

基于多目标强化学习的抗强干扰 Polar 编码优化方法

梁豪 叶淦华* 陆锐敏 王恒 魏鹏

(国防科技大学第六十三研究所 南京 210001)

摘要: 为提升跳频(FH)通信系统信息传输的可靠性和抗干扰能力, 该文基于新型Polar编码的慢跳频抗干扰通信系统模型, 提出一种适应强干扰环境的Polar编码构造优化方法。首先, 面向包含常态和干扰态的混合信道设计多目标强化学习算法, 然后优化编码过程中的信息比特信道序列, 提升码字的纠错性能, 并通过初始化预处理和理论计算回报值降低算法执行复杂度。仿真结果表明, 在包含强干扰的混合信道条件下, 所提编码优化方法的全局误码性能优于传统编码构造方法, 相比于第5代移动通信系统(5G)第3代合作伙伴计划(3GPP)标准方案全局编码增益达0.5 dB, 有效改善Polar编码跳频通信高可靠抗干扰传输性能。

关键词: 信道编码; 抗干扰; Polar码; 强化学习; 可靠性能

中图分类号: TN911.2

文献标识码: A

文章编号: 1009-5896(2023)11-4092-09

DOI: 10.11999/JEIT230572

Anti-strong Jamming Polar Coding Optimization Method with Multiobjective Reinforcement Learning

LIANG Hao YE Ganhua LU Ruimin WANG Heng WEI Peng

(Sixty-third Research Institute, National University of Defense Technology, Nanjing 210001, China)

Abstract: In order to improve the reliability and anti-jamming ability of information transmission for the Frequency-Hopping (FH) communication system, a Polar coding construction optimization method is proposed to adapt to the strong-jamming environment, which is based on a novel Polar coded slow FH communication system model. Firstly, the multi-objective reinforcement learning algorithm is designed for the hybrid channel containing normal state and jamming state, and then the information bit-channel sequence in the coding process is optimized. Consequently the error correction performance of the designed Polar codewords is improved. In addition, the complexity of algorithm is reduced by preprocessing the initialization and theoretically calculating the reward values. The simulation results show that the overall error performance of the proposed coding optimization method is better than those of conventional coding construction methods in the hybrid channel containing strong jamming. Compared with the 3rd Generation Partnership Project (3GPP) standard scheme in Fifth-Generation (5G) mobile communication systems, the obtained overall coding gain is up to 0.5 dB. Therefore the high-reliability and anti-jamming performance of Polar coded FH transmission is effectively improved.

Key words: Channel coding; Anti-jamming; Polar codes; Reinforcement learning; Reliability performance

1 引言

随着第5代\第6代移动通信系统(Fifth\Sixth-Generation, 5G\6G)、低轨卫星互联网以及新一代军事卫星通信系统等演进发展, 信息传输面临信道

衰落、干扰加剧、业务复杂多样等难点挑战。为应对敌方人为干扰而有的跳频(Frequency Hopping, FH)调制技术是将载波频率在伪随机图案的控制下进行跳变从而实现规避干扰的一种技术, 是目前通信抗干扰领域应用最为广泛的一种扩谱通信方式。然而, 干扰方式灵活多变, 例如, 根据跳频体制的特点, 敌方可以采取有效的部分频带干扰, 将干扰功率部署在相对较窄的频带内, 使信号在传输信道中发生突发错误, 传输质量急剧衰退。应对这一问题的重要有效手段是采用非扩谱的先进信道编码技术, 依靠其强有力的纠错能力来对抗干扰, 降低抗

收稿日期: 2023-06-09; 改回日期: 2023-11-03; 网络出版: 2023-11-13

*通信作者: 叶淦华 milsatcom@163.com

基金项目: 国家自然科学基金(62201596), 国防科技大学科研计划(ZK22-45)

Foundation Items: The National Natural Science Foundation of China (62201596), The Research Planning Project of the National University of Defense Technology (ZK22-45)

突发干扰的门限, 实现信息高可靠传输。

Polar码是一种严格证明可达二进制离散无记忆信道香农极限的新型信道编码方案^[1], 因其编译码结构规则且复杂度低、速率适配灵活、纠错性能优良, 成为5G增强移动宽带场景控制信道编码标准^[2,3], 也被视为未来6G最可能应用的关键基础性技术之一^[4-6]。

在抗干扰编码方面, 将可能遭受到强干扰的跳频通信信道称为混合信道, 该信道包含正常、干扰两种通信状态。此条件下, 信道编码方案的编码设计需要根据混合信道状态作出适应性改进。文献^[7]针对部分频带干扰样式兼顾优化低密度奇偶校验码(Low-Density Parity-Check, LDPC)码度分布序列和准循环结构, 提升受干扰时码字性能; 文献^[8,9]通过改进阻塞式干扰状态下LDPC码度分布来获取更多的编码增益, 但这些方案都因未兼顾常态传输条件而造成正常通信时性能损失。Polar编码构造设计, 就是确定用于承载信息的高可靠比特信道位置序列。传统构造方案有密度进化^[10]、高斯近似(Gaussian Approximation, GA)^[11]、极化重量(Polarization Weight, PW)构造^[12]及5G新空口标准给出的可靠度顺序表^[2]等方法。然而, 当信号遭受强干扰时, 部分传输比特被破坏, 接收码字结构样式发生变化, 传统方案针对正常通信条件的构造结果无法保证与受干扰码字高度适配, 从而造成纠错性能损失。因此, 面向混合信道条件优化设计新的编码构造方案, 是提升Polar编码抗干扰能力的一个亟待解决的关键问题。此外, 近年来, 针对一些特殊条件下的编码可靠性需求, 改进的编码构造方法也相继提出, 提供了新的设计思路, 包括面向衰落信道、特定译码器的编码构造^[13,14], 以及基于人工智能(Artificial Intelligence, AI)辅助设计译码算法的尝试^[15,16], 在一定程度上有针对性地改善了码字性能。

鉴于上述分析, 本文基于Polar编码的慢跳频抗干扰通信系统模型, 面向存在强干扰的混合信道提出一种新型Polar编码构造优化方法, 同时兼容

常态和干扰态下的信息可靠传输。首先建立编码构造过程的马尔科夫决策过程模型, 然后设计多目标强化学习驱动的构造算法, 针对多种码型和干扰因子优化信息位比特信道序列, 并进一步提出初始化预处理、理论计算回报值等方法降低学习算法的运行复杂度, 最终实现混合信道中Polar码字纠错性能和抗干扰能力的提升。

2 Polar编码跳频抗干扰通信系统模型

2.1 系统模型

基于Polar编码的跳频通信系统如图1所示, 发送端将待传输信息比特送入Polar编码器进行信道编码, 交织器将输出的码字比特序列打乱, 减弱交织前后数据序列相关性, 降低突发错误影响, 提升抗干扰能力。码字经过二进制相移键控(Binary Phase Shift Keying, BPSK)调制后, 依据伪码发生器对应生成的跳频图案进行跳频调制, 信号经发送天线进入混合传输信道, 这里假设加性高斯白噪声(Additive White Gaussian Noise, AWGN)信道, 与此同时, 跳频通信信道中存在部分频带强干扰。在接收端, 接收信号依次经过解跳、解调、解交织后, 利用Polar译码器进行信道译码恢复得到信源信息。

