

面向5G-NR标准的高效LDPC译码系统设计与FPGA实现

孙源^{1,2}, 易杨¹, 宫丰奎¹, 惠腾飞^{2,1}

(1. 西安电子科技大学空天地一体化综合业务网全国重点实验室, 陕西 西安 710071; 2. 空间微波通信全国重点实验室, 陕西 西安 710100)

摘要: 针对第五代移动通信技术(5G)新空口(NR)对高速率与可靠数据传输的需求, 提出一种高效的低密度奇偶校验(LDPC)译码系统。该系统通过奇偶交织将准循环矩阵分解为两个子矩阵, 结合差分循环移位网络, 显著降低了硬件资源冗余和布线拥塞。在此基础上, 引入3帧交织流水线与离线静态调度策略, 利用预生成的微码指令流取代复杂的冲突检测, 减少了由层间数据依赖引发的流水线停顿。基于现场可编程门阵列(FPGA)的实现与定点测试结果表明, 所提译码系统在5G-NR标准定义的双基图及各种码长、码率和提升因子配置下, 误帧率(FER)曲线均能快速收敛, 且无明显错误平层现象。该系统具备与国际先进商用知识产权(IP)核相当的归一化吞吐率资源比, 并进一步集成了解速率匹配模块, 可满足非地面网络(NTN)等场景下的信道译码需求。

关键词: 5G新空口; 低密度奇偶校验码; 译码器; 解速率匹配; 现场可编程门阵列

中图分类号: TN911.22

文献标志码: A

Design and FPGA implementation of an efficient LDPC decodingsystem for the 5G-NR standard

Sun Yuan^{1,2}, Yi Yang¹, Gong Fengkui¹, Hui Tengfei^{2,1}

1. State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071, China

2. National Key Laboratory of Space Microwave Communications, Xi'an 710100, China

Abstract: To meet the requirements of fifth-generation (5G) new radio (NR) for high-rate and reliable data transmission, an efficient low-density parity-check (LDPC) decoding system was proposed. The system used odd-even interleaving to decompose each quasi-cyclic matrix into two submatrices, while a differential circular-shift network was employed to further reduce hardware redundancy and alleviate routing congestion. To reduce pipeline stalls caused by inter-layer data dependencies, a three-frame interleaved pipeline and an offline static scheduling strategy were introduced. The strategy replaced complex conflict detection with a pre-generated microcode stream. Field-programmable gate array (FPGA) implementation and fixed-point simulation results show that, for both base graphs and various configurations of code length, code rate, and lifting size defined in the 5G-NR standard, the frame error rate (FER) decreases rapidly with no noticeable error floor. The proposed system achieves a normalized throughput-to-resource ratio comparable to state-of-the-art commercial intellectual property (IP) cores and further integrates a rate-dematching module. These results demonstrate that the proposed system can satisfy the channel decoding requirements of non-terrestrial network (NTN) and related scenarios.

Key words: 5G new radio, low-density parity-check code, decoder, rate dematching, field-programmable gate array

收稿日期: 2026-06-11; 修回日期: 2026-09-05

通信作者: 宫丰奎, fkgong@xidian.edu.cn

Foundation Items: The Youth Innovation Team of Shaanxi Universities

0 引言

随着第五代移动通信技术 (fifth-generation mobile communication technology, 5G) 的大规模商用与持续演进, 通信网络正向万物互联的多样化场景扩展。国际电信联盟在 IMT-2020 愿景中明确了 5G 的三大典型应用场景^[1]。其中, 增强型移动宽带 (enhanced mobile broadband, eMBB) 场景对物理层数据链路提出了 20 Gbps 峰值吞吐率与毫秒级端到端时延的严格要求。近年来, 5G 还加速向非地面网络 (non-terrestrial network, NTN) 等星地融合场景拓展, 其中海量数据的宽带传输对底层基带系统的高吞吐率与高可靠性提出了更为严苛的挑战。面对上述挑战, 低密度奇偶校验 (low-density parity-check, LDPC) 码凭借其逼近香农限的优异纠错性能, 以及支持大规模并行处理的稀疏校验矩阵, 被确立为 5G 数据信道的编码方案^[2]。然而, 在基于现场可编程门阵列 (field-programmable gate array, FPGA) 等硬件平台的实现中, 5G LDPC 译码器仍面临硬件效率受限、布线拥塞、多码率兼容困难以及存储器读写冲突等一系列瓶颈^[3-4]。

在译码算法方面, 传统的置信传播 (belief propagation, BP) 算法^[5]是 LDPC 码的理论最优译码算法, 但其涉及双曲正切及其反函数等复杂的非线性运算。为降低计算复杂度, Fossorier 等^[6]提出了最小和 (min-sum, MS) 算法。该算法将校验节点 (check node, CN) 的更新过程简化为寻找最小值和符号异或运算。然而, MS 算法会导致校验节点消息幅度被过度估计, 相较 BP 算法存在约 0.5~1 dB 的性能损失。Chen 等^[7]提出了归一化最小和 (normalized min-sum, NMS) 与偏移最小和 (offset min-sum, OMS) 算法, 二者分别引入乘性衰减因子与加性偏移因子对消息幅度进行修正, 以极小的硬件开销弥补了过度估计造成的性能损失, 现已广泛应用于 LDPC 译码器。在此基础上, 为了更好地适应节点可靠度的动态变化, Fan 等^[8]提出了自适应 NMS 算法, 依据当前迭代次数与节点信息动态调整衰减因子。针对 5G LDPC 码, Zhang 等^[9]指出修正因子也应与层级和 CN 度数相关, 进而提出了二维 NMS/OMS 算法, 显著提升了高码率下的纠错性能。Ren 等^[10]提出了广义调整 MS 算法, 通过分段线性函数拟合 BP 算法中的双曲正切运算, 将性能损失控制在 0.1 dB 以内。Tera 等^[11]引入深度学习

技术, 利用训练后的卷积神经网络对偏移因子进行实时估计与调整, 大幅增强了 OMS 算法在复杂信道环境下的鲁棒性。除改善译码性能外, 还有大量研究聚焦于进一步简化 MS 算法。为避免 CN 更新时寻找前两个最小值的复杂逻辑, 文献[12-14]给最小值添加一个偏移值以模拟次小值。Cheng 等^[15]将树形比较网络中最小值的最后一个竞争者视为概率次小值, 用来代替真正的次小值。为提高概率次小值的可靠度, Wang 等^[16]又将其与最小值按比例合并。文献[17-18]则将相邻变量节点 (variable node, VN) 传递的消息划分为若干组, 先提取各组内的局部最小值, 再从中确定前两个最小值。

