On the Gaussian Approximation in the Analysis of Iterative MIMO Processing

Yibo Jiang, Ralf Koetter, and Andrew C. Singer
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA
email: {yjiang, koetter, acsinger}@uiuc.edu

Abstract— **In this paper, we study the density of extrinsic information at the output of the demodulator of an iterative MIMO processing system. In particular, if the *a priori* information is Gaussian distributed, the conditions under which the extrinsic information is also approximately Gaussian distributed are found.**

I. Introduction

Iterative processing technique has been widely used in communications over multiple-antenna channels, e.g., a near-capacity achieving scheme [1], Turbo-BLAST [2], space-time turbo codes [3], [4], low-density parity-check (LDPC) modulation [5].

The extrinsic information transfer chart (EXIT chart) [6] is a useful analysis and design tool for iterative processing systems, e.g., studying the threshold effects of iterative multi-input multi-output (MIMO) processing [7], designing good LDPC code degree sequences [5].

An implicit assumption in the EXIT chart technique is that, if the conditional density of the *a priori* information is a density function $p(x, \alpha = \alpha_1)$ which has a single parameter α, the conditional density of extrinsic information at the output of a constituent decoder can be reasonably approximated by $p(x, \alpha = \alpha_2)$. Then the decoding process is essentially characterized by the evolution of α. In the analysis of turbo decoding [6], the one parameter density function is

$$\mathcal{N}(\frac{\sigma_L^2}{2}b, \sigma_L^2) \qquad (1)$$

with σ_L as the parameter α and b the corresponding information bit. It was observed in [8] that the conditional density of extrinsic information in turbo decoding is approximately a Gaussian density. Gaussian approximation is also used in other iterative decoding analysis methods, e.g., the SNR approach for turbo decoding [9] and LDPC decoding analysis [10].

In this paper, we will consider the iterative MIMO processing system in Fig. 1 and study the conditional density of extrinsic information at the output of the demodulator. The space time modulation is constrained to be linear in information bits. Our contribution is that we find certain

This work was supported in part by the National Science Foundation under grants NSF CCR 99-79381, NSF ITR 00-85929 and NSF CCR 99-84515, and in part by the Motorola Center for Communications.

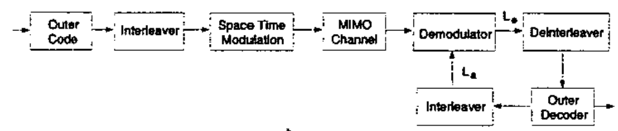

Fig. 1. A bit interleaved serially concatenated coding system in a block fading MIMO channel and the corresponding iterative decoding system.

conditions under which the conditional density of extrinsic information at the output of the demodulator can be well approximated by the type of Gaussian density in Eq. (1) when the *a priori* information is distributed according to Eq. (1) with certain σ_L. If the outer decoder also satisfies such a density approximation, the use of various iterative processing analysis methods which are based on Gaussian approximation is justified.

We first briefly introduce some concepts of iterative decoding and EXIT charts. To decode bit $\mathbf{b}_i$, the soft-input soft-output (SISO) demodulator takes the channel output Y and the *a priori* input $L_a(\mathbf{b}_k)$, $1 \leq k \leq K$, and outputs the extrinsic information $L_e(\mathbf{b}_i)$. The extrinsic information sequence is interleaved and passed to the SISO outer decoder which performs similar computations. Its output sequence is interleaved and passed to the demodulator as the *a priori* input in the next iteration.

The EXIT curve for the demodulator is defined as follows. Define $I_a = I(\mathbf{b}_i; \mathbf{L_a}(\mathbf{b}_i))$ where $\mathbf{L_a}(\mathbf{b}_i)$ is the *a priori* information for bit $\mathbf{b}_i$, and $I_e = I(\mathbf{b}_i; \mathbf{L_e}(\mathbf{b}_i))$ where $\mathbf{L_e}(\mathbf{b}_i)$ is the extrinsic information of bit $\mathbf{b}_i$. The demodulator can be viewed as a nonlinear system mapping the input mutual information to the output mutual information. Given the signal-to-noise ratio (SNR), I_e can be viewed as a function of I_a, $I_e = T_1(I_a)$, which is the EXIT curve for the demodulator. Except for a few special channels, it is intractable in general to derive the analytical form of an EXIT curve.