2.2 Polar编解码

Polar码的理论基础是信道极化, 即将 $N(N=2^q, q=1,2,\dots)$ 个独立的原始物理信道经过迭代的信道重组和拆分操作, 产生 N 个可靠度各异的合成极化比特信道 $W_N^{(n)}(n \in N, N = \{1, 2, \dots, N\})$ 。然后从中选取 K 个最高可靠度的比特信道位置用于传输信息比特, 对应的比特信道位置索引序列 \mathcal{A} 称为信息位序列, 码率为 $R=K/N$; 剩余的比特信道承载双方已知的冻结比特。设 \mathbf{u}_0^{N-1} 表示待编码的源比特集合, \mathbf{x}_0^{N-1} 为编码后的码字集合, 则Polar编码生成过程为

$$\mathbf{x}_0^{N-1} = \mathbf{u}_0^{N-1} \mathbf{G}_N = \mathbf{u}_0^{N-1} \mathbf{F}_2^{\otimes q} \quad (1)$$

其中, \mathbf{G}_N 是 N 维的极化生成矩阵, $\mathbf{F}_2^{\otimes q}$ 表示单位矩阵 \mathbf{F}_2 的 q 阶Kronecker积, $\mathbf{F}_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ 。连续

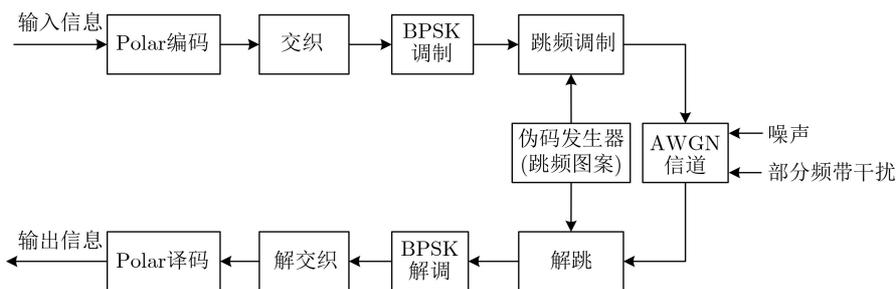


图1 Polar编码跳频通信系统框图

消除(Successive Cancellation, SC)译码算法是Ar1kan最早提出的一种低复杂度串行译码方法,具有良好的渐进性能;改进的连续消除列表(Successive Cancellation List, SCL)译码算法则通过在译码每个比特时最多保留 L 条候选译码路径,避免之前的硬判决错误来提升译码性能。

2.3 混合信道中的编码构造

对于 (N, K) -Polar码的编码构造,就是确定包含 K 个最高可靠度的极化比特信道的索引序列 \mathcal{A} ,这是决定码字纠错性能优劣的基础关键要素。在无干扰的常态信道条件下,通信系统可以利用常规编码构造方案设计信息位序列 \mathcal{A} 进行编码传输。但当受到频繁的强干扰时,传输信号发生突发错误,经过交织和解交织后,受干扰的码字比特随机离散分布,对应出错位置的码字比特被“删除”,码字结构被破坏,接收端译码信息不完整,造成码字纠错能力明显降低。然而,信息位序列 \mathcal{A} 的构造与码字结构密切相关,对于干扰态下多样变化接收码字样式,正常态下针对完整码字设计的传统构造结果已不能很好适配。

因此,在包含强干扰的混合信道中,构造的信息位序列需同时兼容适配正常态和干扰态两种情况下的Polar信道编码。为提升强干扰环境中信息可靠传输能力,本文兼顾常态和干扰态信道条件联合优化编码构造方案,设计高可靠通用信息位序列,提升Polar码字在混合信道中的纠错性能和抗干扰能力。

3 强干扰环境下Polar编码优化设计

本节首先建立基于马尔科夫决策过程(Markov Decision Process, MDP)的编码构造模型,然后提出多目标强化学习驱动的新型编码构造算法,设计适应混合信道的高可靠信息位序列,并进一步优化学习构造算法,降低算法计算复杂度。

3.1 MDP编码构造模型

设计具有嵌套特性的Polar码信息位序列构造过程^[17],可以将编码构造过程转化为马尔科夫决策过程。如图2所示,针对码长为 N 、信息长度为 K 的Polar码,状态(state) \mathbf{s} 是构造过程中渐进的信息位序列 $\{\mathcal{A}_{N,i}\}(1 \leq i \leq K)$,且满足 $\mathcal{A}_{N,1} \subseteq \mathcal{A}_{N,2} \subseteq \dots \subseteq \mathcal{A}_{N,K}$;动作(action) a 表示选取的比特

信道的位置索引,满足 $a_i \notin \mathcal{A}_{N,i-1}, a_i \in \mathcal{A}_{N,i}$;回报(reward) \mathbf{r} 代表对应状态 \mathbf{s} 中生成码字的纠错性能。基于MDP构造模型,用于Polar编码的信息位序列可以逐元素按序选取构造。

3.2 算法设计

部分频带强干扰造成跳频传输信号产生突发错误,交织后错误码字离散分布,在接收端对应受干扰的码字比特被“删除”。定义干扰因子 $\rho = W_J/W_{FH}$,其中 W_J 表示部分频带干扰的带宽, W_{FH} 表示跳频通信总带宽。通过统计分析可得^[7],干扰因子值等于每跳信号受干扰的概率大小,则一个发送的长度为 N 的Polar码块中,被干扰删除的码字比特数量为 $N \cdot \rho$,有效信号的长度为 $M = N - N \cdot \rho$ 。设这些删除比特的位置索引集合为 \mathcal{P} ,对应的码字干扰样式为长度为 N 的二进制矢量 $\mathbf{p} = [1 \ 1 \ 0 \ 1 \dots 0 \ 1 \ 1]$,其中的“0”元素指示了被删除比特的的位置。无干扰时,干扰样式 \mathbf{p} 为全“1”, $\mathbf{p} = \emptyset$ 。

在遭受干扰的跳频系统中,一种干扰因子 ρ_j ,对应生成一类受干扰的接收序列 \mathcal{C}_j 。从统计角度看, \mathcal{C}_j 中被破坏删除的码字比特数量趋于固定,即 $N \cdot \rho_j$,其位置离散随机分布。设 ρ_0 为未受干扰的正常态下的干扰因子,则 $\rho_0 = 0$ 。在跳频抗干扰通信系统中,对应正常态下的接收序列 \mathcal{C}_0 及 $J-1$ 种干扰因子分别对应的接收序列 $\mathcal{C}_j(1 \leq j \leq J-1, J > 1)$,组成目标Polar码字传输后对应的接收序列簇 $\mathcal{C}_{N,K} = \{\mathcal{C}_0, \mathcal{C}_1, \dots, \mathcal{C}_{J-1}\}$,接收序列长度分别为 $M_0 = N, M_j = N \cdot (1 - \rho_j), 1 \leq j \leq J-1$ 。本文提出基于多目标强化学习的编码构造优化方法,面向包含多码字样式的目标序列簇设计最优的信息位序列。