在硬件架构方面, 传统的全并行架构虽能实现极高的吞吐率, 但资源消耗较大且不支持可变的码长与码率; 串行架构虽能显著节省资源, 却难以满足通信系统的低时延要求。为此, Hocevar 等^[19]提出了分层译码算法, 通过将校验矩阵分层并依次更新, 大幅加快了译码收敛速度, 平衡了吞吐率与硬件开销, 目前已成为针对 5G-NR 标准的主流选择。Li 等^[20]重点优化了分层译码过程中的流水线调度机制和节点处理逻辑, 有效缩短了关键路径。Nguyen 等^[21]提出一种深度流水线的层并行架构, 在 FPGA 上实现了超过 10 Gbps 的吞吐率。Mejmaa 等^[22]针对准循环 LDPC (quasi-cyclic LDPC, QC-LDPC) 码特有的宏矩阵结构, 优化了数据路由与存储访问策略, 实现了兼顾高吞吐量与低延迟的译码器。Wei 等^[23]利用层间正交特性对提升因子进行拆分, 提出了更灵活的平衡吞吐量与硬件开销方案。针对分层调度中数据相关性带来的流水线气泡问题, Boncalo 等^[24]提出了基于残差的动态调度策略与离线重排方案。Chen 等^[25]通过优化矩阵映射机制, 有效消除了数据访问冲突, 同时减少了存储资源开销。Jia 等^[26]提出一种半动态分层调度方案, 通过引入核心矩阵延迟更新技术, 以极低硬件开销为代价, 有效减少了迭代次数并改善了纠错性能。Nimara 等^[27]提出面向多码字并行处理的高效架构, 通过多个数据块复用同一物理存储器, 大幅提升了吞吐密度。

上述研究主要通过优化顶层架构与调度策略来减少资源消耗、提升吞吐率。除此之外, 复杂的循环移位网络往往也会限制译码器硬件效率的进一步提升。面向 5G-NR 标准中多种提升因子, Song

等^[28]设计了通用的高效循环移位网络,通过调整置换顺序简化了控制逻辑。Nguyen 等^[29]利用级联的细粒度移位器缓解了多尺寸循环移位网络中控制信号的高扇出和布线拥塞问题。Chen 等^[30]对传统 Benes 网络进行定制化裁剪,提出 QC-LDPC 移位网络 (QC-LDPC shift network, QSN),在简化控制逻辑的同时提升了译码器的时钟频率。综上所述,学术界对 5G LDPC 译码器的研究正向着高吞吐率、低资源消耗的方向演进。然而,面对 5G-NR 标准中定义的各种 LDPC 码,如何在有限的硬件资源约束下降低逻辑与布线复杂度,同时消除流水线停顿,仍需更深入的研究。

此外,以 Xilinx^[31]和 Tannera^[32]为代表的 FPGA 厂商均已推出商用 5G LDPC 译码器知识产权 (intellectual property, IP) 核,代表了目前业界的最高性能水准。然而,商用 IP 核主要提供核心译码功能,并未集成完整解速率匹配模块。

针对上述挑战,本文设计并实现了一种高吞吐率、低硬件开销且支持解速率匹配的 LDPC 译码系统。该系统在 5G-NR 标准定义的各种 LDPC 码下均展现出优秀的译码性能,且具有与国际先进 IP 核相当的硬件效率。本文的主要贡献如下。

1) 采用折叠并行架构,通过奇偶行列交织将循环置换矩阵分解为两个子矩阵,将译码器的最大并行度由 384 缩减至 192。该架构避免了当处理较小提升因子时,大量的节点更新单元处于闲置状态,有效减少了硬件冗余。并行度的降低还能缓解布线拥塞,从而提高了系统的最高工作频率。

2) 提出一种差分循环移位网络,用相邻层间准循环矩阵的移位差值代替真实移位值,并将其分解为粗、中、细粒度三级流水线。更新后的 VN 消息无需反向循环移位至初始状态,只需进行差分移位即可用于下一层 CN 更新,从而避免了冗余移位操作。

3) 采用 3 帧交织流水线技术,在当前帧因层间数据依赖进入等待状态时,执行另外两帧的译码,从而减少流水线气泡对系统吞吐率的影响。

4) 提出了一种离线静态调度策略,首先利用 MATLAB 遍历 3 帧交织译码过程,调整相邻层间共享变量节点的访问顺序,随后生成按时钟周期排列的包含各模块控制信号的微码指令流,以代替译码器运行时复杂的冲突检测逻辑。

5) 设计了一种面向 5G-NR 标准的高效解速率匹配模块,采用分级存储架构与输出移位寄存器,在提高存储资源利用率的同时简化了控制逻辑。

1 最小和译码算法及其改进

令 $L_{\text{ch}}(i)$ 和 $\Gamma_i^{(t)}$ 表示第 i 个 VN 的信道对数似然比 (log-likelihood ratio, LLR) 和经过 t 次迭代的后验 LLR。此外,用 $\alpha_{i \rightarrow j}^{(t)}$ 表示由第 i 个 VN 传递至第 j 个 CN 的 V2C 消息, $\beta_{j \rightarrow i}^{(t)}$ 表示由第 j 个 CN 传递至第 i 个 VN 的 C2V 消息。MS 算法的译码步骤如下:

1) 利用信道 LLR 初始化 V2C 消息。

$$\alpha_{i \rightarrow j}^{(0)} = L_{\text{ch}}(i), \quad \forall j \in \mathcal{N}(i) \quad (1)$$

式中, $\mathcal{N}(i)$ 表示与第 i 个 VN 相邻的 CN 集合。

2) 在第 t 次迭代中, C2V 消息的更新可分为符号异或与寻找最小幅度两部分。

$$\text{sign}(\beta_{j \rightarrow i}^{(t)}) = \prod_{k \in \mathcal{M}(j) \setminus \{i\}} \text{sign}(\alpha_{k \rightarrow j}^{(t-1)}) \quad (2)$$

$$|\beta_{j \rightarrow i}^{(t)}| = \min_{k \in \mathcal{M}(j) \setminus \{i\}} |\alpha_{k \rightarrow j}^{(t-1)}| \quad (3)$$

式中, $\mathcal{M}(j) \setminus \{i\}$ 表示除了第 i 个 VN, 其余与第 j 个 CN 相邻的 VN 集合。为减小式(3)对 C2V 消息幅度的过估计, NMS 算法将其乘以衰减因子 δ , 无衰减 OMS 算法则减去偏移因子 ε , 如式(4)和式(5)所示。二者均能以较低的额外复杂度改善译码性能^[7]。

$$|\beta_{j \rightarrow i}^{(t)}| = \delta \times \min_{k \in \mathcal{M}(j) \setminus \{i\}} |\alpha_{k \rightarrow j}^{(t-1)}| \quad (4)$$

$$|\beta_{j \rightarrow i}^{(t)}| = \max \left(\min_{k \in \mathcal{M}(j) \setminus \{i\}} |\alpha_{k \rightarrow j}^{(t-1)}| - \varepsilon, 0 \right) \quad (5)$$

