I. INTRODUCTION

The analysis of iterative coding systems has been extremely effective in determining the conditions for successful communication. The single most important prediction in this context is the existence of a threshold noise level below which the bit error rate vanishes (as the blocklength and the number of iterations diverge). The threshold can be computed for a large variety of code ensembles using density evolution.

On the other hand, understanding the behavior of these systems above threshold is largely an open issue. Since in this regime the bit error rate remains bounded away from zero, one may wonder about the motivation for such an investigation. We can think of three possible answers: (i) It is intellectually frustrating to have an “half-complete” theory of iterative decoding. Moreover this theory has poor connections with classical issues such as the behavior of the same codes under maximum likelihood (ML) decoding. (ii) Loopy belief propagation has stimulated a considerable interest as a general message-passing philosophy. Below the iterative threshold, its complexity becomes exponential. Its behavior can be analyzed precisely, and provides answers both questions (A) and (B) above (within this circumscribed context). Surprisingly, the resulting picture is most easily conveyed using a well-known information theoretic characterization of the code: the EXIT curve. As a byproduct, we obtain an alternative proof of the area theorem for the BEC.

The connection between the EXIT curve and Maxwell decoder is not a peculiarity of the binary erasure channel, and has instead a rather fundamental origin. The algorithm progressively reduces the uncertainty on the transmitted bits. This can be regarded as an effective change of the noise level of the communication channel. The EXIT curve describe the response of the bits (i.e., the change of the bit uncertainty) to a change in the noise level. The area theorem is obtained when integrating this response: the total bit uncertainty at maximum noise level (the code rate) is thus given by an integral of the EXIT curve.

In Sec. IV, we explain how to generalize these ideas to arbitrary memoryless channels.

Life Above Threshold: From List Decoding to Area Theorem and MSE

Cyril Measson and Rüdiger Urbanke
EPFL, I&C
CH-1015 Lausanne, Switzerland
e-mail: cyril.measson@epfl.ch
ruediger.urbanke@epfl.ch

Andrea Montanari
LPTENS (UMR 8549, CNRS et ENS)
24, rue Lhomond, 75231
Paris CEDEX 05, France
e-mail: montanari@lpt.ens.fr

Tom Richardson
Flarion Technologies
Bedminster, NJ, USA-07921
e-mail: richardson@flarion.com

Abstract — We consider communication over memoryless channels using low-density parity-check code ensembles above the iterative (belief propagation) threshold. What is the computational complexity of decoding (i.e., of reconstructing all the typical input codewords for a given channel output) in this regime? We define an algorithm accomplishing this task and analyze its typical performance. The behavior of the new algorithm can be expressed in purely information-theoretical terms. Its analysis provides an alternative proof of the area theorem for the binary erasure channel. Finally, we explain how the area theorem is generalized to arbitrary memoryless channels. We note that the recently discovered relation between mutual information and minimal square error is an instance of the area theorem in the setting of Gaussian channels.

1Copyright 2004 IEEE. Published in the 2004 IEEE Information Theory Workshop (ITW 2004), scheduled for October 24-29, 2004 at the Riverwalk Marriott in San Antonio, Texas, USA. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works, must be obtained from the IEEE. Contact: Manager, Copyrights and Permissions / IEEE Service Center / 445 Hoes Lane / P.O. Box 1331 / Piscataway, NJ 08855-1331, USA. Telephone: + Intl. 732-562-3066.
bitary memoryless channels. In particular, we define a
generalized EXIT function $\text{GEXIT}$, which has the same important
properties of the usual one. We show that an area theorem
holds for such a function, implying, among other things, an
upper bound on the ML threshold. $\text{GEXIT}$ reduces to $\text{EXIT}$
for the BEC and to the minimal mean-square error (MMSE)
for additive Gaussian channels.