根据信息位序列的MDP构造模型,建立多目标强化学习环境 $(\mathbf{s}, \mathcal{A}, \mathbf{r})_K^N$ 。其中,在每个学习训练周期的第 k 步(step),状态 \mathbf{s}_k 是编码构造过程中的子信息位序列,即 $\mathbf{s}_k: \mathcal{A}_{N,k}^o, k \in \{1, 2, \dots, K\}$,且满足 $\mathcal{A}_{N,1}^o \subseteq \mathcal{A}_{N,2}^o \subseteq \dots \subseteq \mathcal{A}_{N,K}^o$,信息位序列 $\mathcal{A}_{N,k}^o$ 对应生成Polar码字,传输后对应产生一个序列簇 $\mathcal{c}_k^N = \{c_0, c_1, \dots, c_{J-1}\}$,每个序列 c_j 代表了由当前对应的信息位序列 $\mathcal{A}_{N,k}^o$ 编码生成的Polar码字经信道传输后的接收序列,分别对应干扰因子 ρ_j 和长度 M_j 。 c_j 是学习训练过程中的接收序列表示,初始状态中 c_0^N 是全0码字对应的接收序列,对应空集合 $\mathcal{A}_{N,0}^o$ 。

动作 a_k^N 是一个正整数, $a_k^N \in \mathcal{A} \subseteq \mathbb{N}$,表示极化编码中某一合成比特信道的位置索引,满足 $a_k^N \notin \mathcal{A}_{N,k}^o, a_k^N \cup \mathcal{A}_{N,k}^o = \mathcal{A}_{N,k+1}^o$ 。选取的每一个动作 a_k^N 对状态 \mathbf{s}_k 都有一个回报,表征对应状态 \mathbf{s}_k 产生

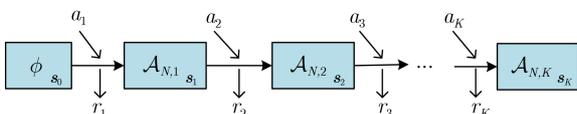


图2 Polar编码的MDP构造模型

的接收序列的纠错性能。接收序列簇 c_k^N 中每个样式的序列 c_j 都有与自己相关的回报 $r_{j,k}$, 回报信号是个矢量 $r_k^N = (r_{0,k}, r_{1,k}, \dots, r_{J-1,k})$, 对应 c_k^N 的回报值为 r_k^N , r_k^N 代表 c_k^N 的综合误码性能, 其值与信息位序列 $\mathcal{A}_{N,k}^0$ 和每个接收序列的回报 $(r_{0,k}, r_{1,k}, \dots, r_{J-1,k})$ 相关。

基于MDP映射关系, 状态转移过程按序执行, 当给定 (s_k, a_k^N) , 则下一状态 s_{k+1} 即确定。路径轨迹 $(s_0, a_0^N, s_1, a_1^N, \dots, s_K, a_K^N)$ 依次对应信息位序列 $(\mathcal{A}_{N,0}^0, \mathcal{A}_{N,1}^0, \dots, \mathcal{A}_{N,K}^0)$ 的构造过程。在多目标强化学习过程中, 代理体(agent)与建立的MDP环境在各个离散阶段相互交互。在阶段 k , 代理体观察状态 s_k , 根据策略 $\pi(a_k^N | s_k)$ 选择相应的动作 a_k^N , 与此同时从环境中获得对应的综合回报 r_k^N , 代理体的目标是优化其策略以最大化综合回报。

为改善目标Polar码接收序列簇的可靠性能, 优化策略需要解决的问题是

$$\left. \begin{aligned} & \min[\varphi_0(\text{bler}_0), \varphi_1(\text{bler}_1), \dots, \varphi_j(\text{bler}_j)] \\ & = \min \Phi(\text{bler}_0, \text{bler}_1, \dots, \text{bler}_j), \\ & \text{s.t. } \text{bler}_j \leq \text{BLER}_j^{\text{th}}, 0 \leq j \leq J-1 \end{aligned} \right\} \quad (2)$$

其中, φ 和 Φ 分别表示单个Polar码接收序列的错误概率的目标函数和接收序列簇的综合错误概率的目标函数; bler_j 表示接收序列 C_j 的误块率(BLock Error Rate, BLER), $\text{BLER}_j^{\text{th}}$ 是对 C_j 预设的误块率门限值。

为解决同时满足多个样式Polar码的纠错性能提升的问题, 本文提出采用多目标强化学习方法^[8], 面向多个不同回报的任务和连续的决策问题, 提供有效的动作策略方案。通过与环境交互, 代理体根据不同的目标同步优化其动作策略 π , 即

$$\begin{aligned} JQ^\dagger(s, a^N) &= \max_{\pi} JQ^\pi(s, a^N) \\ &= \max_{\pi} [Q_0^\pi(s, a^N), Q_1^\pi(s, a^N), \dots, \\ & \quad Q_{J-1}^\pi(s, a^N)] \end{aligned} \quad (3)$$

$JQ^\pi(s, a^N)$ 表示综合的状态-动作矢量函数, JQ^\dagger 表示其最优结果; $Q_j^\pi(s, a^N)$ 表示针对每个目标对象的状态-动作函数; s 和 a^N 分别表示当前的状态和动作, 则最优的策略 π^\dagger 以下面方式获得

$$\pi^\dagger(s) = \arg \max_{a^N} JQ^\dagger(s, a^N) \quad (4)$$

对各个Polar接收序列 c_j 的状态-动作函数依据加权准则处理, 则综合的状态-动作矢量函数 Q 值计算为

$$JQ(s, a^N) = \sum_j^{J-1} \lambda_j Q_j(s, a^N) \quad (5)$$

其中, λ_j 表示第 j 个目标接收序列对象的正权重值; 依据 Q -学习过程, 第 j 个目标接收序列的状态-动作函数 Q 值计算为

$$Q_j(s, a^N) = (1-\alpha)Q_j(s, a^N) + \alpha(r_j + \max_{a^{N'}} Q_j(s', a^{N'})) \quad (6)$$