3) 更新 V2C 消息。

$$\alpha_{i \rightarrow j}^{(t)} = L_{\text{ch}}(i) + \sum_{l \in \mathcal{N}(i) \setminus \{j\}} \beta_{l \rightarrow i}^{(t)} \quad (6)$$

式中, $\mathcal{N}(i) \setminus \{j\}$ 表示排除第 j 个 CN 后的 $\mathcal{N}(i)$ 。

4) 通过累加信道 LLR 与 C2V 消息, 计算后验 LLR 并进行硬判决。

$$\Gamma_i^{(t)} = L_{\text{ch}}(i) + \sum_{j \in \mathcal{N}(i)} \beta_{j \rightarrow i}^{(t)} \quad (7)$$

$$\hat{c}_i^{(t)} = \begin{cases} 0 & \text{if } \Gamma_i^{(t)} \geq 0 \\ 1 & \text{otherwise} \end{cases} \quad (8)$$

当硬判决结果 $\hat{c}^{(t)}$ 满足矩阵中所有奇偶校验方程或达到最大迭代次数时终止译码, 否则继续下一次迭代。

2 5G LDPC 译码器设计

为了满足 5G eMBB 场景对数据传输速率和可靠性的严格要求, 本文首先比较了主流的泛洪调度与分层调度, 以及 NMS 与 OMS 算法。在所仿真的 5G-NR 配置下, 分层调度所需的迭代次数约为泛洪调度的一半, 且 OMS 的性能曲线在瀑布区更陡峭。考虑到衰减操作会引入额外的乘法或移位加法逻辑, 本文采用收敛速度快且性能较优的分层无衰减 (即衰减因子 $\delta = 1$) OMS 算法, 其偏移因子 $\varepsilon = 0.5$ 。此外, 信道 LLR、C2V 和 V2C 消息采用 6 bit 量化、后验 LLR 采用 9 bit 量化, 小数部分位宽均为 2 bit 时, 可以在硬件开销与译码性能间取得最优折中。

进一步分析现有研究可知, 译码器设计的核心瓶颈在于兼容 5G-NR 标准规定的双基图以及多种码长、码率和提升因子导致的硬件效率受限, 以及分层译码中由数据依赖引发的流水线停顿。为突破上述瓶颈, 本文结合折叠并行架构、差分循环移位网络、多帧交织流水线与离线静态调度策略, 设计了一种高吞吐率、低资源消耗的 5G LDPC 译码器架构。

2.1 基于奇偶交织特性的折叠并行架构

传统 5G LDPC 译码器通常采用全尺寸并行架构, 即用 5G-NR 标准定义的最大提升因子 $Z_{c, \max} = 384$ 作为硬件并行度。然而, 这种架构存在严重的资源冗余: 在处理较小的提升因子时, 大量节点更新单元处于闲置状态, 降低了硬件资源利用率。此外, 384 路数据的并行输入和处理会引发严重的布线拥塞, 进而限制系统的最高工作频率。

本文采用了一种折叠并行架构, 通过对行索引与列索引进行奇偶交织重排, 将准循环矩阵分解为两个子矩阵。如图 1 所示, 原始矩阵为 $Z_c \times Z_c$ 的准循环矩阵, 经过奇偶行列变换后被分解为 2 个 $(Z_c/2) \times (Z_c/2)$ 的小型准循环矩阵与 2 个全零矩阵。

采用该折叠并行架构后, 译码器的最大并行度由 384 缩减至 192。当 $Z_c \leq 192$ 时, 处于空闲状态的节点更新单元数量明显减少。对于 $Z_c > 192$ 的高并行度场景, 根据协议规范, 取值集合中的元素恒为偶数, 表明定义在伽罗华域上的准循环矩阵具有可分解的稀疏特性。因此可以将奇偶行列变换后的

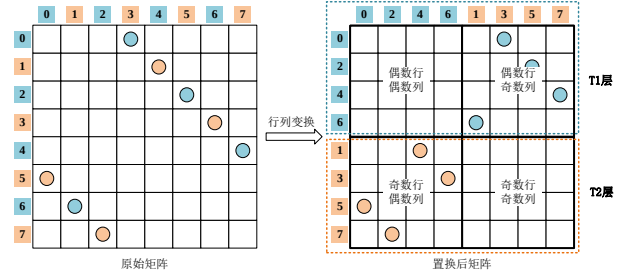


图 1 基于奇偶重排的准循环矩阵分解示意图

上半部分偶数行分配给第一个时隙 T_1 , 下半部分奇数行分配给第二个时隙 T_2 。对于任意准循环矩阵, 假设其对应的非零移位值为 S ($0 \leq S < Z_c$), 采用向下取整与向上取整算子, 将原本基于逻辑判断的参数计算转化为更加简洁的代数表达式。两个子层对应的等效循环移位值 $S^{(i)}$ 为:

$$\begin{cases} S^{(1)} = \left\lfloor \frac{S}{2} \right\rfloor, & T_1 \text{ 层} \\ S^{(2)} = \left\lceil \frac{S}{2} \right\rceil, & T_2 \text{ 层} \end{cases} \quad (9)$$

通过时分复用, 可以将高并行度的节点更新过程分解为两个连续的时隙执行, 从而在硬件效率与吞吐量之间取得最佳平衡。

2.2 差分循环移位网络

为了在有限的硬件资源下实现低延迟、大位宽总线数据的任意循环移位, 本文设计了一种三级差分循环移位网络, 其结构如图 2 所示。

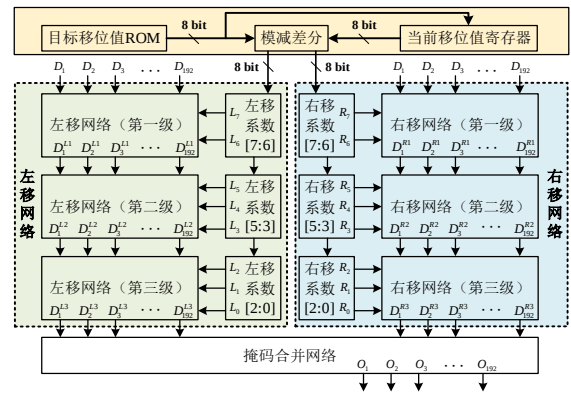


图 2 三级差分循环移位网络结构框图

该循环移位网络首先对目标移位值与当前移位值进行模减运算, 随后将所得 8 bit 差分移位值划分为粗、中、细三个粒度, 依次进行流水线处理:

1) 第一级: 由差分移位值的高 2 位控制, 支持步长为 64 bit 的粗粒度移位。

2) 第二级: 由中间 3 位控制, 支持步长为 8 bit 的中粒度移位。

3) 第三级: 由低 3 位控制, 支持步长为 1 bit 的细粒度移位, 完成按位对齐。

采用上述三级级联结构, 循环移位网络能够以较少的多路选择器资源实现 0 至 192 范围内的任意循环移位。该模块还并行部署了左右移互补与动态掩码逻辑, 确保数据的高效处理, 同时避免提升因子较小时大量无效数据对后续最小和运算的干扰。

为了进一步降低循环移位网络的信号翻转率与动态功耗, 本文还引入了差分对齐策略。在分层译码过程中, 首先从只读存储器 (read-only memory, ROM) 中读取目标移位值, 并与当前移位值寄存器的输出进行模减运算, 得到差分移位值。随后, 将目标移位值写入当前移位值寄存器, 作为后续模减运算的反馈输入。下一层 CN 更新时, 后验 LLR 不必反向循环移位至初始状态, 只需在当前基础上进行差分循环移位即可完成数据对齐, 从而避免了冗余的移位操作。

2.3 多帧交织流水线架构

相较于传统的泛洪调度策略, 分层译码算法虽然显著加快了收敛速度, 但其固有的层间数据依赖性给流水线设计带来了严峻挑战。在分层译码过程中, 第 l 层的节点更新必须依赖第 $l-1$ 层产生的最新消息。当第 l 层更新开始时, 第 $l-1$ 层的节点消息可能尚未完全写入存储器, 这会引发典型的读后写数据冒险。为保证数据一致性, 控制器被迫插入等待周期。在高速 LDPC 译码器的 FPGA 实现中, 由于关键路径的逻辑级数较长, 流水线延迟 D_{pipe} 较大, 导致等待周期在总处理时间中占比较高, 严重限制了系统吞吐量。

为减少流水线停顿, 本文利用不同译码帧之间的数据独立性, 提出了一种多帧交织流水线架构, 如图 3 所示。以 3 帧交织为例, 该架构的核心思想是在当前帧 Frame k 等待层间数据更新的空闲时间内, 交替插入后续两帧 Frame $k+1$ 和 Frame $k+2$, 即通过时分复用机制并行执行三个独立帧的译码。

具体而言, 控制器将流水线划分为连续的时隙: 在第 T 个时隙, 流水线处理 Frame 1 的 Layer i ; 在第 $T+1$ 个时隙, 处理 Frame 2 的 Layer i ; 在第 $T+2$ 个时隙, 处理 Frame 3 的 Layer i 。在第 $T+3$ 个时隙, 返回 Frame 1 处理 Layer $i+1$ 。此时, 经

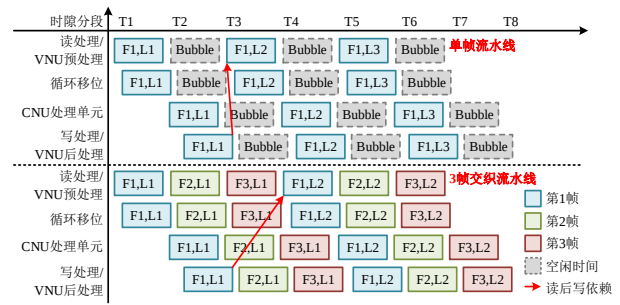


图3 传统单帧流水线与3帧交织流水线对比

过 3 个时隙, Frame 1 的 Layer i 节点消息均已经完成更新并写入随机存取存储器 (random access memory, RAM), 系统可直接读取 Layer i 的最新数据。当交织帧数 M 与单层处理时间 T_{layer} 满足 $M \times T_{layer} \geq D_{pipe}$ 时, 该架构可以利用其它帧的有效译码填充气泡, 从而提高流水线利用率和系统吞吐量。

此外, 为确保多帧数据的无冲突访问, 将存储后验 LLR 与 C2V 消息的 RAM 在物理空间上进行分块以实现严格隔离, 并为每帧配置了独立的迭代计数器。该设计使系统能够灵活处理各帧译码迭代次数不同的情况。

2.4 离线静态调度策略

在高吞吐率的 5G LDPC 译码器中, 控制器承担着协调 VN 与 CN 更新、避免存储器读写冲突以及维护流水线数据一致性等核心任务。传统译码器通常采用在线动态调度方案, 即在 FPGA 内部部署复杂的有限状态机来实时监测层间数据相关性并仲裁流水线停顿。然而, 该方案存在两大缺陷: 一是实时检测多帧交织的 192 路并行数据的读写地址冲突需要构建庞大的比较器网络与仲裁逻辑, 硬件开销巨大; 二是复杂的仲裁逻辑位于控制通路的临界路径上, 显著增加了组合逻辑级数, 限制了系统的最大工作频率。

为此, 本文提出了一种基于软硬协同的离线静态调度策略。其核心思想是将复杂的冲突检测与流水线编排任务从 FPGA 运行时提前至 MATLAB 离线阶段, 从而有效降低译码器运行时的动态调度逻辑开销。

具体而言, 本文首先在 MATLAB 环境中建立流水线模型, 并导入 RTL 中各运算单元的流水线延迟参数。随后遍历基矩阵的行与列, 模拟多帧交织译码过程中的数据流和存储器状态。当检测到相邻层间存在读写冲突时, 交替插入其它帧并调整存储

器访问顺序,以减少流水线停顿。最终,将各模块的控制信号封装为 56 bit 微码并导出为 COE 文件,用于初始化 FPGA 片上控制器 ROM。生成相应微码指令流的离线静态调度算法如下所示。

算法 1 静态流水线调度与微码生成

输入 5G LDPC 码基矩阵集合 $\{H_{bg0}, \dots, H_{bg7}\}$, 移位延迟 $D_{shift} = 5$, CN 更新延迟 $D_{cnu} = 4$, 总流水线深度 $D_{pipe} = 15$, 节点并行度 $P = 192$, 交织帧数 M , 提升因子 Z_c , 基矩阵层数 R_{num}

