高误码率下 LDPC 稀疏校验矩阵重建

吴昭军[1]，张立民[1]，钟兆根[2]，刘仁鑫[1]

（1. 海军航空大学航空作战勤务学院，山东 烟台 264001；2. 海军航空大学航空基础学院，山东 烟台 264001）

摘 要：针对 LDPC 重建问题，提出了一种可直接重建 LDPC 稀疏校验矩阵的算法。首先，根据传统重建算法原理，详细分析了传统重建算法存在的缺陷以及缺陷存在的原因；其次，基于 LDPC 稀疏矩阵的特性，通过多次随机抽取码字中部分比特序列进行高斯消元，同时为了可靠实现抽取的比特序列能包含校验节点，基于一次抽取包含校验节点的概率，确定多次随机抽取的次数；最后，在误码条件下，基于疑似校验向量关系成立的统计特性和最小错误判决准则，实现稀疏校验向量的判定。仿真结果表明，所提算法在误码率为 0.001 的条件下，针对目前 IEEE 802.11 协议中大部分 LDPC 的重建率能达到 95%以上，且噪声稳健性优于传统的重建算法，同时所提重建算法不仅不再需要对校验矩阵稀疏化处理，而且对于双对角线与非双对角线形式的校验矩阵都具有较好的通用性。

关键词：LDPC；稀疏校验矩阵；随机抽取；高斯消元；最小错误判决准则；重建

中图分类号：TN911.7

文献标识码：A

DOI: 10.11959/j.issn.1000−436x.2021009

Reconstruction of sparse check matrix for LDPC at high bit error rate

WU Zhaojun[1], ZHANG Limin[1], ZHONG Zhaogen[2], LIU Renxin[1]

1. School of Aviation Support, Naval Aviation University, Yantai 264001, China

2. School of Basis of Aviation Science, Naval Aviation University, Yantai 264001, China

Abstract: In order to reconstruct the sparse check matrix of LDPC, a new algorithm which could directly reconstruct the LDPC was proposed. Firstly, according to the principle of the traditional reconstruction algorithm, the defects of the traditional algorithm and the reasons for the defects were analyzed in detail. Secondly, based on the characteristics of sparse matrix, some bit sequences in code words were randomly extracted for Gaussian elimination. At the same time, in order to reliably realize that the extracted bits sequence could contain parity check nodes, the multiple random variables were determined based on the probability of containing check nodes in one extraction. Finally, the statistical characteristics of LDPC under the suspected check vector was analyzed. Based on the minimum error decision rule, the sparse check vector was determined. The simulation results show that the rate of reconstruction of most LDPC in IEEE 802.11 protocol can reach more than 95% at BER of 0.001, and the noise robustness of the proposed method is better than that of the traditional algorithm. At the same time, the new algorithm not only does not need sparseness of parity check matrix, but also has the good performance for both diagonal and non-diagonal check matrix.

Keywords: LDPC, sparse check matrix, random extraction, Gauss elimination, minimum error decision rule, reconstruction

---

收稿日期：2020−08−25；**修回日期**：2020−11−16

通信作者：张立民，iamzlm@163.com

基金项目：国家自然科学基金资助项目（No.91538201）；泰山学者工程专项经费基金资助项目（No.ts201511020）；信息系统安全技术重点实验室基金资助项目（No.6142111190404）

Foundation Items: The National Natural Science Foundation of China (No.91538201), Taishan Scholar Special Foundation (No.ts201511020), The Chinese National Key Laboratory of Science and Technology on Information System Security (No.6142111190404)

1 引言

为了对抗信道中噪声的干扰，数字通信系统中广泛采用信道编码技术。受 Turbo 码迭代译码思想的启发，LDPC（low density parity check code）被再次发现，目前，已经被广泛应用于如 IEEE 802.11、DVB-S2 等数据传输协议中。从非合作通信方的角度，研究高误码率条件下 LDPC 稀疏校验矩阵的重建，对于目前大量采用 LDPC 编码的通信协议逆向分析具有非常重要的意义。

目前，针对信道编码参数识别的研究主要集中于分组码[1-2]、卷积码[3]以及 Turbo 码[4-7]等，而对 LDPC 参数识别的文献较少。同时，在有误码条件下 LDPC 稀疏校验矩阵的重建问题一直是一个难点，主要原因在于 LDPC 的码长较长，在有误码条件下，目前针对信道编码识别的算法，如矩阵分析[8]、Walsh-Hadarmd 变换[9]等往往失效。从已发表的论文来看，LDPC 的重建分为闭集识别和开集识别。针对 LDPC 的闭集识别，文献[10-11]引入了对数似然比（LLR, log-likelihood ratio）的概念，将 LDPC 闭集识别问题等价于 LLR 的优势统计问题，即遍历闭集中可能的 LDPC 稀疏校验矩阵，求解出对应于 LLR 值最大的稀疏校验矩阵，从而完成识别。为了提高算法的实时性，文献[12]详细分析了 LLR 的统计特性，基于最小错误判决准则设定最优门限，快速完成闭集中 LDPC 识别，但文献[12]为了简化计算，在计算 LLR 时进行了近似处理，使在低信噪比下基于 LLR 的识别算法性能变差。针对该问题，文献[13]提出了基于余弦符合度的识别方法，该方法计算比较简便，同时未采用近似处理，在低信噪比下性能明显优于基于 LLR 的算法。虽然文献[10-13]能够在低信噪比下完成 LDPC 识别，但是这种识别的前提条件是已知几类稀疏校验矩阵，严格来说这不是真正的盲识别，在某些微波通信或是短波通信系统中，通信协议往往是人为设定的，对于非合作方而言，几乎没有先验信息，故 LDPC 的开集识别更具实用性，同时也更具挑战性。文献[14]最先研究 LDPC 开集识别问题，首先利用截获的数据构造码字矩阵，然后采用高斯消元方法得到校验向量，同时利用校验向量对误码码字进行筛选和剔除，该方法虽然具有一定的容错性，但是在有误码条件下所需数据量较大，同时获得的校验矩阵是非稀疏的，无法用来译码。为了获得稀疏校验向量，文献[15]提出了基于 2 阶行消元和 P 阶行消元变换的校验矩阵稀疏化算法，该算法针对双对角线形式的稀疏矩阵具有较好的实用性，但是对于非双对角线形式的稀疏矩阵，其消元的阶次较高，计算复杂度会急剧增大。为了克服文献[15]的缺点，文献[16]提出基于 Canteaut-Chabaud 算法[17]的随机稀疏化方法，该方法的通用性较好，对于非双对角线形式的稀疏校验矩阵也能较好地重建，但是该方法需要多次的迭代消元，在稀疏化过程中具有一定的盲目性，这使方法的计算复杂度较高。文献[18]从提高 LDPC 校验矩阵重建性能出发，引入了 LDPC 反馈迭代译码方法，利用重建的部分稀疏校验向量对 LDPC 译码纠错，从而改善获取的码字质量，较好地提升了算法的性能，但是该算法需要反复进行高斯消元、稀疏化以及迭代译码，导致其计算复杂度很高。从目前 LDPC 研究现状来看，开集识别仍然是一个棘手的问题，还需要进一步从计算复杂和容错性能 2 个方面改进。

