Convolutional Tanner Structures for Non-Ergodic Wireless Channels

Joseph J. Boutros†, Emanuele Viterbo‡, and Gérard Cohen⋆

†Texas A&M University at Qatar, Doha, Qatar
‡Università della Calabria, Cosenza, Italy
⋆Ecole Nationale Supérieure des Télécommunications, Paris, France

Abstract— **We propose an original technique for the design of convolutional Tanner structures that are full diversity under iterative decoding. The code design is based on the analysis of the local trellis neighborhood and is suitable for transmission over wireless non-ergodic channels. This new technique enables us to split the giant convolutional checknode into multiple smaller checknodes which is a means to mimic the standard analysis of LDPC codes under iterative message passing decoding.**

Index Terms: convolutional codes, iterative decoding, fading channels, erasure channels, generalized low density codes.

I. INTRODUCTION

Let $C[N, K]$ be a linear binary code of length N and dimension K whose codewords are transmitted over a non-ergodic memoryless channel. The channel has n_c states, where the number of states satisfies $2 \leq n_c \ll N$. A channel state may correspond to a fading coefficient, e.g., block-fading channels with single and multiple antennas. The state can also be defined by a cross-over probability of a binary symmetric channel (BSC), or an erasure event as for non-ergodic BEC (block-erasure channel). Let $c = (c_1, c_2, \ldots, c_N) \in C$ denote a codeword. The symbol c_i is assumed to undergo a channel state $h(c_i)$, where $i = 1 \ldots N$. The state $h(c_i)$ is selected from a finite set $\aleph$ defined by

$$\aleph = \{h_1, h_2, \ldots, h_{n_c}\} \qquad (1)$$

The non-ergodic channel has limited diversity because $|\aleph|$ is taken very small when compared to the code length. The problem of code multiplexing [1], also referred to as channel interleaving, is to select $h(c_i)$ among the h_j in order to minimize the word error probability after decoding at the receiver side. In this paper, we restrict our study to the worst case $n_c = |\aleph| = 2$ states, i.e., a channel with minimal diversity order. Half the bits in the codeword c undergo a fading h_1 and the other half undergo a fading h_2. It is assumed that the two fading instances are independent from one codeword to another and are not known by the encoder. Only the state position is controlled by the encoder, i.e., it knows if c_i is transmitted on state 1 or state 2. The generalization to channels with more degrees of freedom should be straightforward. Our objective is to build a rate-1/2 full-diversity Tanner code based

on convolutional constituents [2]. More details on the non-ergodic channel properties and other code constructions can be found in [3] and references cited therein.

II. GENERALIZED LOW DENSITY CODES WITH CONVOLUTIONAL CONSTITUENTS

A Gallager-Tanner construction based on convolutional codes has been proposed in [2]. The construction yields a family of asymptotically good codes, i.e., $d_{Hmin}(C) \geq \delta N$ where $\delta > 0$. Those codes are also suitable for iterative decoding. Briefly, a generalized low density (GLD) code as in [2] interleaves all code symbols whereas only information symbols are interleaved in a parallel turbo code [4]. The mathematical definition of the convolutional GLD-Tanner structure follows.

Let $C_1[N, K_1]$ be a linear binary code of length N and dimension K_1 built from a recursive systematic convolutional code (RSC) properly terminated at both trellis sides. The recursive nature is not essential, we only chose RSC codes because they have a flexible canonical structure for coding rates $r = k/n$, between $\frac{1}{2}$ and 1. The convolutional GLD code is defined by

$$C = C_1 \cap \pi(C_1) \qquad (2)$$

where π is a random permutation of size N. If border effects are neglected, the rate $R = K/N$ of C built from a rate-r convolutional code is $R = 2r - 1$. More precisely, let $r = k/n$ be the original coding rate of the infinite-length convolutional code. Assume that the RSC code has 2^ν states, i.e., the constraint length is $\nu + 1$. In practice, the intersection between C_1 and $\pi(C_1)$ does not include the two trellis terminations. It is easy to show that the overall rate is

$$R = (2r - 1) \times (1 - 2\frac{nT}{N}) \lesssim (2r - 1) \qquad (3)$$

where $T = \lceil \frac{\nu}{k} \rceil$ is the number of transitions required in the trellis termination phase.