输出 56 bit 微码指令流 COE 文件

```

1) 初始化全局时间指针  $T_{curr}$ , 微码数组 Cmd,
结束时间标记  $T_{last}$ 
2) for  $l = 1$  to  $R_{num}$  do
3)  $Col = \{j | H_{bg}(l, j) \neq -1\}$ 
4)  $N_{valid} = \text{length}(Col)$ 
5)  $T_{start} = T_{curr}$ 
6) if  $T_{start} < T_{last} - D_{pipe}$  then
7)  $T_{start} = T_{last} - D_{pipe} + \Delta_{safe}$ 
8) end if
// 相邻层间读写冲突检测
9)  $Cmd[T_{start}] = \{VN_{RdEn} = 1, VN_{Addr} = Col\}$ 
// 读取后验 LLR
10)  $Cmd[T_{start} + 2] = \{Shift_{Val} = f(H_{bg}, Z_c)\}$ 
// 配置差分循环移位网络
11)  $Cmd[T_{start} + 6] = \{CN_{RdEn} = 1, CN_{Addr} = R_{ptr}\}$ 
// 读取 C2V 消息
12)  $Cmd[T_{start} + 10] = \{Temp_{WrEn} = 1\}$ 
// 缓存 V2C 消息
13)  $Cmd[T_{start} + 13] = \{CN_{WrEn} = 1, CN_{Addr} = R_{ptr}\}$ 
// 更新 C2V 消息
14)  $Cmd[T_{start} + 15] = \{VN_{WrEn} = 1, VN_{Addr} = Col\}$ 
// 更新后验 LLR
15)  $T_{curr} = T_{start} + N_{valid} \times M$ 
16)  $T_{last} = T_{start} + D_{pipe} + N_{valid} \times (M + 1)$ 
// 更新时间指针
17) end for
18) for  $t = 0$  to  $T_{last}$  do
19) 将  $Cmd[t]$  中控制信号打包为 56 bit 指令字
20) 写入 COE 文件

```

21) end for

// 封装并导出微码

22) return COE 文件

该算法生成的每条 56 bit 微码包含后验 LLR、C2V 和 V2C 消息存储器的读写使能与地址,以及交织帧编号等信号。利用 COE 文件初始化控制器 ROM 后,译码器只需按时钟顺序读取并解包。

表 1 不同参数下的等效单帧时钟周期数

基图	提升因子	$M = 1$	$M = 2$	$M = 3$	$M = 4$
BG1	$Z_c \leq 192$	1104	566	407.33	344.25
	$Z_c > 192$	2202	1119	803	677.5
BG2	$Z_c \leq 192$	917	460.5	313	245.5
	$Z_c > 192$	1828	916	618.67	484

由表 1 可知,在离线静态调度下,随着交织帧数 M 增加,平均每帧完成单次迭代所需的时钟周期数逐渐减少,但降幅趋于平缓。此外,离线静态调度虽然简化了控制逻辑,但存储开销仍随交织帧数增加。综合考虑系统吞吐率与硬件效率,本文最终采用 3 帧交织。

3 解速率匹配模块设计

速率匹配是 5G-NR LDPC 编码链路的重要组成部分,包含比特选择和比特交织。前者通过打孔、缩短和重复等操作动态调整编码输出比特数,以适配可用传输资源和目标码率;后者根据调制阶数重排输出序列,从而抵抗信号衰落和突发干扰。此外,速率匹配可通过传输码率兼容 LDPC 码的不同冗余版本,进一步支持混合自动重传请求机制。然而,大部分商用 IP 核局限于核心译码器,并未集成完整解速率匹配模块。

为完善译码系统,本文进一步设计了一种面向 5G-NR 标准的高效解速率匹配模块。与发送端的速率匹配对应,该模块由解交织与比特恢复两部分组成。

3.1 解交织模块设计

解交织模块采用不同容量的存储器与特定的读写策略完成数据重组,整体结构如图 4 所示。其中, FIFO1、FIFO2 与 FIFO3 的位宽均为 144 bit,深度分别为 2400、550 和 200。该模块接收 12 路并行的 6 bit LLR 信息,当调制阶数 $Q_m = 2$ 时,仅使用 FIFO1;当 $Q_m = 4$ 时,数据交替写入 FIFO1 与

FIFO2; 当 $Q_m = 6$ 时, 数据循环写入三个 FIFO。这种分级存储方案既确保了不同调制模式下数据存储的一致性, 又提高了存储资源利用率。

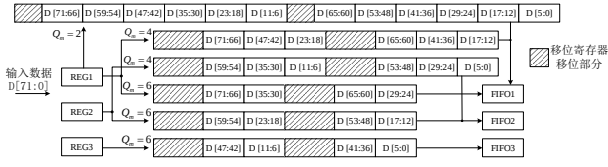


图4 解交织模块数据处理示意图

由于 5G-NR 标准中存在非 72 bit 整数倍位宽的码块, 为确保输出数据的连续性, 本模块采用移位寄存器作为输出缓冲区。从 FIFO 中读出的数据首先写入该移位寄存器, 每次读出有效的 72 bit 并统计已输出的数据量。当达到当前码块长度时自动复位, 并准备接收下一个码块。使用该缓冲区可以有效解决 FIFO 边界之间的数据拼接问题。

3.2 比特恢复模块设计

比特恢复模块首先将解交织后的 12 路并行 LLR 信息按顺序写入 RAM。当速率匹配输出长度 E_r 超过循环缓冲区长度 N_{cb} 时, 将重复比特与 RAM 中相应位置存储的 LLR 信息相加, 并将结果写回原地址。为避免数据溢出, 需要将位宽扩展至 9 bit。

完成重复比特的软信息合并后, 每次从 RAM 中读出 12 路 9 bit LLR 信息, 经饱和限幅恢复为 6 bit。考虑到后续奇偶行列交织单元的并行度为 Z_c , 该模块先将 LLR 信息写入移位寄存器, 再以 Z_c 路为一组并行输出, 并统计数量。当已输出的 LLR 信息数量位于填充比特区间时, 暂停读取 RAM 并插入最大正数, 否则仍按顺序读取并输出。该设计避免了比特恢复过程中的大量碎片化赋值, 从而有效简化了控制逻辑。

4 实现结果与分析

为全面评估本文设计的核心译码器与解速率匹配模块的性能, 在 Xilinx UltraScale+ XCVU13P 上实现了所提 5G LDPC 译码系统, 整体结构如图 5 所示。

在该系统中, 核心 LDPC 译码器采用衰减因子为 1、偏移因子为 0.5 的分层 OMS 算法, 信道 LLR、C2V 和 V2C 消息采用 Q4.2 量化, 后验 LLR 采用 Q7.2 量化, 最大译码迭代次数设为 10。针对 5G-

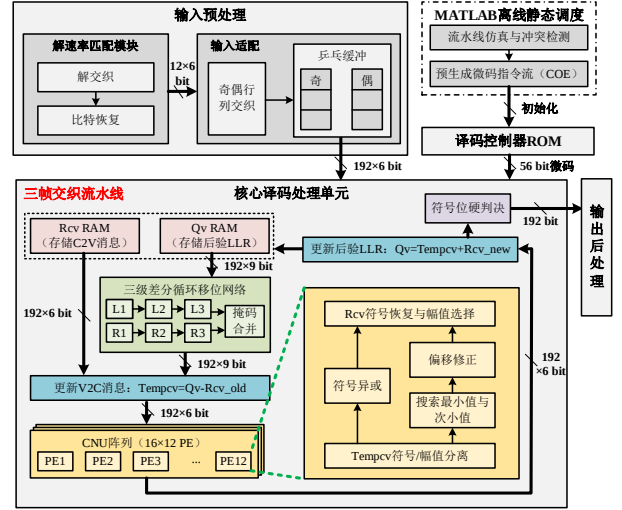


图5 所提 5G LDPC 译码系统整体结构框图

NR 标准规定的两种基图, 分别选取具有代表性的传输块大小 (transport block size, TBS)、码率和提升因子配置进行 FPGA 定点测试与 MATLAB 浮点仿真, 结果如图 6 至图 9 所示。