基于此，本文提出了一种新的 LDPC 稀疏矩阵重建方法，该方法首先多次随机抽取码字中部分比特进行高斯消元，当随机抽取的比特位包含稀疏校验向量的校验位时，消元后的结果张成的空间中一定包含稀疏校验向量，从而完成稀疏校验向量的获取；其次，为了使随机抽取过程中可靠出现抽取的位包含稀疏校验向量校验位，分析了一次随机抽取包含校验位的概率，得到了最小抽取次数；最后，详细分析了有误码条件下疑似稀疏校验向量的统计规律，基于最小错误判决准则，实现了 LDPC 稀疏校验向量的判定，最终完成稀疏校验矩阵的重建。

2 LDPC 原理以及重建问题描述

LDPC 是由稀疏校验矩阵来定义的。1962 年，Gallager 通过稀疏校验矩阵的特性定义了 LDPC；Tanner 在稀疏校验矩阵的基础上结合图论提出了 LDPC 因子图描述方法。下面，给出稀疏校验矩阵以及 LDPC 的定义。

定义 1[19] 一个线性分组码的监督矩阵中元素“1”的个数占总元素个数的比例非常小，则这个监督矩阵被称为稀疏校验矩阵，这个线性分组码被称为 LDPC。

对于一个维度为 $(n-k)\times n$ 的 LDPC 稀疏校验矩阵 $\boldsymbol{H}$，通过初等行变换方式，可将其转化为标准

形式，即

$$\boldsymbol{H}' = \left[\boldsymbol{I}_{(n-k)\times(n-k)} \mid \boldsymbol{P}_{(n-k)\times k} \right] \tag{1}$$

其中，$\boldsymbol{I}_{(n-k)\times(n-k)}$ 为 $(n-k)\times(n-k)$ 维的单位矩阵。由标准形式的校验矩阵，可得到 LDPC 的生成矩阵为

$$\boldsymbol{G} = \left[\boldsymbol{P}^{\mathrm{T}}_{(n-k)\times k} \mid \boldsymbol{I}_{k\times k} \right] \tag{2}$$

其中，$\boldsymbol{I}_{k\times k}$ 为 $k\times k$ 维的单位矩阵，矩阵 $\boldsymbol{P}^{\mathrm{T}}$ 为 $\boldsymbol{P}$ 的转置。

将待编码的信息序列按照 k bit 分成一组，然后与 $\boldsymbol{G}$ 相乘，即可得到 LDPC 的码字。在目前的通信系统中，每一帧数据都有固定的同步码，利用同步码可以快速完成 LDPC 的码长以及码字起点的识别，所以 LDPC 识别的重点是在有误码的条件下利用截获的码字序列重建出 LDPC 的稀疏校验矩阵 $\boldsymbol{H}$。

3　传统重建算法的缺陷

传统的 LDPC 稀疏校验矩阵重建主要分为两步，即非稀疏校验矩阵的获取[14]和校验矩阵稀疏化处理[15]，其中非稀疏校验矩阵的获取是校验矩阵稀疏化处理的前提。传统方法将整个 LDPC 的码字进行高斯消元，然后得到非稀疏的校验向量，不妨设 LDPC 的码长为 n，截获到的序列长度为 l，构造的码字矩阵为 $\boldsymbol{A}$，即

$$\boldsymbol{A} = \begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,n} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{N,1} & c_{N,2} & \cdots & c_{N,n} \end{bmatrix} \tag{3}$$

其中，$N = \left\lfloor \dfrac{l}{n} \right\rfloor$，$\lfloor \cdot \rfloor$ 表示向下取整，$c_{i,j}(1 \leqslant i \leqslant N,\ 1 \leqslant j \leqslant n)$ 为截获码字元序列。

设在高斯消元过程中，$\boldsymbol{A}$ 的初等行变换矩阵为 $\boldsymbol{R}_{N\times N}$，初等列变换矩阵为 $\boldsymbol{S}_{n\times n}$，则经过初等行变换以及初等列变换后，矩阵 $\boldsymbol{A}$ 变为

$$\boldsymbol{R}_{N\times N}\boldsymbol{A}\boldsymbol{S}_{n\times n} = \begin{bmatrix} \boldsymbol{I}_{k'\times k'} & \boldsymbol{0}_{k'\times(n-k')} \\ \boldsymbol{B}_{(N-k')\times k'} & \boldsymbol{D}_{(N-k')\times(n-k')} \end{bmatrix} \tag{4}$$

其中，$\boldsymbol{0}_{k'\times(n-k')}$ 为 $k'\times(n-k')$ 维的全零矩阵。

设列变换矩阵 $\boldsymbol{S}_{n\times n} = [\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_n]$，其中 $\boldsymbol{s}_i(1 \leqslant i \leqslant n)$ 为 $\boldsymbol{S}_{n\times n}$ 的列向量，在无误码条件下，式(4)中 $k' = k$，此时列向量 $\boldsymbol{s}_{k+1}, \boldsymbol{s}_{k+2}, \cdots, \boldsymbol{s}_n$ 正好构成 LDPC 非稀疏校验矩阵，同时 $\boldsymbol{D}_{(N-k')\times(n-k')}$ 为全零矩阵；当存在误码时，部分线性关系受到破坏，此时 $k' > k$，列向量 $\boldsymbol{s}_{k'+1}, \boldsymbol{s}_{k'+2}, \cdots, \boldsymbol{s}_n$ 构成 LDPC 部分非稀疏校验矩阵，同时 $\boldsymbol{D}_{(N-k')\times(n-k')}$ 为一稀疏矩阵。在极端情况下，$k' = n$，矩阵 $\boldsymbol{A}$ 为列满秩矩阵，未能重建出非稀疏校验向量。由上述分析可知，对 $\boldsymbol{A}$ 进行高斯消元时，真正对结果产生影响的是前 n 个码字，只有当前 n 个码字同时满足同一校验关系时，才能得到校验向量。

不妨设 LDPC 中某一校验向量为 $\boldsymbol{v}$，其码重为 w，当误码率为 p_{e} 时，通过对矩阵 $\boldsymbol{A}$ 进行高斯消元，仍能得到校验向量 $\boldsymbol{v}$，则必须满足以下条件：向量 $\boldsymbol{v}$ 中元素为 1 的位置对应于 $\boldsymbol{A}$ 中前 n 个码字中比特没有出现误码或出现误码的个数为偶数，此时通过模 2 运算，误码没有产生影响，即单个码字满足 $\boldsymbol{v}$ 的校验关系概率为

$$P = \sum_{i=0}^{\left\lfloor \frac{w}{2} \right\rfloor} \mathrm{C}_w^{2i} p_{\mathrm{e}}^{2i} \left(1 - p_{\mathrm{e}}\right)^{w-2i} \tag{5}$$