In the sequel, without loss of generality and for simplicity reasons, we restrict our study to a rate-$\frac{3}{4}$ convolutional code obtained by puncturing the 4-state rate-$\frac{1}{2}$ code with generators $g_0(x) = 1 + x + x^2$ and $g_1(x) = 1 + x^2$, i.e., the famous

$(7,5)_8$ code in octal notation. The overall rate is $R = 0.495$ for $N = 800$ and $R = 0.4995$ for $N = 8000$ bits. The decoding of $C = C_1 \cap \pi(C_1)$ is iterative in a way similar to parallel turbo codes [4]. The constituents C_1 and $\pi(C_1)$ are decoded using a forward-backward algorithm [5] while passing extrinsic information on both parity and information bits at each half-iteration.

III. CODE DESIGN FOR BLOCK-FADING CHANNELS BASED ON THE LOCAL TRELLIS STRUCTURE

As stated in section I, our objective is to build a full-diversity convolutional GLD-Tanner code. For the sake of brevity and simplicity, we do not discuss why an ML design approach via the analysis of error events in the two convolutional constituents fails to attain this objective. Despite our belief that iterative decoding of GLD codes is close to ML decoding in terms of error rate performance (only 0.2dB separate iterative and ML decoders for parallel turbo codes as shown in [7]), the diversity order analysis of low weight error events is useless because the GLD code is asymptotically good. The study of iterative decoding on a Tanner graph with two giant convolutional checknodes is intractable. Thus, we propose a new approach based on the local trellis neighborhood.

The full-diversity construction is described for the block-erasure channel. It can be proven that the same structure is full-diversity in presence of Rayleigh-distributed block-fading by considering the partial codewords around the rootbit in the local neighborhood and analyzing the χ^2 distribution within the a posteriori probability message as in [3].
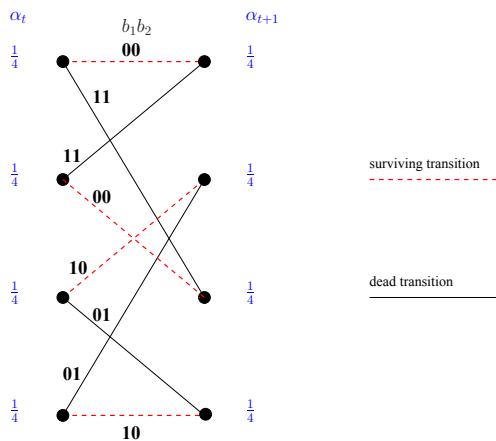

Fig. 1. Forward state probabilities when b_1 is erased and $b_2 = 0$. The surviving transitions are marked with dashed red lines. The configuration $b_1 b_2 = X0$ brings no improvement for reducing the ambiguity.

Let us start by studying how the trellis decoder bahaves locally in presence of block erasures. Two propositions are announced before revieling the full-diversity structure. The standard notations of the Forward-Backward algorithm (BCJR) are used as in [5]. It is always assumed that the

all-zero codeword has been transmitted by the encoder. The symbol X represents a punctured bit. Consider the trellis section shown in Fig. 1. Assume that the forward state probabilities are all equal, i.e., ambiguity is maximal with $\alpha_t(m) = \frac{1}{4}$ for all states m. This maximal ambiguity is equivalent to forgetting all the past in the code trellis. The transition label is denoted $b_1 b_2$. If b_1 is erased and b_2 forced to 0 as in Fig. 1, then the new forward state probabilities are $\alpha_{t+1}(m') = \frac{1}{4}, \forall m'$. Hence a label configured as X0 does not reduce the ambiguity. Similar arguments hold for 0X and XX. Now, we can announce the following proposition.

Proposition 1: The two binary elements of a transition label must be forced to 00 in order to reduce the state ambiguity from 100% to 50%. Illustration is given in Fig. 2.
Proof. Assume that $\alpha_t(m) = \frac{1}{4}, \forall m$. Force the transition label to $b_1 b_2 = 00$. Under erasures, the BCJR transition metric $\gamma(m, m')$ can only take two possible values, 0 or 1. Then, we have $\alpha_{t+1}(0) \propto 1 \times \alpha_t(0) + 0 \times \alpha_t(1) = \frac{1}{4}$, In a similar way, we get $\alpha_{t+1}(1) = 0$, $\alpha_{t+1}(2) \propto \frac{1}{4}$, and $\alpha_{t+1}(3) = 0$. After normalization (divide by $\sum_m \alpha_{t+1}(m)$), the final value of the forward state probabilities is $\alpha_{t+1}(0) = \alpha_{t+1}(2) = \frac{1}{2}$ and $\alpha_{t+1}(1) = \alpha_{t+1}(3) = 0$. QED.
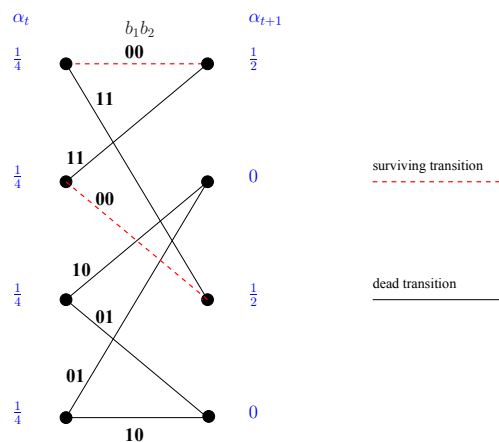