This paper is organized as follows. In Section II, we first describe the signal model, then present necessary and sufficient conditions for the invariance of density and a constant EXIT curve. Later, we will analyze the conditional density of extrinsic information, and find several conditions such that Gaussian approximation is reasonable. Section III contains our conclusions.

Our notation will be as follows: capital letters denote matrices; underscores denote vectors; boldfaced letters denote random objects; daggers denote complex conjugate transpose; asterisks denote conjugation; primes denote transpose;

II. Density Analysis

A. Signal Model

In the MIMO channel in Fig. 1, there are t transmit antennas and r receive antennas. The channel state information (CSI) matrix $\mathbf{H} = [\mathbf{h}_{i,j}] \in \mathbb{C}^{r\times t}$ is known to the receiver only and does not change within a block of T symbols. The coefficient $\mathbf{h}_{i,j}$ denotes the complex channel gain between transmit antenna j and receive antenna i and experiences i.i.d. frequency flat fading for all pairs of i and j. We assume $\mathbf{h}_{i,j}$ is distributed according to $\mathcal{CN}(0,1)$. From block to block, $\mathbf{H}$ changes independently. Let $\mathbf{X} \in \mathbb{C}^{t\times T}$ indicate the transmitted signal matrix, $\mathbf{Y} \in \mathbb{C}^{r\times T}$ be the received signal matrix such that the discrete-time baseband equivalent channel model is

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{N} \tag{2}$$

where $\mathbf{N} \in \mathbb{C}^{r\times T}$ is an additive noise matrix with i.i.d. entries $\mathbf{n}_{i,j} \sim \mathcal{CN}(0, 2\sigma^2)$.

In this paper, for ease of analysis, we focus on a class of linear dispersion (LD) codes [11] whose codeword matrices X are linear in the information bits, i.e.

$$X = \sum_{k=1}^{K} b_k A_k \tag{3}$$

where $\{b_k \in \{-1,+1\}\}_{k=1}^{K}$ are the information bits and $A_k \in \mathbb{C}^{t\times T}, 1 \le k \le K$ are dispersion matrices. For convenience, we call them bit-linear LD (BL-LD) codes. It is easy to verify that with PAM or QAM, space-time block codes (STBC) [12] are BL-LD codes.

We further constrain the dispersion matrices to satisfy the property

$$A_k A_k^\dagger = \frac{\varepsilon}{t} I_t, \quad 1 \le k \le K \tag{4}$$

where ε is the bit energy. The constraint in Eq. (4) is satisfied by various space time codes, e.g., Alamouti scheme [13], STBCs from orthogonal designs [12], LD codes [11]. In order for Eq. (4) to have solutions, we assume $T \ge t$.

B. Necessary and Sufficient Conditions on the Invariance of Density

First we describe the computation of extrinsic information. The extrinsic information is defined as

$$L_e(\mathbf{b}_i) = \ln\frac{P(\mathbf{b}_i = 1|\mathbf{Y} = Y, \mathbf{H} = H)}{P(\mathbf{b}_i = -1|\mathbf{Y} = Y, \mathbf{H} = H)} - L_a(\mathbf{b}_i). \tag{5}$$

In the above computation, the *a priori* probability $P(\underline{\mathbf{b}} = \underline{b}) = \prod_{k=1}^{K} P(\mathbf{b}_k = b_k)$ is updated by the *a priori* information

$$P(\mathbf{b}_k = -1) = 1 - P(\mathbf{b}_k = 1) = \frac{1}{1 + \exp\{L_a(\mathbf{b}_k)\}}. \tag{6}$$

Before analyzing and approximating the density of extrinsic information, we present a theorem which states necessary and sufficient conditions on dispersion matrices for the invariance of density and a constant EXIT curve.