其中，C 表示求组合数运算。

对 $\boldsymbol{A}$ 进行高斯消元，得到非稀疏校验向量 $\boldsymbol{v}$，则至少需要 n 个满足校验关系的码字，即能够通过高斯消元求解得到 $\boldsymbol{v}$ 的概率为 $P_1 = P^n$。

由此可知，对于传统算法而言，在信道误码率一定时，能够获取到校验向量的概率随码长 n 和校验向量码重 w 的增大而呈指数级下降。在实际工程中，LDPC 的码长通常很大，对应于非稀疏校验向量的码重也非常大，故误码率一旦增大，能够通过高斯消元求解校验向量的概率将非常小，同时高斯消元算法的计算复杂度近似为 $O(n^3)$，由此可知，对于码长较长的 LDPC 而言，计算复杂度较大。

对于非稀疏校验向量的稀疏化，传统方法采用 2 阶和 P 阶行变换方式，对于非双对角线形式的稀疏校验矩阵，P 的选取具有一定的盲目性，故计算量比较大，这导致其算法的实时性较差。

从上述分析来看，在有误码条件下，传统的重建算法与工程实用化还有一定的差距。基于此，本文提出了一种新的思路，利用 LDPC 校验矩阵非常稀疏这一特点，通过多次随机抽取码字中部分比特进行高斯消元，当抽取的比特包含稀疏校验向量中校验位时，经过消元后可直接得到涵盖稀疏校验向

量的对偶空间，由于抽取的是码字中部分比特，故该对偶空间维度会很小，通过遍历对偶空间中向量，可以很快提取到稀疏校验向量。由于参与消元的仅为码字的部分比特，故与传统方法相比，计算复杂度以及容错性能均会有明显改善，同时在一定程度上避免了稀疏化步骤。

4 LDPC 稀疏矩阵重建模型建立

4.1 随机抽取次数的确定

从传统识别算法来看，为了获取校验向量，算法将整个 LDPC 的码字进行列消元，但是从 LDPC 稀疏校验矩阵特点来看，构成 LDPC 码字校验关系的比特长度实际上很短，在求解稀疏校验向量过程中，没有必要将整个 LDPC 码字进行消元，如果抽取码字中部分比特进行列消元，当抽取到的比特中涵盖了稀疏校验向量的校验位时，消元的结果构成的对偶空间一定含有稀疏校验向量，同时对偶空间的维度较小，很容易找到稀疏校验向量。由于一次随机抽取并不能保证涵盖一个稀疏校验向量中完整的校验位，故需要多次抽取，为了确定抽取次数，实现可靠涵盖到整个完整的校验位，需要分析一次随机抽取的比特位置能够包含完整校验位的概率。不妨设 LDPC 稀疏校验矩阵中，某一稀疏校验向量为 $\boldsymbol{v}'$，对应于校验节点数目为 w'，码字中随机抽取的比特数目为 s，则对于码长为 n 的 LDPC，抽取出 s 个位置的样本空间数目为 C_n^s，而这 s 个位置中正好包含 w' 的校验节点的个数为 $\mathrm{C}_{n-w'}^{s-w'}$，故得到一次随机抽取可包含 $\boldsymbol{v}'$ 中稀疏校验节点的概率为

$$P_2=\frac{\mathrm{C}_{n-w'}^{s-w'}}{\mathrm{C}_n^s}=\prod_{i=0}^{w'-1}\frac{s-i}{n-i} \tag{6}$$

由式(6)可知，当 LDPC 码长固定后，概率 P_2 与 s 和稀疏校验向量的码重有关，s 越大，校验向量越稀疏，则 P_2 的值就越大。通常情况下，LDPC 重稀疏校验向量码重很小，一般不超过 10；对于 s 而言，虽然取值越大，得到校验向量的可能性就越大，但是获取到的对偶空间的维度也会增大，这不利于稀疏向量的求取，极端情况下，$s=n$，算法退化为传统算法，故 s 的取值不宜过大。为兼顾 P_2 与对偶空间维度，一般参考 LDPC 码率来选择，设 LDPC 码率为 R，则每次随机抽取的比特数目可为 Rn。

确定概率 P_2 后，下面进一步探讨如何确定随机抽取的次数，从而可靠实现在抽取的比特中包含校验节点。不妨设随机抽取的次数为 iter，则在 iter 次的随机抽取中，能够出现包含稀疏校验节点的次数 T 服从二项式分布，即

$$T\sim B\left(\mathrm{iter},P_2\right) \tag{7}$$

当抽取的次数 iter 较大时，由棣莫弗–拉普拉斯定理可得

$$\frac{T-\mathrm{iter}P_2}{\sqrt{\mathrm{iter}P_2\left(1-P_2\right)}}\sim\mathcal{N}\left(0,1\right) \tag{8}$$

其中，$\mathcal{N}\left(0,1\right)$ 表示标准正态分布。

在数理统计过程中，当事件发生的概率大于 0.997 5 时，可将该事件定义为大概率事件，即在随机抽取 iter 次过程中，至少发生一次，则 iter 必须满足式(9)。

$$P\left(\frac{T-\mathrm{iter}P_2}{\sqrt{\mathrm{iter}P_2\left(1-P_2\right)}}\geqslant\frac{1-\mathrm{iter}P_2}{\sqrt{\mathrm{iter}P_2\left(1-P_2\right)}}\right)\geqslant 0.997\,5 \tag{9}$$

通过查询正态分布表可知

$$\frac{1-\mathrm{iter}P_2}{\sqrt{\mathrm{iter}P_2\left(1-P_2\right)}}\leqslant -2.81 \tag{10}$$

求解式(10)，得到 iter 的取值范围为

$$\mathrm{iter}\geqslant\frac{\left(2+2.81^2\left(1-P_2\right)\right)+\sqrt{\left(2+2.81^2\left(1-P_2\right)\right)^2-4}}{2P_2} \tag{11}$$

从而得到可靠包含稀疏校验节点的最小随机抽取次数为

$$\mathrm{iter}_{\min}=\frac{\left(2+2.81^2\left(1-P_2\right)\right)+\sqrt{\left(2+2.81^2\left(1-P_2\right)\right)^2-4}}{2P_2} \tag{12}$$

