Density Evolution for Asymmetric Memoryless Channels[1]

Chih-Chun Wang, Sanjeev R. Kulkarni, H. Vincent Poor

Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, U.S.A.
Phone: (+1) 609 258 1831, Fax: (+1) 609 258 3745
Email: {chihw, kulkarni, poor}@princeton.edu

Abstract: *Density evolution method for general memoryless, symbol-dependent channels (e.g., z-channels, binary asymmetric channels, etc.) is investigated in this paper. Equiprobable codeword averaging is used to circumvent the symbol dependence. A new iterative formula, several underpinning theorems, stability results, and simulations are provided.*

Keywords: Low-density parity-check codes, density evolution, message passing algorithms, symbol-dependent channels, asymmetric channels.

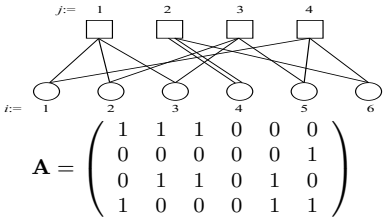
1. Introduction

After the rediscovery of low-density parity-check (LDPC) codes [3], [5] in mid 90's, LDPC codes have gained significant attention because of their capacity-approaching error-correcting ability and the inherent low-complexity ($O(n)$ or $O(n \log n)$), message passing decoding algorithm [5]. And recently, it has been successfully applied to different channels, including binary erasure channels [6], binary symmetric channels, binary-input additive white Gaussian channels [5], [8], Rayleigh fading channels, Markov channels, partial response channels/intersymbol interference channels [4], or even dirty paper coding [1].

The density evolution method proposed in [8] iteratively computes the probability density of the passed messages under a tree-like assumption. With a symmetric channel assumption, the performance concentration theorem and cycle-free convergence theorem presented in [8] provides a solid foundation for density evolution, and helps in finding near optimal LDPC codes [2], [7]. Kavčić *et. al.* in [4] generalized the density evolution methods for partial response channels by introducing the ensemble of *coset codes* and proved the underpinning theorems for the new code ensemble.

In this paper, we develop new formulae and theorems for symbol-dependent channels (e.g., z-channels, binary asymmetric channels, asymmetric Gaussian channels, etc.), under which setting the evolved density is codeword dependent and cannot be iteratively computed as before. Instead of using a larger code ensemble, we circumvent this problem by averaging over all equiprobable codewords, which is straightforward and has practical interpretation as the *averaged* error probability. A *full-rankness* theorem and some underpinning theorems are provided to justify this codeword average approach. We also provide new monotonicity, symmetry, and stability results of the resulting new density evolution method.

[1]This work is supported in part by the Army Research Office under contract number DAAD19-00-1-0466, and the New Jersey Center for Pervasive Information Technologies.



$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

Figure 1: A realization of the code ensemble $\mathcal{C}^6(2,3)$.

This paper is organized as follows. The formulations and basic models are provided in Section 2. In Section 3., the concept of *full rankness*, an iterative formula, underpinning theorems and important properties are discussed. Section 4. gives some simulation results. Section 5. concludes this paper.

2. Formulations

2.1. Symbol-dependent Channels

The memoryless, symbol-dependent channels are modelled as follows. $\mathbf{x}$ and $\mathbf{y}$ denote the transmitted and received codewords of length n. x_i is the i-th transmitted symbol in $GF(2)$, and y_i is the corresponding real-valued output. The channel is specified by the conditional probability density function $f_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} f(y_i|x_i)$.

2.2. LDPC Code Ensembles

Rather than a specific code, we consider the equiprobable regular bipartite code ensemble of length n with variable node degree d_v and check node degree d_c. The ensemble $\mathcal{C}^n(d_v, d_c)$ is constructed from regular bipartite graphs with degree (d_v, d_c) and equiprobable $(nd_v)!$ edge interconnections. One realization of $\mathcal{C}^6(2,3)$ with corresponding parity check matrix $\mathbf{A}$ is illustrated in Figure 1.

For general irregular graphs, let λ denote the variable node *edge degree distribution* polynomial, i.e. $\lambda(x) = \sum_k \lambda_k x^{k-1}$, where λ_k is the fraction of edges connecting to a variable node of degree k. Likewise, let ρ denote the check node *edge degree distribution* polynomial. An irregular code ensemble can then be specified by λ and ρ and denoted as $\mathcal{C}^n(\lambda, \rho)$. For a detailed construction, please refer to [7] and [8].

