Stochastic Models for Turbo Decoding

Minyue Fu
School of EE&CS, University of Newcastle, Australia
Email: minyue.fu@newcastle.edu.au

Abstract— **This paper proposes a stochastic framework for modelling and analysis of turbo decoding. By modelling the input and output signals of a turbo decoder as random processes, we prove that these signals become ergodic when the code block size becomes very large. This basic result allows us to easily model and compute the statistics of the signals in a turbo decoder. Using the ergodicity result and the fact that a sum of lognormal distributions is well approximated using a lognormal distribution, we show that the input-output signals in a turbo decoder, when expressed using the so-called scaled log-likelihood ratios, are well approximated using Gaussian distributions. Combining the two results above, we can model a turbo decoder using two inputs and two outputs (corresponding to the means and variances). Using this model, we have discovered that a typical decoding process is much more intricate than previously known, involving two regions of attractions, several fixed points, and a stable equilibrium manifold at which all decoding trajectories converge.**

Keywords: Turbo decoding, turbo code, iterative decoding, MAP decoding.

I. Introduction

Many papers can be found which attempt to uncover the mystery of the turbo decoding method invented in [1], see, e.g., [2]-[7]. There are two approaches: deterministic and stochastic. The former can address issues such as the existence, uniqueness and stability of equilibrium (or fixed) points; see, e.g., [2]. However, the existence of a fixed point does not imply anything about the convergence of turbo decoding, as the system may exhibit multiple fixed points or even limit cycles.

The stochastic approach, on the other hand, views the input and output of a decoder as random processes and tries to characterize their statistics. Notable examples of the stochastic approach include the methods given in [4]-[7]. These methods all use a single statistical parameter to characterize the input or output signal in each decoding phase or iteration. Using the SISO models, [4]-[7] are able to explain a number of important features of turbo decoding. In particular, the so-called Extrinsic Information Tranfer (EXIT) chart of [4] and similar charts in [5], [6] are found particularly useful in understanding and quantifying the dynamic behavior of turbo decoding. Most SISO models require the density function of the extrinsic information to satisfy the so-called *symmetry condition* [3]. However, as we will point out that when approximate models are used, following the symmetry condition strictly may lead to inadequate modeling.

This paper also embarks on the stochastic approach. Our study, however, aims to answer a number of important questions which are not addressed by the existing work on the stochastic approach. Namely, we want to know the following:

1) When is the stochastic approach meaningful?
2) What statistics are needed to model a turbo decoder?
3) How do we compute these statistics?
4) How does the dynamics of a turbo decoder (or turbo dynamics for short) behave?

We first prove that, when the received signal is subject to additive white Gaussian noise (AWGN) and the interleaver is chosen randomly, the turbo decoding output for each iteration approaches an ergodic random process when the block size approaches infinity. We then show that decoding output for each iteration, when expressed using a scaled logarithmic likelihood ratio (SLLR), is well approximated using a Gaussian distribution. Combining the two results above, we can model a turbo decoder using two input parameters and two output parameters (corresponding to the means and variances of the input and output). Using this model, we have discovered that a typical decoding process is much more intricate than previously known.

II. Ergodic Properties

A. Scaled Log-likelihood Ratios

We first introduce the notion of *scaled log-likelihood ratio* (SLLR). Given any signal s which is a noisy version of a binary signal x with elements $x_i \in \{-1,\ 1\}$, recall that its LLR, denoted by L_s, is defined as

$$L_{s,i} = \log \frac{P(s_i|x_i = 1)}{P(s_i|x_i = -1)}$$

Its SLLR, denoted by $\mathbf{S}$, is defined as $\mathbf{S}_i = x_i L_{s,i}/2$.

When a (received) signal r is subject to AWGN, it can be shown that its SLLR has a Gaussian distribution with mean and variance satisfying a unique relationship: $\mu_r = \sigma_r^2$.

However, as we will see later, the relationship above no longer holds for extrinsic signals (and the *a prior* signals in later iterations). Because of this, we introduce the notion of *mean-to-variance ratio* (MVR): $d = \mu/\sigma^2$.