4.2 疑似校验向量判定

随机抽取码字中部分比特序列构成新的码字矩阵，通过高斯消元法得到的解向量仅能满足前 s 个码字，这类解向量被定义为疑似校验向量，此时需要综合考虑在疑似校验向量下整个码字成立的情况，所以需要利用真实校验向量与非校验向量下，码字校验关系成立的统计特性进行判定。不妨设得到的疑似校验向量为 $\boldsymbol{h}$，对应的码重为 w_h。首先，考虑以下 2 类假设条件：$\mathcal{H}_0$ 表示 $\boldsymbol{h}$ 不是校验向量，

$\mathcal{H}_1$表示 $\boldsymbol{h}$ 是校验向量。

在信道误码率为 p_{e}、假设条件为 $\mathcal{H}_1$ 下，由式(5)可知，校验关系仍然成立的概率为

$$P_{h_1}=\sum_{i=0}^{\left\lfloor\frac{w_h}{2}\right\rfloor}\mathrm{C}_{w_h}^{2i}p_{\mathrm{e}}^{2i}\left(1-p_{\mathrm{e}}\right)^{w_h-2i} \tag{13}$$

对于假设条件 $\mathcal{H}_0$，由于 $\boldsymbol{h}$ 不是校验向量，因此码字校验关系成立概率随机，即 $P_{h_0}=0.5$。将码字成立个数与不成立个数之差 t 作为统计量，当码字个数 N 充分大时，在假设条件 $\mathcal{H}_1$ 下，t 服从均值为 $N(2P_{h_1}-1)$、方差为 $4NP_{h_1}(1-P_{h_1})$ 的正态分布，即

$$\mathcal{H}_1: t\sim\mathcal{N}\left(N\left(2P_{h_1}-1\right),4NP_{h_1}\left(1-P_{h_1}\right)\right) \tag{14}$$

在假设条件 $\mathcal{H}_0$ 下，t 服从均值为 0、方差为 N 的正态分布，即

$$\mathcal{H}_0: t\sim\mathcal{N}\left(0,N\right) \tag{15}$$

为方便描述，不妨记 $\mu_0=0$，$\sigma_0^2=N$，$\mu_1=N(2P_{h_1}-1)$，$\sigma_1^2=4NP_{h_1}(1-P_{h_1})$。设 2 类假设的判决门限为 Λ，则虚警概率 P_{f} 为

$$P_{\mathrm{f}}=\int_{\Lambda}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_0}\mathrm{e}^{-\frac{(x-\mu_0)^2}{2\sigma_0^2}}\mathrm{d}x \tag{16}$$

漏警概率 P_{a} 为

$$P_{\mathrm{a}}=\int_{-\infty}^{\Lambda}\frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\mathrm{d}x \tag{17}$$

综合 2 类错误判决概率，得到平均错误判决概率为

$$P_{\mathrm{er}}=\frac{1}{2}\left(P_{\mathrm{f}}+P_{\mathrm{a}}\right) \tag{18}$$

利用 P_{er} 对 Λ 求导数，并令其等于 0，得到

$$\frac{1}{\sigma_0}\mathrm{e}^{-\frac{(\Lambda-\mu_0)^2}{2\sigma_0^2}}=\frac{1}{\sigma_1}\mathrm{e}^{-\frac{(\Lambda-\mu_1)^2}{2\sigma_1^2}} \tag{19}$$

对式(19)两边取对数，将其化为一元二次方程，求解得到最小错误判决门限 Λ_{opt} 为

$$\Lambda_{\mathrm{opt}}=\frac{\sigma_0^2\mu_1-\sigma_1^2\mu_0-\sigma_0\sigma_1\sqrt{\left(\mu_0-\mu_1\right)^2+\left(\sigma_1^2-\sigma_0^2\right)\ln\left(\frac{\sigma_1}{\sigma_0}\right)}}{\sigma_0^2-\sigma_1^2} \tag{20}$$

当 N 足够大时，Λ_{opt} 可近似为

$$\Lambda_{\mathrm{opt}}\approx\frac{\mu_0\sigma_1+\mu_1\sigma_0}{\sigma_1+\sigma_0} \tag{21}$$

通过高斯消元得到疑似校验向量后，求取对应的统计量 t，然后计算最小错误判决门限 Λ_{opt}，当 $t\geqslant\Lambda_{\mathrm{opt}}$ 时，即可判定为校验向量。需要注意的是，由于每次高斯消元仅利用了前 s 个码字，为了增大疑似校验向量获取的概率，充分利用截获的码字，可以采用多次迭代随机选择 s 个码字进行消元，直到出现疑似校验向量。

4.3　LDPC 稀疏校验矩阵重建算法步骤

所提算法充分利用了 LDPC 校验矩阵非常稀疏这一特点，即对于任一稀疏校验向量而言，码字中实际参与校验的比特位是非常小的，当随机抽取的比特位置中正好涵盖了校验节点时，通过高斯消元可得到包含稀疏校验向量的对偶空间，且对偶空间的维度较小，通过遍历的方式，可以快速获取稀疏向量，具体的算法步骤如下。

步骤 1　初始化参数 s，迭代消元次数 iter_1，$i_1=1$，$i_2=1$，用截获的数据构造 LDPC 码字矩阵 $\boldsymbol{A}_{N\times n}$，其中 N 为码字数目，n 为 LDPC 码长。

步骤 2　计算随机抽取次数 $\mathrm{iter}_{\min}$，从而保证可靠出现涵盖稀疏校验向量节点位置。

步骤 3　随机抽取 $\boldsymbol{A}_{N\times n}$ 中 s 列数据，构造新的码字矩阵 $\boldsymbol{C}_{N\times s}$，同时 $i_1=i_1+1$。

步骤 4　随机选取 $\boldsymbol{C}_{N\times s}$ 中 s 行数据构成方阵 $\boldsymbol{B}_{s\times s}$，同时采用高斯消元法得到 $\boldsymbol{B}_{s\times s}$ 的对偶空间基向量。

步骤 5　若对偶空间为非零空间，则利用对偶空间基遍历整个解空间，寻找稀疏校验向量；否则 $i_2=i_2+1$，重复步骤 4，直到 $i_2>\mathrm{iter}_1$。

步骤 6　计算稀疏向量对应的统计量 t 以及最小错误判决门限 Λ_{opt}，若 $t\geqslant\Lambda_{\mathrm{opt}}$，则保存该稀疏校验向量于集合 $\boldsymbol{H}$ 中，同时重复步骤 3，i_2 置 1；否则 $i_2=i_2+1$，重复步骤 4，直到 $i_2>\mathrm{iter}_1$。

步骤 7　判断 $i_1>\mathrm{iter}_{\min}$ 是否成立，若成立，则输出稀疏校验矩阵 $\boldsymbol{H}$；否则转至步骤 3，直到 $i_1>\mathrm{iter}_{\min}$。

