

LDPC 码的低复杂度译码算法研究^{*})

邓勇强 朱光喜 刘文明

(华中科技大学电子与信息工程系 武汉 430074)

摘要 本文在对 LDPC 码的译码算法分析的基础上,针对校验矩阵中含有的环对译码算法的影响,提出了一种基于置信传播算法基础上的低复杂度译码算法。该算法通过及时切断消息在环上的重传回路,可消除因校验矩阵中的环回传原始信息对译码造成的影响,保证优质的原始信息能尽可能地传播到其能传播的节点,从而提升了 LDPC 码的译码性能。仿真实验表明,该算法具有低于传统算法的复杂度,在低信噪比的信道中可以获得相当于传统算法的性能,在良好的信道条件下可以取得比传统算法更优异的性能。

关键词 低密度校验码,置信传播算法,环,因子图,迭代译码

A Study on Low Complexity Decoding Algorithm of LDPC Codes

DENG Yong-Qiang ZHU Guang-Xi LIU Wen-Ming

(Department of Electronics & Information Engineering, Huazhong University of Science & Technology, Wuhan 430074)

Abstract In this paper, the decoding algorithm of low-density parity-check (LDPC) codes is analyzed, and a new decoding algorithm based on the belief propagation (BP) algorithm to eliminate the influence of cycles in the factor graph is proposed. In the traditional BP algorithm, the cycles of factor graph will send message back to its source, and this will decrease the decoding performance. The new algorithm records each cycle's path and length of each node, and cuts off the path by which message is propagated when the message will come back. It can advance the decoding performance by protect the message of good quality be propagated as widely as possible. Our results of simulation show that the performance of new algorithm is not worse than that of traditional BP algorithm in the low SNR channel and the new algorithm significantly outperform traditional BP algorithm in good channel condition.

Keywords Low-density parity-check (LDPC) codes, Belief propagation (BP) algorithm, Cycle, Factor graph, Iterative decoding

1 前言

LDPC (Low Density Parity-Check, 又称为 Gallager 码) 码是 Robert G. Gallager 于 1962 年提出的编码^[1]。然而因为各种原因, LDPC 码沉睡了 30 来年, 直到 1993 年具有划时代意义的 Turbo 码提出以后, 在 Turbo 码的启发下, 才有 D. J. Mackay, M. Neal 和 N. Wiberg 等人对 LDPC 码重新进行了研究, 他们发现 LDPC 码与 Turbo 一样具有逼近 Shannon 限的性能。自 1999 年以来, 这方面的研究文献开始大量出现, LDPC 码已成为编码领域中的一个研究热点^[2]。

信道编码的译码算法是决定编码性能和应用前景的一个重要因素。LDPC 码的译码算法主要是基于编码因子图结构的迭代置信传播算法 (Belief propagation (BP) algorithm)^[3], 该算法采用信息传递进行迭代译码。在每一次迭代过程中, 所有的信息节点将其估计的信息传递给在因子图上与它相连的校验节点; 同样校验节点也将其估算的解码信息传递给与其相连的信息节点^[4]。该算法在无环因子图上可以实现最优译码。对于有环因子图, 置信传播算法虽然仍可获得较好的译码性能, 但因环的影响, 无法精确计算译码信息以保证译码最优性。本文给出了一种消除校验码因子图中环影响的优化

的置信传播算法, 在减少算法实际计算次数即算法复杂度的情况, 可获得与原算法相当或更优越的性能。

2 置信传播算法

因子图是一个表示因式分解结构的二部图 (bipartite)^[5]。在对 LDPC 码的研究中, 常常使用因子图来表示码字变量节点与校验节点之间的关系, 它的译码算法就可以使用因子图, 直观具体地描述其实过程。