B. Log-MAP Decoding

The well-known Log-MAP decoding algorithm [1] takes an *a priori* signal a and a received signal r, and produces an extrinsic signal e and an *a posteriori* signal d. Their LLR expressions are denoted by L_a, L_r, L_e and L_d respectively. Their SLLR expressions are denoted by $\mathbf{A}$, $\mathbf{R}$, $\mathbf{E}$ and $\mathbf{D}$ respectively, with means and variances μ_a and σ_a^2, etc.

Log-MAP decoding can be interpreted as follows: Let

$$u = \{u_i : i = 1, 2, \ldots, n, u_i \in \{-1,\ 1\}\},$$

$$x = \{x_{ij} : i = 1, 2, \ldots, n; j = 1, 2, \ldots, m, x_{ij} \in \{-1,\ 1\}\}$$

represent the information and code signals respectively. The coding rate is $1/m$ in this case. Let $r = x + w$ be a received signal, where w is a zero-mean AWGN with variance σ_a^2. The Log-MAP decoder computes the following LLR:

$$L_{e,k} = L_{d,k}^{+} - L_{d,k}^{-} - L_{a,k} - \sum_{j \in J_k} L_{r,k,j},$$

where

$$L_{d,k}^{\pm} = \ln \sum_{v: v_k = \pm 1} \exp\Big(\sum_{i: v_i \neq u_i} v_i L_{a,i} + \sum_{\substack{i,j:\\ y_{i,j} \neq x_{i,j}}} y_{i,j} L_{r,i,j}\Big),$$

L_a and L_r are the *a priori* LLR for u and the LLR for r respectively, and J_k corresponds to the systematic bit.

Without loss of generality, we assume that u is an all-one sequence. Using the SLLR expressions, it can be shown that

$$\begin{aligned} \mathbf{E} &= \ln\left(1 + \sum_{t=1}^{T_1} \exp(-C_t^{(1)}\mathbf{A} - C_t^{(2)}\mathbf{R})\right) \\ &\quad - \ln\left(\sum_{t=1}^{T_2} \exp(-C_t^{(3)}\mathbf{A} - C_t^{(4)}\mathbf{R})\right), \qquad (1) \end{aligned}$$

where $C_t^{(i)}$ are row vectors with 0's and 1's and $T_1, T_2 \geq 0$.

C. Asymptotic Behavior of Log-MAP Decoding

From the analysis above, it is clear that the output signal of MAP decoding can be modelled as a random process. The key question we now ask is how to model this random process when the code block size is very large. Our first main result is given below.

Theorem 1: Given a convolutional code with an infinite block size, suppose the SLLR of the received signal $\mathbf{R}$ and the SLLR of the *a priori* signal $\mathbf{A}$ are ergodic, and $\mathbf{R}$ and $\mathbf{A}$ are both independent by themselves and independent of each other. Then, the outputs of the Log-MAP decoder (i.e., $\mathbf{D}$ and $\mathbf{E}$) are both ergodic random processes.

The implications of the property above are important: When the received signal is subject to AWGN (which is ergodic) and $\mathbf{A}$ is ergodic, the result above says that the statistics of $\mathbf{E}_k$ are independent of k and can be computed using a *single* realization of $\mathbf{E}$, i.e., solving only a single (but long) Log-MAP decoding.

D. Stochastic Modelling of Turbo Decoding

We now want to generalize the ergodicity result in Theorem 1 to turbo decoding. Again, we assume $n \to \infty$.

From the above analysis of Log-MAP decoding, we understand that if the received signal is subject to AWGN and the SLLR of the *a priori* signal, $\mathbf{A}$, is an independent ergodic random process, then the SLLR of the extrinsic signal, $\mathbf{E}$, is also ergodic. In turbo decoding, we start with $\mathbf{A} = 0$, which is Gaussian with $\mu_a = \sigma_a = 0$. Therefore, it is natural to conjecture that the SLLR of the extrinsic signal in every iteration is an ergodic random process. It turns out that this is generally incorrect because the extrinsic signal is "locally" correlated. It is easy to imagine that a non-stationary $\mathbf{A}$ is possible if a "bad" interleaver is used.

Fortunately, the correlation in $\mathbf{E}$ decays. Therefore, if the interleaver has a "good" spreading property, the interleaved extrinsic signal, which becomes the *a priori* signal, should be no longer correlated "locally." Since $\mathbf{E}_k$ depends only on those $\mathbf{A}_i$ which are "local" to k, the interleaved extrinsic signal is effectively an uncorrelated signal.