Fig. 2. Forward state probabilities when $b_1 = 0$ and $b_2 = 0$. The surviving transitions are marked with dashed red lines. The configuration $b_1 b_2 = 00$ eliminates half the ambiguity on trellis transitions.

Once the number of surviving states is reduced by forcing the two labeling bits to a zero value, we would like to keep the ambiguity at 50%, i.e., $\alpha_{t+1}(m') \in \{\frac{1}{2}, \frac{1}{2}, 0, 0\}$. This is possible if the two departing states belong to the same butterfly. The trellis section of a 4-state convolutional code has 2 butterflies. The transition matrix for the $(7,5)$ code is

$$A = \begin{bmatrix} 1 & 0 & D_1 D_2 & 0 \\ D_1 D_2 & 0 & 1 & 0 \\ 0 & D_1 & 0 & D_2 \\ 0 & D_2 & 0 & D_1 \end{bmatrix} \quad (4)$$

where a matrix entry at row i and column j represents $D_1^{b_1} D_2^{b_2}$ on the transition from state $m = i$ to state $m' = j$.

The butterflies can be found in the transition matrix by searching rectangular submatrices with four non-zero entries in the corners. In this case, the first butterfly starts at states 0,1 and ends at states 0,2. The second butterfly starts at states 2,3 and ends at states 1,3. The general structure of the trellis (excluding binary labels and not taking into account the state indexation) does not depend on the choice of the generator polynomial (see [8]). A 2^ν-state convolutional code has $2^{\nu-1}$ butterflies in a full trellis section that can be determined from its transition matrix. As depicted in Fig. 3, a butterfly maintains the ambiguity at 50% after erasing both b_1 and b_2.

Proposition 2: Assume that the two binary elements of a transition label are erased. The state ambiguity is maintained at 50% if the surviving states belong to the same butterfly. Otherwise, the ambiguity increases to 100%. Illustration is given in Fig. 3.
Proof. The proof is also based on BCJR equations for the forward recursion. It is left to the reader. Just remember that transition metrics are reduced to 0 and 1. QED.

Full-diversity Gallager-Tanner Construction (Root GLD Code)
Now, a full-diversity Gallager-Tanner construction can be derived from propositions 1 & 2. The convolutional GLD C code is built from (2) where the constituent C_1 follows the multiplexing pattern $11|X1|X2|22|X2|X1$ with a period of 6 transitions as illustrated in Fig. 4. A window including 7 trellis sections is considered. A symbol 1 in the multiplexing pattern represents a binary element sent on the first channel state. A symbol 2 represents transmission on the second channel state. The symbol X corresponds to a punctured bit. A forward BCJR procedure is applied on the first 3 sections. A backward BCJR procedure is applied on the last 3 sections. In Fig. 4, channel state 1 is assumed to correspond to a perfect knowledge of the encoded bit ($|h_1| = +\infty$), channel state 2 is supposed to be an erasure ($h_2 = 0$). After executing the forward and backward procedures, two transitions survive in the middle section. The binary digit b_2 is identical on both labels, its value $b_2 = 0$ is solved. Hence, the multiplexing pattern $11|X1|X2|22|X2|X1$ yields a rate-$\frac{3}{4}$ convolutional constituent capable of solving one erased bit (the second bit) in 11 and one erased bit in 22. The first bit b_1 in C_1 will be placed in the second label position in $\pi(C_1)$. We conclude that the GLD code is capable of solving 2 erased bits out of 4 in the multiplexing pattern $11|X1|X2|22|X2|X1$, those placed on 11 and 22 positions. Those bits, referred to as *rootbits*, have a diversity order equal to 2 and will be considered to be the information bits of C. Parity bits placed on the second position in X1 and X2 will have a diversity order limited to 1 (no diversity).