LDPC 码的译码算法通常被称为置信传播算法或和积算法。在置信传播算法中, 与译码有关的消息沿着边在节点之间进行传播, 如图 1 的因子图所示。图中 $\{z_i; i=1, 2, \dots, M\}$ 表示校验点, $\{x_j; j=1, 2, \dots, N\}$ 表示变量节点, 其中 M 和 N 分别表示校验矩阵 $H = (h_{ij})_{M \times N}$ 的行和列。令 $M(j) = \{i; h_{ij} = 1\}$ 表示与变量节点 x_j 相连的校验节点集合, $N(i) = \{j; h_{ij} = 1\}$ 表示与校验节点 z_i 相连的变量节点的集合。初始消息 f_j^i 是信道转移函数, R_{ij}^j 是校验节点 z_i 向变量节点 x_j 传递的校验信息, 表示 z_i 根据其他变量节点 $\{x_k; k \neq j, k \in N(i)\}$ 的当前状态向 x_j 宣称“ $x_j = a$ 使 z_i 满足”的可信度; Q_{ji}^i 是 x_j 向 z_i 传递的变量信息, 表示 x_j 根据其他校验节点 $\{z_k; k \neq i, k \in M(j)\}$ 以及 f_j^i 向 z_i 宣称“ $x_j = a$ ”的可信度。

^{*}) 国家自然科学基金重大项目“未来移动通信系统基础理论与技术研究”(No. 60496315)、国家高技术研究发展计划 (863 计划) (No. 2003AA12331005)。邓勇强 博士研究生; 朱光喜 教授, 博导。

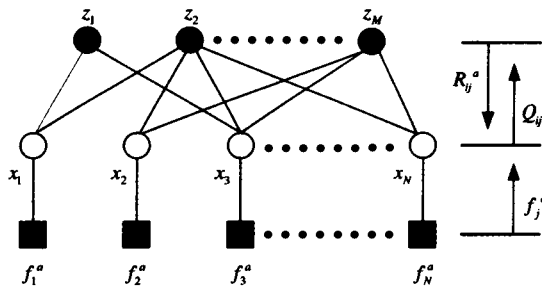


图1 置信传播译码的因子图表示

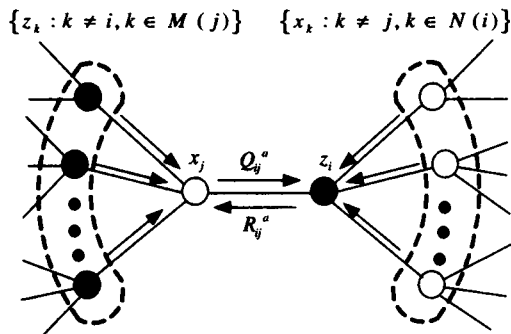


图2 消息传递修正规则

图2是描述消息传递和修正规则的局部因子图。置信传播算法将在所有的变量节点与校验节点之间计算、传递及不断修正其对译码的估计信息,其主要过程如下:

- 初始化:计算初始消息 $\{f_j^a\}$ 和初始化 $\{R_{ij}^a\}$;
- 迭代消息传递和修正:

阶段1:各个 f_j 和校验节点 z_i 向其相连的变量节点 x_j 分别传递消息 f_j^a 和 R_{ij}^a ,各个变量节点 x_j 根据接收到的消息修正 Q_{ij}^a ;

阶段2:各个变量节点 x_j 向其相连的所有校验节点 z_i 传递已经更新的消息 Q_{ij}^a ,各个校验节点 z_i 根据接收消息修正 R_{ij}^a ;

• 试译码: z_i 和 x_j 节点根据所有接收消息分别计算最终消息 R_{ij}^a 和 Q_{ij}^a ,并对 $\{x\}$ 值进行判断,然后试译码。若满足 $H \cdot x^T = 0$,则译码成功,退出;否则,继续进行消息迭代和修正;

• 迭代终止:迭代次数达到设定次数上限,对 $\{x\}$ 值进行判断,然后试译码。若满足 $H \cdot x^T = 0$,则译码成功,退出;否则,译码失败,退出。

根据上述关于置信传播译码算法的描述,在概率测度下,对于给定的信道, $\{x\}$ 和 $\{y\}$ 分别表示信道的输入和输出, a 表示信号的取值,在二元信道上, $a \in \{0, 1\}$ 。