To understand how well an interleaver works, we introduce the notion of a *spreading factor*. Given an interleaver T of size n, its spreading factor S_T is given by:

$$S_T = \min_S \{S : 1 \leq i, j \leq n; |i - j| \leq S \Rightarrow |T_i - T_j| > S\}$$

Lemma 1: Given any $S > 0$, if an interleaver T of size n is chosen randomly, then

$$P(S_T \geq S) \to 1, \quad \text{as } n \to \infty. \qquad (2)$$

The lemma above leads us to our second main result.

Theorem 2: Given a turbo code with block size n, suppose a random interleaver T is used and the received signal is subject to AWGN. Denote by $\mathbf{E}(\ell, n)$ the SLLR of the extrinsic signal from the ℓ-th iteration of Log-MAP decoding. Then, for any $\ell \geq 1$, $\mathbf{E}(\ell, n)$ approaches an ergodic random process as $n \to \infty$.

In the above and the rest of the paper, the number of iterations refers to the number of Log-MAP decoding processes, rather than the number of turbo cycles.

To demonstrate the ergodicity of $\mu_e(\ell, n)$, we simulate a 1/3-rate turbo code with $G(D) = (1,\ 5/7)$, $E_b/N_0 = 0.5$dB and pseudo-random interleaver. For each n and ℓ, many runs of $\mu_e(\ell, n)$ are simulated. These values are used to compute a lower bound and upper bound for $\mu_e(\ell, n)$. The lower bound is the average of these $\mu_e(\ell, n)$ values minus their standard deviation, whereas the upper bound is the the average of these $\mu_e(\ell, n)$ values plus their standard deviation. The size of the gap between the lower and upper bound shows how well $\mu_e(\ell, n)$ converges as $n \to \infty$. The simulation results are shown in Fig. 1.

III. Gaussian Approximations

In this section, we study Gaussian approximations for Log-MAP decoding and turbo decoding.

A. Log-Sum of Lognormal Distributions

Given a set of Gaussian-distributed random variables X_i with means μ_i and variances σ_i^2, $i = 1, 2, \ldots, n$, we define

$$Z = \ln \sum_{i=1}^{n} \exp(X_i)$$

Then, each $\exp(X_i)$ is a lognormal distribution and $\exp(Z)$ is a sum of lognormal distributions (SLND). We will call Z a log-sum of lognormal distributions (LSLND).

The statistical properties of SLND have been well studied. It is well known that the distribution of a SLND can be

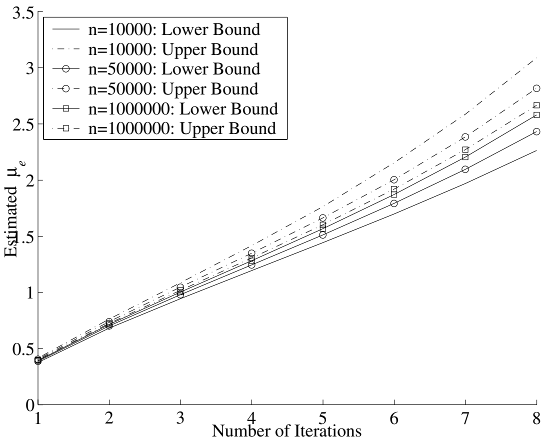

Fig. 1. Convergence of the means of $\mathbf{E}(\ell, n)$

closely approximated using a lognormal distribution when X_i are independent with the same mean and variance. Correspondingly, Z is well approximated by a Gaussian distribution. Although no closed-form description is given on the distribution of a SLND or LSLND, a number of methods are available for computing the mean and variance (or equivalently the first and second moments) of Z; see, e.g., [8] for a summary. The Gaussian approximation works well when X_i are weakly correlated and their statistical parameters are not significantly different.

