}$ 。

2.3 码长识别

根据 2.2 节分析可知, 当码字矩阵设定的码长与真实码字长度相同时, 利用码字矩阵 R 与矩阵 $F^{\otimes n}$ 的运算结果, 可以实现信息比特和冻结比特置的识别。但实际接收中, 无法事先知道极化码码长, 因此本节将继续对码长识别进行分析。

首先, 考虑无误码情况。当码字矩阵设定的码长大于真实码长时, 即 $N' = LN = 2^{n+m}$ ($m > 1$), 可构造码字矩阵 R_L 。此时码长为 $N = 2^n$ 、码率为 k/N 的极化码, 被拓展为码长为 LN ($L = 2^m$) 的新极化码。接收到的码字矩阵可表示为 $R_L = U_L G_N^L(A)$, 其中 U_L 为对应的信息矩阵。记码字矩阵 R_L 与矩阵 $F^{\otimes(n+m)}$ 相乘所得矩阵为 Φ_L 。

命题 3 在无误码条件下, 矩阵 Φ_L 中零列向量的数量为 $(n-k)L$, 且周期性出现, 周期为 N 。

证明 根据命题 1 与式(17)可得

$$\begin{aligned} \Phi_L &= R_L F^{\otimes(n+m)} = U_L G_N^L(A) F^{\otimes(n+m)} = \\ &U_L \begin{bmatrix} G_N(A) & & \\ & \ddots & \\ & & G_N(A) \end{bmatrix} F^{\otimes m} \otimes F^{\otimes n} = \\ &U_L \left((I_L \otimes G_N(A)) (F^{\otimes m} \otimes F^{\otimes n}) \right) = \\ &U_L (F^{\otimes m} \otimes G_N(A) F^{\otimes n}) = U_L (F^{\otimes m} \otimes \Gamma) \quad (19) \end{aligned}$$

根据 2.2 节分析可知, 矩阵 Γ 的任意一行 γ_i 的元素, 只有在 $\alpha(i)$ 的位置为 1, 其他位置均为 0。因此, 矩阵 $F^{\otimes m} \otimes \Gamma$ 将周期性出现零列向量, 周期为 N 。证毕。

根据命题 3 可知, 矩阵 $F^{\otimes(n+m)}$ 中与码字矩阵 R_L 满足校验关系的列将周期性出现, 其周期与码长、矩阵 $F^{\otimes(n+m)}$ 存在严格约束关系。因此, 当接收码字不存在误码时, 选择合适的码长并计算矩阵 Φ_L , 根据矩阵中零列向量的分布情况可以得到码字矩阵设定的码长与真实码长之间的关系, 计算出真实码长。

但当接收序列存在误码时, 码字矩阵 R_L 与矩阵 $F^{\otimes(n+m)}$ 之间的校验关系将被破坏。另一方面, 对于矩阵 $F^{\otimes(n+m)}$ 中码重较重的列向量而言, 由于码重较大, 也更容易受误码影响, 导致校验关系失效。

根据矩阵 $F^{\otimes n}$ 构造原理可知, $F^{\otimes n}$ 满足性质

$$F^{\otimes n} = \begin{bmatrix} F^{\otimes(n-1)} & \mathbf{0} \\ F^{\otimes(n-1)} & F^{\otimes(n-1)} \end{bmatrix} \quad (20)$$

因此, 将 $F^{\otimes n}$ 的后 2^k 列 ($k = 2, \dots, n$) 记为 P_k , 则有

$$P_k = \begin{bmatrix} \mathbf{0} \\ (2^n - 2^k) \times 2^k \\ F^{\otimes k} \end{bmatrix} \quad (21)$$

即码字矩阵 R_L 与矩阵 $F^{\otimes(m+n)}$ 进行校验关系检测时, 其后 2^k 个校验值, 等价于码字矩阵 R_k 与矩阵 $F^{\otimes k}$ 进行校验关系的检验。因此, 可将矩阵 $F^{\otimes k}$ 中未通过正交检验的列的数量与 2^k 的比值, 作为码字矩阵码长为 2^k 时的等效极化码码率估计值, 记为 $\hat{\eta}$ 。根据文献[21]中定理 1 可知, 在无误码情况下, 当码字矩阵设定的码长小于真实码长时, 等效极化码码率 $\hat{\eta}$ 将大于真实码率; 当码字矩阵设定的码长大于或等于真实码长时, 其 $\hat{\eta}$ 将等于真实码率。但又由于误码的影响, 部分校验关系将被破坏, 使通过检验的冻结比

特减少，码率增加。因此在误码条件下，当码字矩阵设定的码长等于真实码长时，所求得码率 $\hat{\eta}$ 为最低码率，可利用该性质判断极化码码长。

2.4 校验关系检测

在实际信息传输过程中，码字会受到噪声的影响而产生误码，由于硬判决序列只包含序列判决结果，而软判决序列不仅能提供 0/1 比特数据，还能额外提供判决序列的可靠性信息。因此，本文采用软判决序列进行处理，对编码参数进行识别。为与前文区分，由软判决序列构成的码字矩阵记为 \mathbf{T}_L 。

假设信号采用二进制相移键控 (BPSK, binary phase-shift keying) 调制，在加性白高斯噪声 (AWGN, additive white Gaussian noise) 信道条件下， $t_{i,j} = s_{i,j} + n_{i,j}$ ，其中 $n_{i,j}$ 为均值为 0、方差为 σ^2 的高斯白噪声序列， $s_{i,j}$ 为 BPSK 调制符号，与原始码字 $c_{i,j}$ 之间满足

$$s_{i,j} = 2c_{i,j} - 1 \quad (22)$$

引入对数似然比 (LLR, log-likelihood ratio)，定义为^[24]

$$l_{i,j} = \ln \left(\frac{P_r(c_{i,j} = 0 | r_{i,j})}{P_r(c_{i,j} = 1 | r_{i,j})} \right) = -\frac{2t_{i,j}}{\sigma^2} \quad (23)$$