在步骤 5 中，由于随机抽取了实际码字中部分比特进行高斯消元，故其对偶空间维度远小于实际 LDPC 对偶空间维度，通过遍历解空间中向量，能

很容易地求解对偶空间中稀疏向量。

在实际仿真中，算法对每次得到的校验向量都要进行筛选和判定，当向量的码重与码长之比小于 0.03 时，将其判定为稀疏校验向量，保存的校验向量不仅要满足稀疏条件，同时还必须满足与已保存的稀疏向量相互线性不相关的条件，故稀疏校验向量的个数等于重建矩阵的秩。在后续算法性能验证过程中，本文将重建的稀疏向量个数与定义的稀疏矩阵秩之比定义为重建率，在不同误码条件下，利用重建率这一指标来表征算法的性能。由于 LDPC 译码依赖于稀疏校验矩阵，故在信道较恶劣的条件下，利用本文重建的部分稀疏校验矩阵可以对 LDPC 进行迭代译码，从而改善截获的码字质量，进一步提升算法的性能。

4.4　计算复杂度分析

设随机抽取的列数为 s，随机抽取的次数为 $\mathrm{iter}_{\min}$，构造方阵 $\boldsymbol{B}_{s\times s}$ 并进行高斯消元的最大次数为 $\mathrm{iter}_{\mathrm{I}}$；由于进行一次高斯消元的计算复杂度为 $O(s^3)$，则在最不利条件下，本文算法的最大计算量为 $O(\mathrm{iter}_{\min}\mathrm{iter}_{\mathrm{I}}s^3)$。由于本文算法仅抽取码字中部分比特进行消元，故得到的对偶空间维度较小，可以很容易地找到稀疏校验向量，而且在绝大多数情况下，能直接得到稀疏校验向量。对于传统算法而言，针对码率为 k/n 的 LDPC，需要将整个码字进行高斯消元，其最大计算量为 $O(\mathrm{iter}_{\mathrm{I}}n^3)$，由于得到的是非稀疏校验向量，故还需要进行稀疏化处理，在稀疏化处理过程中，采用 P 阶行变化处理，其计算复杂度为 $O\left(n\sum_{i=2}^{P}\mathrm{C}_{n-k}^{i}\right)$，故传统算法[14-15]总的计算复杂度为 $O\left(\mathrm{iter}_{\mathrm{I}}n^3+n\sum_{i=2}^{P}\mathrm{C}_{n-k}^{i}\right)$，对于非双对角线形式的 LDPC，$P$ 的取值一般会比较大，此时传统算法的复杂度会急剧增加，由此可知，本文算法的通用性要优于传统方法。

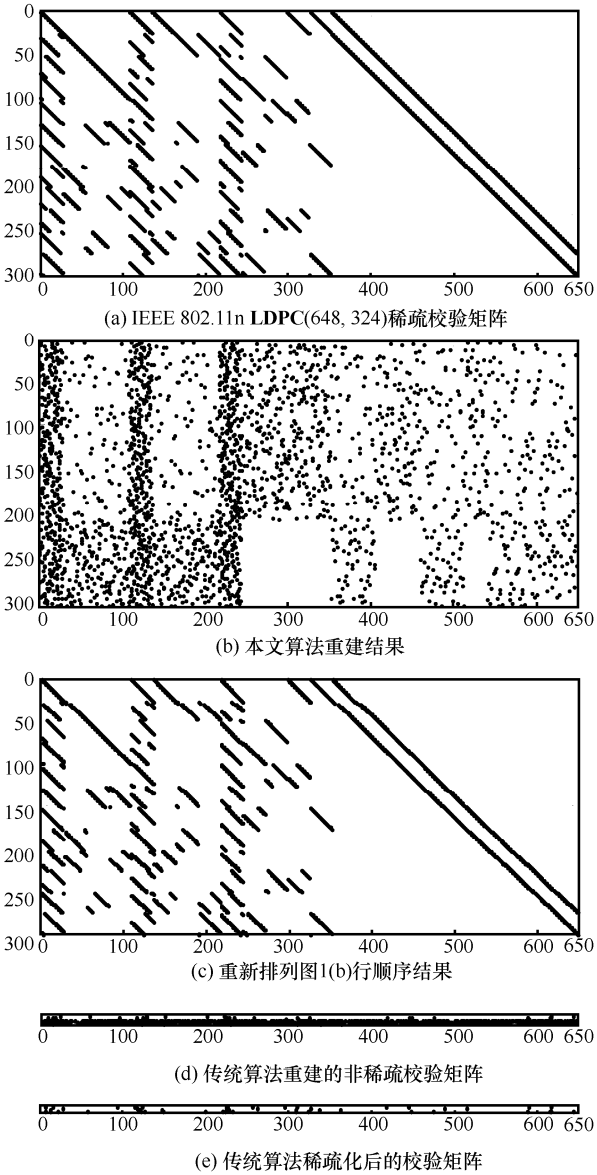
5　仿真验证

本节首先验证本文算法的有效性，即能够在有误码条件下，完成双对角以及非双对角线形式的稀疏校验矩阵重建；其次考察在不同迭代消元次数、不同截获码字个数、不同码长以及码率条件下，算法的重建性能；最后将本文算法与传统的 LDPC 重建方法[14-15]进行对比。

5.1　算法有效性验证

5.1.1　双对角线稀疏校验矩阵重建

本节设定 LDPC 的稀疏校验矩阵为协议 IEEE 802.11n 中定义的 **LDPC**(648, 324)，其稀疏校验矩阵具有明显的双对角线，如图 1(a)所示。设截获的码字个数为 5 000 个，误码率为 0.001 5，高斯消元迭代次数为 20 次，本文算法重建结果如图 1(b)所示。由于重建稀疏矩阵行顺序是随机的，为了方便与图 1(a)对比，将图 1(b)的行顺序参照图 1(a)的行顺序重新排列，得到结果图 1(c)。为了与本文算法进行对比，在同等条件下，传统算法也对该稀疏校验矩阵进行了重建，图 1(d)为传统算法重建的非稀疏校验矩阵，而图 1(e)为稀疏化后的结果。



(a) IEEE 802.11n **LDPC**(648, 324)稀疏校验矩阵

(b) 本文算法重建结果

(c) 重新排列图1(b)行顺序结果

(d) 传统算法重建的非稀疏校验矩阵

(e) 传统算法稀疏化后的校验矩阵

图 1　2 种算法针对双对角线稀疏校验矩阵的重建过程

从图 1 可以看出，当误码率为 0.0015 时，本文算法直接重建出 306 个稀疏校验向量，即重建率达到 94.44%。与原始稀疏校验矩阵对比发现，重建的稀疏校验向量与原始稀疏校验矩阵中向量完全一致（重新排列得到的图 1(c)与图 1(a)是一致的），这说明本文算法能够很好地重建稀疏矩阵；相反，传统重建算法在误码率为 0.0015 时，仅得到 8 个非稀疏的校验向量，经过稀疏化处理后，得到 6 个稀疏校验向量，重建率仅为 1.85%，这说明传统算法对误码的稳健性较差。