2.3. Belief Propagation Algorithm

The belief propagation algorithm is an instance of a message passing algorithm derived from Pearl's inference network, which is based on the cycle-free assumption and is actually channel model independent. Although most our results can be generalized to message passing algorithms, henceforth we only consider belief propagation algorithms.

3. Density Evolution

The density evolution computes the density of passed messages during a belief propagation algorithm. Suppose the initial message $m_0 := \log\left(\frac{\mathsf{P}(y_i|x_i=1)}{\mathsf{P}(y|x=0)}\right)$, $P^{(0)}$ is the density of m_0, $P^{(l)}$ is the density of the passed massages from variable nodes to check nodes after the l-th iteration, and $Q^{(l)}$ is for the messages from check nodes to variable nodes. For $\mathcal{C}^n(d_v, d_c)$, the iterative formula is as follows,

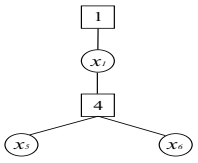
$$\begin{aligned} P^{(l)} &= P^{(0)} \otimes \left(Q^{(l-1)}\right)^{\otimes(d_v-1)} \\ Q^{(l-1)} &= \Gamma^{-1}\left(\left(\Gamma\left(P^{(l-1)}\right)\right)^{\otimes(d_c-1)}\right), \end{aligned}$$

where $\otimes$ stands for convolution operator, Γ is the probability transformation generated from certain nonlinear function $\gamma(m)$, and Γ^{-1} is its inverse. For details, please refer to [7].

Benefited from the symmetry of parity check constraints and the assumption of symmetric channels, for different transmitted codeword $\mathbf{x}$, the evolved densities of the passed messages only differ in parity, but have the same shape. In the general symbol-dependent setting, the shape also differs, which means the input messages are independent (from cycle free assumption) but *NOT* identically distributed anymore. So the iteration fails. A straightforward solution is to take average over the equiprobable codeword ensemble, which nevertheless takes prohibitively $2^{2^{O(l)}}$ times more computations with l being the number of iterations. We need new tools.

3.1. Full Rankness

With the assumption that the corresponding graph is tree-like till depth $2l$, we define the following quantities.



$$\mathbf{X}^1_{(1,1)} := \{x_1x_5x_6 : x_1x_5x_6 = 000, 011, 101, 110\}$$

Figure 2: Illustrations of $\mathbf{X}^1_{(1,1)}$ and $\mathcal{N}^{2*1}_{(1,1)}$.

1. Let $\mathcal{N}^{2l}_{(i,j)}$ be the depth $2l$, tree-like subset spanned from edge (i,j) and let $\left|\mathcal{N}^{2l}_{(i,j)}\right|_c$ denote the number of check nodes in $\mathcal{N}^{2l}_{(i,j)}$ (excluding check node j).

2. $\mathbf{X} = \{\mathbf{x} \in \{0,1\}^n : \mathbf{Ax} = \mathbf{0}\}$ is the codebook.

3. $\mathbf{x}|_i$ and $\mathbf{x}|_{\mathcal{N}^{2l}_{(i,j)}}$ are the projections of $\mathbf{x}$ on bit i and on the $\mathcal{N}^{2l}_{(i,j)}$ tree.

4. Let $\mathbf{X}^l_{(i,j)}$ denote the set of all strings that subject to the $\left|\mathcal{N}^{2l}_{(i,j)}\right|_c$ check node constraints in $\mathcal{N}^{2l}_{(i,j)}$. Note that the length of these strings is determined by the number of variable nodes in $\mathcal{N}^{2l}_{(i,j)}$. $\mathbf{x}^l$ denotes an element of $\mathbf{X}^l_{(i,j)}$.

Figure 2 demonstrates the above quantities for the codebook in Figure 1. Note that for any $\mathbf{x} \in \mathbf{X}$, $\mathbf{x}|_{\mathcal{N}^{2l}_{(i,j)}} \in \mathbf{X}^l_{(i,j)}$, which implies $\mathbf{X}|_{\mathcal{N}^{2l}_{(i,j)}} \subseteq \mathbf{X}^l_{(i,j)}$. However, the equality may or may not hold. As illustrated in Figure 1, the second row of $\mathbf{Ax} = \mathbf{0}$ implies $x_6 = 0$ so that two of the four elements in $\mathbf{X}^1_{(1,1)}$ in Figure 2 are invalid projections. Hence, we introduce the notion of *full rankness* in $\mathcal{N}^{2l}_{(i,j)}$.