Permutation Embedded in the Root GLD Code
The class of information bits sent on h_1 (resp. h_2) is denoted $1i$ (resp. $2i$). Similarly, the class of parity bits transmitted on h_1 is denoted $1p$ and the other denoted by $2p$. The permutation

π of the GLD code C built from the multiplexing pattern $11|X1|X2|22|X2|X1$ shown in Fig. 4 must interleave the binary elements within the same class without mixing the four classes. Such a GLD code is similar to multi-edge type LDPC codes [9] and its density evolution should follow the same rules. It is called *root GLD code* in reference to root LDPC codes proposed in [3]. The information bits of classes $1i$ and $2i$ are referred to as rootbits because they are the root of a checknode that guarantees a full-diversity outgoing message.

The permutation π of size N is built as follows:

- A random permutation π_1 of size $N/4$ is applied on class $1i$, $\pi_1 : 1i \rightarrow 1i$. The permutation π_1 is the direct sum of $\pi_{1,1}$ and $\pi_{1,2}$ each of size $N/8$. The first permutation $\pi_{1,1}$ interleaves information bits from even positions to odd positions. The second permutation $\pi_{1,2}$ interleaves information bits from odd positions to even positions.
- Similarly, a randomly selected permutation π_2 of size $N/4$ is applied on class $2i$, $\pi_2 : 2i \rightarrow 2i$.
- A random permutation π_3 of size $N/4$ is applied on class $1p$, $\pi_3 : 1p \rightarrow 1p$.
- A random permutation π_4 of size $N/4$ is applied on class $2p$, $\pi_4 : 2p \rightarrow 2p$.
- The overall permutation π of the root GLD code is the direct sum of π_1, π_2, π_3, and π_4.

A fully random GLD code has a permutation π selected from an ensemble of cardinality $N!$. As described above, a full-diversity root GLD code has a permutation selected from an ensemble of cardinality $(\frac{N}{4}!)^2(\frac{N}{8}!)^4$.

Proposition 3: Let $C = C_1 \cap \pi(C_1)$ be a rate-1/2 binary GLD code of length N and C_1 the rate-3/4 punctured (7,5) convolutional code. Let the permutation π be randomly selected from the ensemble of cardinality $(\frac{N}{4}!)^2(\frac{N}{8}!)^4$ as described previously. If C_1 has binary digits ordered according to the periodic multiplexing pattern $11|X1|X2|22|X2|X1$, then C is a full-diversity GLD code on both block-fading and block-erasure channels under iterative probabilistic decoding.