Theorem 1: Let C be a BL-LD code

$$C = \{X : X = \sum_{k=1}^{K} b_k A_k, b_k \in \{-1,+1\}, \forall k\}$$

with dispersion matrices $A_k \in \mathbb{C}^{t\times T}, 1 \le k \le K$. If and only if

$$A_i A_j^\dagger + A_j A_i^\dagger = \mathbf{0}, \forall i, \forall j, i \ne j, \tag{7}$$

the following two statements are true:

1) For all k, the density of extrinsic information $\mathbf{L_e}(\mathbf{b_k})$ conditioned on the information bit $\mathbf{b_k} = b_k$ is invariant with respect to the *a priori* information.
2) The EXIT curve $I_e = T_1(I_a)$ is constant.

Proof: We first prove the direct part of the claim.

Assume $A_i A_j^\dagger + A_j A_i^\dagger = \mathbf{0}$. Without loss of generality, we consider the SISO decoding of the bit $\mathbf{b}_1$. Define the vector $\underline{b}$ as $\underline{b} = [b_1, \cdots, b_K]$. Define the vector $\tilde{\underline{b}}$ as $\tilde{\underline{b}} = [b_2, \cdots, b_K]$. Define the matrix $\tilde{X}$ as $\tilde{X} = \sum_{k=2}^{K} b_k A_k$. This implies that the codeword matrix X for $\underline{b}$ satisfies $X = \tilde{X} + b_1 A_1$. We have

$$\begin{aligned} L_e(\mathbf{b}_1) &= \ln\frac{P(\mathbf{b}_1 = 1|\mathbf{Y} = Y, \mathbf{H} = H)}{P(\mathbf{b}_1 = -1|\mathbf{Y} = Y, \mathbf{H} = H)} \\ &= \ln\frac{\sum_{\underline{b}:b_1=1} P(\underline{\mathbf{b}} = \underline{b}) p(\mathbf{Y} = Y|\mathbf{H} = H, \underline{\mathbf{b}} = \underline{b})}{\sum_{\underline{b}:b_1=-1} P(\underline{\mathbf{b}} = \underline{b}) p(\mathbf{Y} = Y|\mathbf{H} = H, \underline{\mathbf{b}} = \underline{b})} \\ &= \ln\frac{\sum_{\underline{b}:b_1=1} P(\underline{\mathbf{b}} = \underline{b}) \exp\{-\frac{\mathrm{tr}\{(Y-HX)(Y-HX)^\dagger\}}{2\sigma^2}\}}{\sum_{\underline{b}:b_1=-1} P(\underline{\mathbf{b}} = \underline{b}) \exp\{-\frac{\mathrm{tr}\{(Y-HX)(Y-HX)^\dagger\}}{2\sigma^2}\}}. \end{aligned} \tag{8}$$

Because $A_1 A_j^\dagger + A_j A_1^\dagger = \mathbf{0}$, $\forall j \ne 1$, we have

$$\begin{aligned} &HXX^\dagger H^\dagger = H(\tilde{X} + b_1 A_1)(\tilde{X}^\dagger + b_1 A_1^\dagger) H^\dagger \\ &= H[\tilde{X}\tilde{X}^\dagger + b_1(\tilde{X}A_1^\dagger + A_1\tilde{X}^\dagger) + A_1 A_1^\dagger] H^\dagger \\ &= H[\tilde{X}\tilde{X}^\dagger + b_1 \sum_{k=2}^{K} b_k (A_k A_1^\dagger + A_1 A_k^\dagger) + A_1 A_1^\dagger] H^\dagger \\ &= H[\tilde{X}\tilde{X}^\dagger + A_1 A_1^\dagger] H^\dagger. \end{aligned} \tag{9}$$

Since the information bits are independent and $P(\mathbf{b}_1 = 1) = P(\mathbf{b}_1 = -1) = 0.5$, after some manipulation, we obtain

$$\begin{aligned} L_e(\mathbf{b}_1) &= \frac{\mathrm{tr}\{Y(HA_1)^\dagger + (HA_1)Y^\dagger\}}{\sigma^2} \\ &+ \ln\frac{\sum_{\tilde{\underline{b}}} P(\tilde{\underline{\mathbf{b}}} = \tilde{\underline{b}}) \exp\{\frac{\mathrm{tr}\{-H\tilde{X}\tilde{X}^\dagger H^\dagger + Y(H\tilde{X})^\dagger + (H\tilde{X})Y^\dagger\}}{2\sigma^2}\}}{\sum_{\tilde{\underline{b}}} P(\tilde{\underline{\mathbf{b}}} = \tilde{\underline{b}}) \exp\{\frac{\mathrm{tr}\{-H\tilde{X}\tilde{X}^\dagger H^\dagger + Y(H\tilde{X})^\dagger + (H\tilde{X})Y^\dagger\}}{2\sigma^2}\}} \\ &= \frac{\mathrm{tr}\{Y(HA_1)^\dagger + (HA_1)Y^\dagger\}}{\sigma^2}. \end{aligned} \tag{10}$$