5.1.2　非双对角线稀疏校验矩阵重建

本节设 LDPC 为 **QC-LDPC**(600, 300)，该稀疏校验矩阵不再具有双对角线形式，如图 2(a)所示，生成的 LDPC 码字个数为 5 000，信道误码率为 0.001，高斯迭代消元次数为 20。本文算法重建结果如图 2(b)所示，为了方便与原始稀疏校验矩阵对比，同样将图 2(b)的行顺序参照图 2(a)的行顺序进行重排列，得到图 2(c)结果。在同等条件下，利用传统重建方法首先得到如图 2(d)所示的非稀疏校验矩阵，然后稀疏化处理得到如图 2(e)的结果。

从图 2 可以看出，本文算法有效恢复出了 295 个稀疏校验向量，虽然实际的稀疏校验向量个数为 300 个，但由于该稀疏校验矩阵是非满秩矩阵，其秩为 295，故本文算法重建的 LDPC 稀疏校验矩阵和原始矩阵等效；同时由图 2(c)可知，本文重建的稀疏校验向量和原始的稀疏校验向量是完全一致的，这进一步说明本文算法能够较好地重建出稀疏矩阵。虽然传统的重建方法能够得到 295 个非稀疏的校验向量，但是通过多次行消元稀疏，得到的结果仍然不够稀疏，同时存在大量 4 环（稀疏矩阵的设计应避免出现 4 环），这说明传统算法针对非双对角形式的稀疏校验矩阵重建效果不佳。

5.2　算法容错性验证

本节主要考察消元迭代次数、截获码字个数、

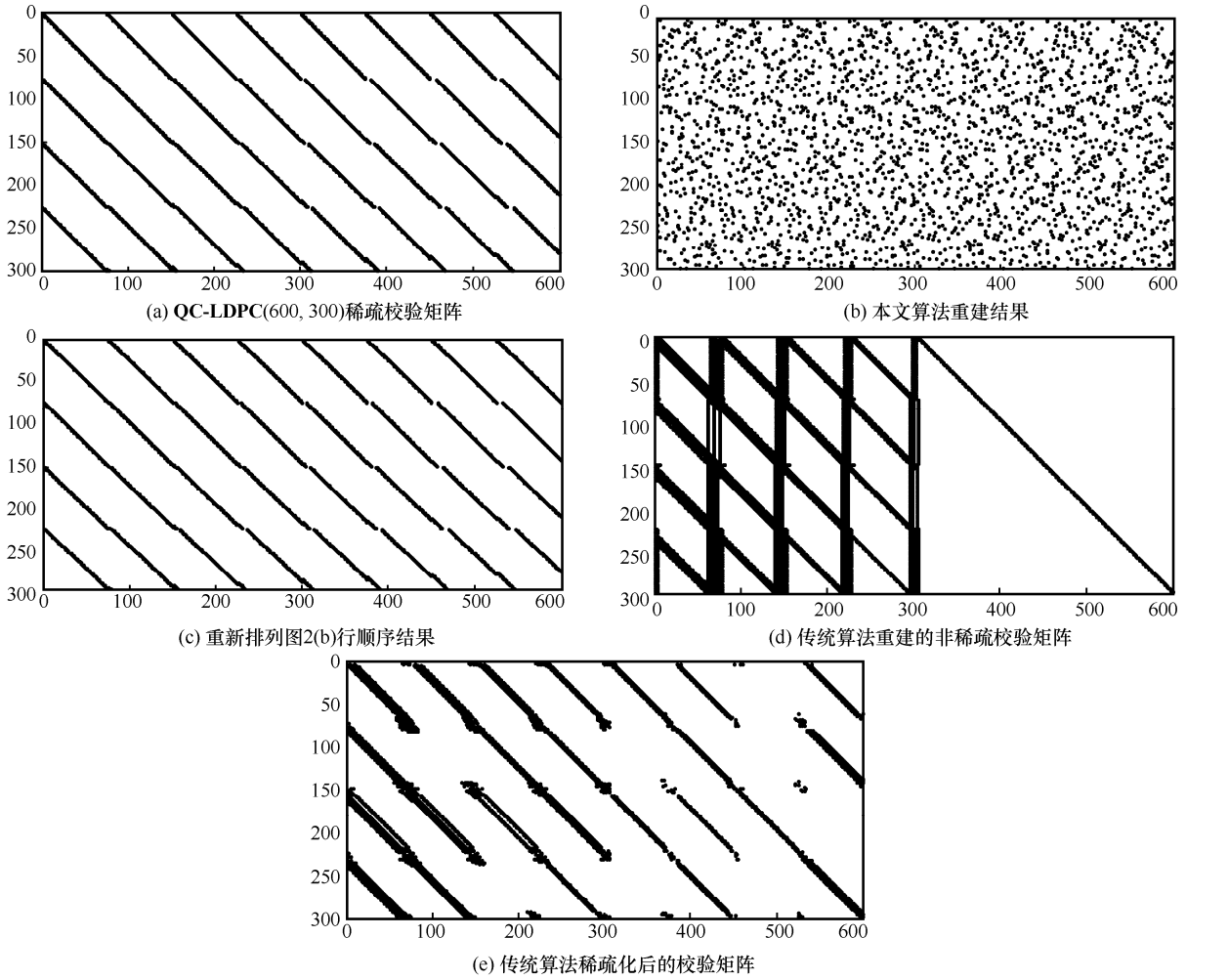


图 2　2 种算法针对非双对角线形式的 LDPC 重建过程

码长以及码率几个因素对算法性能的影响，记录在某一因素、不同误码率下的重建率。

5.2.1　迭代次数的影响

仿真设定 LDPC 稀疏校验矩阵为协议 IEEE 802.11e 中定义的 **LDPC**(576, 288)，设截获的码字个数为 1 500 个，在高斯消元过程中，迭代次数分别设定 1 次、5 次、10 次、15 次、20 次；设定误码率范围为 0～0.005，取值间隔为 0.000 25，统计在不同迭代次数以及不同误码率下，LDPC 稀疏校验矩阵重建率，结果如图 3 所示。