记矩阵 $\mathbf{F}^{\otimes n}$ 中第 j 列为 \mathbf{f}^j ，为了研究码字与 \mathbf{f}^j 的校验关系，单独取码字矩阵 \mathbf{T}_L 第 i 行 \mathbf{t}_i 进行研究，此时对应的码字为 \mathbf{c}_i ，可以将式(23)进行推广得到 \mathbf{c}_i 与 \mathbf{f}^j 正交的对数似然比为

$$L_{i,j} = \ln \frac{P_r \left(\sum_{t=1}^N \oplus c_{i,t} f_{t,j} = 0 | \mathbf{t}_i \right)}{P_r \left(\sum_{t=1}^N \oplus c_{i,t} f_{t,j} = 1 | \mathbf{t}_i \right)} = 2 \arctan \left(\prod_{t=1}^{d_j} \tanh \left(-\frac{t_{i,l_{j,t}}}{\sigma^2} \right) \right) \quad (24)$$

其中， $\mathbf{f}^j = [f_{1,j}, f_{2,j}, \dots, f_{N,j}]^T$ ， $\mathbf{t}_i = [t_{i,1}, t_{i,2}, \dots, t_{i,N}]$ ， $\mathbf{c}_i = [c_{i,1}, c_{i,2}, \dots, c_{i,N}]$ ， $l_{j,t}$ 为 \mathbf{f}^j 第 t 个非零值的索引。在计算 LLR 时，需要多次进行较复杂的双曲正切及其反函数的运算。为降低运算量，引入似然差 LD 来替代 LLR，其定义为^[25]

$$LD_{i,j} = P_r \left(\sum_{t=1}^N \oplus c_{i,t} f_{t,j} = 0 | \mathbf{t}_i \right) - P_r \left(\sum_{t=1}^N \oplus c_{i,t} f_{t,j} = 1 | \mathbf{t}_i \right) = \prod_{t=1}^{d_k} \tanh \left(-\frac{t_{i,l_{j,t}}}{\sigma^2} \right) \quad (25)$$

考虑以下两类假设检验。

H_1 ： \mathbf{f}^j 与极化码码字构成校验关系。

H_0 ： \mathbf{f}^j 与极化码码字不构成校验关系。

设 $w_i = \text{weight}(\mathbf{f}^j)$ ，即向量 \mathbf{f}^j 中含有 w_i 个“1”。当校验关系成立时，意味着码字 \mathbf{c}_i 中下标为 $\{l_{j,t}\}$ 的码元中应有偶数个码元为“1”，共有 $\sum_{k=0}^{\lfloor \frac{w_i}{2} \rfloor} \binom{w_i}{2k}$ 种可能。因此，当 H_1 成立时， LD_j 的均值和方差分别为

$$\begin{aligned} \mu(LD_j | H_1) &= \frac{1}{2^{w_i-1}} \sum_{k=0}^{\lfloor \frac{w_i}{2} \rfloor} \binom{w_i}{2k} \cdot \\ &\left(\int \tanh \left(-\frac{x}{\sigma^2} \right) P(x | c=1) dx \right)^{2k} \cdot \\ &\left(\int \tanh \left(-\frac{x}{\sigma^2} \right) P(x | c=0) dx \right)^{w_i-2k} \end{aligned} \quad (26)$$

$$\begin{aligned} \sigma_D^2(LD_j | H_1) &= -\mu^2(LD | H_1) + \\ &\frac{1}{2^{w_i-1}} \sum_{k=0}^{\lfloor \frac{w_i}{2} \rfloor} \binom{w_i}{2k} \left(\int \left(\tanh \left(-\frac{x}{\sigma^2} \right) \right)^2 P(x | c=1) dx \right)^{2k} \cdot \\ &\left(\int \left(\tanh \left(-\frac{x}{\sigma^2} \right) \right)^2 P(x | c=0) dx \right)^{w_i-2k} \end{aligned} \quad (27)$$

当 H_0 成立时，由于码字与向量 \mathbf{f}^j 之间不存在校验关系， LD_j 的均值和方差分别为

$$\mu(LD_j | H_0) = 0 \quad (28)$$

$$\begin{aligned} \sigma_D^2(LD_j | H_0) &= \frac{1}{2^{w_i}} \sum_{k=0}^{w_i} \binom{w_i}{k} \cdot \\ &\left(\int \left(\tanh \left(-\frac{x}{\sigma^2} \right) \right)^2 P(x | c=1) dx \right)^k \cdot \\ &\left(\int \left(\tanh \left(-\frac{x}{\sigma^2} \right) \right)^2 P(x | c=0) dx \right)^{w_i-k} \end{aligned} \quad (29)$$

通过定义可知，若函数 $\tanh \left(-\frac{x}{\sigma^2} \right)$ 为奇函数， $P(x | c=1)$ 与 $P(x | c=0)$ 则互为对称函数，可得

$$\begin{aligned} \zeta &= \int \tanh \left(-\frac{x}{\sigma^2} \right) P(x | c=0) dx = \\ &-\int \tanh \left(-\frac{x}{\sigma^2} \right) P(x | c=1) dx \end{aligned} \quad (30)$$