其中, α 表示强化学习中的学习率, $s', a^{N'}$ 分别表示代理体与环境交互过程中的未来状态和动作。另外, 在基于多目标强化学习的编码构造中, 每个阶段 k 所获得的回报值是矢量, 综合回报值与每个目标对象的回报值相关, 反映了所选取动作的优劣。若假设学习率 $\alpha = 1$, $\text{bler}_{j,k}$ 表示第 k 步 c_k^N 中序列 c_j 的误块率值, 则依据加权准则, 综合回报值 r_k^N 相应地计算为

$$r_k^N = \sum_j^{J-1} -\lambda_j \lg(\text{bler}_{j,k}) \quad (7)$$

联合考虑干扰因子及接收序列样式对混合信道中码字综合性能的影响, 设 $\lambda_0 = 1 - \max_j(\rho_j)$, $\lambda_j = \rho_j \cdot (N/M_j) = \rho_j / (1 - \rho_j)$, 即 λ_j 与干扰因子 ρ_j 正相关、与受干扰码字的接收序列长度 M_j 负相关, 代表Polar码字受干扰越严重、码长越短时, 干扰态相比常态的回报值权重越大, 弥补受干扰时的性能损失。基于计算获得的综合 JQ 值, 依据贪婪准则从构造的学习环境中依次确定最佳的动作选择 a_k^N , 进而输出最优的信息位序列 $\mathcal{A}_{N,K}^0$, 用于强干扰环境下混合信道中的Polar编码。

3.3 学习算法优化

为降低所提构造算法的计算复杂度, 便于快速实施得出优化结果, 联合Polar编码特性优势, 从下面两个方面优化学习算法。

3.3.1 GA辅助的初始化预处理

对于给定的干扰因子 ρ_j , 受干扰码字被删除的比特数量一定, 干扰样式 p 各异。设在学习训练中取对应干扰因子 ρ_j 的 S 个不同干扰样式 p_i 的接收序列 $c_j^{p_i}$ ($1 \leq i \leq S$), 对应每种码字干扰样式 p_i , 分别利用高斯近似(GA)算法^[10]重构获得对应 $c_j^{p_i}$ 的信息位序列 $\mathcal{A}_{N,i}^{\text{GA}}$ 。虽然接收序列 $c_j^{p_i}$ 样式多样, 但 $\{\mathcal{A}_{N,i}^{\text{GA}}, 1 \leq i \leq S\}$ 和未干扰码字对应的信息位序列 $\mathcal{A}_{N,0}^{\text{GA}}$ 包含相当数量的共同信息位比特信道元素。鉴于此, 令学习环境中的初始状态 s_0 : $\mathcal{A}_{N,\text{in}}^0 = \{\mathcal{A}_{N,0}^{\text{GA}} \cap \mathcal{A}_{N,1}^{\text{GA}} \cap \mathcal{A}_{N,2}^{\text{GA}} \cap \dots \cap \mathcal{A}_{N,S}^{\text{GA}}\}$, 则一部分具有较高可靠度的信息比特信道被预先保留下来; 初始的动作空间为

$a^N \in \mathbb{T}_{\text{act}}, \mathbb{T}_{\text{act}} = \left\{ \mathcal{A}_N^{\text{GA}} \cup \mathcal{A}_{N,1}^{\text{GA}} \cup \dots \cup \mathcal{A}_{N,S}^{\text{GA}} \right\} \setminus \left\{ \mathcal{A}_N^{\text{GA}} \cap \mathcal{A}_{N,1}^{\text{GA}} \cap \dots \cap \mathcal{A}_{N,S}^{\text{GA}} \right\}$ 。利用该方式, 动作空间由 \mathbb{N} 被压缩至 \mathbb{T}_{act} , 学习轨迹空间由 $N!$ 降为 $|\mathbb{T}_{\text{act}}|!$, 每个学习训练周期的最大步数降为 $(K - |\mathcal{A}_{N,\text{in}}^{\text{O}}|)$ 。状态-动作空间不超过 $O(N^2)$ 。若 $J > 2$, 即混合信道中包含多种干扰因子时, 可针对每一个 ρ_j 获得的 $\mathcal{A}_{N,\text{in}}^{\text{O}}, \mathbb{T}_{\text{act}}$ 继续相互间取交集。表1给出了不同码长、不同干扰因子时初始化预处理后的动作空间大小 $|\mathbb{T}_{\text{act}}|$ 与信息位序列维度 K 的比值结果, 设 $J=2, K=1/3N$ 。可以看出, 预处理后的动作空间占比显著降低, 且随干扰因子增大而增大。通过预先筛选高可靠的初始信息位序列, 缩减状态-动作空间, 减少每个学习训练周期中的步数, 显著降低算法执行复杂度。

3.3.2 回报率高效快速获取

根据式(7), 学习训练中计算回报率 r_k^N 需要先获得第 k 步 c_k^N 中接收序列 c_j 对应的误块率 $\text{bler}_{j,k}$ 。为获取采用信道编码方案后的误块率值, 现有方案往往采用大量蒙特卡洛数值仿真的方式, 尤其当训练周期较大时, 误块率仿真数量大、时间长, 不便于实际操作。鉴于上述问题, 本文提出一种理论计算回报率的方法。

根据3.3.1节中所述, 对于干扰因子 ρ_j, c_j 对应不同的干扰样式 \mathbf{p}_i , 即接收序列 $c_j^{\mathbf{p}_i}$ ($1 \leq i \leq S$), Polar码字中被删除比特的位置索引集合为 \mathcal{P}_i 。在学习训练中码字的干扰样式对于接收端已知。假设传输信道为AWGN信道, 接收端译码时接收信号 $\{\mathbf{y} : y_n, n \in \mathbb{N}\}$ 的对数似然比(Log-Likelihood Ratio, LLR)计算为

$$L(y_n) = \begin{cases} \frac{2}{\sigma^2} y_n, & n \in \mathbb{N} \setminus \mathcal{P}_i \\ 0, & n \in \mathcal{P}_i \end{cases} \quad (8)$$

其中, σ^2 为信道噪声方差。对于 $N=2$ 的情况, 两个物理信道的LLR均值分别表示为 $E[L(y_1)], E[L(y_2)]$, 根据 $N=2$ 的Polar码基本构造单元, 利用GA算法, 重新计算极化变换后两个比特信道的LLR均值 $E[L_2^{(1)}], E[L_2^{(2)}]$

$$\left. \begin{aligned} E[L_2^{(1)}] &= \phi^{-1} \left(1 - \left(1 - \phi(E[L(y_1)]) \right) \right. \\ &\quad \left. \left(1 - \phi(E[L(y_2)]) \right) \right) \\ E[L_2^{(2)}] &= E[L(y_1)] + E[L(y_2)] \end{aligned} \right\} \quad (9)$$

$$\left. \begin{aligned} E[L_N^{(2n-1)}] &= \phi^{-1} \left(1 - \left(1 - \phi \left(E[L_{N/2}^{(n)}] \right) \right) \right. \\ &\quad \left. \cdot \left(1 - \phi \left(E[L_{N/2}^{(n+N/2)}] \right) \right) \right) \\ E[L_N^{(2n)}] &= E[L_{N/2}^{(n)}] + E[L_{N/2}^{(n+N/2)}] \end{aligned} \right\} \quad (10)$$

其中, 函数

$$\phi(x) = \begin{cases} 1 - \frac{1}{\sqrt{4\pi x}} \int_{-\infty}^{\infty} \tanh\left(\frac{u}{2}\right) e^{-\frac{(u-x)^2}{4x}} du, & x > 0 \\ 1, & x = 0 \end{cases}$$