B. Gaussian Approximations for Log-MAP Decoding

We now analyze the distribution of $\mathbf{E}$ (the SLLR of the extrinsic signal) for Log-Map decoding. Consider the expression $\mathbf{E}_k$ from (1). Recall that $\mathbf{R}$ is a vector of (independent) Gaussian distributions when the received signal r is subject to AWGN. Suppose $\mathbf{A}$ is also Gaussian distributed. Then, $\mathbf{E}$ is the difference between the two LSLNDs. This observation is summarized below:

For a convolutional binary code with an infinite block size, if the received signal is subject to AWGN and the SLLR of the a priori *signal is a Gaussian distribution, then the SLLR of the extrinsic signal can be well approximated using a Gaussian distribution.*

Although the result above says that the SLLR of the extrinsic signal can be approximated using a Gaussian distribution, its MVR is no longer equal to 1 in general. Therefore, it is insufficient to characterize the output signal by its SNR. Instead, two parameters, the mean and variance of the SLLR need to be used. We conclude the following:

A Log-MAP decoder can be approximated as a mapping $\mathcal{M}$ from (μ_r, μ_a, σ_a) to (μ_e, σ_e). If μ_r is suppressed, the decoder is simply a mapping from (μ_a, σ_a) to (μ_e, σ_e).

C. Simplified Model for Log-MAP Decoding at High SNRs

To help understand the behavior of Log-Map decoders, we derive a simplified model at high SNRs.

Lemma 2: Suppose $\mathbf{A}$ is approximately Gaussian distributed with $d_a \geq 1$, the received signal r is subject to AWGN and $\mu_a \gg \mu_r$. Also suppose that the convolutional code is recursive, $n \to \infty$, and u is an all-one sequence. For each k, denote by $V_{k,2}$ the set of weight-2 information sequences v with $v_k = -1$ which are terminating. Then,

$$\mathbf{E}_k \approx -\ln \sum_{v:\ v \in V_{k,2}} \exp(-\mathbf{A}_t - \sum_{\substack{i,j: j \notin J_k \text{if} i=k, \\ y_{i,j}=-1}} \mathbf{R}_{i,j}) \quad (3)$$

where t is the index for the other information bit with $v_t = -1$ and J_k is as defined earlier.

D. Gaussian Approximations for Turbo Decoding

From the analysis of Log-MAP decoding, we understand that if $\mathbf{A}$ and $\mathbf{R}$ are independent Gaussian white noises, then $\mathbf{E}$ is well approximated using a stationary process with a Gaussian distribution. In turbo decoding, we start with $\mathbf{A} = 0$. Therefore, $\mathbf{E}$ from the first iteration is well approximated using a Gaussian distribution. If a random interleaver is used, $\mathbf{A}$ for the next iteration will become effectively independent when the block size is large. Hence, Gaussian approximations can continue, i.e., $\mathbf{E}$ in every iteration is well approximated using a Gaussian distribution.

To formalize our analysis, we define the *Gaussian approximation model* for a turbo decoder as follows:

For each decoding iteration, the SLLR of the a priori *signal is well approximated using an uncorrelated Gaussian distribution and the SLLR of the extrinsic signal is well approximated using a (locally correlated) Gaussian distribution.*

To check the validity of the Gaussian approximation model, we compare it to turbo decoding using a $1/3$-rate turbo code with $G(D) = (1, 15/13)$, $n = 500{,}000$, and a pseudo-random interleaver. We take $E_b/N_0 = 0.3$dB. Twelve iterations are used. For the method using the Gaussian approximation model, $\mathbf{A}$ in each iteration is chosen to be a Gaussian distribution with the same mean and variance as those for the $\mathbf{A}$ fed into the corresponding iteration of the turbo decoding. Fig. 2 compares the means and variances of $\mathbf{E}$ in the two cases.

IV. Dynamics of Turbo Decoding

Recall that when $n \to \infty$, the decoding output for each iteration becomes an ergodic random process. Each decoding instance is a realization of the random process and the decoded signal has the same statistics (with probability 1). As $n \to \infty$, the state of the decoded signal either converges at a finite *stable fixed point* or diverges. The former scenario occurs only at a low SNR value, leading to a large BER. In contrast, the latter scenario leads to an ever increasing SNR and thus arbitrarily low BER. However, when the block size is finite, this trend can not be sustained indefinitely, causing the decoding process to converge at a high SNR point.