代入式(26)可得

$$\mu(\text{LD}_j | H_1) = \frac{1}{2^{w_j-1}} \sum_{k=0}^{\lfloor \frac{w_j}{2} \rfloor} \binom{w_j}{2k} (-\zeta)^{2k} \zeta^{w_j-2k} = \zeta^{w_j} \quad (31)$$

同理可得

$$\sigma_D^2(\text{LD}_j | H_1) = \zeta^{w_j} - \zeta^{2w_j} \quad (32)$$

$$\sigma_D^2(\text{LD}_j | H_0) = \zeta^{w_j} \quad (33)$$

式(31)~式(33)只考虑了第*i*个码字的似然差,将码字矩阵的所有行考虑进去,则可得到平均似然差,即

$$\overline{\text{LD}}_j = \frac{1}{M} \sum_{i=1}^M \text{LD}_{i,j} \quad (34)$$

本文将 $\overline{\text{LD}}_j$ 作为 $\mathbf{F}^{\otimes n}$ 中第*j*列 \mathbf{f}^j 与码字构成校验关系的检验量。根据中心极限定理,平均似然差 $\overline{\text{LD}}_j$ 的分布可以近似为正态分布,其均值与方差分别为

$$\begin{cases} \mu_1 = \mu(\overline{\text{LD}}_j | H_1) = \mu(\text{LD}_j | H_1) \\ \sigma_1^2 = \sigma^2(\overline{\text{LD}}_j | H_1) = \frac{\sigma_D^2(\text{LD}_j | H_1)}{M} \\ \mu_0 = \mu(\overline{\text{LD}}_j | H_0) = \mu(\text{LD}_j | H_0) = 0 \\ \sigma_0^2 = \sigma^2(\overline{\text{LD}}_j | H_0) = \frac{\sigma_D^2(\text{LD}_j | H_0)}{M} \end{cases} \quad (35)$$

设两类假设的判决门限为 Λ ,由于在非合作通信系统中无法获得两类假设的先验信息,因此本文采用极小化极大准则来求解门限 Λ 。同时,假定正确判断的代价因子取值为0,错误判断代价因子取值为1。根据极小化极大准则可知,门限 Λ 需满足

$$\int_{\Lambda}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma_0} e^{-\frac{(x-\mu_0)^2}{2\sigma_0^2}} dx = \int_{-\infty}^{\Lambda} \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} dx \quad (36)$$

根据 $\frac{x-\mu_0}{\sigma_0} \sim N(0,1)$, $\frac{x-\mu_1}{\sigma_1} \sim N(0,1)$,以及标

准正态分布上下分位点的性质关系可知

$$\frac{\Lambda - \mu_0}{\sigma_0} + \frac{\Lambda - \mu_1}{\sigma_1} = 0 \quad (37)$$

解得门限为

$$\Lambda = \frac{\sigma_0 \mu_1 + \sigma_1 \mu_0}{\sigma_0 + \sigma_1} \quad (38)$$

根据设定的码长,利用截获的数据构造出码字矩阵,然后依次利用矩阵 $\mathbf{F}^{\otimes n_{\max}}$ 中列向量与码字矩阵进行校验,计算出统计校验量 $\overline{\text{LD}}_j$ 与对应的门限 Λ 。当统计校验量大于门限时,则认为该位置为冻

结比特,反之则为信息比特。

2.5 识别算法流程

由前文分析可知,合理设定接收码字矩阵的码长值,构建码字矩阵 \mathbf{R}_L 与矩阵 $\mathbf{F}^{\otimes n_{\max}}$,依次对码字矩阵 \mathbf{R}_L 与矩阵 $\mathbf{F}^{\otimes n_{\max}}$ 后 2^k ($k=2,3,\dots$)列之间的校验关系进行检验,当某一个列向量的校验关系成立时,矩阵 \mathbf{G}_N 中与其对应的行即冻结子信道的所在行。然后计算等效极化码码率的估计值,对码长、码率进行估计,即通过一次计算完成参数的识别,而不需要像现有算法需要对所有可能码长进行遍历,有效降低运算量,具体如算法1所示。

算法 1 基于编码矩阵结构特征的极化码参数识别算法

输入 判决序列 \mathbf{r} , 码字矩阵的码长 $2^{n_{\max}}$

输出 码长 \hat{n} 、信息比特集合 \hat{A}

1) 构造矩阵 \mathbf{F} 的 n_{\max} 次克罗内克矩阵 $\mathbf{F}^{\otimes n_{\max}}$

2) $B = \emptyset$, $\hat{\eta} = \mathbf{0}_{1 \times n_{\max}}$, $k = 1$

3) while 1

4) 对码字矩阵与矩阵 $\mathbf{F}^{\otimes n_{\max}}$ 的第 $2^{\max} - 2^k + 1$ 至 $2^{\max} - 2^{k-1} + 1$ 列向量之间的关系进行检验;

5) if \mathbf{f}^j 不符合校验关系

6) $B = [B, j]$

7) end if

8) 计算向量后 2^k 元素中不符合校验关系的列向量的占比 $\hat{\eta}(k)$;

9) if $(\hat{\eta}(k) \geq \hat{\eta}(k-1)) \& (k > 2)$

10) index = $k - 1$

11) break;

12) else if $k = n_{\max}$

13) index = k

14) break;

15) end if

16) $k = k + 1$

17) end

18) $\hat{N} = 2^{\text{index}}$

19) $\hat{A} = \{\pi^{-1}(b - 2^{\hat{N}}) | b \in B, 2^{n_{\max}} - 2^{\hat{N}} + 1 \leq b \leq 2^{n_{\max}}\}$

