

# 具有逆向边增长结构LT码的设计与分析

宋 鑫<sup>1</sup>,程乃平<sup>2</sup>,廖育荣<sup>2</sup>,倪淑燕<sup>2</sup>,雷拓峰<sup>1</sup>

(1. 航天工程大学研究生院,北京 101416; 2. 航天工程大学电子与光学工程系,北京 101416)

**摘要:** 为了解决传统卢比变换(Luby Transform,LT)码在加性高斯白噪声(Additive White Gaussian Noise,AWGN)信道中存在高误码平台的问题,提出了一种具有逆向边增长(Reverse Edge Growth,REG)结构的LT码. REG-LT码会标记度数数值相对较小的信息节点,并使其逆向精准地选取足够数量的校验节点,从而获得不低于期望目标的度数. 首先,设计了REG-LT码的编码算法,分析了校验节点和信息节点的选取原则. 其次,推导了REG-LT码的信息节点度分布和校验节点度分布的数学表达式,分析了REG-LT码的收敛性. 最后,引入增益损失比(Gain Loss Ratio,GLR)以衡量参数 $T_v$ 对收敛性的影响,将最大化GLR作为求解最优 $T_v$ 值的设计原则. 仿真结果表明,与传统LT码相比,本文的REG-LT码可将误码平台降低3至5个数量级,且至少可获得2 dB的编码增益;此外,REG-LT码也能够实现优于现有LT码的误比特率(Bit Error Rate,BER)性能.

**关键词:** 喷泉码; LT码; 高斯信道; 误比特率; 度分布; 编码算法; 外信息传递图

**中图分类号:** TN911.22 **文献标识码:** A **文章编号:** 0372-2112(2022)05-1201-10

**电子学报 URL:**http://www.ejournal.org.cn

**DOI:**10.12263/DZXB.20210595

## Design and Analysis of LT Code with Reverse Edge Growth Framework

SONG Xin<sup>1</sup>,CHENG Nai-ping<sup>2</sup>,LIAO Yu-rong<sup>2</sup>,NI Shu-yan<sup>2</sup>,LEI Tuo-feng<sup>1</sup>

(1. Department of Graduate Management, Space Engineering University, Beijing 101416, China;

2. Department of Electronic and Optical Engineering, Space Engineering University, Beijing 101416, China)

**Abstract:** To solve the problem of high error floor for the conventional Luby transform(LT) codes in additive white gaussian noise(AWGN) channels, the LT code with reverse edge growth(REG) framework is proposed. For the REG-LT code, those information nodes with relatively small degrees will be marked, and these nodes will accurately select enough check nodes in reverse to obtain a degree that is not lower than the desired target. The coding scheme of the REG-LT code is designed, and the selection principle of check node and information node is analyzed. The mathematical expressions of the information degree distribution and check degree distribution of the REG-LT code are derived, and the convergence behavior of the REG-LT code is analyzed. The parameter gain loss ratio(GLR) is designed to measure the impact of the parameter  $T_v$  on convergence behavior, and maximize the GLR is used as a design principle to solve the optimal  $T_v$ . The simulation results show that compared with the conventional LT code, the REG-LT code in this paper can reduce the error platform by 3 to 5 orders of magnitude, and at least 2dB of coding gain can be obtained. Besides, the designed LT code outperforms the existing LT codes in terms of bit error rate(BER) performance.

**Key words:** fountain codes; LT codes; Gaussian channel; bit error rate; degree distribution; coding algorithm; extrinsic information transfer chart

## 1 引言

喷泉码<sup>[1]</sup>最初是针对二进制删除信道(Binary Erasure Channel, BEC)设计的,旨在为大规模数据分发和可靠广播场景提出一种理想的解决方案<sup>[2]</sup>. 以卢比变换(Luby Transform,LT)码<sup>[3]</sup>为代表的喷泉码,具有天然

的信道自适应特性,这使得它非常适合应用于传输视频、音频的广播场景<sup>[4]</sup>,协作中继场景<sup>[5]</sup>,水声通信场景<sup>[6]</sup>,自由空间光通信场景<sup>[7]</sup>等.

尽管喷泉码最初是面向BEC进行设计的,但其在加性高斯白噪声(Additive White Gaussian Noise,AWGN)信道中也具有潜在的应用前景<sup>[8,9]</sup>. 文献[10]研究

结果表明,LT码在二进制对称信道(Binary Symmetric Channel, BSC)和 AWGN 信道中存在明显的误码平台. 为了改善LT码的误码平台现象,文献[11]将LT码与高码率的低密度奇偶校验(Low Density Parity Check, LDPC)码<sup>[12]</sup>级联,构成了Raptor码. 但这也存在显著的缺点:AWGN信道中LDPC码和LT码均采用和积译码算法<sup>[13]</sup>,具有较高的译码复杂度. 因此,学者们在研究Raptor码的同时,也在探索设计独立的无速率码<sup>[14-16]</sup>以及改进的低误码平台LT码.

作为第一种实用的无速率码,LT码的优化设计大多是针对校验度分布和编码算法进行的. 文献[17~20]提出了几种设计LT码校验度分布的方法. 文献[17]针对二进制输入AWGN(Binary Input AWGN, BIAWGN)信道,设计了一种基于线性规划模型的校验度分布优化方法. 文献[18]针对极低信噪比条件提出了一种改进的度分布设计方式. 文献[19]给出了一种适用于大范围信噪比的度分布函数设计策略,以保持足够高的码率效率. 文献[20]针对系统LT码,引入了误比特率(Bit Error Rate, BER)下界作为新的约束条件,实现了以更小的译码开销进入瀑布区的效果. 但是,在BIAWGN信道中,优化校验度分布函数对降低误码平台的效果有限. 其原因主要在于:(1)误码平台的形成与校验度分布不存在因果关系;(2)在利用线性规划方法设计度分布时,往往考虑的都是无限码长和甚高迭代次数时的理想状态,因此,所设计的度分布一般只在码长较长时才能表现出优良的渐进性能,但对中短码长BER性能改善却十分有限.

针对上述问题,本文考虑改进LT码的编码算法以提升其BER性能. 文献[21]指出误码平台主要是由译码成功概率较低的小度数值信息节点造成的. 为此,文献[21]提出对信息节点按照度数值大小进行分类,并迫使校验节点优先选取度数最小的信息节点. 文献[22]则将文献[21]的编码思想应用到了不等差错保护场景中,以略微增加译码开销为代价,将重要数据(Most Important Bits, MIB)的BER平台降低了将近3个数量级. 文献[23]的改进算法,是在校验节点选择每个信息节点时,都先从所有的信息节点中随机选取若干个节点作为一组,然后再从这一组中选择度数最小的信息节点进行连接. 文献[24]的改进算法引入了衡量校验节点度数值大小的参数 $d^*$ ,并根据该参数与校验节点度数的大小关系,判断选取信息节点的方式. 仿真结果显示,这几种改进算法在BEC或者AWGN信道中都能够显著地降低误码平台. 但是这几类改进算法也存在下述问题:(1)没有对改进算法的信息节点度分布进行公式化表述;(2)没有给出算法所涉及参数的优化设计方法;(3)算法可达的最低信息节点度数无法预测.