对于一般的 $N > 2$ 的码字, 基于蝶形极化编码结构, 根据式(10)迭代计算获得LLR均值 $E[L_N^{(i)}]$, 则对应源比特 u_n 的比特信道 $W_N^{(n)}$ 错误概率计算为

$$p_\varepsilon(W_N^{(n)}; N, \mathbf{p}_i) = Q \left(\sqrt{E[L_N^{(n)}] / 2} \right), \quad n \in \mathbb{N} \quad (11)$$

其中, 函数 $Q(x) = 1 / \sqrt{2\pi} \int_x^{+\infty} \exp(-\lambda^2/2) d\lambda$ 。因此, 接收序列 $c_j^{\mathbf{p}_i}$ 对应于干扰样式 \mathbf{p}_i 的误块率估计值计算为

$$\begin{aligned} \text{bler}_{j,k}(N, \mathcal{A}_{N,k}^{\text{O}}, \mathbf{p}_i) \\ \simeq 1 - \prod_{n \in \mathcal{A}_{N,k}^{\text{O}}} \left(1 - p_\varepsilon(W_N^{(n)}; N, \mathbf{p}_i) \right) \end{aligned} \quad (12)$$

对 S 个干扰样式的误块率估计值取均值, 得到接收序列 c_j 理论计算的误块率值 $\text{bler}_{j,k}$, 进而通过计算式(7)快速获得回报率, 避免大量重复仿真带来的计算资源和时间消耗, 降低所提编码优化算法的执行复杂度。

基于上述分析及优化, 多目标强化学习驱动的Polar编码构造算法描述如算法1所示。在实际应用中, 根据Polar码型配置和干扰检测识别结果, 所设计的信息位序列构造算法在线下执行, 输出优化的 $\mathcal{A}_{N,K}^{\text{O}}$ 并存储于发送端和接收端, 用于编译码过程。

4 仿真结果与分析

4.1 仿真设置

为验证本文所设计的新型Polar编码构造优化算法的码字抗干扰能力, 对Polar编码的慢跳频BPSK系统在部分频带干扰的混合信道中进行抗干扰性能仿真与分析。仿真条件设置如下: (1) Polar码型: 码率1/3, 信息位序列分别采用所提优化算法构造的 $\mathcal{A}_{N,K}^{\text{O}}$ 和未干扰码字对应的 $\mathcal{A}_N^{\text{GA}}$ 、PW算法构造的 \mathcal{A}^{PW} 等现有构造方案, 译码算法采用低复杂

表1 初始化预处理后的动作空间占比结果 $|\mathbb{T}_{\text{act}}|/K$

N	ρ_j				
	0.1	0.2	0.3	0.4	0.6
256	0.05	0.09	0.12	0.18	0.25
512	0.03	0.05	0.08	0.10	0.16
1024	0.10	0.12	0.15	0.17	0.20

算法1 基于多目标强化学习的Polar编码优化构造算法

- (1) 初始化设置Polar码长 N , 信息比特长度 K , 干扰因子 $\{\rho_j, 1 \leq j \leq J - 1\}$;
- (2) 对未干扰接收序列和对应干扰因子 ρ_j 的 S 个不同干扰样式 ρ_j 的接收序列, 分别重构信息位序列 $\{\mathcal{A}_N^{\text{GA}}, \mathcal{A}_{N,1}^{\text{GA}}, \mathcal{A}_{N,2}^{\text{GA}}, \dots, \mathcal{A}_{N,S}^{\text{GA}}\}$;
- (3) 确定初始动作空间 $\mathbb{T}_{\text{act}} = \{\mathcal{A}_N^{\text{GA}} \cup \mathcal{A}_{N,1}^{\text{GA}} \cup \dots \cup \mathcal{A}_{N,S}^{\text{GA}}\} \setminus \{\mathcal{A}_N^{\text{GA}} \cap \mathcal{A}_{N,1}^{\text{GA}} \cap \dots \cap \mathcal{A}_{N,S}^{\text{GA}}\}$, 初始状态的信息位序列
 $s: \mathcal{A}_{N,\text{in}}^{\text{O}} = \{\mathcal{A}_N^{\text{GA}} \cap \mathcal{A}_{N,1}^{\text{GA}} \cap \dots \cap \mathcal{A}_{N,S}^{\text{GA}}\}$; 若 $J > 2$, 则对每个 ρ_j 获得的 $\mathcal{A}_{N,\text{in}}^{\text{O}}$, \mathbb{T}_{act} 相互间取交集。设周期(episode)数的最大值为 E ;
- (4) 随机初始化 $TQ(s, a^N)$;
- (5) 对于每个周期 e ($1 \leq e \leq E$), 重复下述(6)~(15)操作;
- (6) 初始化状态 s ;
- (7) 对于每个周期的阶段 k , 重复下述操作;
- (8) 选取动作 a_k^N , 估计误块率值 $\text{bler}_{j,k}$, 计算回报值 $r_{j,k}^N$, 获得 $r_{1,k}^N, r_{2,k}^N, \dots, s'$;
- (9) 对于 $j = 0, 1, \dots, J - 1$, 根据 $r_{1,k}^N, r_{2,k}^N, \dots, s'$ 依次计算对应接收序列 c_j 的 Q 值
 $Q_j(s, a^N) = (1 - \alpha)Q_j(s, a^N) + \alpha(r_j + \max_{a^{N'}} Q_j(s', a^{N'}))$;
- (10) 计算对应接收序列簇 $c_k^N = \{c_0, c_1, \dots, c_{J-1}\}$ 的综合 Q 值 $TQ(s, a^N)$;
- (11) 基于 $TQ(s, a^N)$ 确定动作 a_k^N ;
- (12) 更新 $\mathcal{A}_{N,k+1}^{\text{O}} = a_k^N \cup \mathcal{A}_{N,k}^{\text{O}}$, $\mathbb{T}_{\text{act}} \leftarrow \mathbb{T}_{\text{act}} \setminus a_k^N$;
- (13) 状态转移: $s \leftarrow s'$;
- (14) 判断当前状态 s 是否截止, 否, 则转到第7步; 是, 则继续执行下一步;
- (15) 判断是否满足 $e=E$, 否, 转到第(5)步; 是, 继续执行下一步;
- (16) 输出构造的最优信息位序列 $\mathcal{A}_{N,K}^{\text{O}}$ 。

度SC译码。(2)交织器：采用伪随机交织方式。(3)跳频系统：每个Polar码块的跳频数为8跳。(4)信道环境：假设传输信道为AWGN信道，含有部分频带强干扰，按条件设置相应的信噪比(Signal-to-Noise Ratio, SNR)以及干扰因子，仿真BLER的错误码块数至少100。(5)学习环境：贪婪准则探索率 $\varepsilon = 0.8$, 学习率 $\alpha = 10^{-2}$, 初始化预训练中对应干扰因子 ρ 的不同干扰样式接收序列个数 $S=10^2$, 训练周期数 $E=10^4$ 等。假设通信中混合信道包含1种大于0的干扰因子，即 $J=2$ 。在不同码长、干扰因子、信噪比等设置条件下，所提Polar信息位序列优化构造方案及对比方案在混合信道中的误码性能结果及分析如下。