Definition 1 (Full Rankness of $\mathcal{N}^{2l}_{(i,j)}$) *$\mathcal{N}^{2l}_{(i,j)}$ is of full rank, if for any $\mathbf{x}^l \in \mathbf{X}^l_{(i,j)}$,*

$$\frac{\left|\{\mathbf{x} \in \mathbf{X} : \mathbf{x}|_{\mathcal{N}^{2l}_{(i,j)}} = \mathbf{x}^l\}\right|}{|\mathbf{X}|} = \frac{1}{\left|\mathbf{X}^l_{(i,j)}\right|}.$$

That is if we only look at the projections of all codewords on $\mathcal{N}^{2l}_{(i,j)}$, the equiprobable appearance of all $\mathbf{x}^l$ is as if there are only $\left|\mathcal{N}^{2l}_{(i,j)}\right|_c$ check node constraints and no other else. The example in Figure 2 is obviously not *full rank*.

3.2. Iterative Formula

Since now the densities of the messages are codeword-dependent, we have to append subscripts and input argument to denote densities of the passed messages, which are as $P^{(l)}_{(i,j)}(\mathbf{x})$ and $Q^{(l)}_{(j,i)}(\mathbf{x})$ where (i,j) is the concerned edge and $\mathbf{x}$ is the transmitted codeword.

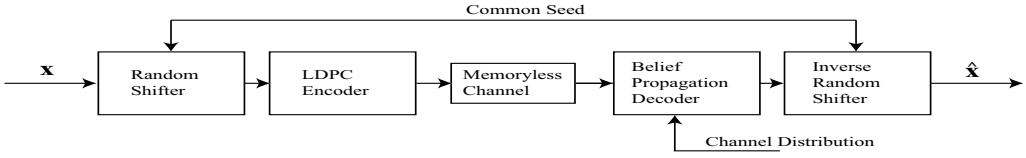


Figure 3: Randomized codeword transmission scheme for general memoryless channels.

For any fixed (i_0, j_0), we denote the average as (dropping indices)

$$P^{(l)}(x) := \frac{1}{|\{\mathbf{x} \in \mathbf{X} : x = \mathbf{x}|_{i_0}\}|} \sum_{\{\mathbf{x} \in \mathbf{X}: x = \mathbf{x}|_{i_0}\}} P^{(l)}_{(i_0,j_0)}(\mathbf{x})$$

$$Q^{(l)}(x) := \frac{1}{|\{\mathbf{x} \in \mathbf{X} : x = \mathbf{x}|_{i_0}\}|} \sum_{\{\mathbf{x} \in \mathbf{X}: x = \mathbf{x}|_{i_0}\}} Q^{(l)}_{(j,i_0)}(\mathbf{x}).$$

The immediate result is that the averaged bit error probability is

$$p_e^{(l)} = \frac{1}{2}\left(\int_{0^+}^{\infty} P^{(l)}(0) + \int_{-\infty}^{0^-} P^{(l)}(1)\right). \qquad (1)$$

With the assumption that $\mathcal{N}^{2l}_{(i,j)}$ is of full rank, we have a new iterative formula:

$$\begin{aligned}
& P^{(l)}(x) = P^{(0)}(x) \otimes \left(Q^{(l-1)}(x)\right)^{\otimes(d_v-1)} \\
& Q^{(l-1)}(x) \\
\overset{(a)}{=}\ & \Gamma^{-1}\left(\frac{1}{2^{d_c-2}} \sum_{\{(-1)^{c+x}=1\}} \binom{d_c-1}{c} S_c\right) \\
\overset{(b)}{=}\ & \Gamma^{-1}\left(\left(\Gamma\left(\frac{P^{(l-1)}(0)+P^{(l-1)}(1)}{2}\right)\right)^{\otimes(d_c-1)}\right. \\
& \left. +(-1)^x\left(\Gamma\left(\frac{P^{(l-1)}(0)-P^{(l-1)}(1)}{2}\right)\right)^{\otimes(d_c-1)}\right) \\
S_c\ :=\ & \Gamma\left(P^{(l-1)}(0)\right)^{\otimes(d_c-1-c)} \otimes \Gamma\left(P^{(l-1)}(1)\right)^{\otimes c}.
\end{aligned}$$