针对现有改进算法存在的不足,本文设计了一种具有逆向边增长(Reverse Edge Growth, REG)结构的LT码. REG-LT码中具有正向、逆向选取邻居节点的编码结构,从而可使所有信息节点的度数值均不小于预先设计的参数 $T_v$ . 此外,还分析了 $T_v$ 对收敛性的影响,给出了最优参数的求解原则. 仿真结果显示,与传统LT码相比,本文的REG-LT码能够显著地降低误码平台,且BER性能优于文献[21, 23]中的改进算法,证明了算法的正确性.

## 2 REG-LT码的编码算法

### 2.1 传统LT码的编译码算法

对于传统LT码,编码器会对 $K$ 个信息节点 $v_1, v_2, \dots, v_K$ 进行不定速率编码,生成 $N$ 个校验节点 $c_1, c_2, \dots, c_N$ . LT码的Tanner图如图1所示.

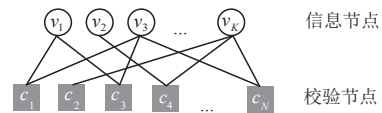


图1 LT码的Tanner图

在图1中,定义每个校验节点连接的信息节点的个数为该校验节点的度数值,每个信息节点连接的校验节点的个数为该信息节点的度数值. 定义校验节点的度分布函数为 $\Omega(x) = \sum_{j=1}^{d_c} \omega_j x^j$ ,其中, $\omega_j$ 表示度数为 $j$ 的校验节点出现的概率, $d_c$ 表示校验节点的最大度数值. 在编码过程中,校验度分布是预先给定的,度数为 $j$ 的校验节点会从所有的信息节点中随机选取 $j$ 个相连,对这 $j$ 个被选中信息节点进行异或操作,就得到了当前校验节点的比特值. 定义信息节点度分布为 $A(x) = \sum_{i=0}^{d_v} \zeta_i x^i$ ,其中, $\zeta_i$ 表示度数为 $i$ 的信息节点出现的概率, $d_v$ 表示信息节点的最大度数值. 与校验度分布不同,信息节点的度分布由 $\Omega(x)$ 、 $K$ 、 $N$ 共同决定,只有在编码结束后才能得到. 尽管LT码是无速率码,但定义其瞬时码率值为 $R = K/N$ . 定义校验度分布的平均度数为 $\beta = \sum_{j=1}^{d_c} \omega_j j$ ,则信息节点的平均度数值为 $\alpha = N\beta/K$ . 根据文献[17],信息节点度分布的系数 $\zeta_i$ 可近似计算为

$$\zeta_i \approx \frac{\alpha^i}{i!} e^{-\alpha} \quad (1)$$

图1所示的Tanner图中,在信息节点和校验节点之间任选一条边,定义该条边连接到度数为 $j$ 的校验节点的概率为 $\rho_j$ ,连接到度数为 $i$ 的信息节点的概率为 $\lambda_i$ . 在此基础上,进一步定义校验节点边的度数分布为 $\rho(x) =$

$\sum_{j=1}^{d_c} \rho_j x^{j-1}$ , 信息节点边的度数分布为  $\lambda(x) = \sum_{i=1}^{d_v} \lambda_i x^{i-1}$ . 其中, 系数  $\rho_j$  和  $\lambda_i$  的计算方法如下<sup>[25]</sup>:

$$\lambda_i = \frac{i \zeta_i K}{d_v} = \frac{i \zeta_i}{\sum_{i=1}^{d_v} i \zeta_i} \quad (2)$$

$$\rho_i = \frac{j \omega_j N}{d_c} = \frac{j \omega_j}{\sum_{j=1}^{d_c} j \omega_j} \quad (3)$$

LT码在AWGN信道中采用置信传播(Belief Propagation, BP)算法进行迭代译码. 该算法将对数似然比(Log-Likelihood Ratio, LLR)信息在信息节点和校验节点之间进行迭代更新, 使LLR信息逐渐收敛于稳定值并据此进行判决.

令  $L_{c_j \rightarrow v_i}$  表示迭代过程中第  $j$  个校验节点传递给第  $i$  个信息节点的LLR信息, 定义为:

$$\tanh\left(\frac{L_{c_j \rightarrow v_i}}{2}\right) = \tanh\left(\frac{L_{\text{ch}}^i}{2}\right) \prod_{i \in N(j) \setminus (i)} \tanh\left(\frac{L_{v_i \rightarrow c_j}}{2}\right) \quad (4)$$

其中,  $L_{\text{ch}}^i$  为每个校验节点获得的来自信道的初始LLR信息.  $N(j) \setminus (i)$  表示除第  $i$  个信息节点之外, 与第  $j$  个校验节点相连的所有信息节点的集合. 式(4)中,  $L_{v_i \rightarrow c_j}$  表示迭代过程中第  $i$  个信息节点传递给第  $j$  个校验节点的LLR信息, 定义为:

$$L_{v_i \rightarrow c_j} = \sum_{j \in N(i) \setminus (j)} L_{c_j \rightarrow v_i} \quad (5)$$

其中,  $N(i) \setminus (j)$  表示除第  $j$  个校验节点之外, 与第  $i$  个信息节点相连的所有校验节点的集合.

对于信息节点  $v_i$ , 其LLR判决值为

$$L(v_i) = \sum_{j \in N(i)} L_{c_j \rightarrow v_i} \quad (6)$$

## 2.2 REG-LT码的编码算法

为了克服传统方案、文献[21]中的等度数(Equal Degree, ED)编码方案和文献[23]中的选择连接(Connection Choice, CC)编码方案的不足, 需要设计满足以下条件的改进方案:(1)通过优化编码结果以降低误码平台. 这样可以更准确地为低度数信息节点连接最佳数量的校验节点;(2)改进方案的信息度分布和校验度分布能够公式化表达. 这是因为, 只有准确的度分布才可以用来估计LT码的误比特率下界以及预测LT码是否能够收敛;(3)编码结果能够匹配任意信道参数.

综上分析, 本文设计了一种具有逆向边增长结构的LT码, 其编码过程如算法1所示.

从算法1可以看出, REG-LT码与传统LT码的最大区别在于两点:(1)信息节点可逆向精准随机选取校验节点. 这样做的优点是易于使所有信息节点的度数

### 算法1 REG-LT码的编码算法

**初始化:** 给定  $K$  个信息节点  $v_1, v_2, \dots, v_K$ , 待生成校验节点个数  $N$  和校验度分布  $\Omega(x)$ . 给定信息节点期望达到的最低度数  $T_v$ .

步骤1: 按照传统LT编码方式生成  $N$  个校验节点.

步骤2: 更新所有信息节点的度数. 筛选出度数小于期望最低度数  $T_v$  的信息节点, 将其记为集合  $S$ . 记集合  $S$  中的元素个数为  $M$ . 定义  $m=1$ .

步骤3: 标记集合  $S$  中第  $m$  个信息节点的度数值为  $d_m$ , 此处  $1 \leq m \leq M$ .