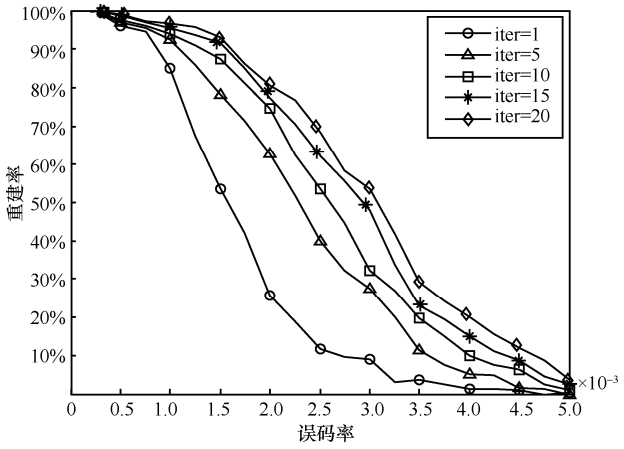


图 3　不同迭代次数对算法的影响

从图 3 可以看出，增加迭代次数可以有效提高 LDPC 稀疏校验矩阵的重建率。主要原因在于，在高斯消元过程中，迭代次数一旦增加，满足同一校验关系的码字被抽到的概率就会增加，此时重建出的稀疏校验向量也会相应增多。此外，本文提出的重建算法具有较好的容错性，在误码率为 0.001 的条件下，稀疏校验矩阵重建率能达到 95%以上。

5.2.2　截获码字数目影响

同样设定 LDPC 稀疏校验矩阵为协议 IEEE 802.11e 中的 **LDPC**(576, 288)，设信道误码率为 0.000 5、0.001、0.001 5、0.002、0.002 5，高斯消元迭代次数为 20 次，截获的 LDPC 码字个数范围为 500～2 000，取值为间隔 100，统计 5 种信道误码率情况下，不同截获码字个数对应的稀疏矩阵重建率，结果如图 4 所示。

从图 4 可以看出，增加码字个数能够有效增大稀疏校验矩阵的重建率，原因在于，当码字个数增多时，满足同一校验关系的码字会随之增加，所以在迭代消元过程中，随机抽取到这一类码字的可能性就会增大；当码字增加后，计算的判决门限会更加准确，对疑似校验向量的误判就会减少。

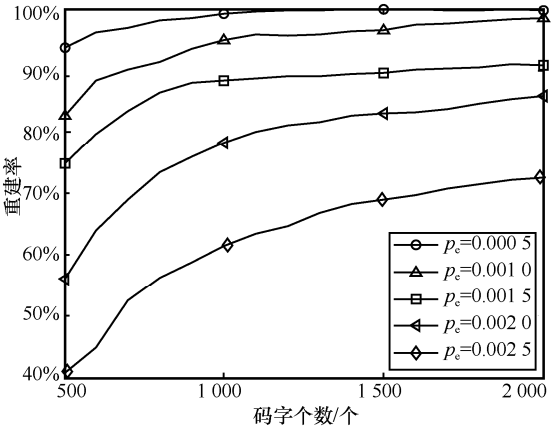


图 4　不同码字个数对算法的影响

5.2.3　码长和码率的影响

仿真设定 LDPC 码率为 1/2 和 2/3，每种码率对应的码长为 576、648 以及 768，选取的稀疏校验矩阵都为 IEEE 802.11 协议中定义的校验矩阵，设截获的码字个数为 2 000，误码率范围为 0～0.004，取值间隔为 0.000 25，高斯消元迭代次数为 20 次，统计在不同 LDPC 编码类型下，不同误码率对应的稀疏校验矩阵重建率，结果如图 5 所示。

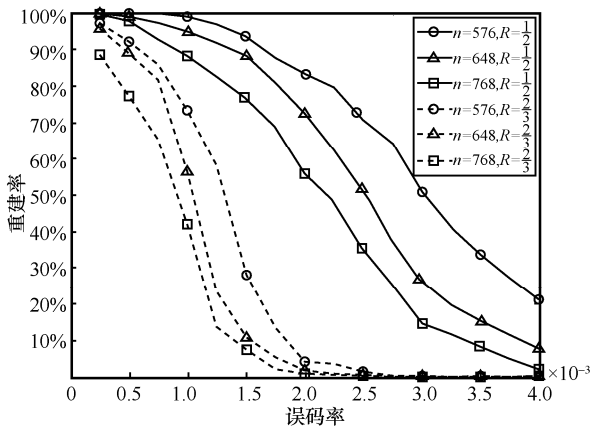


图 5　不同码长和码率对算法影响

从图 5 结果来看，在同一码长条件下，算法重建率随码率的增加而降低；在同一码率条件下，重建率随码长的增加而下降。主要原因在于，首先，当码长增加时，抽取参与消元的比特相应增加，此时误码会随比特消元过程而扩散，参与消元的比特越多，扩散情况越严重；其次，当码率增加时，稀疏校验矩阵中向量码重增加，此时能够抽取到包含

校验节点的数据比特的概率会降低，同时码重增加后，疑似校验向量的误判概率也会增加，这 2 个因素综合起来会导致算法性能降低。

5.3　与传统算法比较

本文算法与传统算法比较时，选取 IEEE 802.11 协议中 **LDPC**(576, 288)（码率为 1/2）、**LDPC**(648, 324)（码率为 1/2）以及 **LDPC**(648, 432)（码率为 2/3）。设截获的码字个数为 2 000，信道误码率范围为 0～0.005，取值间隔为 0.000 25，统计不同误码率下 2 种算法的重建率，结果如图 6 所示。

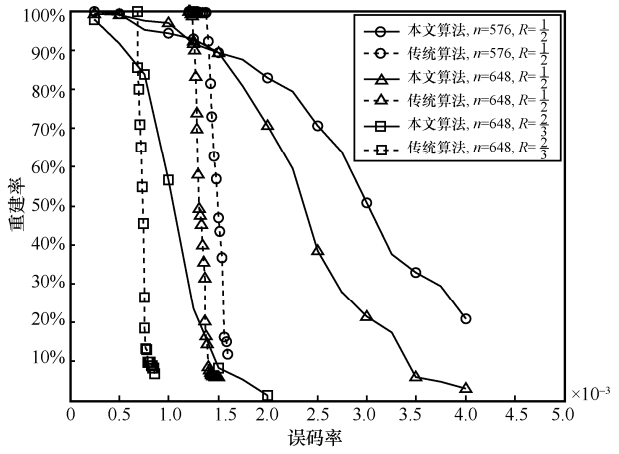


图 6　本文算法与传统算法重建率对比

从图 6 可以看出，本文算法重建率明显优于传统算法。主要原因在于，首先，本文算法通过随机抽取码字中部分比特进行消元，实际参与消元的比特数目远小于码字长度，所以本文算法对于信道噪声具有更强的稳健性；其次，传统算法的重建性能较差，信道误码率稍有增加，算法重建率就急剧下降。由此可见，本文算法在高误码率下具有更好的工程实用性。

6　结束语

本文基于 LDPC 稀疏校验矩阵的特征，提出了可直接重建稀疏校验矩阵的算法。算法首先通过多次随机抽取码字中部分比特序列进行高斯消元，当抽取的比特序列包含稀疏校验节点时，从消元的结果中可直接获取 LDPC 稀疏校验向量；其次分析了在有误码条件下，码字校验关系成立的统计特性，基于最小错误判决准则，完成疑似校验向量的判定，最终完成 LDPC 稀疏校验矩阵的重建。与传统重建算法相比，本文算法不再需要单独进行稀疏化步骤，不仅具有较好的容错性能，而且对于双对角线与非双对角线形式的稀疏校验矩阵都具有很好的通用性。

参考文献：

[1] 吴昭军, 张立民, 钟兆根, 等. 基于平均余弦符合度下的本原 BCH 码盲识别[J]. 通信学报, 2020, 41(1): 15-24.
WU Z J, ZHANG L M, ZHONG Z G, et al. Blind recognition of primitive BCH code based on average cosine conformity[J]. Journal on Communications, 2020, 41(1): 15-24.