在所述典型配置下, FPGA 定点实测与 MATLAB 浮点仿真的误帧率 (frame error rate, FER) 曲线均在瀑布区迅速收敛。在 $FER = 10^{-4}$ 时, 二者间差距不到 0.4 dB, 且会随码率的提升逐渐缩小。该性能损失主要源于定点化时引入的量化噪声, 处于工程允许的合理范围内。此外, 在测试信噪比区间内, FER 曲线均未出现明显的错误平台现象, 证明该系统能在有限的迭代次数与量化位宽约束下稳定收敛。

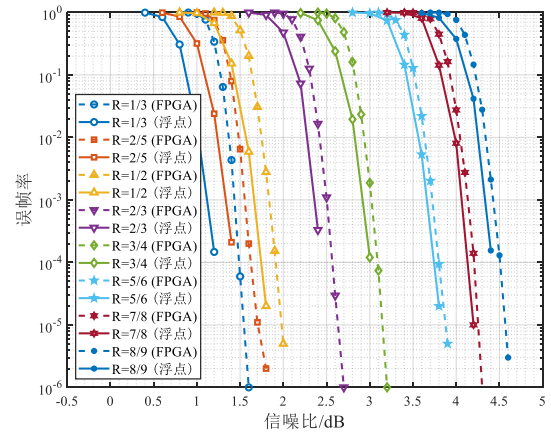
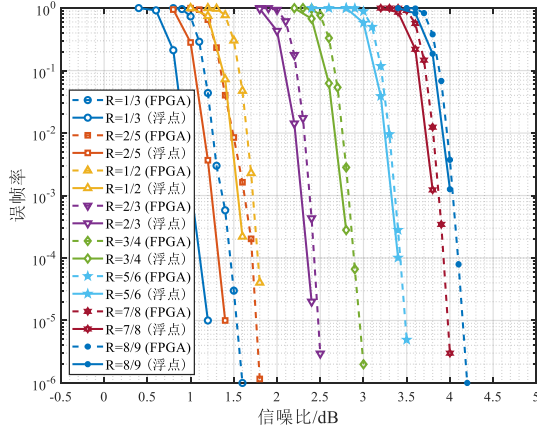
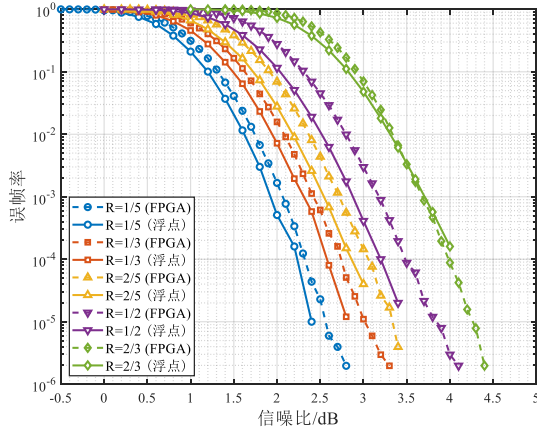
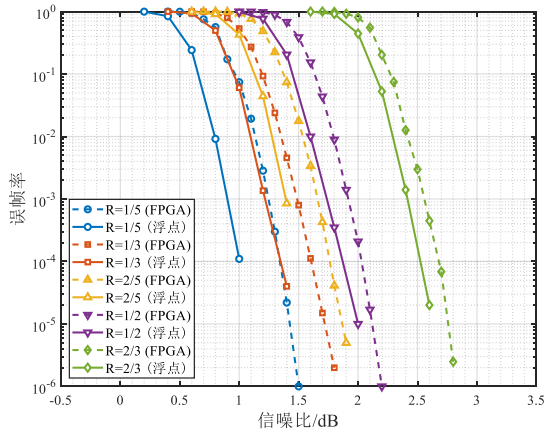


图6 BGI, TBS = 3840, $Z_c = 176$, FPGA 实测与浮点仿真对比

采用 Vivado 2018.3 完成综合、布局与布线后,

图7 BG1, TBS = 8192, $Z_c = 384$, FPGA实测与浮点仿真对比图8 BG2, TBS = 304, $Z_c = 40$, FPGA实测与浮点仿真对比图9 BG2, TBS = 3824, $Z_c = 384$, FPGA实测与浮点仿真对比

系统最高工作频率 $f_{\max}$ 为300 MHz, 相应的信息吞吐率 T_{info} 为:

$$T_{\text{info}} = \frac{N_{\text{info}} \times f_{\max}}{\text{Iter}_{\max} \times T_{\text{frame}}} \quad (10)$$

式中, N_{info} 表示信息位长度, $\text{Iter}_{\max}$ 为最大迭代次数, 取值为10, T_{frame} 为平均每帧完成单次迭代所需的时钟周期数。

在两种基图 and 不同提升因子 Z_c 下, 信息吞吐率随码率的变化如图10所示。对于BG2中 $Z_c = 4$ 时低码率的(208, 40) LDPC码, 信息吞吐率为4.52 Mbps; 对于BG1中 $Z_c = 384$ 时高码率的(10368, 8448) LDPC码, 系统达到1447.4 Mbps的峰值吞吐率。上述差异主要由码率和 Z_c 共同决定, 前者决定了有效信息位的长度, 后者则影响节点处理单元的利用率。当 $Z_c \leq 192$ 时, 每个循环置换矩阵均占用一个时隙, 实际参与运算的节点处理单元数量随 Z_c 变化。文献[20]和[23]中兼容5G-NR标准的LDPC译码器也呈现出类似的规律。继续进行奇偶行列交织可以将译码器的最大并行度降至64甚至更低, 从而进一步提高节点处理单元的利用率, 但并未改善短码长、低码率LDPC码的信息吞吐率, 反而会降低系统峰值吞吐率。因此, 在兼容5G-NR标准规定的两种基图和各种提升因子的前提下, 本文采用192作为译码并行度, 在硬件复杂度与吞吐率之间取得了合理折中。