Note: (a) embodies the equivalence of averaging the evolved density and evolving the averaged density, which is our major result, and (b) further simplifies the computation to a constant factor (independent of (d_v, d_c)) over the symmetric case. For the irregular code ensemble $\mathcal{C}^n(\lambda, \rho)$, we have

$$\begin{aligned}
P^{(l)}(x) &= P^{(0)}(x) \otimes \lambda\left(Q^{(l-1)}(x)\right) \\
Q^{(l-1)}(x) &= \Gamma^{-1}\left(\rho\left(\Gamma\left(\frac{P^{(l-1)}(0)+P^{(l-1)}(1)}{2}\right)\right)\right. \\
&\quad \left. +(-1)^x\rho\left(\Gamma\left(\frac{P^{(l-1)}(0)-P^{(l-1)}(1)}{2}\right)\right)\right). \qquad (2)
\end{aligned}$$

3.3. Underpinning Theorems

Since the codeword averaging approach focuses on the same ensemble, the [Cycle-Free Convergence Theorem] and [Concentration Theorem] in [8] still hold. Besides those theorems, we need the following theorems to justify our approach.

Theorem 1 (Full Rankness Convergence) *For any $\mathcal{C}^n(d_v, d_c)$, with fixed l, i_0, and j_0, we have*

$$\lim_{n\to\infty} \mathsf{P}\left(\mathcal{N}^{2l}_{(i_0,j_0)} \text{ is of full rank}\right) = 1.$$

Theorem 2 (Validity of Density Evolution) *Let Z denote the number of wrong messages (those m's such that $m(x - \frac{1}{2}) < 0$). Consider any $\mathcal{C}^n(d_v, d_c)$ with fixed l. The probability over equiprobable codeword ensemble $\mathbf{x}$, the code ensemble $\mathcal{C}^n(d_v, d_c)$, and the channel realizations $\mathbf{y}$, satisfies*

$$\lim_{n\to\infty} \mathsf{P}\left(\left|\frac{Z}{nd_v} - p_e^{(l)}\right| > \epsilon\right) = 0, \forall \epsilon > 0,$$

where $p_e^{(l)}$ is computed from (1) and our iterative formula.

3.4. Important Properties

Theorem 3 (Monotonicity) *Let $f(y|x)$ and $g(y|x)$ denote two different memoryless channels, where $g(y|x)$ is physically degraded with respect to $f(y|x)$. The corresponding decoding error probabilities as defined in (1) are denoted as $p^{(l)}_{e,f}$ and $p^{(l)}_{e,g}$. For any fixed l, we have*

$$p^{(l)}_{e,f} \le p^{(l)}_{e,g}.$$

Definition 2 (Symmetric Probability Measures) *Two probability measures P and Q form a symmetric pair if for any integrable function h, we have*

$$\int h(\omega) d\mathsf{P}(\omega) \quad = \int e^{-\omega} h(-\omega) d\mathsf{Q}(\omega).$$

Theorem 4 (Symmetry) *The resulting densities $P^{(l)}(0)$ and $P^{(l)}(1)$ of (2) satisfy that $(P^{(l)}(0) \circ I^{-1},\ P^{(l)}(1))$ is a symmetric pair for all l, where $I(\omega) = -\omega$ reverses the parity.*

Corollary 1

$$\langle P^{(l)} \rangle := \frac{P^{(l)}(1) + P^{(l)}(0) \circ I^{-1}}{2}$$

is self symmetric, i.e. $(\langle P^{(l)} \rangle, \langle P^{(l)} \rangle)$ is a symmetric pair.

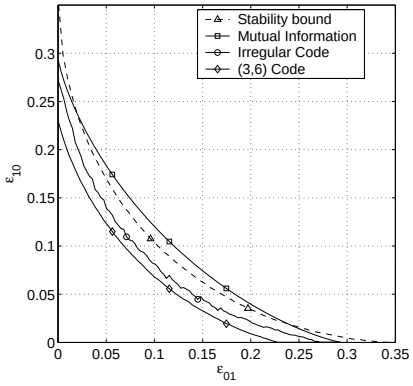


Figure 4: Capacity region of binary channels including: symmetric information rate, irregular code with $\max d_v = 12$ in [7], (3,6) code, and the stability upper bound derived for that specific irregular code.