3 仿真分析

考虑到5G通信中增强移动宽带场景下控制信道编码方案,确定采用的极化码最大母码长度在下

行控制信道中为 512，在上行控制信道中为 1 024，最低码率为 1/8。因此，仿真设定极化码码长分别为 32、64、128、256、512、1 024，最大码长 $2^{n_{\max}}$ 为 1 024。不失一般性，设噪声环境为 AWGN 信道。

3.1 统计模型验证

本节主要对前文推导的似然差统计特性进行验证。前文在推导似然差分布函数的过程中进行了一定假设，其正确性是算法有效求解判决门限的前提条件。根据矩阵 $\mathbf{F}^{\otimes n_{\max}}$ 的构造规律可知，其列向量的码重取值集合为 $\{2^k | k \in \{1, 2, \dots, n_{\max}\}\}$ 。仿真设定了 4 种码重值，分别为 32、64、256、1 024。设定的信噪比范围为 $-10 \sim 12$ dB，步长为 0.5 dB。利用均值与方差定义，仿真求得在不同条件下平均似然差的均值与方差，并和理论推导计算得到的均值与方差进行对比。

从图 3 中可以看出，在不同假设条件下，利用定义所求得的均值与方差和理论计算得到的值几乎完全重合，证明了前文中对于似然差分布函数推导的正确性。其中，图 3(b) 中的方差曲线在某信噪比下出现峰值，其原因是当 H_1 成立时，矩阵 $\mathbf{F}^{\otimes(m+n)}$ 中任意列向量 \mathbf{f}^j 的似然差均值为 ζ ，其方差为 $\zeta^{w_i} - \zeta^{2w_i}$ 。 ζ 可由数值计算求得，其取值范围为 $(0, 1)$ 。同时，对于函数 $f(\zeta) = \zeta^{w_i} - \zeta^{2w_i}$ ，当 $\zeta \in (0, 1)$ 时，函数 $f(\zeta)$ 在 $\zeta = 2^{-1/w_i}$ 时取得极大值 0.25。图 3(b) 的横坐标为信噪比，而随着信噪比的提高， ζ 值逐渐趋近于 1，因此方差曲线会在某一信噪比时取得峰值。图 4 为 w_i 取不同值时函数 $f(\zeta) = \zeta^{w_i} - \zeta^{2w_i}$ 在 $(0, 1)$ 的取值。

3.2 算法有效性验证

本节主要利用仿真对命题 1 及算法的有效性进行验证。本节所采用极化码码率为 1/2，码长为 64，设定的最长码长 $2^{n_{\max}}$ 为 128。信噪比分别取 5 dB 和 0，设定码字个数为 2 000 个。在 2 种信噪比条件下，记录不同列的平均似然差值和门限，以及不同码字矩阵码长的等效码率。不同信噪比下的平均似然差与门限及判决结果分别如图 5 和图 6 所示。

图 5(a) 给出了不同列的平均似然差和对应的判决门限。当似然差大于门限时，判定校验关系成立，该列向量对应的 $\pi^{-1}(i)$ 为冻结比特，否则为信息比特。图 5(b) 则给出了似然差判决结果。从图 5(b) 可以看出，当信噪比为 5 dB 时，接收信号的误码率较低，校验关系检测结果的序列中存在 2 个完整的周期。同时，图 5(c) 中等效码率也随着码字矩阵码长

的增加而降低，当 $n=5$ 时降到最低，与仿真设定值相对应，且与 2.3 节分析结果一致。

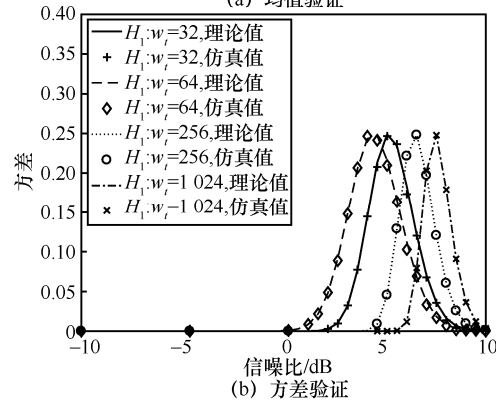
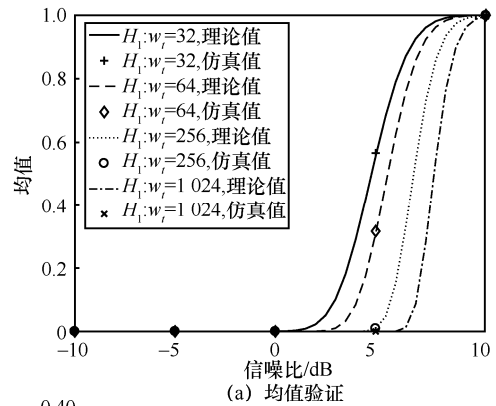