[2] 刘杰, 张立民, 钟兆根. 基于二元域等效的 RS 码编码参数盲识别[J]. 电子学报, 2018, 46(12): 2888-2895.
LIU J, ZHANG L M, ZHONG Z G. Blind parameter identification of RS code based on binary field equivalence[J]. Acta Electronica Sinica, 2018, 46(12): 2888-2895.

[3] 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.

[4] 钟兆根, 吴昭军, 张立民, 等. 基于对数符合度下的RSC码识别[J]. 通信学报, 2018, 39(10): 79-86.
ZHONG Z G, WU Z J, ZHANG L M, et al. Blind recognition of RSC based on logarithmic conformity[J]. Journal on Communications, 2018, 39(10): 79-86.

[5] 吴昭军, 张立民, 钟兆根. 基于最大序列相关性的 Turbo 码交织器识别[J]. 航空学报, 2019, 40(6): 262-273.
WU Z J, ZHANG L M, ZHONG Z G. Blind recognition of interleaver for Turbo codes based on maximum sequence correlation[J]. Acta Aeronautica et Astronautica Sinica, 2019, 40(6): 262-273.

[6] 陈泽亮, 李静, 彭华, 等. 利用 Gibbs 采样进行优化的 Turbo 码交织器识别[J]. 电子学报, 2018, 46(1): 15-23.
CHEN Z L, LI J, PENG H, et al. An optimization method using gibbs sampler for turbo-code interleaver identification[J]. Acta Electronica Sinica, 2018, 46(1):15-23.

[7] 刘骏, 李静, 于沛东. 一种 Turbo 码随机交织器的迭代估计方法[J]. 通信学报, 2015, 36(6): 201-206.
LIU J, LI J, YU P D. Iterative estimation method for random interleaver of Turbo codes[J]. Journal on Communications, 2015, 36(6): 201-206.

[8] 刘杰, 张立民, 占超. 基于矩阵分析的线性分组码盲识别[J].系统工程与电子技术, 2017, 39(2): 404-409.
LIU J, ZHANG L M, ZHAN C. Blind recognition of linear block codes based on matrix analysis[J]. Systems Engineering and Electronics, 2017, 39(2): 404-409.

[9] 姚智刚, 解辉, 韩壮志, 等. 基于分段 Walsh-hadamard 变换的卷积码盲重构算法[J]. 电子与信息学报, 2019, 41(9): 2047-2054.
YAO Z G, XIE H, HAN Z Z, et al. Blind reconstruction of convolutional code based on partitioned Walsh-Hadamard transform[J]. Journal of Electronics & Information Technology, 2019, 41(9): 2047-2054.

[10] XIA T, WU H C. Novel blind identification of LDPC codes using average LLR of syndrome a posteriori probability[J]. IEEE Transactions on Signal Processing, 2014, 62(3): 632-640.

[11] XIA T, WU H C. Blind identification of nonbinary LDPC codes using

average LLR of syndrome a posteriori probability[J]. IEEE Communications Letters, 2013, 17(7): 1301-1304.

[12] 包昕, 王达, 刘婉月. 利用软解调序列的 LDPC 码闭集识别方法[J]. 电讯技术, 2015, 55(1): 55-60.
BAO X, WANG D, LIU W Y. A finite set recognition algorithm of LDPC coding by using soft-demodulation sequence[J]. Telecommunication Engineering, 2015, 55(1): 55-60.

[13] 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.

[14] 包昕, 周磊硐, 何可, 等. 误码条件下的 LDPC 码盲识别算法[J].西安电子科技大学学报, 2015, 49(12): 53-58.
BAO X, ZHOU L K, HE K, et al. A recognition algorithm for LDPC codes in a noisy environment[J]. Journal of Xi'an Jiaotong University, 2015, 49(12): 53-58.

[15] 包昕, 周磊硐, 何可, 等. LDPC 码稀疏校验矩阵的重建方法[J].电子科技大学学报, 2016, 45(2): 191-196.
BAO X, ZHOU L K, HE K, et al. A method of restructuring LDPC parity-check matrix[J]. Journal of University of Electronic Science and Technology of China, 2016, 45(2): 191-196.

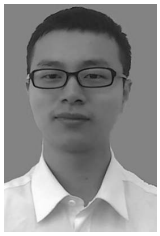
[16] 于沛东, 彭华, 巩克现, 等. 基于寻找小重量码字算法的 LDPC 码开集识别[J]. 通信学报, 2017,38(6):108-117.
YU P D, PENG H, GONG K X, et al. LDPC code reconstruction based on algorithm of finding low weight code-words[J]. Journal on Communications, 2017, 38(6): 108-117.

[17] CANTEAUT A, CHABAUD F. A new algorithm for finding minimum-weight words in a linear code: application to primitive narrow-sense BCH codes of length 511[J]. IEEE Transactions on Information Theory, 1998, 44(1): 367-378.

[18] 陈泽亮, 彭华, 巩克现, 等. 误码条件下 LDPC 码参数的盲估计[J]. 电子学报, 2018, 46(3): 652-658.
CHEN Z L, PENG H, GONG K X, et al. A method for blind recognition of LDPC codes in a noisy environment[J]. Acta Electronica Sinica, 2018, 46(3): 652-658.

[19] 刘玉君. 信道编码（第三版）[M]. 郑州: 河南科学技术出版社, 2006.
LIU Y J. Channel codes (the third edition)[M]. Zhengzhou: Henan Science and Technology Press, 2006.

[作者简介]



吴昭军（1992− ），男，四川遂宁人，海军航空大学博士生，主要研究方向为信道编码盲识别。

张立民（1966− ），男，辽宁开原人，博士，海军航空大学教授，主要研究方向为卫星信号处理及应用。

钟兆根（1984− ），男，江西南昌人，博士，海军航空大学副教授，主要研究方向为通信信号盲分离与统计信号处理。

刘仁鑫（1995− ），男，山东临沂人，海军航空大学硕士生，主要研究方向为信道编码盲识别。