II. Area Theorem for the Binary Erasure
Channel

Consider a degree distribution pair $(\lambda, \rho)$ and ensembles
LDPC$(\alpha, \lambda, \rho)$ of increasing length $n$. Figure shows a typical
asymptotic $\text{EXIT}^2$ function. Its main characteristics (for a
regular ensemble with left degree at least 3) are as follows:
The function is zero below the ML threshold $\epsilon_{\text{ML}}$. It jumps at
$\epsilon_{\text{ML}}$ to a non-zero value and continues then smoothly until it
reaches one for $\epsilon = 1$. The area under the $\text{EXIT}$ curve equals
the rate of the code, see [5]. Compare this to the equivalent
$\epsilon$ areas are indicated in dark gray in the picture). This unique
line and is enclosed by the line and the iterative curve (thes e
iterative curve is equal to the area which lies to the right of the
left of this straight line and is enclosed by the line and t he

now the curve which is zero to the left of the threshold and
equals the iterative curve to the right of this threshold. In
other words, the ML threshold is determined by a balance
between two areas$^3$.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig1}
\caption{Left: The EXIT curve of the ML decoder for the degree
distribution pair $(\lambda(x) = x^2, \rho(x) = x^5)$. The curve is zero until
$\epsilon_{\text{ML}}$ at which point it jumps. It then continuous smoothly until
it reaches one at $\epsilon = 1$. Also shown is the equivalent curve under
iterative decoding. Right: The full iterative EXIT curve including the
“spurious branch”. This corresponds to an unstable fixed point
$x > \epsilon$. The ML threshold is determined by the balance of the two
dark gray areas.}
\end{figure}

III. Maxwell Decoder

The balance condition described above, cf. Fig. is
strongly reminiscent of the so-called ‘Maxwell construction’ in
statistical mechanics [6]. This allows, for instance, to deter-
mine the location of a liquid-gas phase transition, by bal-
ancing two areas in the pressure-volume phase diagram. The
Maxwell construction is derived by considering a reversible
transformation between the liquid and vapor phases. The
balance condition follows from the observation that the net
work exchange along such a transformation must vanish at the
phase transition point.

Inspired by the statistical mechanics analogy, we shall ex-
plain the balance condition determining the ML threshold
by analyzing an algorithm which moves from the non zero-
entropy branch to the zero-entropy branch of the EXIT curve.
To this end we construct a fictitious decoder, which for obvio-
sus reasons we name the Maxwell decoder. Instead of explain-
ing the balance between the areas as shown in Fig. we will ex-
plain the balance of the two areas shown in Fig. Note that
these two areas differ from the previous ones only by a com-
mon term so that the condition for balance stays unchanged.

Let us now introduce the decoder: Given the received word
which was transmitted over the BEC$(\epsilon)$, the decoder proceeds
iteratively as does the standard message passing decoder. At
any time the iterative decoding process gets stuck in a non-
empty stopping set the decoder randomly chooses a position
$i \in [n]$. If this position is not known yet the decoder splits any
running copy of the decoding process into two, one which pro-
ceed with the decoding process by assuming that $x_i = 0$ and
one which proceeds by assuming that $x_i = 1$. This splitting
procedure is repeated any time the decoder gets stuck and

$^3$The ML threshold was first determine by the replica method in [6]. Further, in [7] a simple counting argument leading to an
upper bound for this threshold was given. In this paper we take as a
starting point the point of view taken in [6].
we say that the decoder guesses a bit. During the decoding it can happen that contradictions occur, i.e., that a variable node receives inconsistent messages. Any copy of the decoding process which contains such contradictions terminates. From the above description it follows that at any given point of the decoding process there are $2^{h(\ell)}$ copies alive, where $h(\ell)$ is a natural number which evolves with time $\ell$. Eventually, each surviving copy will have determined all the erased bits, and outputs the corresponding word of size $n$. It is hopefully clear from the above description that the final list of surviving copies is in one-to-one correspondence with the list of codewords that are compatible with the received message. In other words, the Maxwell decoder performs a complete list decoding of received message.