图 3 似然差统计特性对比

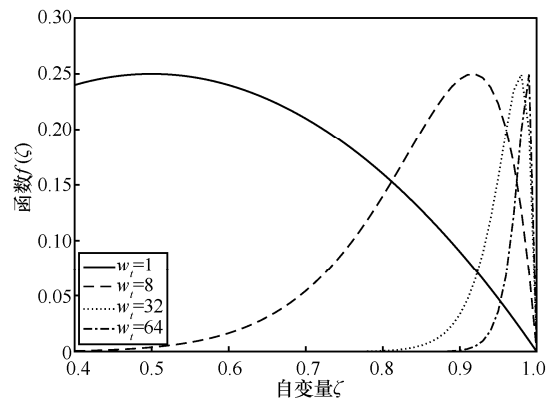
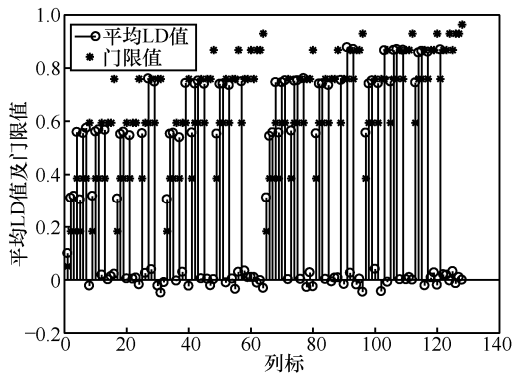
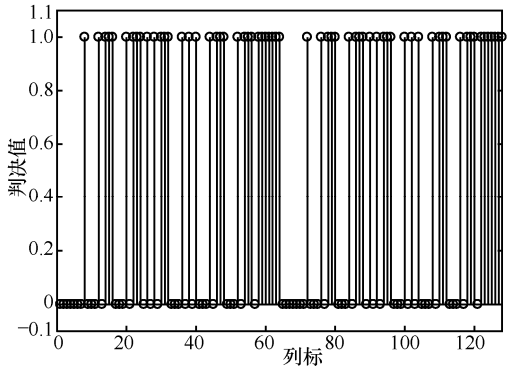


图 4 w_i 取不同值时函数 $f(\zeta)$ 在 $(0, 1)$ 的取值

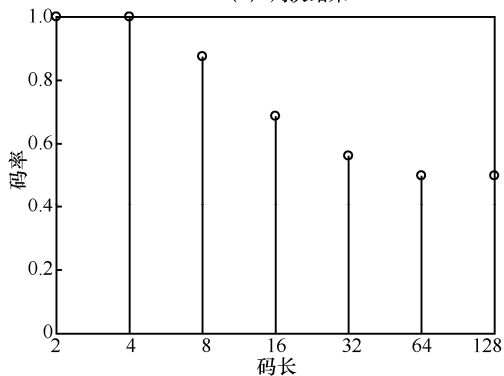
当信噪比降到 0 时，接收信号的质量下降，导致了更多的误判。从图 6(b) 中可以看出，判决结果前后两段已经不同，即由于误码的影响，判决结果出现了错误，周期性被破坏。从图 6(c) 可以看出，当码字矩阵码长小于真实码长时，码率大于真实码长的码率；当码字矩阵码长大于真实码长时，码率同样大于真实码长的码率，与 2.3 节的分析一致。



(a) 似然差与门限

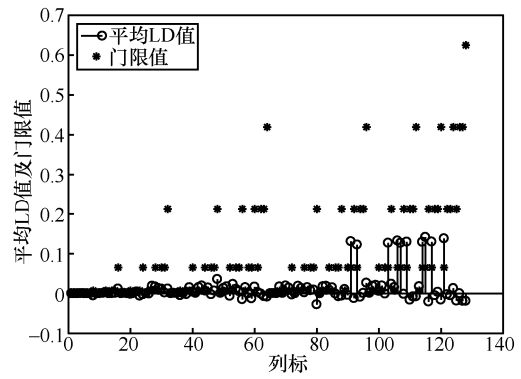


(b) 判决结果

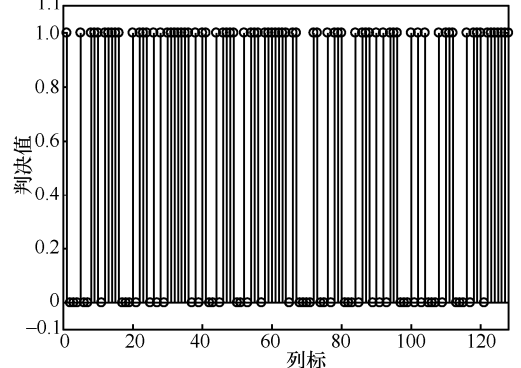


(c) 不同码长对应的等效码率

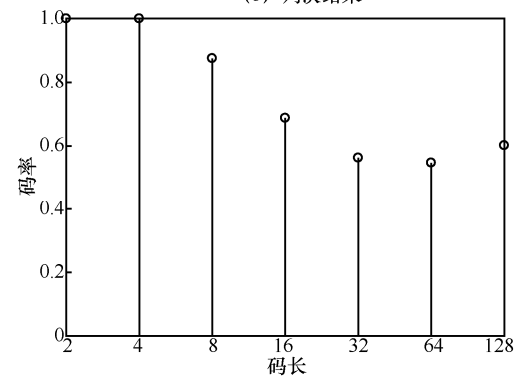
图 5 $N=64$, $SNR=5$ dB 条件下判决结果



(a) 似然差与门限



(b) 判决结果



(c) 不同码长对应的等效码率

图 6 $N=64$, $SNR=0$ 条件下判决结果

3.3 容错能力验证

本节主要考察不同码字个数、码率对本文算法识别性能的影响，并对算法容错能力进行验证。

1) 码字个数对识别性能的影响

本节主要考察累计码字数量对算法的影响。仿真设定极化码码长为 64，码率为 1/2，设收到的码字数量分别为 1 000、2 000、3 000、4 000，信噪比为 -2~4 dB，间隔为 0.5 dB，蒙特卡罗次数为 500，统计不同信噪比下算法的识别率，结果如图 7 所示。从图 7 中可以看出，随着码字数量的增加，算法性能也有所增加，主要是因为随着码字数量的增加，所求得平均似然差逐渐逼近理论值，使判决结果正确率更高，算法性能也随之提升。