步骤4: 从度数大于2的校验节点中随机选取  $T_v - d_m$  个.

步骤5: 如果步骤4中选中的  $T_v - d_m$  个校验节点, 与第  $m$  个信息节点已连接的  $d_m$  个校验节点不存在交集, 则将选中的  $T_v - d_m$  个校验节点连接至第  $m$  个信息节点. 如果存在交集, 则返回步骤4.

步骤6: 如果  $m=M$ , 则进行步骤7. 如果  $m < M$ , 则令  $m=m+1$ , 并返回步骤3.

步骤7: 更新Tanner图, 重新计算度数大于2的校验节点比特值. 输出  $N$  个校验节点  $c_1, c_2, \dots, c_N$ .

均不低于期望值  $T_v$ , 并且维持了Tanner图的随机性以便于快速译码收敛;(2)编码算法可逆向改变信息度分布和校验度分布. 这样做的优点在于, 其一是在编码的同时即优化了信息度分布, 其二是能够对预先给定的校验度分布进行自适应调整以匹配信道参数.

## 2.3 校验节点的选取原则

在算法1中, 并不是将所有校验节点都作为待选节点供信息节点选取, 待选校验节点应满足如下条件:(1)度数值不等于1和2;(2)与当前待处理的信息节点不存在连接关系.

满足条件(2)的原因, 是为了避免边重复连接, 导致无效编码. 满足条件(1)的原因, 是缺乏度数值为1和2的校验节点会导致译码失败, 具体如下.

度数值为1的校验节点使得译码过程能够正常启动. 2.1节介绍了LT码在AWGN信道中的BP译码算法, 本节根据此译码算法, 给出几个实例对LT码的详细译码过程进行分析. 图2中, 在第一次迭代时, 所有信息节点传递给校验节点的LLR信息均为0; 校验节点  $c_4$  则按照式(4), 将其获得的来自信道的非零LLR信息  $L_{\text{ch}}^4$  传递给唯一相连的信息节点  $v_2$ , 其余校验节点传递给信息节点的LLR信息均为0. 第二次迭代时,  $v_2$  将第一次迭代时获得的来自  $c_4$  的LLR信息传递给与之相连接的  $c_1, c_2$  和  $c_3, c_1, c_2$  和  $c_3$  则将接收到的LLR信息按式(4)进行处理, 之后传递给  $v_1, v_3$  和  $v_4$ . 自此, 所有信息节点均获得了非0值LLR信息.

可以看出, 正是由于校验节点  $c_4$  的度数为1, 才能将自身的LLR信息传递给  $v_2$ , 从而开启了译码进程. 在图3中, 不存在度数为1的校验节点, 因此, 根据式(4)可知, 每次迭代时信息节点和校验节点之间传递的LLR信息均为0, 译码失败.

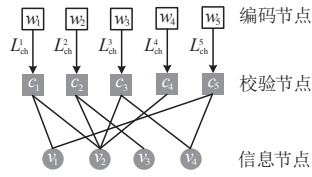


图2 可顺利译码的Tanner示例

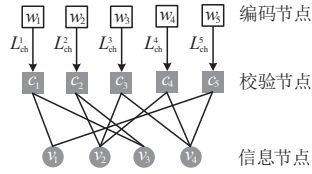


图3 缺少度数为1的校验节点

度数值为2的校验节点使得LLR信息能够传播出去。在图4中,不存在度数为2的校验节点。尽管按照式(5), $v_2$ 可以将获得的来自 $c_4$ 的LLR信息传递其余校验节点,但是按照式(4)的传递规则,除 $c_4$ 之外的校验节点传递给信息节点 $v_1$ 、 $v_3$ 和 $v_4$ 的LLR信息均为0。因此,来自信道的LLR信息无法传播并有效迭代,译码失败。

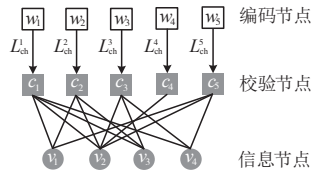


图4 缺少度数为2的校验节点

综上,对于所提出的编码算法,若要在接收端译码器采用2.1节给出的LT码在AWGN信道中的译码算法,则应满足条件(1)与条件(2)。

## 2.4 算法复杂度分析

相较于传统算法,本文算法增加了分类信息节点以及逆向边增长的操作,因而不可避免地增加了算法的复杂度,本节将对其进行详细分析。

算法1中,每个校验节点选取信息节点时,新增操作为:节点按度数值大小排序、信息节点度数值和阈值数值比较、逆向随机选取校验节点。为了便于对比,表1给出了生成 $N$ 个校验节点时,算法1、ED方案、CC方案所产生的新增操作及其次数。

表2给出了这三种改进算法的编码运行时间。仿真所采用的计算机硬件配置为如下,处理器为Intel Xe-

表1 生成 $N$ 个校验节点时的新增操作对比

编码操作	传统算法	本文算法	ED算法	CC算法
排序/次	0	1	$N-1$	$N\beta$
数值比较/次	0	$K$	$N-1$	$N\beta-N$
随机选取/次	$N$	$<N+K$	$N$	$N\beta$

on CPU E3-1240 v5@3.5 GHz,单颗CPU,4核8线程, RAM内存为32 GB,仿真软件为MATLAB2020a.设计并行仿真次数为1000次, $K=2048$ , $N=4096$ ,信噪比为 $E_s/N_0=1$  dB,采用BP译码算法,最大迭代次数为50次。其中,本文算法阈值为 $T_v=12$ ,文献[23]算法中 $T=3$ ,校验度分布采用文献[26]的设计结果,记为

$$\Omega_1(x) = 0.025x + 0.495x^2 + 0.167x^3 + 0.082x^4 + 0.071x^5 + 0.05x^8 + 0.044x^9 + 0.043x^{19} + 0.023x^{66} \quad (7)$$

表2 不同算法的平均运行时间

算法	平均编码时间/s	平均译码时间/s	总计/s
传统算法	100.7045	146.8158	247.5203
本文算法	157.2001	131.9716	289.1717
ED算法	227.8599	117.8297	345.6896
CC算法	614.2804	119.5156	733.7960

可以看出,与传统算法相比,三种改进算法的平均编码运行时间均增加了,这是因为改进算法均增加了额外的编码操作。但与其他两种算法相比,本文算法的编码运行速度则相对较快。还可以看出,本文算法的译码时间略高于其他两种算法,这是因为本文算法为校验节点添加了额外的边数,使得LLR信息的传递、更新次数多于其他两种算法。不过,总体而言,本文算法的编码、译码时间之和仍然少于其他两种参考算法,这体现了本文算法的优势。

## 3 REG-LT码的度分布分析

### 3.1 REG-LT码的信息节点度分布

在算法1的逆向边增长阶段,度数值小于 $T_v$ 的信息节点会主动选取足够数量的校验节点相连,而度数值大于 $T_v$ 的信息节点则不作任何处理。因此,一方面精准剔除了度数小于 $T_v$ 的信息节点,且该类信息节点所占比例值 $\xi_i$  ( $0 \leq i \leq T_v - 1$ )将全部累加至 $\xi_{T_v}$ 处。另一方面,度数值大于阈值的信息节点其比例不变,即与传统LT码的度分布相同。