初始概率消息:

$$f_j^a = P(x_j = a | y_j) = \alpha_j f(y_j | x_j = a) \quad (1)$$

其中, α_j 是归一化常数, $f_j^0 + f_j^1 = 1$ 保证。

变量消息:

$$Q_{ij}^a = P(x_j = a | y_j, \{z_k : k \in M(j) \setminus i\}) \quad (2)$$

校验消息:

$$R_{ij}^a = P(z_i | x_j = a) \quad (3)$$

R_{ij}^a 值的初始化为1。

变量消息 Q_{ij}^a 和校验消息 R_{ij}^a 的修正公式推导如下:

$$Q_{ij}^a = P(x_j = a | y_j, \{z_k : k \in M(j) \setminus i\})$$

$$\begin{aligned} &= \frac{P(x_j = a, y_j, \{z_k : k \in M(j) \setminus i\})}{P(y_j, \{z_k : k \in M(j) \setminus i\})} \\ &= \frac{P(x_j = a | y_j) P(\{z_k : k \in M(j) \setminus i\} | y_j, x_j = a)}{P(y_j, \{z_k : k \in M(j) \setminus i\})} \\ &= \frac{P(x_j = a | y_j) P(\{z_k : k \in M(j) \setminus i\} | x_j = a)}{P(y_j, \{z_k : k \in M(j) \setminus i\})} \\ &= \frac{P(x_j = a | y_j) \prod_{k \in M(j) \setminus i} P(z_k | x_j = a)}{P(y_j, \{z_k : k \in M(j) \setminus i\})} \\ &= \alpha_{ij} f_j^a \prod_{k \in M(j) \setminus i} R_{kj}^a \end{aligned} \quad (4)$$

$$\begin{aligned} R_{ij}^a &= P(z_i | x_j = a) = \sum_{x_i : x_j = a} P(z_i, x | x_j = a) \\ &= \sum_{x_i : x_j = a} P(z_i | x_j = a, x) P(x | x_j = a) \\ &= \sum_{x_i : x_j = a} P(z_i | x) P(x | x_j = a) \\ &= \sum_{x_i : x_j = a} P(z_i | x) \prod_{k \in N(i) \setminus j} P(x_k) \\ &= \sum_{x_i : x_j = a} P(z_i | x) \prod_{k \in N(i) \setminus j} Q_{ik}^a \end{aligned} \quad (5)$$

其中,式(4)中的 α_{ij} 为归一化常数, R_{ij}^a 是从上一次迭代中传递来的校验消息,首次迭代时为初始值1。式(5)中, $P(x_k)$ 是指求和组合 x 中各分量 x_k 取相应值的概率。用于逐符号判决的最终变量消息为:

$$Q_j^a = \alpha \cdot f_j^a \prod_{i \in M(j)} R_{ij}^a \quad (6)$$

其中, α 是归一化因子,使得 $Q_j^0 + Q_j^1 = 1$ 。相应的判决准则是:如果 $Q_j^0 > 0.5$,则 $\hat{x}_j = 0$,否则 $\hat{x}_j = 1$ 。

3 消除环影响的置信传播算法

LDPC码的置信传播算法通过迭代计算校验节点与变量节点的译码信息,在校验矩阵的Tanner图没有环的情况下可以获得很优越的译码性能。但对于LDPC的中短码而言,没有环的情况基本上是不存在的。特别是对于列重不小于3的规则码,我们可以证明其Tanner图中必然存在环。当码长不是特别大时,就会存在大量的短长度的环。对于这种情况下的环的分布,前人已有研究。Gallager在文[6]中给出了规则码的分组长度 N 、环的长度 g 与校验矩阵中的列重 d 和行重 k 之间的关系:

1)对于 $g = 4m + 2$,分组长度 N 满足:

$$N \geq \sum_{i=1}^{m+1} S_i \quad (7)$$

$$S_i = \begin{cases} 1 & i=1 \\ d(d-1)^{(i-2)}(k-1)^{(i-1)} & i \geq 2 \end{cases}$$

2)对于 $g = 4m$,分组长度 N 满足:

$$N \geq \sum_{i=1}^m L_i \quad (8)$$

$$L_i = k(d-1)^{(i-1)}(k-1)^{(i-1)} \quad \forall i$$

根据Gallager给出的结论,我们可以得到在给定行重量为3、列重量为6、码率为1/2的LDPC码其分组长度和校验矩阵所具有的环长度的下限之间的关系,如图3所示。