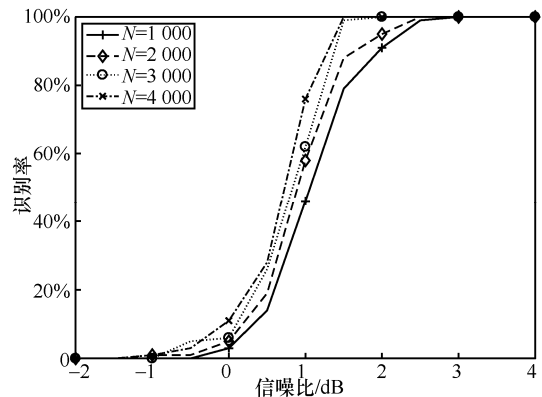


图 7 $N=64$ 时不同信噪比下算法的识别率

2) 码率对识别性能的影响

本节主要考察码率对算法性能的影响。仿真设

定极化码码长为 64，码率分别设定为 1/4、1/2、5/8、3/4，码字数量为 2 000，信噪比为 -2~4 dB，间隔为 0.5 dB，蒙特卡罗次数为 500，统计不同信噪比下算法的识别率，结果如图 8 所示。从图 8 可以看出，码率对于算法的识别性能影响不大。通过对极化码编码过程的研究发现，这主要是因为随着码率的变化，虽然信息比特数和冻结比特数随之改变，但高码率对应的对偶向量是低码率对偶向量的子集，且是低码率对偶向量中码重较重的部分，更易受误码影响。也就意味着，随着码率的降低，对偶向量集合元素不断增加，但增加的对偶向量码重较低，受误码影响较小，更易识别。综上所述，对识别性能影响较大的对偶向量是不同码率条件下对偶向量的交集，因此算法识别性能受码率的影响也较小。

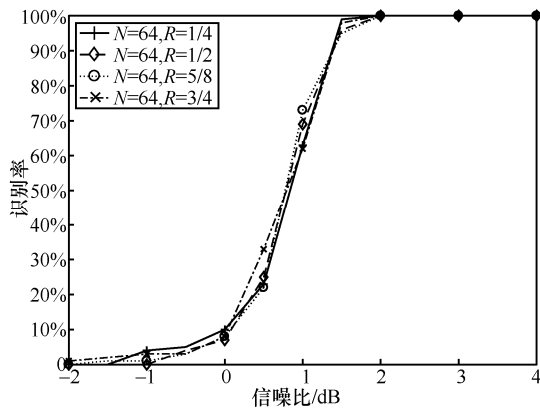


图 8 $M=2\ 000$ 时不同信噪比下算法的识别率

3.4 性能对比

本节主要考察码长对于识别性能的影响，并与现有算法的性能比较。仿真设定的目标码长为 32、64、128、256、512、1 024，设定极化码码率为 1/2，码字数量为 2 000，信噪比范围为 -2~8 dB，间隔为 1 dB，蒙特卡罗仿真次数设定为 500 次，统计不同算法在不同信噪比下的识别率，结果如图 9 所示。

从图 9 可以看出，随着码长的增加，算法识别性能逐渐下降。这主要是因为当码长增加时，码重也随之增加，同时需要识别的信息比特和冻结比特增加，更易受到误码的影响，单个码字中出现误码的概率也就更高。因此在相同条件下，随着码长的增加，识别难度逐渐增大，误判概率也就越大。从识别结果来看，当信噪比为 6 dB 时，对于码长为 1 024 的极化码，所提算法能够达到接近 100% 的识别率，说明算法具有较好的误码适应能力。

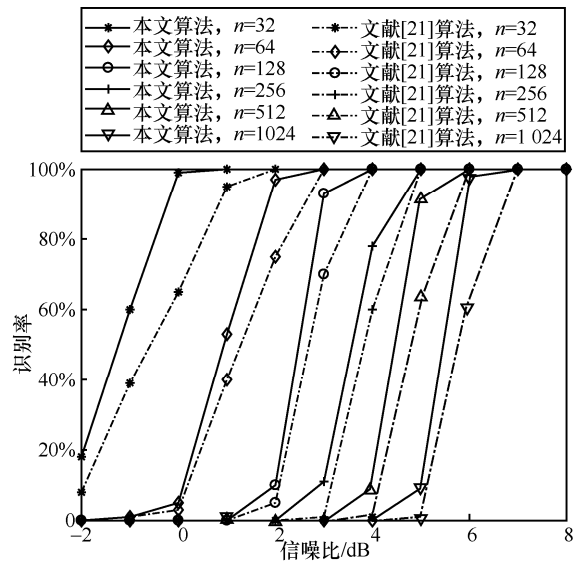


图 9 不同算法在不同信噪比下的识别率

同时，图 9 给出了与文献[21]算法的仿真对比结果，文献[21]算法的识别性能是现有最优的。从图 9 中可以看出，本文算法在性能上有明显的提升，尤其是在信噪比较高时，本文算法能够较快地达到 100% 的识别率。文献[21]同样采用软序列作为处理对象，并采用对数似然系数作为统计判决量，但在计算对数似然系数时进行了近似处理，降低了运算精度。本文对于似然差的计算未做近似处理，且从 3.1 节可以看出，本文推导的统计特性与实际情况一致，因此具有更好的识别性能。同时，在计算量方面，本文算法在计算似然差的过程中，有效利用了极化码生成矩阵的结构性质，与现有算法相比，不需要进行码长、码率等参数的遍历，有效降低了运算量，具体见 3.5 节分析。