定义REG-LT码的信息度分布为 $A_n(x) = \sum_{i=0}^{d_v} \xi_i^{(n)} x^i$ 。

其中,度数值为 $T_v$ 的信息节点所占的比例为:

$$A_{T_v}^{(n)} = \sum_{i=0}^{T_v} \xi_i \quad (8)$$

则REG-LT码的信息度分布为

$$A_n(x) = \sum_{i=0}^{T_v} \xi_i x^{T_v} + \sum_{i=T_v+1}^{d_v} \xi_i x^i \quad (9)$$

可以看出,度分布 $A_n(x)$ 的系数 $\xi_i^{(n)}$ 只取决于初始度分布系数 $\xi_i$ 和阈值 $T_v$ 。其中, $\xi_i$ 的计算方式参照式(1)。

### 3.2 REG-LT 码的校验节点度分布

定义算法 1 在步骤 1 完成之后,得到的校验节点度分布为  $\Omega(x) = \sum_{j=1}^{d_c} \omega_j x^j$ . 算法 1 的剩余步骤完成之后,度数为 1 和 2 的校验节点所占比例保持不变,其余度数校验节点的占比均会改变. 定义算法 1 结束之后,校验节点的度分布为

$$\Omega_n(x) = \sum_{j=1}^{d_{c(n)}} \omega_j^{(n)} x^j \quad (10)$$

其中,  $d_{c(n)}$  是 REG-LT 码校验节点的最大度数.

算法 1 中,需要进行逆向边增长操作的信息节点平均个数为  $K_a = K \sum_{i=0}^{T_v-1} \zeta_i$ . 此  $K_a$  个信息节点一共包含  $T_v$  种度数值,分别为  $0, 1, \dots, T_v - 1$ . 能够参与逆向边增长操作的校验节点的平均个数为  $N_a = N(1 - \omega_1 - \omega_2)$ .

对于 REG-LT 码而言,度数值小于  $T_v$  的信息节点会随机选取校验节点与之相连. 这个过程恰好是传统编码算法的逆向操作,因此,考虑引入泊松分布来近似计算新的校验度分布. 显然,在 REG-LT 码的逆向边增长阶段,信息节点和校验节点的角色进行了互换. 每种度数信息节点的平均个数为  $K\zeta_0, K\zeta_1, \dots, K\zeta_{T_v-1}$ , 且这些信息节点需要连接的校验节点的个数分别为  $T_v, T_v - 1, \dots, 1$ . 因此,定义这  $K_a$  个信息节点满足逆向信息度分布:

$$\Omega_A(x) = \sum_{j=1}^{T_v} \omega_j^{(A)} x^j \quad (11)$$

其中,  $\omega_j^{(A)} = K\zeta_{j-1}/K_a$ .

逆向边增长操作结束后,此  $K_a$  个信息节点增加的平均度数值为  $\beta_A = \sum_{j=1}^{T_v} j \omega_j^{(A)}$ . 则,待选取  $N_a$  个校验节点增加的平均度数为  $\alpha_\Omega = \beta_A K_a / N_a$ . 若初始化此  $N_a$  个校验节点的度数为 0,则编码结束后,其度数分布可采用均值为  $\alpha_\Omega$  的泊松分布近似代替,定义为

$$A_\Omega(x) = \sum_{i=0}^{d_{c(\Omega)}} \zeta_i^{(\Omega)} x^i \quad (12)$$

其中,  $d_{c(\Omega)} = [K_a]$ ,  $[\cdot]$  表示向下取整函数. 则 REG-LT 码校验节点的最大度数  $d_{c(n)}$  可计算为  $d_{c(n)} = d_c + [K_a]$ . 系数  $\zeta_i^{(\Omega)}$  可计算为

$$\zeta_i^{(\Omega)} = \frac{\alpha_\Omega^i e^{-\alpha_\Omega}}{i!} \quad (13)$$

$\zeta_i^{(\Omega)}$  表示此  $N_a$  个校验节点度数值增加  $i$  的概率,但这是考虑其初始度数均为 0 的情况. 事实上,在逆向边增长操作之前,其初始度数并不 0,因此,还需将  $A_\Omega(x)$  叠加至  $\Omega(x)$  中,以获得最终的校验度分布.

由于度数为 1 和 2 的校验节点并不需要参与逆向操作,故其度分布系数保持不变,即  $\omega_1^{(n)} = \omega_1, \omega_2^{(n)} = \omega_2$ . 算法 1 的步骤 1 完成之后,对于任意一个度数为  $m(3 \leq m \leq d_c)$  的校验节点而言,其在逆向操作中度数值变为  $m+i$  的概率为  $\omega_m \zeta_i^{(\Omega)}$ ,其中  $0 \leq i \leq d_{c(\Omega)}$ . 逆向操作结束后,此  $N_a$  个校验节点的最大度数为  $d_c + d_{c(\Omega)}$ ,最低度数为 3. 分析易得,编码结束后,任意一个度数值为  $j(3 \leq j \leq d_c + d_{c(\Omega)})$  的校验节点,其产生来源包括以下两方面:(1) 度数原本就是  $j$  的校验节点,其在逆向操作过程中没有新增连接边;(2) 原本边数小于  $j$  的校验节点,其度数值增加至  $j$ . 故而,逆向操作之后,度数值为  $j$  的校验节点所占比例为

$$\omega_j^{(n)} = \sum_{m=3}^j \omega_m \zeta_{j-m}^{(\Omega)} \quad (14)$$

则可得,REG-LT 码的最终校验度分布为

$$\Omega_n(x) = \omega_1 x + \omega_2 x^2 + \sum_{j=3}^{d_{c(n)}} \left( \sum_{m=3}^j \omega_m \zeta_{j-m}^{(\Omega)} \right) x^j \quad (15)$$

## 4 REG-LT 码的最优最低度数设计

REG-LT 码的关键环节是信息节点的逆向边增长操作,直接影响其 BER 性能的因素是信息节点的期望最低度数值  $T_v$ . 一方面,信息节点的最低度数值直接决定了 LT 码的误码平台下界<sup>[21]</sup>;另一方面,  $T_v$  又决定了 REG-LT 码的信息度分布和校验度分布,进而影响其收敛性能. 因此,有必要分析参数  $T_v$  的设计原则,并给出最优  $T_v$  的设计方法.