Turbo dynamics are in fact much more complex than indicated by the two stable fixed points. The complete picture is illustrated in Fig. 3. The turbo code used here is

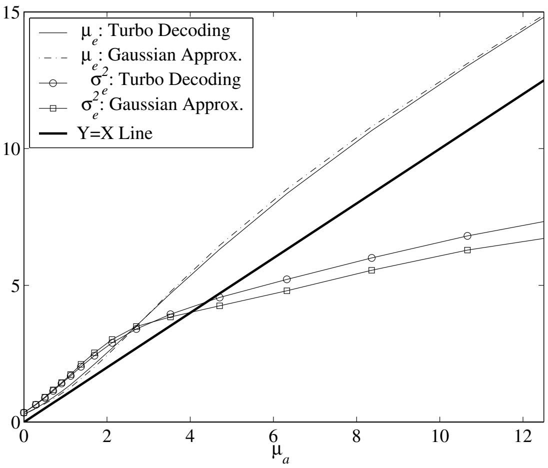

Fig. 2. Validity of Gaussian Approximation

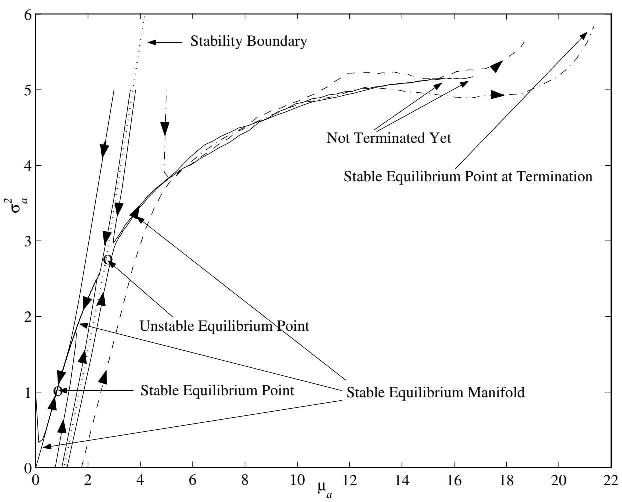

Fig. 3. Dynamics of Turbo Decoding

a 1/3-rate code with $G(D) = (1,\ 5/7)$, $E_b/N_0 = -0.1$dB, $n = 10^6$, and a pseudo-random interleaver. We see from the figure that there is a *stable equilibrium manifold* at which every turbo decoding trajectory converges, regardless of the initial point. On this manifold, there is a stable fixed point with a low SNR which is paired with an unstable fixed point above it. The whole state space (i.e., the space of $(\mu_a,\ \sigma_a^2)$) is divided into two regions by a *stability boundary* which intersects the unstable fixed point. When the initial state is to the left of the stability boundary, the decoding trajectory quickly moves to the stable equilibrium manifold and then converges at the stable fixed point with a low SNR. When the initial state is to the right of the stability boundary, the decoding trajectory again approaches the stable equilibrium manifold very quickly, then moves to the right right along the manifold for a while but eventually converges at a stable fixed point or region with a high SNR.

The scenario in Fig. 3 happens when E_b/N_0 is below a certain threshold. If E_b/N_0 exceeds this threshold, the stable and unstable equilibrium points coalesce and disappear and the decoding trajectory moves to a high SNR region.

$G(D)$ in octal	E_b/N_0 Threshold
(1, 5/7) (4-state)	0.04 dB
(1, 15/13) (8-state)	-0.04 dB
(1, 33/23) (16-state)	-0.01 dB
(1, 35/23) (16-state)	-0.01 dB
(1, 37/25) (16-state)	-0.15 dB
(1, 41/77) (32-state)	-0.21 dB
(1, 113/111) (64-state)	-0.22 dB

TABLE I

E_b/N_0 THRESHOLDS FOR AVOIDING LOW SNR EQUILIBRIUM POINTS

A. Low SNR Analysis

The low SNR stable fixed point is invariant when n becomes very large due to the ergodicity of $\mathbf{E}$. Thus, they can be easily found by simulating a single long block code. By adjusting the value of E_b/N_0, we can easily search for the threshold at which the low SNR equilibrium point vanishes; see Table I for turbo codes with coding rate $1/3$.

B. High SNR Analysis

In order to understand the behavior of turbo decoding at a high SNR, it is necessary to study the simplified model (3). We have the following result.