In Fig. 4 we depict an instance of the decoding process is shown from the perspective of the various simultaneous copies. The initial phase coincides with standard message passing: a single copy of the process decodes a bit at a time. After three steps, belief propagation gets stuck in stopping set and several steps of guessing follow. During this phase $h(\ell)$ (the associated entropy, i.e., the log$_2$ of the number of simultaneously running copies) increases. After this guessing phase, the standard message passing phase resumes. More and more copies will terminate due to inconsistent messages (incorrect guesses). At the end, only one copy survives, which shows that the example has a unique ML solution.

In Fig. 4 we plot the entropy $h(\ell)$ as a function of the number of iterations for several code and channel realizations (here we consider a (3, 6) ensemble with blocklength $n = 10^4$ and erasure probability $\epsilon = 0.47$). It can be shown that the rescaled entropy $h(\ell)/n$ concentrates around a finite limiting value if we take the large blocklength limit $n \to \infty$, with $\ell/n$ fixed. Moreover the limiting curve can be computed exactly. Here we limit ourselves to outline the connection with the various areas highlighted in Fig. 2 and to explain why these areas should be in balance at the ML threshold. To simplify matters consider only channel parameters $\epsilon$ with $\epsilon \geq \epsilon_{\text{IT}}$. We claim that the total number of guesses one has to venture during the guessing phase of the algorithm is equal to the dark gray area shown in the left picture of Fig. 2, i.e., it is equal to the integral under the iterative curve from $\epsilon_{\text{IT}}$ up to $\epsilon$.

The effect of the guesses is to bring the effective erasure probability down from $\epsilon$ to $\epsilon_{\text{IT}}$. At this point the standard message passing decoder can resume. The guesses are now resolved in the following manner. Assume that at some point in time there is a variable node which has $d$ connected check nodes of degree one. The corresponding incoming messages have to be consistent. This gives rise to $d-1$ constraints, or in other words, only a fraction $2^{d-1}$ of the running copies survive. It can now be shown that the total number of such constraints which are imposed is equal to the area in the left picture of Fig. 2. At the ML threshold all guesses have to be resolved at the end of the decoding process. This implies that the total number of required guesses has to equal the total number of resolved guesses which implies an equality of the areas as promised!

Notice that the Maxwell decoder plays the same role as a reversible transformation in thermodynamics.

IV. GENERAL CHANNEL

Three important lessons can be learned from the BEC example treated in the previous Sections. First of all: the EXIT curve gives the change in conditional entropy of the transmitted message when the channel noise level is incremented by an infinitesimal amount. Second: in a search algorithm reconstructing all the typical input codewords, this change...
has to be compensated for by an increase of the algorithm entropy. This is the fundamental reason of the equality between the area under the stable branch of the EXIT curve and the number of guesses made by the Maxwell decoder. Third: the fact that the iterative EXIT curve extends below the maximum-likelihood one implies that the corresponding additional guesses must be eventually resolved. The unstable branch of the EXIT curves yield the number of contradictions found in this resolution stage.

The first step towards a generalization of this scenario for an arbitrary memoryless channel consist in finding the appropriate generalization of the EXIT curve. We obtain such a generalization by enforcing the first of the above properties. For the sake of definiteness, we assume both the input and output alphabets to be finite and denote by $Q(y|x)$, $x \in X$, $y \in Y$ the transition probability. Formulae for continuous alphabets are easily obtained by substituting integrals $\int dx$, $\int dy$, to sums $\sum_x \sum_y$. We moreover denote by $w$ a generic noise-level parameter and assume $Q(y|x)$ to be differentiable with respect to $w$. In analytical calculations, it is convenient to distinguish the noise levels for each channel use $w_i$, $i \in [n]$. The time-invariant channel is recovered by setting $w_1 = \ldots = w_n = w$. Finally, we denote by $X \equiv X^n$ the channel input and $Y \equiv Y^n$ the channel output. Our definition of a generalized EXIT curve is

$$\text{GEXIT} \equiv \frac{1}{n} \frac{d}{d w_i} H(X|Y) .$$

(2)