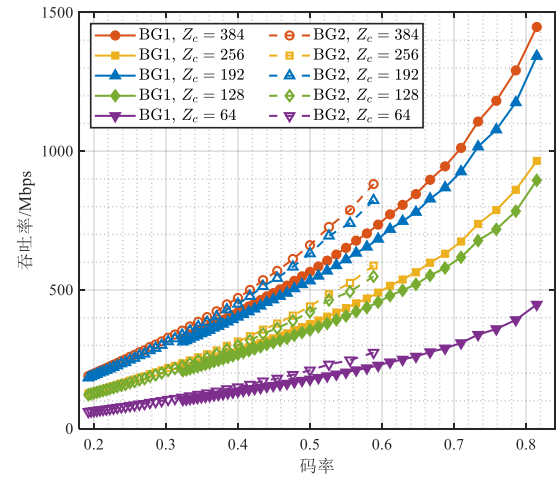


图10 双基图与不同提升因子下所提译码系统的吞吐率随码率变化

为进一步评估所提译码系统的硬件效率, 选取近年来国内外代表性的研究成果及商用IP核进行对比, 具体参数如表2所示。其中, 文献[20]和[33]均支持解速率匹配; 文献[23]所列实现结果和Tanner IP核^[32]仅包含比特恢复模块; Xilinx IP核^[31]仅提供核心LDPC译码功能。为直观比较不同设计, 以查找表数量表征硬件开销, 同时定义归一化

吞吐率 T_{norm} 为:

$$T_{\text{norm}} = \frac{T_{\text{info}} \times \text{Iter}_{\text{max}}}{f_{\text{max}}} = \frac{N_{\text{info}}}{T_{\text{frame}}} \quad (11)$$

通过将 T_{info} 乘以 Iter_{max} 并除以 f_{max} , 所得 T_{norm} 表示系统在单次迭代下平均每个时钟周期处理的信息位数, 从而消除不同工作频率和迭代次数的影响。

通过图 11 可以看出, 本文译码系统的归一化吞吐率资源比优于[20]、[23]和[33]。其中, 核心译码器实现了更高的硬件效率, 其归一化吞吐率资源比约为 Xilinx IP 核的两倍, 并接近 Tannera IP 核。

此外, 所提译码系统的动态功耗为 3.252 W, 其中核心译码器和解速率匹配模块分别占用 2.819 W 和 0.361 W。在峰值吞吐率下, 单位信息比特的能耗为 2.247 nJ/bit。上述结果为评估所提译码系统在基站、星载终端等资源与功耗受限平台上的适用性提供了定量依据。

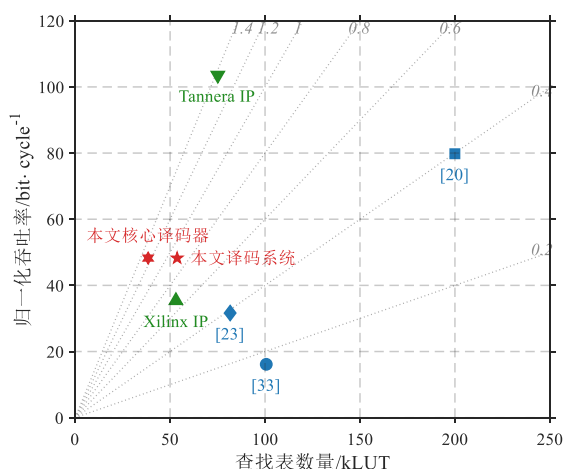


图 11 不同 5G-NR LDPC 译码方案的归一化吞吐率与硬件开销对比

5 结束语

本文针对 5G-NR 标准的严苛要求, 提出了一种高吞吐率、低资源消耗且集成解速率匹配模块的 5G LDPC 译码系统, 可以兼容双基图以及各种码长、码率和提升因子。在核心译码器设计中, 采用折叠并行架构与差分循环移位网络, 有效减少了冗余的节点更新单元与移位操作, 并缓解了布线拥塞。结合 3 帧交织流水线架构与离线静态调度策略, 在减少层间数据相关性造成的流水线气泡的同时, 利用预生成的微码指令流代替了复杂的冲突检测逻辑。此外, 解速率匹配模块采用了分级存储架构与输出移位寄存器, 在提高存储资源利用率的同时简化了控制逻辑。

基于 Xilinx UltraScale+ XCVU13P 的硬件实现与基带链路仿真结果表明, FER 曲线在瀑布区快速收敛且未出现明显错误平层。此外, 该系统可在 300 MHz 的时钟频率下实现 1447.4 Mbps 的峰值吞吐率, 吞吐率资源比为 26.94 Mbps/kLUT。其中, 核心译码器的吞吐率资源比可达 37.65 Mbps/kLUT。综上所述, 本文提出的 5G LDPC 译码系统不仅具有较高的硬件效率, 还集了解速率匹配模块, 为非地面网络等场景下的信道译码提供了一种高效的硬件实现方案。

表 2

不同 5G-NR LDPC 译码方案的 FPGA 实现结果

设计	解速率匹配	FPGA 平台	工作频率/MHz	迭代次数	译码算法	资源消耗	信息吞吐率/Mbps
[20]	支持	Virtex-7	180	10	ANMS	LUT:199959 BRAM:524.5	1436.37
[23]	部分支持	Virtex-7	180	10	NMS	LUT:81634 BRAM:209	569.53
Xilinx IP ^[31]	不支持	UltraScale+	400	8	OMS/NMS	LUT:53062 BRAM:119	1770
Tannera IP ^[32]	部分支持	UltraScale+	425	8	—	LUT:75124 BRAM:120.5	5500
[33]	支持	UltraScale+	200	25	OMS	LUT:100642 BRAM:175	129
本文	支持	UltraScale+	300	10	OMS	LUT:53720 BRAM:219	1447.4

参考文献:

- [1] 3GPP. Study on scenarios and requirements for next generation access technologies (Release 14): TR 38.913 V14.3.0[R]. Valbonne, France: 3GPP Organizational Partners, 2017.
- [2] 3GPP. Multiplexing and channel coding (Release 15): TS 38.212 V15.3.0[S]. Valbonne, France: 3GPP Organizational Partners, 2018.
- [3] Li M, Derudder V, Bertrand K, et al. High-speed LDPC decoders towards 1 Tb/s[J]. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2021, 68(5): 2224-2233.
- [4] Tran-Thi B N, Nguyen-Ly T T, Hoang T. An FPGA design with high memory efficiency and decoding performance for 5G LDPC decoder[J]. *Electronics*, 2023, 12(17): 3667.
- [5] Hu X Y, Eleftheriou E, Arnold D M. Regular and irregular progressive edge-growth Tanner graphs[J]. *IEEE Transactions on Information Theory*, 2005, 51(1): 386-398.
- [6] Fossorier M P C, Mihaljevic M, Imai H. Reduced complexity iterative decoding of low-density parity check codes based on belief propagation [J]. *IEEE Transactions on Communications*, 1999, 47(5): 673-680.
- [7] Chen J, Fossorier M P C. Density evolution for two improved BP-based decoding algorithms of LDPC codes[J]. *IEEE Communications Letters*, 2002, 6(5): 208-210.
- [8] Fan L, Pan C, Peng K, et al. Adaptive normalized min-sum algorithm for LDPC decoding[C]//*Proceedings of the 2013 9th International Wireless Communications and Mobile Computing Conference (IWCMC)*. Piscataway: IEEE Press, 2013: 1081-1084.
- [9] Zhang J, Fossorier M, Gu D. Two-dimensional correction for min-sum decoding of irregular LDPC codes[J]. *IEEE Communications Letters*, 2006, 10(3): 180-182.
- [10] Ren Y, Harb H, Shen Y, et al. A generalized adjusted min-sum decoder for 5G LDPC codes: algorithm and implementation[J]. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2024, 71(6): 2911-2924.
- [11] Tera S P, Chinthajinjala R, Al-Turjman F, et al. Deep learning assisted LDPC decoding for 5G IoT networks in fading environments[J]. *Scientific Reports*, 2025, 15(1): 37469.
- [12] Darabiha A, Carusone A, Kschischang F. A bit-serial approximate min-sum LDPC decoder and FPGA implementation[C]//*Proceedings of the 2006 IEEE International Symposium on Circuits and Systems (ISCAS)*. Piscataway: IEEE Press, 2006: 1-4.
- [13] Angarita F, Valls J, Almenar V, et al. Reduced-complexity min-sum algorithm for decoding LDPC codes with low error-floor[J]. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2014, 61(7): 2150-2158.
- [14] Lopez H, Chan H W, Chiu K L, et al. A 75-Gb/s/mm² and energy-efficient LDPC decoder based on a reduced complexity second minimum approximation min-sum algorithm[J]. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 2020, 28(4): 926-939.
- [15] Cheng C C, Yang J D, Lee H C, et al. A fully parallel LDPC decoder architecture using probabilistic min-sum algorithm for high-throughput applications[J]. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2014, 61(9): 2738-2746.
- [16] Wang C X, Lin C H. Improved normalized probabilistic minimum summation algorithm for LDPC decoding[C]//*Proceedings of the 2020 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-Taiwan)*. Piscataway: IEEE Press, 2020: 1-2.
- [17] Tsatsaragkos I, Paliouras V. Approximate algorithms for identifying minima on min-sum LDPC decoders and their hardware implementation[J]. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2015, 62(8): 766-770.
- [18] Zhang Q, Chen K, Qu Y, et al. Hierarchical approximate min-sum and multi-frame parallel QC-LDPC decoder for FPGA optical links[J]. *IEEE Photonics Technology Letters*, 2026, 38(10): 635-638.
- [19] Hocevar D E. A reduced complexity decoder architecture via layered decoding of LDPC codes[C]//*Proceedings of the 2004 IEEE Workshop on Signal Processing Systems*. Piscataway: IEEE Press, 2004: 107-112.
- [20] 李文乐. 高吞吐 5G LDPC 码编译码器的研究与实现[D]. 西安: 西安电子科技大学, 2024.
- Li W L. Research and implementation of high-throughput 5G LDPC encoder and decoder[D]. Xi'an: Xidian University, 2024.
- [21] Nguyen T T B, Tan T N, Lee H. Low-complexity high-throughput QC-LDPC decoder for 5G new radio wireless communication[J]. *Electronics*, 2021, 10(4): 516.
- [22] Mejmaa B, Akharraz I, Ahaitouf A. A field-programmable gate array-based quasi-cyclic low-density parity-check decoder with high throughput and excellent decoding performance for 5G new-radio standards [J]. *Technologies*, 2024, 12(11): 215.
- [23] 魏瑞. 基于FPGA的5G LDPC译码系统研究与实现[D]. 西安: 西安电子科技大学, 2024.
- Wei R. Research and implementation of 5G LDPC decoding system based on FPGA[D]. Xi'an: Xidian University, 2024.
- [24] Boncalo O, Kolumban-Antal G, Amaricai A, et al. Layered LDPC decoders with efficient memory access scheduling and mapping and built-in support for pipeline hazards mitigation[J]. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2019, 66(4): 1643-1656.
- [25] Chen W, Liu D. Conflict-free parallel data access technology for matrix calculation in memory system of ASIP of 5G/6G macro base stations[J]. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2024, 43(1): 99-112.
- [26] Jia X, Zhou Y, Xie T, et al. An improved scheduling scheme for layered decoding of 5G LDPC codes[J]. *IEEE Transactions on Vehicular Technology*, 2026, 75(7): 15270-15274.
- [27] Nimara S, Boncalo O, Amaricai A, et al. FPGA architecture of multi-codeword LDPC decoder with efficient BRAM utilization[C]//*Proceedings of the 2016 IEEE 19th International Symposium on Design and Diagnostics of Electronic Circuits & Systems (DDECS)*. Piscataway: IEEE Press, 2016: 1-4.
- [28] Song S, Cui H, Wang Z. A universal efficient circular-shift network for reconfigurable quasi-cyclic LDPC decoders[J]. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 2022, 30(10): 1553-1557.
- [29] Nguyen T T B, Tan T N, Lee H. Low-complexity multi-size circular-shift network for 5G new radio LDPC decoders[J]. *Sensors*, 2022, 22(5): 1792.
- [30] Chen X, Lin S, Akella V. QSN—a simple circular-shift network for reconfigurable quasi-cyclic LDPC decoders[J]. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2010, 57(10): 782-786.
- [31] AMD. 5G NR LDPC encoder/decoder[EB/OL]. [2026-06-11]. <https://www.amd.com/en/products/adaptive-socs-and-fpgas/intellectual>

property/ef-di-ldpc-enc-dec.html.

[32] Tannera. 5G NR LDPC decoder IP core[EB/OL]. [2026-06-11]. <https://tannera.io/5g-nr-ldpc-decoder-ip-core/>.

[33] 刘柏文, 白宝明. 5G LDPC 编码链设计与实现[J]. 无线电工程, 2025, 55(6): 1306-1317.

Liu B W, Bai B M. Design and implementation of 5G LDPC coding chain[J]. Radio Engineering, 2025, 55(6): 1306-1317.

宫丰奎 (1979-), 男, 博士, 西安电子科技大学教授, 主要研究方向为宽带卫星互联网、自动化卫星地面检测设备研制和先进数字视频传输。



孙源 (2001-), 男, 西安电子科技大学博士生, 主要研究方向为信道编码和卫星通信。



惠腾飞 (1983-), 男, 中国空间技术研究院西安分院高级工程师, 主要研究方向为卫星通信系统设计和信号处理。



易杨 (2001-), 男, 西安电子科技大学硕士生, 主要研究方向为信道编码。