Lemma 3: Given a 1/2-rate systematic, recursive, binary convolutional code $G(D) = (1,\ P(D)/Q(D))$ with an infinitely long block size, suppose $P(D)$ and $Q(D)$ are coprime and monic polynomials with the same degree. Denote by w_0 the weight of the parity sequence corresponding to the shortest weight-2 information sequence. Suppose the SLLR of the received signal $\mathbf{R}$ and SLLR of the *a priori* signal $\mathbf{A}$ are independent Gaussian white noises with means μ_r and μ_a and variances σ_r^2 and σ_a^2, respectively. Then, the simplified model (3) of the SLLR of the extrinsic signal $\mathbf{E}$, when only the first N weight-2 sequences on each side of the trellis are included, can be rewritten as

$$\mathbf{E}_k = -\ln(\exp(-L_+) + \exp(-L_-)) \tag{4}$$

where

$$\begin{aligned} L_\pm &= \rho_\pm + L_\pm^{(1)} \\ L_\pm^{(i)} &= \delta_\pm^{(i)} - \ln(\exp(-\eta_\pm^{(i)}) + \exp(-L_\pm^{(i+1)})) \\ &\qquad\qquad i = 1, 2, \ldots, N-1 \\ L_\pm^{(N)} &= \delta_\pm^{(N)} + \eta_\pm^{(N)} \end{aligned} \tag{5}$$

where $\rho_\pm, \delta_\pm^{(i)}, \eta_\pm^{(i)}$ are independent Gaussian variables with the following distributions:

$$\begin{aligned} \rho_\pm &\sim \mathcal{N}(\mu_r,\ \sigma_r^2) \\ \delta_\pm^{(i)} &\sim \mathcal{N}((w_0-2)\mu_r,\ (w_0-2)\sigma_r^2) \\ \eta_\pm^{(i)} &\sim \mathcal{N}(2\mu_r + \mu_a,\ 2\sigma_r^2 + \sigma_a^2) \end{aligned} \tag{6}$$

Fig. 4 demonstrates a typical decoding behavior at a high SNR. This example uses $G(D) = (1,\ 5/7)$. We consider the case of $E_b/N_0 = 0.5$ first. We see that when σ_a is small, $\sigma_e > \sigma_a$. As σ_a becomes large, $\sigma_e < \sigma_a$. This implies that there is a crossover point which is a stable fixed point for σ_a where $\sigma_e = \sigma_a \approx 2$. At this point, $\mu_e - \mu_a \approx 1.1$.

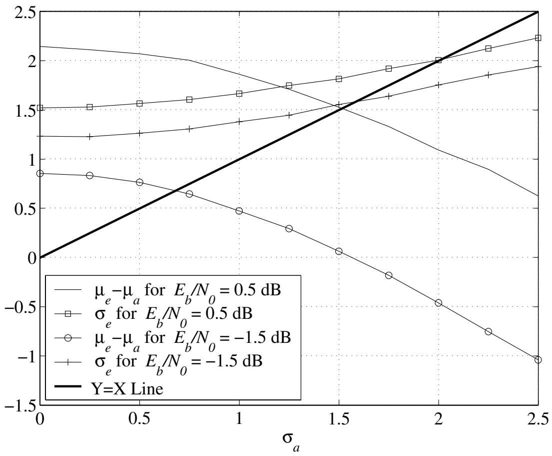

Fig. 4. **E** vs. **A** at a High SNR

Similar observations apply to the case of $E_b/N_0 = -1.5$. In this case, the stable fixed point occurs at $\sigma_a \approx 1.6$ and at this point, $\mu_e - \mu_a \approx 0$. This is summarized below:

As the iteration number increases, σ_e and $\mu_e - \mu_a$ always approach constant values for any initial values of σ_e and μ_a. Moreover, if $\mu_e - \mu_a$ approaches a positive constant, the SNR of the extrinsic signal (and thus the a posteriori *signal) will increase indefinitely.*

C. The Inadequacy of SISO Models

In this section, we show the inadequency of SISO models based on SNR or mutual information of $\mathbf{E}$. More precisely, we point out the "strange" phenomenon that a Log-MAP decoder can take an input with a very high SNR or very good mutual information and produce an output with a much lower SNR or poorer mutual information.