4.2 性能分析

图3—图5给出了不同编码码长($N=128/256/512$)时所提信息位序列优化构造方案 $\mathcal{A}_{N,K}^{\text{O}}$ 和现有构造方案 $\mathcal{A}_N^{\text{GA}}$, \mathcal{A}^{PW} 在混合信道中的BLER结果, 干扰因子设置为 $\rho_1 = 0.1$ 。由图3可知, 所提优化方案在混合信道中均优于采用传统未干扰态对应的 $\mathcal{A}_N^{\text{GA}}$ 和PW构造的 \mathcal{A}^{PW} 方案。观察图4和图5发现, 当码长增大时, PW构造方案在较低SNR区间内性能表现更好, 优于所提优化方案。然而随着SNR增大, PW构造的 \mathcal{A}^{PW} 方案纠错性能逐渐恶化, 在较高SNR区间内BLER明显高于其他构造方案, 而所提优化方案带来的编码增益不管在正常态(ρ_0)还是在干扰态(ρ_1)下都显著高于对比方案。

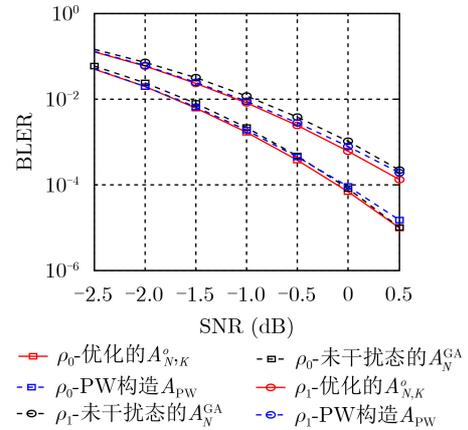


图3 $N=128$ 时不同信息位序列方案的BLER性能结果

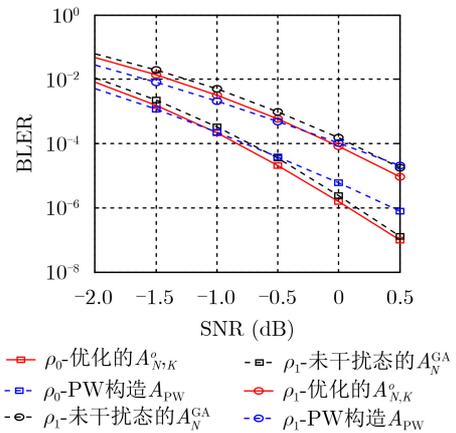


图4 $N=256$ 时不同信息位序列方案的BLER性能结果

为进一步清晰表示所提编码优化算法在混合信道中的综合性能优势,设计新的性能衡量指标进行比较。定义 $G_{\text{dB}}[\mathcal{A}_1, \mathcal{A}_2]_{\rho_j}$ 表示在给定的BLER门限 BLER^{th} 、干扰因子 ρ_j 下,信息位序列 \mathcal{A}_1 对应的码字相比信息位序列 \mathcal{A}_2 对应的码字的性能增益差。基于图3—图5的仿真条件设置,表2分别给出了混合信道中所提优化方案相比于现有构造方案在不同BLER门限时全局性能增益差。从结果中看出,大多数条件下的全局增益差 $G_{\text{dB}} > 0$,证明本文所设计的编码优化方法能够在强干扰环境中获得更优的全局编码增益,对抗干扰的纠错能力更强。

为研究干扰因子对所提多目标强化学习驱动的编码优化方案的影响,图6给出混合信道中干扰因子分别为 $\rho_1 = 0.1, 0.2, 0.3$ 时不同信息位序列构造方案的误块率性能,码长设为 $N=128$ 。观察仿真结果可知,所提信息位序列构造方案在各干扰因子情况下均优于其他对比构造方案,因此证明所提方案能够同时获得正常态与干扰态时的性能提升,适应于干扰环境的能力更强。

图7给出了SCL译码算法下的各种信息位序列构造方案的BLER结果,码长设为 $N=256$,SCL列表大小 $L=8$,干扰因子分别设为 $\rho_1 = 0.1, 0.3$ 。由仿

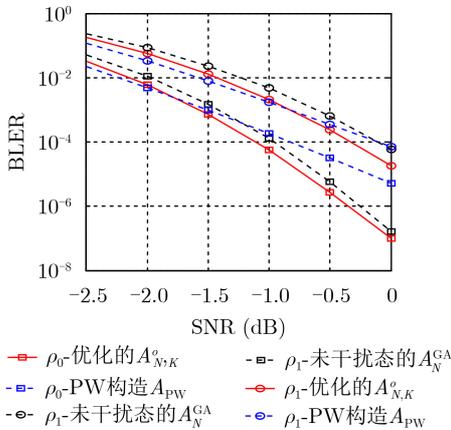


图5 $N=512$ 时不同信息位序列方案的BLER性能结果

表2 混合信道中所提优化方案与对比方案的全局性能增益差 (dB)

码长 N	全局增益差 G_{dB}	BLER th		
		10^{-2}	10^{-3}	10^{-4}
128	$\sum_j^{J-1} G_{\text{dB}}[\mathcal{A}_{N,K}^o, \mathcal{A}_N^{\text{GA}}]_{\rho_j}$	0.27	0.28	0.18
	$\sum_j^{J-1} G_{\text{dB}}[\mathcal{A}_{N,K}^o, \mathcal{A}^{\text{PW}}]_{\rho_j}$	0.08	0.13	0.27
256	$\sum_j^{J-1} G_{\text{dB}}[\mathcal{A}_{N,K}^o, \mathcal{A}_N^{\text{GA}}]_{\rho_j}$	0.26	0.28	0.26
	$\sum_j^{J-1} G_{\text{dB}}[\mathcal{A}_{N,K}^o, \mathcal{A}^{\text{PW}}]_{\rho_j}$	-0.34	-0.16	0.13
512	$\sum_j^{J-1} G_{\text{dB}}[\mathcal{A}_{N,K}^o, \mathcal{A}_N^{\text{GA}}]_{\rho_j}$	0.32	0.35	0.37
	$\sum_j^{J-1} G_{\text{dB}}[\mathcal{A}_{N,K}^o, \mathcal{A}^{\text{PW}}]_{\rho_j}$	-0.17	0.08	0.45