### 4.1 REG-LT 码的收敛性分析

LT 码在 AWGN 信道中采用 BP 算法进行迭代译码,分析 LT 码译码收敛性常用的工具为外信息传递 (Extrinsic Information Transfer, EXIT) 图法. 参照文献[27],定义单调递增函数  $J(\theta)$  为

$$J(\theta) = 1 - \frac{1}{\sqrt{2\pi}\theta} \int_{-\infty}^{+\infty} e^{-\frac{(\gamma - \theta^2/2)^2}{2\theta^2}} \log_2(1 + e^{-\gamma}) d\gamma \quad (16)$$

$J(\theta)$  具有唯一的反函数  $\theta = J^{-1}(I)$ . 关于  $J(\theta)$  和  $\theta = J^{-1}(I)$ ,文献[28]给出了一种近似的计算方法. 为便于分析,将 LT 码的译码器分为校验节点译码器 (Check Node Decoder, CND) 和信息节点译码器 (Information Node Decoder, IND). 在 BIAWGN 信道下,LT 码 CND 的 EXIT 公式为

$$I_{E,C} = 1 - \sum_{j=1}^{d_c} \rho_j J \left( \sqrt{(j-1)\sigma_c^2 + \sigma_{ch}^2} \right) \quad (17)$$

其中,  $\sigma_c^2 = [J^{-1}(1 - I_{A,C})]^2$ ,  $I_{A,C}$  是校验节点的输入先验信息;  $\sigma_{ch}^2 = [J^{-1}(1 - J(\sqrt{4\sigma_n^2}))]^2$ ,  $\sigma_n^2$  是 Gaussian 白噪声的方差.  $I_{E,C}$  是校验节点输出外信息.

LT码 IND 的 EXIT 公式为

$$I_{E,I} = \sum_{i=1}^{d_v} \lambda_i J \left( \sqrt{(i-1)\sigma_i^2} \right) \quad (18)$$

其中,  $\sigma_i = J^{-1}(I_{A,I})$ ,  $I_{A,I}$  是信息节点的输入先验信息.

REG-LT码的输出外信息可计算为

$$I_{E,C}^{(n)} = 1 - \sum_{j=1}^{d_c+d_{c(n)}} \rho_j^{(n)} J \left( \sqrt{(j-1)\sigma_c^2 + \sigma_{ch}^2} \right) \quad (19)$$

$$I_{E,I}^{(n)} = \sum_{i=1}^{d_v} \lambda_i^{(n)} J \left( \sqrt{(i-1)\sigma_i^2} \right) \quad (20)$$

其中,  $\lambda_i^{(n)}$  表示 REG-LT码的信息节点边度分布系数,  $\rho_j^{(n)}$  表示校验节点边度分布系数.

相比于传统 LT 码, REG-LT 码的 CND 曲线和 IND 曲线均发生了偏移, 图 5 给出了一组仿真结果进行说明. 其中仿真参数为:  $R^{-1}=2$ ,  $E_s/N_0=2$  dB, 校验度分布为  $\Omega_1(x)$ . 图 5 中, 定义 IND 曲线和 CND 曲线之间的空隙为“译码收敛区”, 其含义为: 只要两条曲线之间不存在交点, 则当码长  $K$  逼近无限长时, 译码器总能通过多次迭代成功恢复出信息节点.

#### 4.2 最优参数的设计方法

理论而言, 信息节点的平均度数越高, LT 码的误码平台越低, 因此, 从这个角度而言, 参数  $T_v$  应越大越好. 但是, 从 4.1 节的分析中可知,  $T_v$  过大可能会导致无法成功收敛. 因此, 参数  $T_v$  应存在临界值, 即使得误码平台最低的同时还能正确收敛, 我们将其定义为最优值.

为合理的量化参数  $T_v$  对译码收敛性的影响, 引入

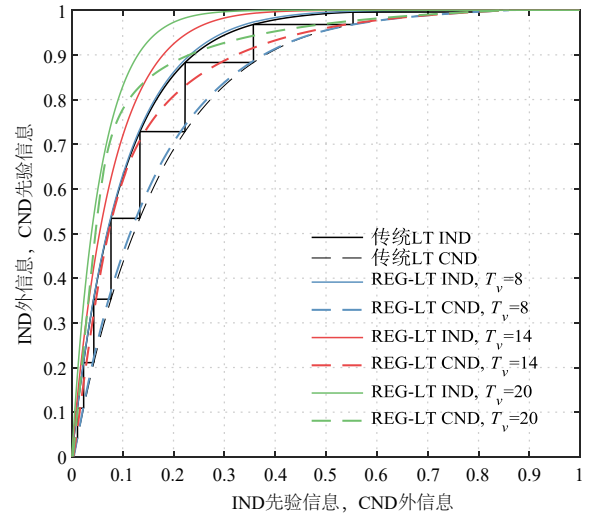


图 5 不同  $T_v$  值下的外信息曲线

外信息增益损失比 (Gain Loss Ratio, GLR) 作为评价指标. 定义校验节点译码器的外信息损失为:

$$L_{E,C} = \int_0^1 I_{E,C} d(I_{A,C}) - \int_0^1 I_{E,C}^{(n)} d(I_{A,C}) \quad (21)$$

定义信息节点译码器的外信息增益为

$$G_{E,I} = \int_0^1 I_{E,I}^{(n)} d(I_{A,I}) - \int_0^1 I_{E,I} d(I_{A,I}) \quad (22)$$

在此基础上, 定义 GLR 为

$$\text{GLR} = \frac{G_{E,I}}{L_{E,C}} \quad (23)$$

表 3 给出了不同参数条件下的 GLR 计算结果, 采用的校验度分布为  $\Omega_1(x)$ . 可得出以下结论:

表 3 不同信噪比和码率值下的 GLR 值

GLR	$E_s/N_0=0$ dB				$E_s/N_0=1$ dB				$E_s/N_0=2$ dB			
	$R^{-1}=2.0$	$R^{-1}=2.1$	$R^{-1}=2.2$	$R^{-1}=2.3$	$R^{-1}=2.0$	$R^{-1}=2.1$	$R^{-1}=2.2$	$R^{-1}=2.3$	$R^{-1}=2.0$	$R^{-1}=2.1$	$R^{-1}=2.2$	$R^{-1}=2.3$
$T_v=8$	0.7034	0.6698	0.6391	0.6107	0.6412	0.6106	0.5825	0.5567	0.5955	0.5670	0.5409	0.5169
$T_v=9$	0.7115	0.6782	0.6477	0.6196	0.6486	0.6182	0.5904	0.5648	0.6024	0.5742	0.5483	0.5245
$T_v=10$	0.7166	0.6838	0.6536	0.6257	0.6533	0.6233	0.5958	0.5704	0.6068	0.5789	0.5533	0.5297
$T_v=11$	0.7196	0.6872	0.6573	0.6298	0.6561	0.6265	0.5993	0.5741	0.6095	0.5819	0.5566	0.5332
$T_v=12$	0.7212	0.6891	0.6596	0.6324	0.6576	0.6283	0.6013	0.5765	0.6109	0.5836	0.5586	0.5355
$T_v=13$	0.7219	0.6899	0.6607	0.6337	0.6583	0.6291	0.6024	0.5778	0.6116	0.5845	0.5596	0.5367
$T_v=14$	0.7219	0.6901	0.6611	0.6343	0.6584	0.6294	0.6028	0.5784	0.6118	0.5848	0.5601	0.5373
$T_v=15$	0.7217	0.6900	0.6610	0.6344	0.6583	0.6293	0.6028	0.5785	0.6117	0.5847	0.5601	0.5375
$T_v=16$	0.7214	0.6896	0.6607	0.6342	0.6580	0.6290	0.6026	0.5784	0.6115	0.5845	0.5599	0.5374
$T_v=17$	0.7210	0.6892	0.6603	0.6338	0.6578	0.6287	0.6023	0.5781	0.6113	0.5843	0.5597	0.5371
$T_v=18$	0.7208	0.6889	0.6598	0.6333	0.6576	0.6284	0.6019	0.5777	0.6112	0.5840	0.5594	0.5368

(1) 给定信噪比和码率值时, GLR 会随着  $T_v$  的增加而增加. 这说明, 尽管从总体上而言, IND 的增益量不足以弥补 CND 的损失量, 但其增益量的增长速度高于损失量. 换言之, 较高的  $T_v$  值能够获得总面积更大的译码收敛区. 因此, 尽管 REG-LT 码以略微损失收敛性

的代价换取了低误码平台, 但其损失程度是可以接受的, 且便于调控.