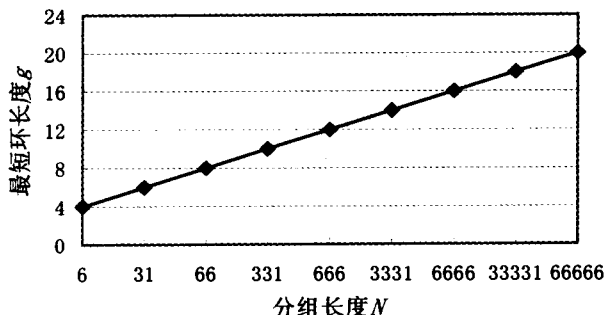


图3 行重为3、列重为6的LDPC码分组长度与最短环长度关系

从图3中我们可以看到,在一定的分组长度 N 情况下,最短环的长度 g 远远小于 N 。即使在分组长度达到66666的长码时,其最短环长也只有20。因而,对于一般长度,特别是中短长度的LDPC码,在译码时不考虑其校验矩阵的Tanner图中存在环的情况是不现实的。

Tanner图中存在环对译码的影响主要是因为多次的迭代运算后,由一个校验节点或变量节点发出去的译码信息通过环又回来了,并再次参与迭代计算,我们把这种由自己发出去、经过多次迭代后又返回的信息称为重复信息,那种从信道传输中带来和经过没有重复信息参与的计算得到的信息称为原始信息。虽然重复信息并不是简单地等于发出去的信息,当初发出去的信息经过多次的迭代运算后,返回来的信息与初始信息相比已经有了很大的变化,但是当这些信息与从其他节点传来的新信息一起参与计算,由于期间的部分信息已经计算过,故其必然会对计算结果产生干扰,影响了该节点挖掘新信息中隐藏的有用部分。因此,尽可能地减少重复信息对迭代运算的影响,可以有效地增加节点对其邻接点传来信息的充分利用,减少迭代运算的次数,提高译码的效率和效果。

我们设计了一种译码算法,可以有效地减少环对译码的影响。算法的主要思想是在译码的过程中,阻止重复信息返回原节点参与计算,使得原始信息获得充分的传播,将其未受干扰的信息传递到尽可能多的地方;在所有的原始信息传递到它能到达的节点后,重复信息才重新参与迭代计算,直到译码成功或是迭代次数达到上限。

算法的具体步骤是:

(1)先对校验矩阵进行扫描,找出以每个节点的为出发点的每一条环路,记录每一条环路的起始点及其长度。这里,起始点是指该环路上与该节点直接相连的节点。

(2)开始迭代运算。对每一个节点,在接受相连节点的信息时,比较该邻节点所在的环路长度与当前迭代的次数。若环路长度小于迭代次数,则表示由该邻节点发出去的信息通过环路后又回到了起点,则不接收该邻节点发来的信息,除该邻节点以外的其他节点传来的信息参与运算。

(3)当所有的节点都停止接收信息,即所有的原始信息都传递给环路上的所有节点以后,每个节点恢复接收所有邻节点的信息,继续迭代运算。

(4)译码成功,退出;或是达到迭代次数上限,译码失败,退出。

在实现算法的过程中,由于对所有节点而言,它参与传递信息的过程是双向,即既从其相连节点接收信息,又向其相连节点发送信息;而且,对校验节点与变量节点而言,在这个迭代运算的过程中,其地位是平等的。由于整个译码的迭代运算从校验节点开始的,当我们停止接收环回的校验节点的消息后,校验节点向其相邻的变量节点发送的消息就不会有重复消息了。因此,我们采用只对数量较少的校验节点进行操作,就可以达到对所有节点都进行操作的作用。

在上面的第2步和第3步的过程中,同样也可以进行译码判断。若满足 $H \cdot x^T = 0$,则译码成功,退出计算。

4 仿真实验及性能分析

在AWGN信道下,采用BPSK调制,我们对本文中所述的算法构造的LDPC码进行了计算机仿真,其性能及与随机构造算法的比较如图7所示。仿真参数为:码长 $N=200$,列