Consider $\mathbf{A}u = L_a/2$ which is the true input signal. Using Gaussian approximation on $\mathbf{A}$, the probability density function for $\mathbf{A}u$ is given by

$$p_a(\xi|u_i = \pm 1) = \frac{1}{\sqrt{2\pi}\sigma_a}\exp(-(\xi \mp \mu_a)^2/2\sigma_a^2)$$

The mutual information between u and L_a is given by [4]:

$$\begin{aligned} I_a &= \frac{1}{2}\sum_{u_i \in \{1,-1\}} \int_{-\infty}^{\infty} p_a(\xi|u_i) \\ &\qquad \log_2 \frac{2p_a(\xi|u_i)}{p_a(\xi|u_i = 1) + p_a(\xi|u_i = -1)} d\xi \\ &= 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-(t - \frac{\mu_a}{\sigma_a})^2/2) \\ &\qquad \log_2(1 + \exp(-\frac{\mu_a}{\sigma_a} 2t))dt \end{aligned}$$

Since I_a is a function of μ_a/σ_a (the SNR of $\mathbf{A}$), we will denote I_a as $I_a(\mu_a/\sigma_a)$. It can be shown that $I_a(\cdot)$ is monotonically increasing with $I_a(0) = 0$ and $I_a(\infty) = 1$.

To demonstrate the "strange" phenomenon mentioned above, we consider $(\mu_a, \sigma_a^2) = (1, 0)$ which has the input SNR$= \infty$ and $I_a = 1$. However, the output of Log-MAP decoding has $\mu_e < 2$ and $\sigma_e^2 > 1$, meaning that the output SNR or mutual information is much worse than the input. If we continue with more iterations, the decoding trajectory will eventually lead to the low SNR fixed point; see Fig. 3.

In [3], it is proved that the LLRs in a belief propagation algorithm obeys the so-called *symmetry condition*. This condition states that the density function f of an LLR obeys

$$f(x) = e^x f(-x), \quad \forall x \in (-\infty, \ \infty)$$

When applied to an SLLR with a Gaussian distribution, the symmetry condition requires $\sigma^2 = m$. It seems that the use of two parameters (σ and m) contradicts the symmetry condition. However, the answer lies in the fact that Gaussian distributions are only approximations. That is, the Gaussian approximation aims to give a good approximate model by scarifying the symmetry condition. We argue that the violation of the symmetry condition does not create a serious problem. To see this, we note that the distribution of a (scaled) LLR is mostly one-sided and decays exponentially fast as $|x| \to \infty$. Thus, it is not important to enforce $f(x) = e^x f(-x)$ when either $f(x)$ or $f(-x)$ is very small. What is much more important is how to capture the portion of the density with a significant mass distribution using a simple model, which is achieved by Gaussian approximation.

V. Conclusions

In this paper, we have proposed a stochastic approach to the modelling and analysis of turbo decoding. Two key results, ergodicity and Gaussian approximations, have been established which lead to some new understanding of turbo decoding. In particular, we are able to build a simple dynamic model for turbo decoding and reveal the intricate behavior of turbo decoding unknown previously.

References

[1] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," in *Proc. Int. Conf. Commun.*, Geneva, Switzerland, pp. 10641070, May 1993.

[2] T. Richardson, "The geometry of turbo-decoding dynamics," *IEEE Trans. Inform. Theory*, vol. 46, no. 1, pp. 923, 2000.

[3] 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, 2001.

[4] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 17271737, 2001.

[5] D. Divsalar, S. Dolinar, and F. Pollarar, "Iterative turbo decoder analysis based on density evolution," *IEEE J. Selected Areas in Commun.*, vol. 19, no. 5, pp. 891907, 2001.

[6] H. El Gamal and A. R. Hammons, "Analyzing the turbo decoder using the Gaussian approximation," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 671-686, Feb. 2001.

[7] J. W. Lee and R. E. Blahut, "Generalized EXIT chart and BER analysis of finite-length turbo codes," *Proc. GlobeCom 2003*, San Fransisco, Dec. 2003.

[8] N. C. Baulieu, A. Abu-Dayya and P. J. McLane, "Estimating the distribution of independent lognormal random variables," *IEEE Trans. Commun.*, vol. 43, no. 12, pp. 2869-2873, 1995.