(2) 不同信噪比和码率值时, 相同  $T_v$  值时的 GLR 并不相等. 这意味着, 无法给定一个普适于任何参数条件的 GLR 临界数值, 换言之, 仅仅从数值上并不一定能

判断得到最优的 GLR 和  $T_v$  值,而是需要从 GLR 的变化趋势进行分析.

(3) 当  $T_v$  达到一定值后, GLR 呈现出了下降的趋势. 这说明, 达到临界值之后继续增加  $T_v$ , 增益量的增长要缓于损失量. 因此, 译码收敛区的总面积将变小, 不利于快速译码迭代.

本文拟将 GLR 的极大值点对应的  $T_v$  值作为最优  $T_v$ . 但在求解最优  $T_v$  时, 需给定其上限值. 为此, 将使得 IND 曲线和 CNL 曲线恰好不存在交点时的  $T_v$  值作为上界, 记为  $T_{v(\max)}$ . 即  $T_v = T_{v(\max)}$  时, EXIT 图中的译码收敛区处于打开状态, 而  $T_v = T_{v(\max)} + 1$  时, IND 曲线与 CNL 曲线相交.

综上, 将求解最优  $T_v$  的方法总结为算法 2.

**算法 2 REG-LT 码的最优参数求解算法**

**初始化:** 给定信息节点的个数  $K$ , 采用的度分布为  $\Omega(x)$ , 待生成校验节点的个数  $N$ .

步骤 1: 计算给定参数条件下的最大参数值  $T_{v(\max)}$ .

步骤 2: 计算度分布, 以及相应的 GLR 值, 如下.

for  $m=1: T_{v(\max)}$

(1) 根据 3.1 节、3.2 节的结论, 求解  $T_v = m$  时, 对应的信息度分布、校验度分布.

(2) 根据 4.1 节、4.2 节的结论, 求解  $T_v = m$  时, 对应的 GLR 值并记录为  $GLR_m$ .

end

步骤 3: 令  $m=1$ , 求解上述 GLR 值的局部最大值点, 如下.

for  $m=1: T_{v(\max)} - 1$

if  $GLR_{m+1} > GLR_m$

则令  $m=m+1$ .

elseif  $GLR_{m+1} \leq GLR_m$

记录当前  $m$  值, 结束 for 循环.

end

end

(2) 根据 4.1 节、4.2 节的结论, 求解  $T_v = m$  时, 对应的 GLR 值并记录为  $GLR_m$ .

end

步骤 4: 输出最优  $T_v$  值为  $m$ .

根据算法 2, 求解出不同参数对应的最优  $T_v$  值, 如表 4 所示. 采用的度分布为  $\Omega_1(x)$ .

**5 仿真结果及分析**

本节对 REG-LT 码的 BER 性能进行仿真, 并与传统 LT 码和参考文献的改进算法进行比较. 为便于对照, 将文献 [21] 的算法记为 ED Code, 文献 [23] 的算法记为 CC Code. 仿真条件为: 发送端采用 BPSK 调制方式, 接收端均采用 BP 译码算法, 最大译码迭代次数为 50 次. 码长  $K=2048$ . 所有结果均通过多次蒙特卡洛仿真得到, 当错误码字数达到 10 个时停止. 采用的度分布分

表 4 不同信噪比和码率值下的最优  $T_v$  值

$E_s/N_0$	$R^{-1}$	$T_{v(\text{opt})}$	$E_s/N_0$	$R^{-1}$	$T_{v(\text{opt})}$
0 dB	2.0	14	1.5 dB	2.0	14
0 dB	2.1	14	1.5 dB	2.1	14
0 dB	2.2	14	1.5 dB	2.2	15
0 dB	2.3	15	1.5 dB	2.3	15
1 dB	2.0	14	2.0 dB	2.0	14
1 dB	2.1	14	2.0 dB	2.1	14
1 dB	2.2	14	2.0 dB	2.2	15
1 dB	2.3	15	2.0 dB	2.3	15

别记为  $\Omega_1(x)$  和  $\Omega_2(x)$ , 其中,  $\Omega_1(x)$  如式 (7) 所示,  $\Omega_2(x)$  为<sup>[26]</sup>

$$\begin{aligned} \Omega_2(x) = & 0.015x + 0.495x^2 + 0.167x^3 \\ & + 0.082x^4 + 0.071x^5 + 0.049x^8 \\ & + 0.048x^9 + 0.05x^{19} + 0.023x^{66} \end{aligned} \quad (24)$$

**5.1 不同信噪比时的 BER 性能**

图 6 和图 7 给出了采用不同校验度分布时, REG-LT 码的 BER 性能随信噪比的变化情况, 其中  $R^{-1}=2$ . 从中可以得出如下结论:

(1) REG-LT 码的 BER 性能优于传统 LT 码, 达到了预期的设计目的. 例如, 当  $E_s/N_0=0$  dB 时, 采用最优  $T_v$  的 REG-LT 码可将误码平台降低 3 个数量级以上; 而当  $E_s/N_0=1.5$  dB 时, 两者的差距已经逼近 5 个数量级. 若以  $10^{-5}$  为参照标准, REG-LT 码至少可以获得 3 dB 的编码增益, 这验证了本文算法的正确性.