3.5 算法计算复杂度分析

假设接收到的软判决序列长度为 L ，最大码长为 $2^{n_{\max}}$ 。首先，需要对每个判决值计算其似然差，在实际计算中已有多种高精度近似快速算法，其计算速度与乘法运算处于同一数量级^[26]。本文仿真采用 MATLAB 2016b 软件，在 i7-9750H 处理器平台进行。在实际仿真平台运算中，测得计算似然差计算所需时间约为乘法的 5 倍。在对 $\mathbf{F}^{\otimes n}$ 的每个列向量进行检验时，需要进行元素相乘和向量加法，共需要 $\frac{L}{2^n}(2^n - 1)$ 次乘法运算和 $\left(\frac{L}{2^n} - 1\right)2^n$ 次加法运算。同时，本文将门限计算等价于 5 次乘法运算。对于本文算法，当真实码长为 2^n 时，只需要计算 2^{n+1} 个统计量，因此本文算法最大需要 $5L +$

$5 \times 2^{n+1} + (2^{n+1} - 1)L$ 次乘法运算和 $\left(\frac{L}{2^{n+1}} - 1\right)2^{2n+2}$ 次加法运算。

文献[21]中所提统计量需要对每个判决值进行判断正负号和取模运算, 各等价于一次加法运算, 共需要 $2L$ 次加法。计算每个统计量时, 需要进行元素相乘、计算向量元素最小值和向量加法运算, 其中比较大小可等价于一次加法, 共需要 L 次乘法运算和 $\frac{L}{2^n}(2^n - 1) + \left(\frac{L}{2^n} - 1\right)2^n$ 次加法运算。对于不同的码长 2^n , 均需要计算 2^n 个统计量。同时, 其门限计算可等价于 10 次乘法运算。最后, 对不同码长的计算量进行求和, 最终为 $\sum_{i=n_{\min}}^{n_{\max}} ((L+10)2^i)$ 乘法运算和 $2L + \sum_{i=n_{\min}}^{n_{\max}} \left(L(2^i - 1) + \left(\frac{L}{2^i} - 1\right)2^{2i} \right)$ 次加法运算。不论码长大小, 所需计算量为固定值。

综上, 本文算法计算复杂度更低, 尤其是在码长较小的情况下。

4 结束语

本文对标准非删余极化码盲识别问题进行了研究, 推导证明了生成矩阵、码字矩阵分别与克罗内克矩阵之间的约束关系, 以及不同参数对该校验关系的影响, 并基于以上性质提出了一种极化码盲识别算法。所提算法选择软判决序列进行处理, 引入平均似然差作为判决统计量, 能够有效利用软判决序列中的置信度信息, 具有良好的误码适应能力。同时, 算法有效利用了极化码生成矩阵的结构性质, 采用了一种新的识别策略, 不需要对所有可能的码长进行遍历, 所需计算量低于现有软判决处理算法。仿真结果表明, 所提算法识别性能优于已有算法, 工程实用性更强。后续研究主要针对非标准删余极化码, 并考虑冻结比特不为零的场景, 以完善极化码盲识别问题的研究。

参考文献:

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

[2] GAMAGE H, RAJATHEVA N, LATVA-AHO M. Channel coding for enhanced mobile broadband communication in 5G systems[C]//*Proceedings of 2017 European Conference on Networks and*

Communications (EuCNC). Piscataway: IEEE Press, 2017: 1-6.

[3] BARBIER J, LETESSIER J. Forward error correcting codes characterization based on rank properties[C]//*Proceedings of 2009 International Conference on Wireless Communications & Signal Processing*. Piscataway: IEEE Press, 2009: 1-5.

[4] 李歆昊, 张旻, 史英春, 等. 基于游程特征的线性分组码与卷积码类型识别[J]. *数据采集与处理*, 2015, 30(6): 1205-1214.

LI X H, ZHANG M, SHI Y C, et al. Blind identifying of type of linear block code and convolutional code based on run feature[J]. *Journal of Data Acquisition and Processing*, 2015, 30(6): 1205-1214.

[5] 杨晓静, 闻年成. 基于秩函数和 Euclidean 算法的循环码盲识别[J]. *电路与系统学报*, 2012, 17(5): 120-123, 129.

YANG X J, WEN N C. A blind method of cyclic codes based on rank function and Euclidean arithmetic[J]. *Journal of Circuits and Systems*, 2012, 17(5): 120-123, 129.

[6] YARDI A D, VIJAYAKUMARAN S, KUMAR A. Blind reconstruction of binary cyclic codes from unsynchronized bitstream[J]. *IEEE Transactions on Communications*, 2016, 64(7): 2693-2706.

[7] 张天骐, 王俊霞, 江晓磊, 等. 基于校验矩阵匹配的循环码参数盲识别算法[J]. *电子与信息学报*, 2017, 39(4): 901-907.

ZHANG T Q, WANG J X, JIANG X L, et al. Blind recognition of cyclic code based on check matrix match algorithm[J]. *Journal of Electronics & Information Technology*, 2017, 39(4): 901-907.

[8] 杨晓静, 闻年成. 基于码根信息差熵和码根统计的 BCH 码识别方法[J]. *探测与控制学报*, 2010, 32(3): 69-73.

YANG X J, WEN N C. Recognition method of BCH codes based on roots information dispersion entropy and roots statistic[J]. *Journal of Detection & Control*, 2010, 32(3): 69-73.

[9] JO D, KWON S, SHIN D J. Blind reconstruction of BCH codes based on consecutive roots of generator polynomials[J]. *IEEE Communications Letters*, 2018, 22(5): 894-897.

[10] 刘杰, 张立民, 钟兆根, 等. 一种软判决下的本原 BCH 码盲识别方法[J]. *西安交通大学学报*, 2017, 51(6): 59-65.