真结果发现,SCL译码时PW构造方案的性能损失大,主要由于其构造选取的信息位序列与SCL译码算法适配性低,纠错能力下降。然而所提编码优化方案构造的信息位序列对应的误码性能显著优于PW构造方案,在 $\rho_1 = 0.1$, $\text{BLER}^{\text{th}} = 10^{-3}$ 时,两种方案的全局性能增益差甚至可达 $\sum_j^{J-1} G_{\text{dB}}[\mathcal{A}_{N,K}^o, \mathcal{A}^{\text{PW}}]_{\rho_j} = 1.75$ dB。另外,由于采用了纠错能力更强、复杂度更高的SCL译码,采用未干扰态对应的信息位序列与所提优化方案的性能表现相近。图8进一步给出了采用循环冗余校验(Cyclic Redundancy Check, CRC)级联编码和CRC辅助SCL译码下的不同构造方案的性能结果,设码长 $N=256$, $\rho_1 = 0.2$,CRC长度为6,SCL列表大小 $L=4$ 。可以看出,所提编码构造优化方案在混合信道中性能优于其他构造方案。

Polar码作为5G控制信道编码方案,第3代合作伙伴计划(3rd Generation Partnership Project,

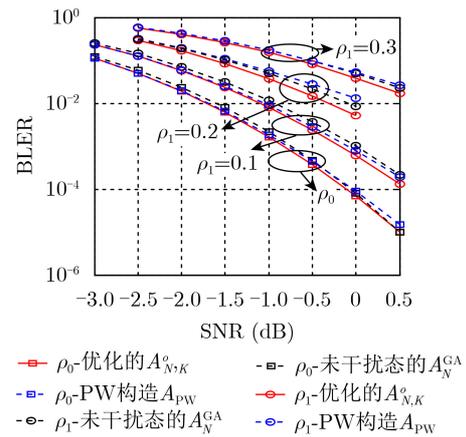


图6 不同干扰因子时各信息位序列构造方案的BLER性能结果 $\rho_1 = 0.1, 0.2, 0.3$

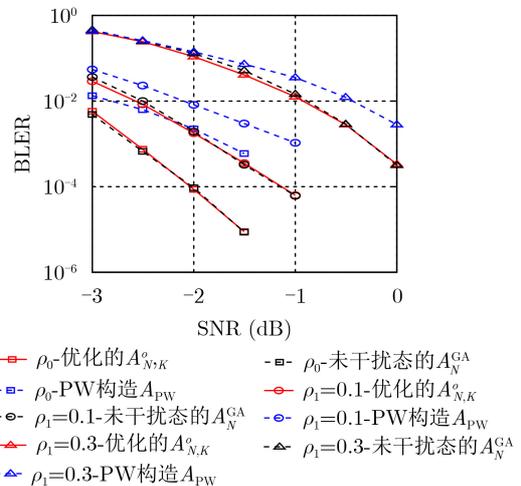


图7 SCL译码时不同信息位序列构造方案的BLER结果

3GPP)标准中给出了经验性的极化比特信道可靠度顺序表, 据此可以获得编码时的信息位序列 \mathcal{A}^{3GPP} 。为验证在混合信道中本文所提信息位序列构造方案与3GPP标准构造方案的纠错性能差异, 两种方案的误块率对比结果如图9所示, 设干扰因子 $\rho_1 = 0.2$, 码长分别为 $N=128, 512$, 采用低复杂度的SC译码。从仿真结果中可以看出, 所提优化方案在混合信道中的全局性能整体上优于3GPP标准方案, 当 $N=128$, $\text{BLER}^{\text{th}} = 10^{-2}$, 全局性能增益差 $\sum_j^{J-1} G_{\text{dB}}[\mathcal{A}_{N,K}^o, \mathcal{A}^{3GPP}]_{\rho_j} = 0.5 \text{ dB}$; 当 $N=512$, $\text{BLER}^{\text{th}} = 10^{-5}$, 可达 $\sum_j^{J-1} G_{\text{dB}}[\mathcal{A}_{N,K}^o, \mathcal{A}^{3GPP}]_{\rho_j} = 0.48 \text{ dB}$ 。所设计的优化构造算法在实际应用时, 可依据码型配置和检测的干扰因子在线下执行, 得到优化后的信息位序列用于Polar编译码。

5 结束语

基于Polar编码的慢跳频通信系统模型, 本文面向包含常态和干扰态的混合信道设计一种新型Polar编码构造优化方法。建立信息位序列构造过程的马

尔科夫决策过程模型, 提出基于多目标强化学习的构造算法, 针对多种码型和干扰因子优化用于编码的信息位序列, 并进一步设计初始化预处理、理论计算回报值等方法降低学习算法的复杂度。与其他主流编码构造方案相比, 所提优化方法的全局编码增益更高, 在强干扰环境下的信息可靠传输性能和抗干扰能力得到提升, 同时在实际通信系统中具备较好的应用价值, 为智能化抗干扰信道编码提供一种新的思路和手段。

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. doi: [10.1109/TIT.2009.2021379](https://doi.org/10.1109/TIT.2009.2021379).
- [2] ETSI. ETSI TS 38 212 V15.2. 0 5G; NR; Multiplexing and channel coding[S]. 2018.
- [3] NIU Kai, ZHANG Ping, DAI Jincheng, *et al.* A golden decade of polar codes: From basic principle to 5G applications[J]. *China Communications*, 2023, 20(2): 94–121. doi: [10.23919/JCC.2023.02.015](https://doi.org/10.23919/JCC.2023.02.015).
- [4] 白宝明, 马啸, 陈文, 等. 面向B5G/6G的信息传输与接入技术专题序言[J]. *西安电子科技大学学报*, 2020, 47(6): 1–4. doi: [10.19665/j.issn1001-2400.2020.06.001](https://doi.org/10.19665/j.issn1001-2400.2020.06.001).
BAI Baoming, MA Xiao, CHEN Wen, *et al.* Editorial: Introduction to the special issue on information transmission and access technologies for B5G/6G[J]. *Journal of Xidian University*, 2020, 47(6): 1–4. doi: [10.19665/j.issn1001-2400.2020.06.001](https://doi.org/10.19665/j.issn1001-2400.2020.06.001).
- [5] DONG Yanfei, DAI Jincheng, NIU Kai, *et al.* Joint source-channel coding for 6G communications[J]. *China Communications*, 2022, 19(3): 101–115. doi: [10.23919/JCC.2022.03.007](https://doi.org/10.23919/JCC.2022.03.007).
- [6] 魏浩, 张梦洁, 王东明. 6G极化码低时延译码方案[J]. *移动通信*, 2022, 46(6): 64–71. doi: [10.3969/j.issn.1006-1010.2022.06.010](https://doi.org/10.3969/j.issn.1006-1010.2022.06.010).
WEI Hao, ZHANG Mengjie, and WANG Dongming. Low-latency decoding algorithm of polar codes for 6G wireless systems[J]. *Mobile Communications*, 2022, 46(6): 64–71. doi: [10.3969/j.issn.1006-1010.2022.06.010](https://doi.org/10.3969/j.issn.1006-1010.2022.06.010).
- [7] 王任之, 潘克刚, 赵瑞祥. 跳频抗干扰通信系统中LDPC码的编码优化设计[J]. *系统工程与电子技术*, 2022, 44(11): 3548–3555. doi: [10.12305/j.issn.1001-506X.2022.11.31](https://doi.org/10.12305/j.issn.1001-506X.2022.11.31).
WANG Renzhi, PAN Kegang, and ZHAO Ruixiang. Optimal design of LDPC Codes in frequency hopping anti-jamming communication system[J]. *Systems Engineering and Electronics*, 2022, 44(11): 3548–3555. doi: [10.12305/j.issn.1001-506X.2022.11.31](https://doi.org/10.12305/j.issn.1001-506X.2022.11.31).