The proof of proposition 3 on the block-erasure channel is a direct result of the Gallager-Tanner construction described above. We omit the proof in the presence of Rayleigh-distributed block-fading. It is based on showing that the log-ratio message of a rootbit includes a χ^2 distribution of order 4. This can be done by a direct application of the BCJR algorithm on the local neighborhood (the window with 7 trellis sections) or by the full enumeration of all partial codewords via the product of 7 transition matrices.
Monte Carlo simulation of the code from proposition 3 and the comparison with information theoretical limits are given in the next section. Notice that proposition 3 states that the code is full diversity when the binary digits ordering is well selected, but it says nothing about its coding gain. Fortunately, the next section shows a performance at 1dB distance from the outage probability limit [10][6].
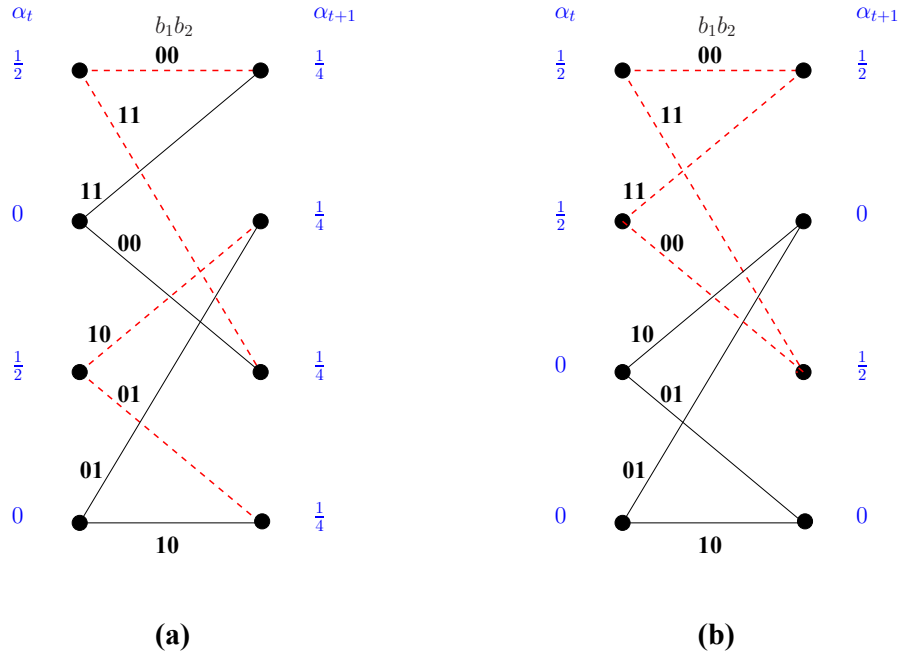
Fig. 3. Forward state probabilities when both b_1 and b_2 are erased. The surviving transitions are marked with dashed red lines. Figure (a) shows how the ambiguity spreads from 50% to 100%. Figure (b) shows the butterfly structure that inhibits ambiguity increase.
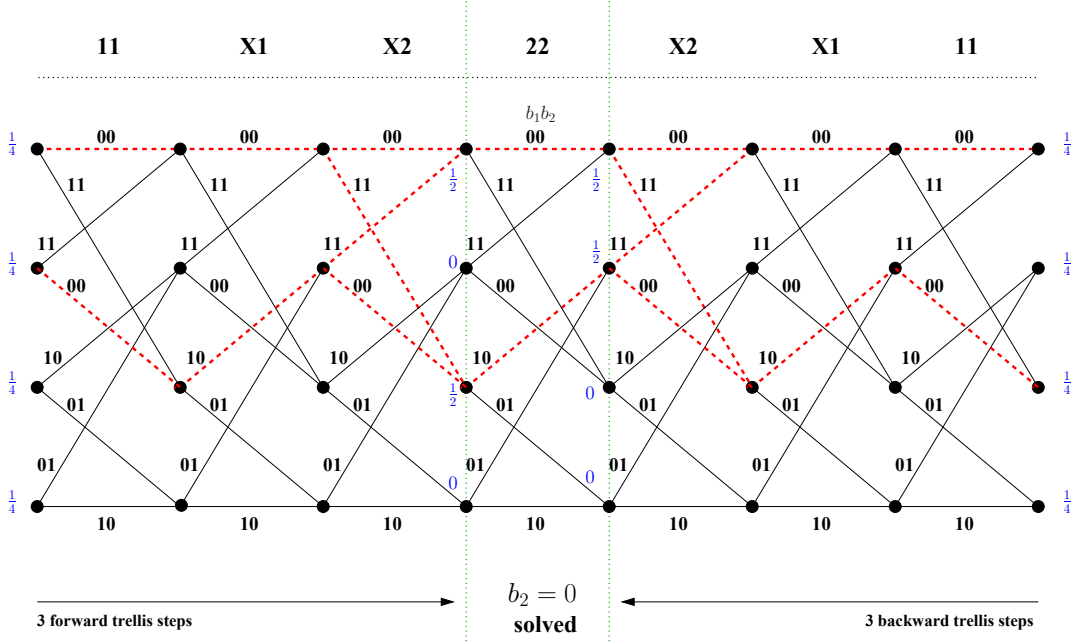