Theorem 5 (Necessary Condition of Stability) *Let*

$$r := -\ln \int_{\bar{\mathbf{R}}} \langle P^{(0)} \rangle(\omega) e^{-\frac{\omega}{2}} d\omega.$$

Assume that $\int_{\bar{\mathbf{R}}} \langle P^{(0)} \rangle(\omega) e^{s\omega} d\omega < \infty$ *for all* s *in some neighborhood of zero. If* $\lambda'(0)\rho'(1) > e^r$*, then there exists a constant* $\epsilon > 0$ *such that for all* l*,* $p_e^{(l)} > \epsilon$*.*

Remark: The above complies with the intuition that the decoding of a noisier symbol can be helped by other less noisy symbols, since the stability result is for the average $\langle P^{(0)} \rangle$ rather than individual $P^{(0)}(0)$ and $P^{(0)}(1)$.

4. Implementations & Simulations

By implementation as in Figure 3, uniform error probability can be obtained with a fixed encoder and decoder pair, while leaving the codeword randomization to the shifter. Thus a uniform bit error rate can be achieved for all codewords. The capacity region derived from density evolution is illustrated in Figure 4. We note that the irregular code in [7] with $\max d_v = 12$ (optimized for BiAWGN channels) outperforms the (3,6) code even in binary channels, and is a good candidate for z-channels. Figure 5 displays the bit error rate of fixed specific codebooks, and demonstrates that our new density evolution indeed predicts the upper bound of a finite-length code.

5. Conclusions

We have shown that from both theoretical and experimental points of view, the codeword average approach is a successful generalization of the density evolution method for symbol-dependent memoryless channels. In addition to its value in implementations, we believe the concept of codeword-averaged density of passed messages will lead to various new applications and research directions.

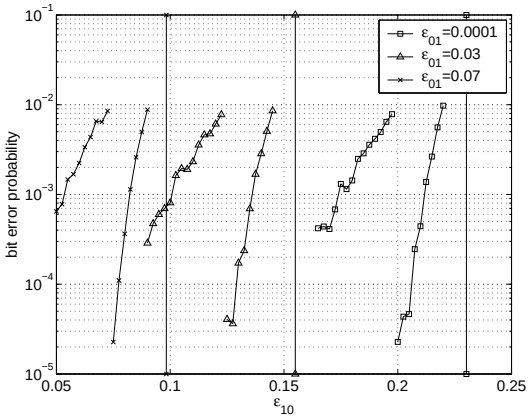


Figure 5: Performance and predicted thresholds for specific (3,6) codes with $n = 1000, 10000$, and different $(\epsilon_{01}, \epsilon_{10})$.

REFERENCES

[1] G. Caire, D. Burshtein, and S. Shamai, "LDPC Coding for Interference Mitigation at the Transmitter," *not yet published.*

[2] S.Y. Chung, G.D. Forney, T.J. Richardson, and R.L. Urbanke, "On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit," *IEEE Commun. Letters*, 5 (2): 58–60 Feb. 2001.

[3] R.G. Gallager, *Low-Density Parity-Check Codes*, no. 21 in Research Monograph Series. Cambridge, MA: MIT Press, 1963

[4] A. Kavčić, X. Ma, and M. Mitzenmacher, "Binary Intersymbol Interference Channels: Gallager Codes, Density Evolution and Code Performance Bound," submitted to *IEEE Trans. Inform. Theory.*

[5] D.J.C. MacKay, "Good Error-Correcting Codes Based on Very Sparse Matrices," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 399–431, Mar. 1999.

[6] P. Oswald and A. Shokorollahi, "Capacity-Achieving Sequences for the Erasure Channel," *IEEE Trans. Inform. Theory*, vol. 48, no. 12, pp. 3017–3028, Dec. 2002.

[7] T.J. Richardson, M.A. Shokrollahi, and R.L. Urbanke, "Design of Capacity-Approaching Irregular Low-Density Parity-Check Codes," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001.

[8] T.J. Richardson and R.L. Urbanke, "The Capacity of Low-Density Parity-Check Codes," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 599–618, Feb. 2001.