Since $Y = HX + N = \sum_{k=1}^{K} b_k H A_k + N$, we further observe that

$$
\begin{aligned}
&L_e(\mathbf{b}_1)\\
&= \frac{\mathrm{tr}\{\sum_{k=1}^{K} b_k H(A_k A_1^\dagger + A_1 A_k^\dagger)H^\dagger + N A_1^\dagger H^\dagger + H A_1 N^\dagger\}}{\sigma^2}\\
&= \frac{\mathrm{tr}\{2 H A_1 A_1^\dagger H^\dagger\}}{\sigma^2} b_1 + \frac{\mathrm{tr}\{N A_1^\dagger H^\dagger + H A_1 N^\dagger\}}{\sigma^2}.
\end{aligned} \tag{11}
$$

In the second step, we used the conditions $A_1 A_j^\dagger + A_j A_1^\dagger = \mathbf{0}$, $\forall j \neq 1$. Hence the extrinsic information $\mathbf{L_e}(\mathbf{b}_1)$ is a function of $\mathbf{b}_1$, $\mathbf{H}$ and $\mathbf{N}$, and is independent of the *a priori* information $\mathbf{L_a}(\mathbf{b}_k)$, $2 \leq k \leq K$ which are used to determine the priors of $\mathbf{b}_k$, $2 \leq k \leq K$. Thus, the first statement is proved for bit $\mathbf{b}_1$. For other bits $\mathbf{b}_k$, $2 \leq k \leq K$, the proof is essentially same. It follows that the extrinsic mutual information $I_e = I(\mathbf{b}_k; \mathbf{L_e}(\mathbf{b}_k)|I_a)$ is constant for all values of *a priori* mutual information I_a. Thus the EXIT curve $I_e = T_1(I_a)$ is constant.

Next we prove the only if part of the claim.

Let $\underline{b}\backslash b_k$ denote $[b_1, \cdots, b_{k-1}, b_{k+1}, \cdots, b_K]$. If the conditional density of extrinsic information is invariant with respect to the *a priori* information, the EXIT curve is constant. Assume the EXIT curve $I_e = T_1(I_a)$ is constant. Thus the heights of left and right endpoint are same. Since for $1 \leq k \leq K$

$$I(\mathbf{b}_k; \mathbf{L_e}(\mathbf{b}_k)|I_a = 0) = I(\mathbf{b}_k; \mathbf{Y}|\mathbf{H}),$$

and

$$I(\mathbf{b}_k; \mathbf{L_e}(\mathbf{b}_k)|I_a = 1) = I(\mathbf{b}_k; \mathbf{Y}|\mathbf{H}, \underline{\mathbf{b}}\backslash\mathbf{b_k}),$$

we have

$$I(\mathbf{b}_k; \mathbf{Y}|\mathbf{H}) = I(\mathbf{b}_k; \mathbf{Y}|\mathbf{H}, \underline{\mathbf{b}}\backslash\mathbf{b_k}). \tag{12}$$

By the chain rule for mutual information, we have

$$
\begin{aligned}
&I(\mathbf{b}_k; \mathbf{Y}, \underline{\mathbf{b}}\backslash\mathbf{b_k}|\mathbf{H})\\
=&I(\mathbf{b}_k; \mathbf{Y}|\mathbf{H}) + I(\mathbf{b}_k; \underline{\mathbf{b}}\backslash\mathbf{b_k}|\mathbf{H}, \mathbf{Y})\\
=&I(\mathbf{b}_k; \underline{\mathbf{b}}\backslash\mathbf{b_k}|\mathbf{H}) + I(\mathbf{b}_k; \mathbf{Y}|\mathbf{H}, \underline{\mathbf{b}}\backslash\mathbf{b_k})\\
=&I(\mathbf{b}_k; \mathbf{Y}|\mathbf{H}, \underline{\mathbf{b}}\backslash\mathbf{b_k}).
\end{aligned} \tag{13}
$$