Notice that GEXIT satisfies the area theorem by construction: our purpose is to get a manageable expression for it. It is convenient to think of the above differentiation as acting on each channel separately

$$\text{GEXIT}_i \equiv \frac{1}{n} \frac{d}{d w_i} H(X_i|Y) \equiv \frac{1}{n} \sum_{i=1}^n \text{GEXIT}_i ,$$

(3)

with GEXIT$_i$ defined as the derivative with respect to $w_i$. In order to compute GEXIT$_i$, it is convenient to isolate the contribution of $X_i$ to the conditional entropy. If we denote by $Z_i$ the extrinsic information at $i$, and use the shorthand $X^{[i]} \equiv \{X_j : j \in [n]\setminus i\}$, $Y^{[i]} \equiv \{Y_j : j \in [n]\setminus i\}$, we get

$$H(X|Y) = H(X_i|Z_i, Y) + H(X^{[i]}|X_i, Y^{[i]}).$$

(4)

This is obtained by a standard application of the entropy chain rule

$$H(X|Y) = H(X_i|Y) + H(X^{[i]}|X_i, Y) = H(X_i|Y, Z_i) + H(X^{[i]}|X_i, Y^{[i]}) = H(X_i|Y_i, Y^{[i]}|Z_i) + H(X^{[i]}|X_i, Y^{[i]}) = H(X_i|Y_i, Z_i) + H(X^{[i]}|X_i, Y^{[i]}).$$

We remark at this point that only the first term of the decomposition (4) depends upon the channel at position $i$. Therefore

$$\text{GEXIT}_i \equiv \frac{d}{d w_i} H(X_i|Z_i, Y_i) .$$

(5)

It is convenient to obtain a more explicit expression for the above formula. To this end we write

$$H(X_i|Z_i, Y_i) = - \sum_{x_i, y_i, z_i} P(z_i) P(x_i|z_i) Q(y_i|x_i) \cdot$$

$$\log \left\{ \frac{P(x_i|z_i) Q(y_i|x_i)}{\sum_{x_i'} P(x_i'|z_i) Q(y_i|x_i')} \right\} .$$

The dependence of $H(X_i|Z_i, Y_i)$ upon the channel at position $i$ is completely explicit and we can differentiate. The terms obtained by differentiating with respect to the channel inside the log vanish. For instance, when differentiating with respect to the $Q(y_i|x_i)$ at the numerator, we get

$$- \sum_{x_i, y_i, z_i} P(z_i) P(x_i|z_i) \frac{d}{d w_i} Q(y_i|x_i)$$

$$= - \sum_{x_i, z_i} P(z_i) P(x_i|z_i) \frac{d}{d w_i} \sum_{y_i} Q(y_i|x_i) = 0 .$$

We thus proved the following

**Theorem 1** With the above definitions

$$\text{GEXIT}_i = \sum_{x_i, y_i, z_i} P(x_i) P(z_i|x_i) Q'(y_i|x_i) \cdot$$

$$\log \left\{ \sum_{x_i'} P(x_i'|z_i) Q(y_i|x_i) \right\} ,$$

(7)

where we denoted by $Q'(y|x)$ the derivative of the channel transition probability with respect to the noise level $w$.

The interest of the above result is that it encapsulates all our ignorance about the code behavior into the distribution of extrinsic information $P(z_i)$. This is in turn the natural object appearing in message passing algorithms and in density evolution analysis. In order to fully appreciate the meaning of Eq. (7), it is convenient to consider a couple of more specific examples.