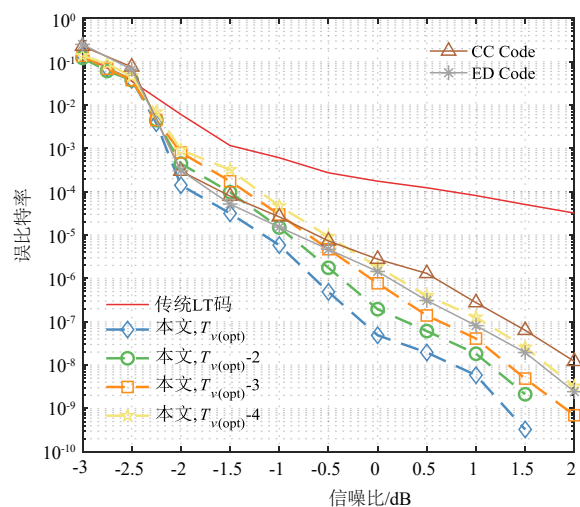
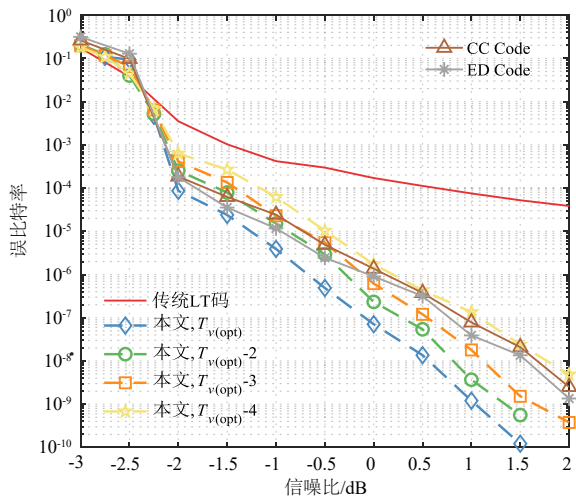


图 6 采用  $\Omega_1(x)$  时, REG-LT 码的 BER 性能随信噪比的变化情况

(2) REG-LT 码优于参考文献的编码算法. 从图中可以看出, 即使不采用最优  $T_v$ , 例如  $T_{v(\text{opt})} - 2$  时, REG-LT 码在高信噪比区域依然可以获得比 ED 算法和 CC 算

图7 采用 $\Omega_2(x)$ 时,REG-LT码的BER性能随信噪比的变化情况

法更低的误码平台;而在低信噪比区域内,尽管BER略高于参考算法,但相差甚小.类似地,若以 $10^{-7}$ 为参考标准, $T_{v(\text{opt})}$ 时的REG-LT码可分别获得近1.1 dB、0.8 dB的编码增益,这体现了本文算法的优势.

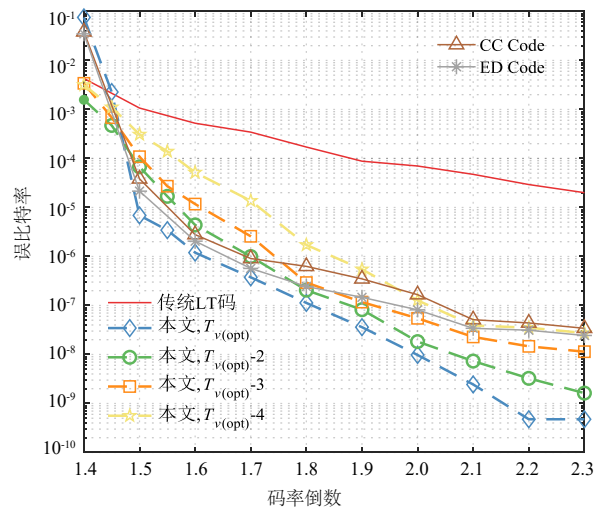
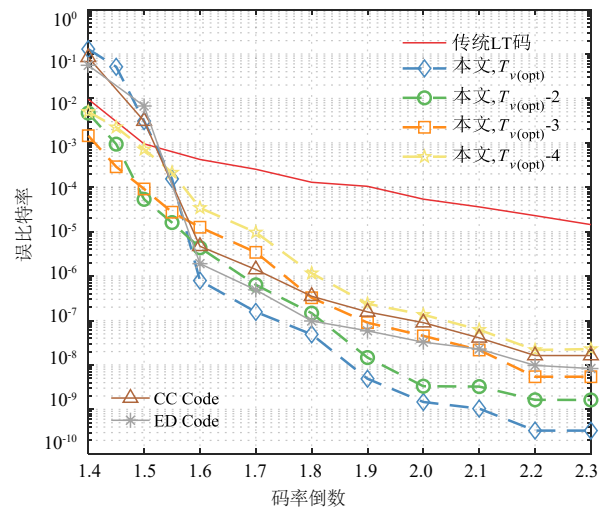
(3) 通过改变参数 $T_v$ ,可以便捷地调控REG-LT码能够达到的BER下界.本文算法引入了可供调整的参数 $T_v$ ,并通过改变校验度分布,实现了精准地调控信息节点最低度数值的目的,进一步可以预测算法的性能,这是ED算法和CC算法均不具备的.此外,在当前码率值下,信噪比为1.5 dB时,REG-LT码的BER比参考算法低了2个数量级;这说明,从优化编码结果的角度进行改进设计,性能提升效果要优于控制过程的改进算法.

## 5.2 不同码率值时的BER性能

图8和图9给出了采用不同校验度分布时,REG-LT码的BER性能随码率值的变化情况,其中 $E_s/N_0=1$  dB.为了便于观察,将横坐标置为码率倒数 $R^{-1}$ .从中可以得出如下结论:

(1) 当信噪比固定不变时,REG-LT码的BER性能依然是最优的.例如,当采用 $T_{v(\text{opt})}-2$ 时,REG-LT码便达到了比ED算法、CC算法更低的误码平台,且差距高达1个数量级以上.此外,根据算法2设计得到的最优 $T_v$ ,其对应的误码平台仍是最低的,这说明4.2节所设计的参数求解方法是可行的.

(2) 切换校验度分布时,REG-LT码仍然能够实现降低误码平台的效果.这是因为,通过设定参数 $T_v$ ,REG-LT码不仅可以提高信息节点的平均度数值,还能够自适应地优化校验度分布,使其最优匹配当前信道参数和编码倍率.换言之,REG-LT码不会受到校验度分布的限制,具有较强的灵活性.

图8 采用 $\Omega_1(x)$ 时,REG-LT码的BER性能随码率值的变化情况图9 采用 $\Omega_2(x)$ 时,REG-LT码的BER性能随码率值的变化情况

## 6 结论

针对传统LT码在AWGN信道中存在的严重误码平台问题,本文设计了具有逆向边增长结构的LT码.通过理论推导和仿真分析,可得出以下结论:(1)REG-LT码能够显著降低误码平台.本文算法可将BER下界降低近3个数量级以上,且可获得超过2 dB的编码增益;(2)REG-LT码的度分布可公式化表达,并通过参数GLR建立起了 $T_v$ 与收敛性的联系,这为设计最优 $T_v$ 值提供了依据;(3)REG-LT码的最优阈值设计方法仍待改进.文中仅借助GLR的单调性来寻找最优阈值,略显不足,后续研究中可考虑设计具有严谨数学理论支撑的寻优方法;(4)REG-LT码的收敛性略有损失.度分布的变化使得收敛区出现局部压缩的情况,不利于快速收敛.因此,下一步研究中可以考虑设计具有无损收

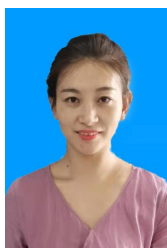
敛性的编码算法.