Fig. 4. Illustrating how the multiplexing pattern $11|X1|X2|22|X2|X1$ is capable of solving one erased bit on a block-erasure channel with two states. Illustration is given for the (7,5) rate-1/2 convolutional code. The all-zero codeword is assumed to be transmitted.
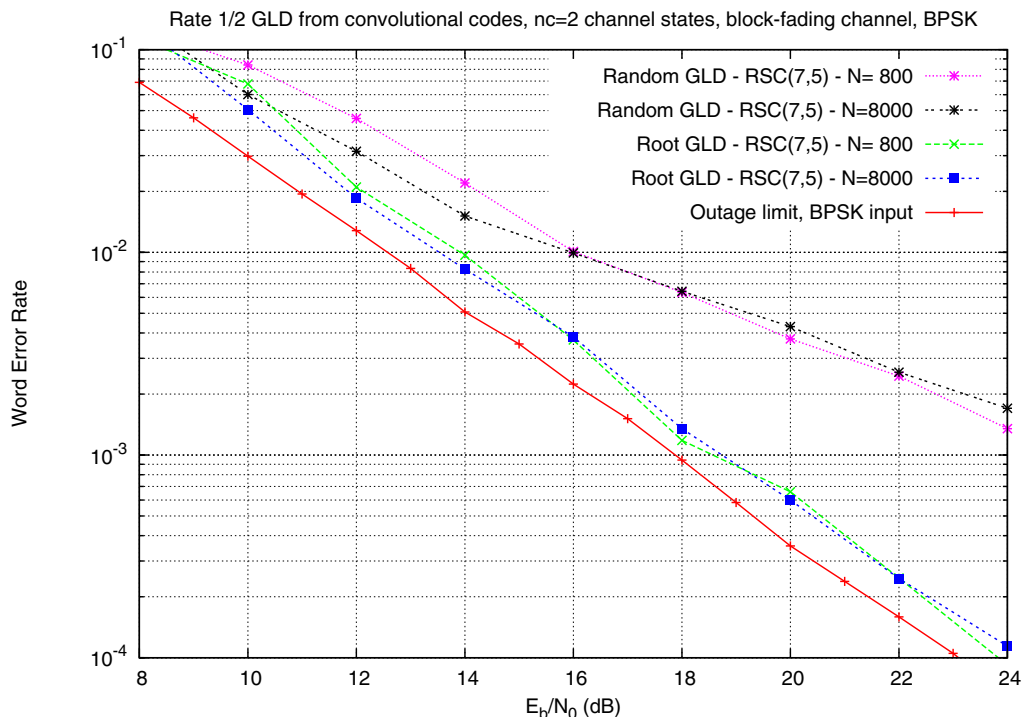
Rate 1/2 GLD from convolutional codes, nc=2 channel states, block-fading channel, BPSK

Fig. 5. Computer simulations of convolutional GLD codes for length $N = 800$ and $N = 8000$ bits on a Rayleigh block-fading channel.

IV. EXPERIMENTAL RESULTS

We performed Monte Carlo simulations of Generalized low-density codes based on the punctured $(7,5)_8$ convolutional rate-1/2 constituent. Results of word error rate versus signal-to-noise ratio per bit are depicted in Fig. 5. GLD codes with a random permutation are not full diversity, i.e., the word error rate varies as $P_{ew} \propto 1/(E_b/N_0)$ (diversity order is equal to 1). Root GLD codes from proposition 3 are full diversity, $P_{ew} \propto 1/(E_b/N_0)^2$ (diversity order is 2). We also plotted the outage probability limit computed by evaluating the probability that the instantaneous mutual information is less than 0.5 bits per channel use. Root GLD codes based on the (7,5) constituent achieve a near-outage performance. As shown in [3] for LDPC codes, the coding gain can be improved by introducing some irregularity in the convolutional GLD code structure in order to minimize the outage area near the ergodic line.

REFERENCES

[1] J.J. Boutros, E. Calvanese Strinati, and A. Guillén i Fàbregas, "Turbo code design for block fading channels," *Allerton's Conference*, Monticello, Illinois, Sept 2004.

[2] S. Vialle and J.J. Boutros, "A Gallager-Tanner construction based on convolutional codes," Workshop on Coding and Cryptography, pp. 393-404, Paris, January 1999. Click to download.

[3] J.J. Boutros, A. Guillén i Fàbregas, E. Biglieri, and G. Zémor, "Low-density parity-check codes for nonergodic block-fading channels," *Submitted to the IEEE Transactions on Information Theory*, Oct 2007. Click to download.

[4] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: turbo-codes," Proceedings of ICC'93, Geneva, pp. 1064-1070, May 1993.

[5] L.R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding for linear codes for minimizing symbol error rate," *IEEE Transactions on Information Theory*, vol. 20, pp. 284-287, March 1974.

[6] D.N.C. Tse and P. Viswanath, *Fundamentals of Wireless Communications*, Cambridge University Press, 2005.

[7] T. Hsueh and D. Shiu, "Theory and Performance of ML Decoding for Turbo Codes using Genetic Algorithm," *IEEE International Symposium on Information Theory 2007*, Nice, France, Jun. 2007.

[8] A.J. Viterbi and J.K. Omura, Principles of Digital Communication and Coding, McGraw-Hill, New York, 1979.

[9] T. J. Richardson and R. L. Urbanke, "Multi-edge type LDPC codes," *IEEE Trans. on Information Theory*, to appear. Click to download.

[10] L.H. Ozarow, S. Shamai (Shitz), and A.D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. on Vehicular Tech.*, vol. 43, no. 2, pp. 359-378, May 1994.