**Linear Codes over BMS Channels**

We assume the code to be linear and to be used over a binary-input memoryless output-symmetric (BMS) channel. We furthermore denote the channel input by $X = \{0, 1\}$. This is the most common setting in the analysis of iterative coding systems. Exploiting the channel symmetry we can fix $x_i = 0$ in Eq. (7) and get

$$\text{GEXIT}_i = \sum_{y_i, z_i} P_0(z_i) Q'(y_i|x_i) \log \left\{ 1 + \frac{P(1|z_i) Q(y_i|1)}{P(0|z_i) Q(y_i|0)} \right\} ,$$

(8)

where we defined $P_0(z_i)$ to be the distribution of the extrinsic information at $i$ under the condition that the all-zero codeword has been transmitted. Recall that $Z_i$ is a function of $Y^{[i]}$ and $P_0(z_i)$ is the distribution induced on $Z_i$ by the distribution of $Y^{[i]}$.

It is convenient to encode the extrinsic information $z_i$ as an extrinsic log-likelihood ratio $l_i \equiv \log[P(1|z_i)/P(0|z_i)]$. Analogously, we define $L_Q(y) \equiv \log[Q(y|0)/Q(y|1)]$. Finally, we denote by $a^{(i)}(l)$ the density of $l_i$ with respect to the Lebesgue measure. We thus get the following handy expression

**Corollary 1** For a linear code over a BMS channel

$$\text{GEXIT}_i = \int_{-\infty}^{+\infty} a^{(i)}(l) k_L(l) dl ,$$

(9)

where we introduced the GEXIT kernel

$$k_L(l) \equiv \sum_y Q'(y|0) \log(1 + e^{-L_Q(y) - l}) .$$

(10)
It is worth recalling that the usual EXIT curve has a similar expression. In fact

$$\text{EXIT}_i = \int_{-\infty}^{+\infty} a^{(i)}(l) \tilde{k}_{L}(l) \, dl,$$

with the channel-independent EXIT kernel \( \tilde{k}_{L}(l) \equiv \log(1 + e^{-l}) \). Finally, we notice that it is possible to use alternative encodings for the extrinsic information. One important possibility is to work in the so-called ‘difference domain’ \( z = \frac{1}{2} \tanh(l/2) \). The new kernel will be given by \( k_D(z) \equiv k_L(2 \tanh^{-1}(z)) \). It is moreover possible to exploit the symmetry property of \( a^{(i)}(l) \) to get

$$\text{GEXIT}_i = \int_{0}^{+\infty} a^{(i)}(l) \tilde{k}_{L}(l) \, dl,$$

where \( \tilde{k}_{L}(l) = k_L(l) + e^{-l} k_L(l) \). Analogously, one can consider an ‘absolute difference’ kernel \( k_D(z) \).

Let us work out a couple of examples. In order to compare the different cases, it is useful to define a unified convention for the noise level parameter \( w \). We choose \( w \) to be the channel entropy, or, in other words, one minus the channel capacity: \( w = 1 - C(Q) \) (in bits).

For the BEC we have \( \mathcal{V} = \{0, 1, \ast\} \) and the transition probabilities read \( Q(0|0) = 1 - \epsilon, Q(0|0) = \epsilon, Q(1|0) = 0 \). Obviously \( w = \epsilon \) and \( Q'(0|0) = -1, Q'(0|0) = 1, Q'(1|0) = 0 \). We get

$$L^\text{BEC}_{L}(l) \equiv \log(1 + e^{-l}) .$$

Therefore \( k^\text{BEC}_{L}(l) = \tilde{k}_{L}(l) \) and \( \text{GEXIT}_i = \text{EXIT}_i \). We thus recovered a well known result: the EXIT curve verifies the area theorem for the BEC.

Consider now the BSC with flip probability \( p \). Proceeding as above, we get

$$k^\text{BSC}_{L}(l) = \frac{1}{\log(1-p)} \left[ \log(1 + \frac{1-p}{p} e^{-l}) - \log(1 + \frac{p}{1-p} e^{-l}) \right] .$$

In Figs. 5 and 6 we plot the GEXIT kernel in the difference and absolute difference domains, comparing it with the usual EXIT one. In Fig. 6 we plot the EXIT and GEXIT curves for a few regular LDPC ensembles over the BSC.