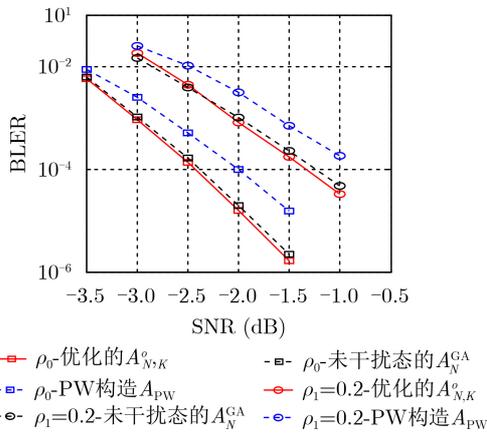


图 8 CRC辅助SCL译码时不同信息位序列构造方案的BLER结果

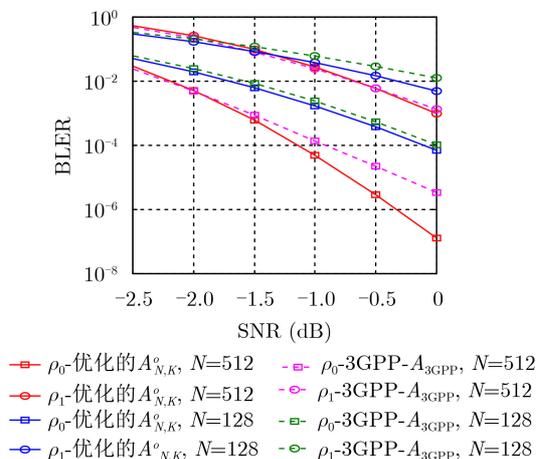


图 9 与5G 3GPP标准信息位序列构造方案的BLER对比结果

- [8] 孙康宁, 马林华, 茹乐, 等. 混合信道下LDPC码稳定条件分析及度序列优化[J]. 通信学报, 2016, 37(9): 168–174. doi: [10.11959/j.issn.1000-436x.2016188](https://doi.org/10.11959/j.issn.1000-436x.2016188).
SUN Kangning, MA Linhua, RU Le, *et al.* Analysis of stability condition for LDPC codes and optimizing degree sequences over mixed channel[J]. *Journal on Communications*, 2016, 37(9): 168–174. doi: [10.11959/j.issn.1000-436x.2016188](https://doi.org/10.11959/j.issn.1000-436x.2016188).
- [9] 刘士平, 马林华, 孙康宁, 等. 阻塞式干扰环境下LDPC编码跳频通信优化方案[J]. 火力与指挥控制, 2019, 44(2): 32–36, 40. doi: [10.3969/j.issn.1002-0640.2019.02.007](https://doi.org/10.3969/j.issn.1002-0640.2019.02.007).
LIU Shiping, MA Linhua, SUN Kangning, *et al.* Optimizing of LDPC coded frequency-hopping communication over blocking interference[J]. *Fire Control & Command Control*, 2019, 44(2): 32–36, 40. doi: [10.3969/j.issn.1002-0640.2019.02.007](https://doi.org/10.3969/j.issn.1002-0640.2019.02.007).
- [10] MORI R and TANAKA T. Performance of polar codes with the construction using density evolution[J]. *IEEE Communications Letters*, 2009, 13(7): 519–521. doi: [10.1109/LCOMM.2009.090428](https://doi.org/10.1109/LCOMM.2009.090428).
- [11] TRIFONOV P. Efficient design and decoding of Polar codes[J]. *IEEE Transactions on Communications*, 2012, 60(11): 3221–3227. doi: [10.1109/TCOMM.2012.081512.110872](https://doi.org/10.1109/TCOMM.2012.081512.110872).
- [12] HE Gaoning, BELFIORE J C, LAND I, *et al.* Beta-expansion: A theoretical framework for fast and recursive construction of Polar codes[C]. Proceedings of 2017 IEEE Global Communications Conference, Singapore, 2017: 1–6. doi: [10.1109/glocom.2017.8254146](https://doi.org/10.1109/glocom.2017.8254146).
- [13] LI Jianxiu and CHENG Wenchi. Stacked denoising autoencoder enhanced Polar codes over Rayleigh fading channels[J]. *IEEE Wireless Communications Letters*, 2020, 9(3): 354–357. doi: [10.1109/LWC.2019.2954907](https://doi.org/10.1109/LWC.2019.2954907).
- [14] LIAO Yun, HASHEMI S A, CIOFFI J M, *et al.* Construction of polar codes with reinforcement learning[J]. *IEEE Transactions on Communications*, 2022, 70(1): 185–198. doi: [10.1109/TCOMM.2021.3120274](https://doi.org/10.1109/TCOMM.2021.3120274).
- [15] TENG C F and WU A Y A. Convolutional neural network-aided tree-based bit-flipping framework for polar decoder using imitation learning[J]. *IEEE Transactions on Signal Processing*, 2021, 69: 300–313. doi: [10.1109/TSP.2020.3040897](https://doi.org/10.1109/TSP.2020.3040897).
- [16] LU Yang, ZHAO Mingmin, LEI Ming, *et al.* Deep learning aided SCL decoding of polar codes with shifted-pruning[J]. *China Communications*, 2023, 20(1): 153–170. doi: [10.23919/JCC.2023.01.013](https://doi.org/10.23919/JCC.2023.01.013).
- [17] KORADA S B and URBANKE R L. Polar codes are optimal for lossy source coding[J]. *IEEE Transactions on Information Theory*, 2010, 56(4): 1751–1768. doi: [10.1109/TIT.2010.2040961](https://doi.org/10.1109/TIT.2010.2040961).
- [18] LIU Chunming, XU Xin, and HU Dewen. Multiobjective reinforcement learning: A comprehensive overview[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2015, 45(3): 385–398. doi: [10.1109/TSMC.2014.2358639](https://doi.org/10.1109/TSMC.2014.2358639).
- 梁 豪: 男, 助理研究员, 研究方向为信道编码, 卫星通信.
叶淦华: 男, 副研究员, 研究方向为卫星通信抗干扰.
陆锐敏: 男, 研究员, 研究方向为卫星通信, 抗干扰通信.
王 恒: 男, 高级工程师, 研究方向为卫星通信中的信号处理.
魏 鹏: 男, 副研究员, 研究方向为卫星通信抗干扰, 智能抗干扰.

责任编辑: 余 蓉