重量 $d=3$,行重量 $k=4$,译码迭代次数的上限为60次。

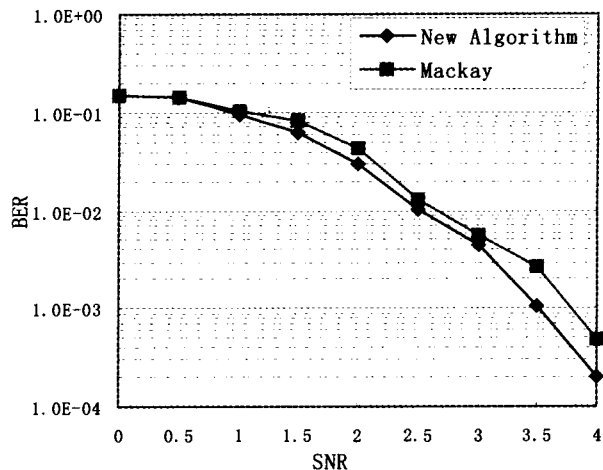


图4 两种译码算法的性能比较

图中对比的是Mackay实现的传统的置信传播译码算法,除译码算法以外其他的条件全部一致。从图中可以看出,在中短码长的情况下,经过改进的译码算法在信道情况比较差的条件下,即信噪比较小,性能与传统的置信传播算法大致相当;在信道条件比较好的情况下,改进算法具有比传统算法更好的性能。

改进算法的更好性能还可以从译码迭代次数直方图和分析。图5所示的是改进算法与传统算法的译码迭代次数的直方图对比情况,实验条件与限制同上文的一样。直方图的横轴表示一组译码迭代的次数,纵轴表示该迭代次数出现的次数。图5(a)和(b)分别表示在信道的信噪比为2.4dB时,用改进算法与传统算法各计算了800组数据所得到的迭代次数的分布,图5(c)和(d)分别表示在信噪比为2.6dB时,用改进算法与传统算法各计算1000组数据时迭代次数的分布情况。从图中可以看出,在相同信噪比情况下,改进算法译码时的迭代次数分布更为集中;当信道条件变好时,两种算法的迭代次数都有集中的趋势,但改进算法表现得更为明显些。图中右边的竖条表示译码过程中出现迭代次数达到设置的上限60次时的情况。一般而言,当译码过程中迭代次数达到设置的上限时,此次译码都会出现误码。我们注意到,在信噪比相同的情况下,改进的算法出现迭代次数达上限的情况要少于传统算法。而且,随着信道状况的改善,改进算法的优势也表现得更为明显。

从图5的直方图对比中,我们可以大致看出,在图中所示的信道条件下的实验中,改进算法总的译码迭代次数要多于传统算法的译码迭代次数,统计的数据也证实了这一点。但并不是在所有的情况下总是这样的。表1给出了改进算法与传统算法在不同信噪比的条件下,译码所需的平均迭代次数的对比情况。当信道条件比较差时,改进算法比传统算法具有更少的平均迭代次数。随着信道信噪比的增大,两种算法的迭代次数都在减少,但传统算法减少速度更快一些。在图中所示的数据中,当信噪比大于2.4dB时,传统算法比改进算法具有更少的迭代次数。

这种改进算法迭代次数随信道状况的变好的收敛速度慢于传统算法的情况之所以会出现,是因为在低信噪比的情况下,数据的误码比较严重。因此在译码过程中,一般需要经过多次迭代运算才能求解,迭代次数一般多于校验矩阵中存在

的环的长度。改进的算法通过切断环对传出信息的重复计算避免了对原始优质信息的破坏,因而译码所需要的迭代次数要更少一些。在高信噪比的情况下,由于误码比较少,往往只需要很少的迭代次数就可以译码成功。在这种情况下,通过