From these examples it should be clear that computing the GEXIT curve is not harder than computing the EXIT one. The difference between these two curves is often quantitatively small (cf. for instance Fig. 6). Nevertheless such a difference is definitely different from zero and it is not hard to show that an area theorem cannot hold for the EXIT curve. Finally, several qualitative properties remain unchanged. In particular

**Lemma 1**

Given a density \( a(l) \) over the reals, let

$$\text{GEXIT}_{\text{BSC}}(a) \equiv \int_{-\infty}^{\infty} a(l) \tilde{k}_{L}^{\text{BSC}}(l) \, dl .$$

If the density \( b(l) \) is physically degraded with respect to \( a(l) \), then \( \text{GEXIT}_{\text{BSC}}(b) \geq \text{GEXIT}_{\text{BSC}}(a) \).

An important application of the above Lemma consists in approximating the correct extrinsic LLR densities with the site-averaged belief propagation density. This yields an upper bound on the GEXIT curve:

$$\text{GEXIT} = \frac{1}{n} \sum_{i=1}^{n} \text{GEXIT}_i \leq \text{GEXIT}_{\text{BSC}}(a^{\text{BP},k}) .$$
where we denoted by $a^{BP,k}$ the belief propagation density after $k$ iterations. We can now take the $n \to \infty$ limit and (afterwards) the $k \to \infty$ limit to get

$$GEXIT \leq GEXIT_{BSC}(a^{DE,*})$$

where $a^{DE,*}$ is the density at the density evolution fixed point. We obtain therefore the following

**Corollary 2** Consider communication over the BSC using LDPC$(n, \lambda, \rho)$ ensembles of rate $R$, and let $p_{ML,DE}$ be defined by

$$R = \int_{p_{ML,DE}} GEXIT_{BSC}(a^{DE,*}) \, dp,$$

with $a^{DE,*}$ the density at the density evolution fixed point at flip probability $p$. Let moreover $p_{ML}$ be the maximum likelihood threshold defined as the smallest noise level such that the ensemble-averaged conditional entropies $\mathbb{E} H(X|Y)$ is linear in the blocklength. Then

$$p_{ML} \leq p_{ML,DE}.$$

**Example 1** For the $(3, 6)$ ensemble and the BSC, the previous method gives $p_{ML} \leq 0.101$.

**Gaussian Channels**

We assume $X = Y = R$ and

$$Q(y|x) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left( y - \sqrt{\text{snr}} x \right)^2 \right\}. \quad (18)$$

Notice that, in this case, $Q(y|x)$ should be interpreted as a density with respect to Lebesgue measure. An alternative formulation of the same channel model consists in saying that $Y = \sqrt{\text{snr}} X + W$ with $W$ a standard Gaussian variable. It is also useful to define the minimal mean square error $\text{MMSE}_i$, in estimating $X_i$ as follows

$$\text{MMSE}_i \equiv \mathbb{E}_{Y_i,Z_i} \left\{ \mathbb{E}_{X_i} [x_i^2 | y_i, z_i] - \mathbb{E}_{X_i} [x_i | y_i, z_i]^2 \right\}, \quad (19)$$

where $\mathbb{E}_{X,Y,Z,...}$ denotes expectation with respect to the variables $\{X,Y,...\}$. Finally, we take the signal-to-noise ratio as the noise parameter entering in the definition of the GEXIT curve: $w = \text{snr}$. The reader will easily translate the results to other choices of $w$ by a change of variable.