#### 参考文献

- [1] BYERS J W, LUBY M, MITZENMACHER M. A digital fountain approach to asynchronous reliable multicast[J]. *IEEE Journal on Selected Areas in Communications*, 2002, 20(8): 1528-1540.
- [2] HUANG J X, FEI Z S, CAO C Z, et al. Reliable broadcast based on online fountain codes[J]. *IEEE Communications Letter*, 2021, 25(2): 369-373.
- [3] LUBY M. LT codes[C]//*Proceedings of the IEEE Symposium on Foundations of Computer Science*. Vancouver: IEEE, 2002: 271-280.
- [4] ZHANG Z Y, ZHANG H Z, DAI H Y, et al. Fountain-coded file spreading over mobile networks[J]. *IEEE Transactions on Wireless Communications*, 2017, 16(10): 6766-6778.
- [5] ABBAS W B, CASARI P, ZORZI M. Controlled flooding of fountain codes[J]. *IEEE Transactions on Wireless Communications*, 2017, 16(7): 4698-4710.
- [6] 王丽娟, 杜秀娟, 李冲. 面向水声网络可靠传输的FDR编解码算法[J]. *通信学报*, 2020, 41(4): 81-91.  
WANG L J, DU X J, LI C. FDR coding and decoding algorithm for reliable transmission in underwater acoustic network[J]. *Journal on Communications*, 2020, 41(4): 81-91. (in Chinese)
- [7] 刘建航, 武万萍, 王树庆, 等. 基于Raptor10码的自由空间光通信系统设计[J]. *红外与激光工程*, 2019, 48(9): 173-178.  
LIU J H, WU W P, WANG S Q, et al. Design of free space optical communication system based on Raptor10 code[J]. *Infrared and Laser Engineering*, 2019, 48(9): 173-178. (in Chinese)
- [8] SHANG J, WU W J, LEE C H, et al. REF codes: intermediate performance oriented fountain codes with feedback [J]. *IEEE Transactions on Vehicular Technology*, 2020, 69(11): 13148-13164.
- [9] ZHANG K, JIAO J, HUANG Z X, et al. Finite block-length analog fountain codes for ultra-reliable low latency communications[J]. *IEEE Transactions on Communications*, 2020, 68(3): 1391-1404.
- [10] PALANKI R, YEDIDIA J S. Rateless codes on noisy channels[C]//*Proceedings of the IEEE International Symposium on Information Theory*. Chicago: IEEE, 2004: 271-280.
- [11] SHOKROLLAHI A. Raptor codes[J]. *IEEE Transactions on Information Theory*, 2006, 52(6): 2551-2567.
- [12] GALLAGER R. Low-density parity-check codes[J]. *IRE Transactions on Information Theory*, 1962, 8(1): 21-28.
- [13] 李锦明, 王国栋, 刘梦欣, 等. CCSDS标准下LDPC码的编译码算法研究[J]. *电子学报*, 2020, 48(11): 2114-2121.  
LI J M, WANG G D, LIU M X, et al. Research on encoding and decoding algorithm of LDPC codes based CCSDS standard [J]. *Acta Electronica Sinica*, 2020, 48(11): 2114-2121. (in Chinese)
- [14] TIAN S, LI Y H, SHIRVANIMOGHADDAM M, et al. A physical-layer rateless code for wireless channels[J]. *IEEE Transactions on Communications*, 2013, 61(6): 2117-2127.
- [15] SHI P C, WANG Z Y, LI D Z, et al. Zigzag decodable online fountain codes with high intermediate symbol recovery rates[J]. *IEEE Transactions on Communications*, 2020, 68(11): 6629-6641.
- [16] HU Y M, LIU R K, BIAN H X, et al. Design and analysis of a low-complexity decoding algorithm for spinal codes [J]. *IEEE Transactions on Vehicular Technology*, 2019, 68(5): 4667-4679.
- [17] ETESAMI O, SHOKROLLAHI A. Raptor codes on binary memoryless symmetric channels[J]. *IEEE Transactions on Information Theory*, 2006, 52(5): 2033-2051.
- [18] SHIRVANIMOGHADDAM M, JOHNSON S. Raptor codes in the low SNR regime[J]. *IEEE Transactions on Communications*, 2016, 64(11): 4449-4460.
- [19] KUO S H, GUAN Y L, LEE S K, et al. A design of physical-layer Raptor codes for wide SNR ranges[J]. *IEEE Communications Letter*, 2014, 18(3): 491-494.
- [20] XU S K, XU D Z. Optimization design and asymptotic analysis of systematic Luby transform codes over BIAWGN channels[J]. *IEEE Transactions on Communications*, 2016, 64(8): 3160-3168.
- [21] HUSSAIN I, XIAO M, RASMUSSEN L K. Error floor analysis of LT codes over the additive white gaussian noise channel[C]//*Proceedings of the IEEE Global Communications Conference*. Houston: IEEE, 2011: 1-5.
- [22] HUSSAIN I, XIAO M, RASMUSSEN L K. Design of LT codes with equal and unequal erasure protection over binary erasure channels[J]. *IEEE Communications Letter*, 2013, 17(2): 261-264.
- [23] CHEN C M, CHEN Y P. Connection choice code[J]. *IEEE Transactions on Communications*, 2014, E97-B(7): 1350-1357.
- [24] CHANG L J, WANG C H, ZAO J K. An error-floor re-

duction technique for short-length LT codes[C]//Proceedings of the International Symposium on Information Theory and its Applications. Victoria: IEEE, 2014: 279-283.

- [25] KHAREL A, CAO L. Improved fountain codes for BI-AWGN channels[C]//Proceedings of the IEEE Wireless Communications and Networking Conference. San Francisco: IEEE, 2017: 1-6.
- [26] ZHANG W Z, HRANILOVIC S, SHI C. Soft-switching hybrid FSO/RF links using short-length Raptor codes: design and implementation[J]. IEEE Journal on Selected Areas in Communications, 2009, 27(9): 1698-1708.
- [27] BRINK S T. Convergence behavior of iteratively decoded parallel concatenated codes[J]. IEEE Transactions on Communications, 2001, 49(10): 1727-1737.
- [28] BRINK S T. Design of low-density parity-check codes for modulation and detection[J]. IEEE Transactions on Communications, 2004, 52(4): 670-678.



**倪淑燕(通讯作者)** 女,1981年出生于河北省邢台市. 现为航天工程大学副教授. 主要研究方向为阵列信号处理,正交频分复用技术等.

E-mail: daninini@163.com



**雷拓峰** 男,1998年出生于陕西省西安市. 现为航天工程大学博士研究生. 主要研究方向为正交频分复用技术,稀疏码分多址技术,信道编码等.

E-mail: tuofenglei@163.com

#### 作者简介



**宋 鑫** 男,1995年出生于山西省临汾市. 现为航天工程大学博士研究生. 主要研究方向为信道编码,信道均衡,稀疏码分多址技术等.

E-mail: singlersx@163.com



**程乃平** 男,1962年出生于陕西省宝鸡市. 现为航天工程大学研究员. 主要研究方向为卫星信号处理,正交频分复用技术,信道编码等.

E-mail: cheng\_np@163.com



**廖育荣** 男,1972年出生于四川省德阳市. 现为航天工程大学研究员. 主要研究方向为航天测控技术,卫星信号处理技术等.

E-mail: 406573420@qq.com