LIU J, ZHANG L M, ZHONG Z G, et al. A blind recognition method for primitive BCH codes in soft decision situations[J]. *Journal of Xi'an Jiaotong University*, 2017, 51(6): 59-65.

[11] 李灿, 张天骐, 刘瑜. 基于伽罗华域高斯列消元法的 RS 码盲识别[J]. *电讯技术*, 2014, 54(7): 926-931.

LI C, ZHANG T Q, LIU Y. Blind recognition of RS codes based on Galois field columns Gaussian elimination[J]. *Telecommunication Engineering*, 2014, 54(7): 926-931.

[12] LIU P T, PAN Z P, LEI J. Parameter identification of reed-Solomon codes based on probability statistics and Galois field Fourier transform[J]. *IEEE Access*, 2019, 7: 33619-33630.

[13] 吴昭军, 张立民, 钟兆根, 等. 一种软判决下的 RS 码识别算法[J]. *电子与信息学报*, 2020, 42(9): 2150-2157.

WU Z J, ZHANG L M, ZHONG Z G, et al. Blind recognition of RS codes based on soft decision[J]. *Journal of Electronics & Information Technology*, 2020, 42(9): 2150-2157.

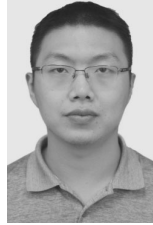
[14] 于沛东, 彭华, 巩克现, 等. 基于最小二乘代价函数的卷积码盲识别方法[J]. *电子学报*, 2018, 46(7): 1545-1552.

YU P D, PENG H, GONG K X, et al. Blind recognition of convolutional codes based on least-square cost-function[J]. *Acta Electronica Sinica*, 2018, 46(7): 1545-1552.

[15] HAN S N, ZHANG M, LI X H. A fast method for blind identification

- of punctured convolutional codes[J]. Radioengineering, 2018, 27(2): 572-579.
- [16] 刘钰, 王方刚, 张静文, 等. 多径信道下基于 EM 算法的盲 LDPC 编码器识别研究[J]. 通信学报, 2018, 39(9): 43-48.
LIU Y, WANG F G, ZHANG J W, et al. EM-based blind LDPC identification in multipath channels[J]. Journal on Communications, 2018, 39(9): 43-48.
- [17] WU Z J, ZHANG L M, ZHONG Z G, et al. Blind recognition of LDPC codes over candidate set[J]. IEEE Communications Letters, 2020, 24(1): 11-14.
- [18] DEBESSU Y G, WU H C, JIANG H. Novel blind encoder parameter estimation for turbo codes[J]. IEEE Communications Letters, 2012, 16(12): 1917-1920.
- [19] WU Z J, ZHANG L M, ZHONG Z G. A maximum cosinoidal cost function method for parameter estimation of RSC turbo codes[J]. IEEE Communications Letters, 2019, 23(3): 390-393.
- [20] 张天骐, 胡延平, 冯嘉欣, 等. 基于零空间矩阵匹配的极化码参数盲识别算法[J]. 电子与信息学报, 2020, 42(12): 2953-2959.
ZHANG T Q, HU Y P, FENG J X, et al. Blind identification algorithm of polarization code parameters based on null space matrix matching[J]. Journal of Electronics & Information Technology, 2020, 42(12): 2953-2959.
- [21] 吴昭军, 钟兆根, 张立民, 等. 基于软判决下的非删余极化码参数识别[J]. 通信学报, 2020, 41(12): 60-71.
WU Z J, ZHONG Z G, ZHANG L M, et al. Recognition of non-drilled polar codes based on soft decision[J]. Journal on Communications, 2020, 41(12): 60-71.
- [22] 邵堃, 雷迎科. 基于离散度分析的帧同步快速盲识别算法[J]. 信号处理, 2020, 36(3): 361-372.
SHAO K, LEI Y K. Fast blind recognition algorithm of frame synchronization based on dispersion analysis[J]. Journal of Signal Processing, 2020, 36(3): 361-372.
- [23] 李歆昊, 张旻, 韩树楠. 基于多重分形谱的物理层帧结构检测算法研究[J]. 电子与信息学报, 2018, 40(2): 400-407.
LI X H, ZHANG M, HAN S N. Study on frame structure detection in physical layer based on multi-fractal spectrum[J]. Journal of Electronics & Information Technology, 2018, 40(2): 400-407.
- [24] XIA T, WU H C. Novel blind identification of LDPC codes using average LLR of syndrome a posteriori probability[C]//Proceedings of IEEE Transactions on Signal Processing. Piscataway: IEEE Press, 2013: 632-640.
- [25] YU P D, PENG H, LI J. On blind recognition of channel codes within a candidate set[J]. IEEE Communications Letters, 2016, 20(4): 736-739.
- [26] MULLER J M. Elementary functions and approximate computing[J]. Proceedings of the IEEE, 2020, 108(12): 2136-2149.

[作者简介]



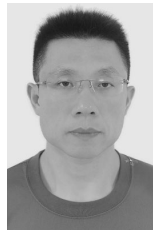
王垚 (1991-), 男, 河北石家庄人, 国防科技大学博士生, 主要研究方向为信道编码盲识别、协议分析等。



王翔 (1985-), 男, 福建福州人, 博士, 国防科技大学副教授, 主要研究方向为航天电子侦察、信号处理、模式识别等。



杨国东 (1993-), 男, 云南祥云人, 中国人民解放军 92001 部队助理工程师, 主要研究方向为电子信号分析、多源数据挖掘等。



黄知涛 (1976-), 男, 湖北荆州人, 博士, 国防科技大学教授、博士生导师, 主要研究方向为航天电子侦察、雷达/通信信号处理、综合电子战系统与技术等。