As recently shown by Guo, Shamai and Verdu [9], the derivative with respect to the signal-to-noise ratio of the mutual information across a gaussian channel is related to the minimal mean-square error. Adopting their result to the present context, we immediately obtain the following

**Corollary 3** For the additive Gaussian channel defined above, we have

$$GEXIT_i = -\frac{1}{2} \text{MMSE}_i. \quad (20)$$

For greater convenience of the reader we briefly recall the derivation of this result from the expression (17). In order to keep things simple, we shall consider here the case of a simple symbol with input density $P(x)$ transmitted uncoded through the channel. The generalization is immediate. We rewrite Eq. (17) in the single symbol case as

$$\text{GEXIT} = \int \int P(x) Q^i(y|x) \log \left[ \frac{P(x) Q^i(y|x)}{P(x) Q(y|x)} \right] \, dx \, dy. \quad (21)$$

It is convenient to group at this point a couple of remarks which simplify the calculations. First

$$Q'(y|x) = -\frac{x}{2\sqrt{\text{snr}} y} \frac{Q(y|x)}{dy}. \quad (22)$$

Second

$$\frac{1}{\sqrt{\text{snr}}} \frac{dy}{dy} \mathbb{E}_X [y|x] = \mathbb{E}_X [x^2 | y] - \mathbb{E}_X [x | y]^2. \quad (23)$$

Both of these formulae are obtained through simple calculus. In order to prove Eq. (20) we use (22) in Eq. (21) and integrate by parts with respect to $y$. After re-ordering the various terms, we get

$$\text{GEXIT} = \frac{1}{2\sqrt{\text{snr}}} \int \int \mathbb{E}_X [y|x] P(x) \frac{dy}{dy} Q(y|x) \, dx \, dy. \quad (24)$$

At this point we integrate by parts once more with respect to $y$ and use (20) to get the desired result.

Notice that the strikingly simple relation (20) was recently used in an iterative coding setting by Bhattacharyya and Narayanan [10].

**Acknowledgments**

The work of A. Montanari was partially supported by the European Union under the project EVERGROW.

**References**

[1] C. Méasson and R. Urbanke, “An upper-bound for the ML threshold of iterative coding systems over the BEC”, Proc. of the 41st Allerton Conference on Communications, Control and Computing, Allerton House, Monticello, USA, October 1–3, 2003.

[2] S. ten Brink, “Convergence behavior of iteratively decoded parallel concatenated codes”, IEEE Trans. on Communications, vol. 49, no. 10, October 2001.

[3] M.G. Luby, M. Mitzenmacher, M.A. Shokrollahi and D.A. Spielman, “Efficient erasure correcting codes”, IEEE Trans. on Information Theory, vol. 47, no. 2, pp. 569–584, February 2001.

[4] T.J. Richardson, M.A. Shokrollahi and R. Urbanke, “Design of capacity-approaching irregular low-density parity-check codes”, IEEE Trans. on Information Theory, vol. 47, no. 2, pp. 619–637, February 2001.

[5] A. Ashikhmin, G. Kramer and S. ten Brink, “Code rate and the area under extrinsic information transfer curves”, Proc. of the 2002 ISIT, Lausanne, Switzerland, June 30–July 5, 2002.

[6] S. Franz, M. Leone, A. Montanari and F. Ricci-Tersenghi, “The dynamic phase transition for decoding algorithms”, Phys. Rev. E 66, 046120, 2002.

[7] A. Montanari, “Why “practical” decoding algorithms are not as good as “ideal” ones?”, DIMACS Workshop on Codes and Complexity, Rutgers University, Piscataway, USA, December 4–7, 2001.

[8] C. Kittel and H. Kroemer, *Thermal physics*, 2nd Edition, W. H. Freeman and Co., New York, March, 1980.

[9] D. Guo, S. Shamai and S. Verdu, “Mutual information and MMSE in Gaussian channels”, Proc. 2004 ISIT, Chicago, IL, USA, June 2004, p. 347.

[10] K. Bhattacharyya and K.R. Narayanan, “An MSE Based Transfer Chart for Analyzing Iterative Decoding”, Proc. of the 42nd Allerton Conference on Communications, Control and Computing, Allerton House, Monticello, USA, September 29–October 3, 2004.