In the third step, we use the independence of the information bits. Combining Eq. (12) and (13), we obtain

$$I(\mathbf{b}_k; \underline{\mathbf{b}}\backslash\mathbf{b_k}|\mathbf{H}, \mathbf{Y}) = 0. \tag{14}$$

Therefore, for $1 \leq k \leq K$, we have

$$
\begin{aligned}
&p(\mathbf{b}_1 = b_1, \cdots, \mathbf{b}_K = b_K|H, Y)\\
=&p(\mathbf{b}_k = b_k|H, Y) p(\underline{\mathbf{b}}\backslash\mathbf{b_k} = \underline{b}\backslash b_k|H, Y).
\end{aligned} \tag{15}
$$

Hence it must be true that

$$p(\mathbf{b}_1 = b_1, \cdots, \mathbf{b}_K = b_K|Y, H) = \prod_{k=1}^{K} p(\mathbf{b}_k = b_k|Y, H). \tag{16}$$

By Theorem 1 of [14], we require $A_i A_j^\dagger + A_j A_i^\dagger = \mathbf{0}$, $\forall i$, $\forall j$, $i \neq j$. ■

One consequence of a constant EXIT curve is that iterative decoding can be done in one iteration.

Next we will give a BL-LD code example such that Theorem 1 is satisfied. As mentioned in Section II-A, if the modulation scheme is PAM or QAM, STBCs from orthogonal designs are BL-LD codes. Furthermore, it is easy to verify that with BPSK or QPSK, we have $A_i A_j^\dagger + A_j A_i^\dagger = \mathbf{0}$, $\forall i$, $\forall j$, $i \neq j$. Therefore, if a STBC from orthogonal designs with BPSK or QPSK is used as the space time modulation scheme in Fig. 1, the density of extrinsic information will be invariant with respect to the *a priori* information, and hence only one iteration of demodulation and decoding is needed.

Now we move on to analyze and approximate the conditional density of extrinsic information. Motivated by Theorem 1, we will first study the density of extrinsic information when the *a priori* information is perfect, i.e., $I_a = I(\mathbf{b}_i; \mathbf{L_a}(\mathbf{b}_i)) = 1$.

C. Density of Extrinsic Information when $I_a = 1$

Without loss of generality, we consider the decoding of b_1. When $I_a = 1$, the decoder has perfect *a priori* information of $b_2, \cdots, b_K$. Therefore, the contribution from $b_2, \cdots, b_K$ can be subtracted from the channel output Y and the signal model essentially reduces to

$$\mathbf{Y} = \mathbf{b_1} \mathbf{H} A_1 + \mathbf{N}, \tag{17}$$

where according to the constraint Eq. (4), matrix A_1 satisfies $A_1 A_1^\dagger = \frac{\epsilon}{t} I_t$. For notational simplicity, we will omit the subscript in Eq. (17).

The extrinsic information for bit b is computed as

$$
\begin{aligned}
L_e(\mathbf{b}) &= \ln \frac{\exp\{-\frac{\mathrm{tr}\{(Y-HA)(Y-HA)^\dagger\}}{2\sigma^2}\}}{\exp\{-\frac{\mathrm{tr}\{(Y+HA)(Y+HA)^\dagger\}}{2\sigma^2}\}}\\
&= \frac{\mathrm{tr}\{Y(HA)^\dagger + HAY^\dagger\}}{\sigma^2}\\
&= 2\frac{\mathrm{tr}\{HAA^\dagger H^\dagger\}}{\sigma^2} b + \frac{\mathrm{tr}\{N(HA)^\dagger + HAN^\dagger\}}{\sigma^2}\\
&\triangleq \gamma(H) b + \eta(N, H).
\end{aligned} \tag{18}
$$