切断环以避免重复计算原始信息,同时也减缓了有用信息在校验节点与变量节点之间的传播速度,因此成功译码所需要的迭代次数就会更多一些。

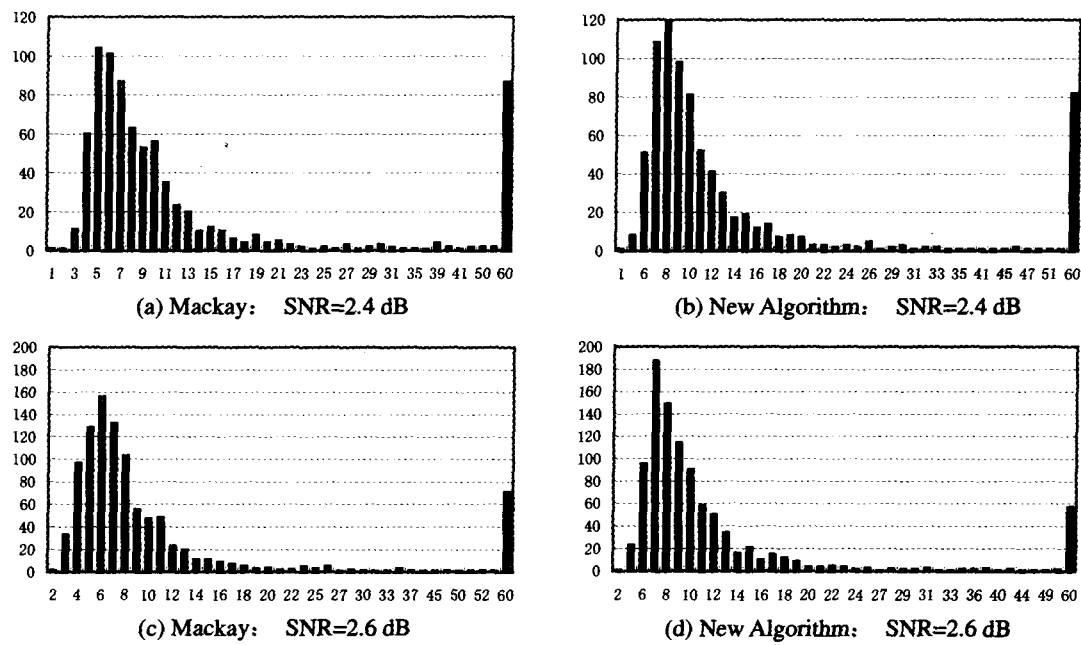


图 5 译码迭代次数直方图

表 1 两种译码算法迭代次数对比表

dB	New Algorithm	Mackay Algorithm
1.6	29.2270	32.6958
2.2	16.3525	17.0357
2.4	15.0938	14.9586
2.6	13.6560	12.0150
2.8	11.6429	10.0932

但是,更少的迭代次数并不一定意味着更少的计算量。表 2 给出了两种译码算法在不同信道条件下平均每次迭代所需要的对式(4)和(5)的计算次数。实验中使用的 LDPC 码的校验矩阵大小为 200×150 ,我们每次迭代运算过程中计算一次 Q_i 或 R_i 称为一次计算。改进算法由于使用了切断环的方式,因而平均每次迭代的计算次数要少于传统的算法,并且随着信道状况的改善,其计算次数越来越少。而传统算法每一次迭代的计算次数只与校验矩阵的大小及其行重量和列重量有关。

表 2 两种译码算法每次迭代计算次数比较

dB	New Algorithm	Mackay Algorithm
1.6	1179.8	1224
2.2	1170.1	
2.4	1145.0	
2.6	1132.1	
2.8	1116.7	

通过以上直方图与表 1 和表 2 的比较,我们可以得知,就复杂度和计算量而言,改进算法在所有的信道情况下都优于传统的算法,其性能在低信噪比的情况下与传统算法相当;当信道条件变好时,改进算法具有比传统算法更优越的性能。

结束语 置信传播算法是 LDPC 码译码的基本算法,该算法在表现优异性能的同时也具有一定的局限性,本文即是在对该算法分析的基础上,针对校验矩阵中含有的环对译码算法的影响,提出了一种改进算法。改进算法通过扫描校验矩阵,对所有节点找出它自身为起点的环,并在迭代译码的过程中,根据消息传递的进展及时切断消息在环上的重传回路,这样就可消除因校验矩阵中的环回传原始信息对译码造成的影响,保证优质的原始信息能尽可能地传播到其能传播的节点,从而提升了 LDPC 码的译码性能。仿真实验表明,该算法在计算复杂度低于传统算法的情况下,在低信噪比的信道中具有相当于传统算法的性能,在良好的信道条件下可以取得比传统算法更优异的性能。

参考文献

- Gallager R G. Low-Density Parity-Check Codes. IRE Trans. Inform. Theory, 1962, 8: 21~28
- Moura J M F, Lu Jin, Zhang Haotian. Structured Low Density Parity-Check Codes. IEEE Signal Processing Magazine, 2004; 42~55
- Mackay D J C. Good Error-Correcting Codes Based on Very Sparse Matrices. IEEE Trans. Inform. Theory, 1999, 45: 399~431
- Kaschischang F R, Frey B J, Loeliger H-A. Factor Graphs and the Sum-Product Algorithm. IEEE Trans. Inform. Theory, 2001, 47: 498~519
- 贺玉成. 基于图模型的低密度校验码理论及应用研究[D]. 西安: 西安电子科技大学, 2002
- Gallager R G. Low-Density Parity-Check Codes. Cambridge, MA: MIT Press, 1963. 87~88