Conditioned on $\mathbf{H} = H$

$$
\begin{aligned}
\eta &= \frac{\mathrm{tr}\{\mathbf{N}(HA)^\dagger + HA\mathbf{N}^\dagger\}}{\sigma^2}\\
&= \frac{1}{\sigma^2} \sum_{j=1}^{r} \sum_{k=1}^{T} [\mathbf{n}_{j,k} (HA)_{j,k}^* + (HA)_{j,k} \mathbf{n}_{j,k}^*]\\
&= \frac{1}{\sigma^2} \sum_{j=1}^{r} \sum_{k=1}^{T} 2 \cdot \mathrm{Re}\{\mathbf{n}_{j,k} (HA)_{j,k}^*\}.
\end{aligned} \tag{19}
$$

Thus conditioned on $\mathbf{H} = H$, η is a Gaussian random variable with zero mean and variance $2\gamma(H)$.

Since $AA^\dagger = \frac{\epsilon}{t} I_t$, we have

$$\gamma(H) = 2\frac{\mathrm{tr}\{HAA^\dagger H^\dagger\}}{\sigma^2} = \frac{2\epsilon}{\sigma^2 t} \mathrm{tr}\{HH^\dagger\}. \tag{20}$$

Conditioned on $\mathbf{b} = b$, the density of extrinsic information can be written as

$$p(x) = \int p(\mathbf{H} = H) \frac{1}{\sqrt{4\pi\gamma(H)}} \exp\left\{-\frac{[x - \gamma(H)b]^2}{4\gamma(H)}\right\} \mathrm{d}H. \tag{21}$$

Because H is a $r \times t$ matrix, H are irrelevant to block length T. From Eq. (20) and (21), it follows that the conditional density of extrinsic information $p(x)$ is invariant with respect to block length T. Therefore, in this subsection, we focus on the density analysis with $T = t$.

It is computationally difficult to find the analytical form of the conditional density from Eq. (21). From Eq. (18) and Eq. (19), we obtain

$$\begin{aligned}\mathbf{L_e(b)} = &\frac{2\epsilon}{\sigma^2 t}\mathrm{tr}\left\{\left(\mathbf{H}\sqrt{\frac{t}{\epsilon}}A\right)\left(\mathbf{H}\sqrt{\frac{t}{\epsilon}}A\right)^\dagger\right\}\mathbf{b} \\ &+\frac{2}{\sigma^2}\sqrt{\frac{\epsilon}{t}}\sum_{j=1}^{r}\sum_{k=1}^{t}\mathrm{Re}\left\{\mathbf{n}_{j,k}\left(\mathbf{H}\sqrt{\frac{t}{\epsilon}}A\right)^*_{j,k}\right\}.\end{aligned} \tag{22}$$

Since $(\sqrt{\frac{t}{\epsilon}}A)(\sqrt{\frac{t}{\epsilon}}A)^\dagger = I_t$, A is a square matrix and $\mathbf{H}$ is an i.i.d. circular symmetric complex Gaussian random matrix, $\mathbf{H}(\sqrt{\frac{t}{\epsilon}}A)$ has the same distribution as $\mathbf{H}$. Therefore, $\mathbf{L_e(b)}$ has the same distribution as $\tilde{\mathbf{L}}$ which is defined as

$$\begin{aligned}\tilde{\mathbf{L}} &= \frac{2\epsilon}{\sigma^2 t}\mathrm{tr}\{\mathbf{HH}^\dagger\}\mathbf{b} + \frac{2}{\sigma^2}\sqrt{\frac{\epsilon}{t}}\sum_{j=1}^{r}\sum_{k=1}^{t}\mathrm{Re}\{\mathbf{n}_{j,k}\mathbf{H}^*_{j,k}\} \\ &= \sum_{j=1}^{r}\sum_{k=1}^{t}\left[\frac{2\epsilon}{\sigma^2 t}|\mathbf{H}_{j,k}|^2\mathbf{b} + \frac{2}{\sigma^2}\sqrt{\frac{\epsilon}{t}}\mathrm{Re}\{\mathbf{n}_{j,k}\mathbf{H}^*_{j,k}\}\right] \\ &\triangleq \sum_{j=1}^{r}\sum_{k=1}^{t}\tilde{\mathbf{L}}_{j,k}.\end{aligned} \tag{23}$$

Since conditioned on $\mathbf{b} = b$, $\tilde{\mathbf{L}}_{\mathbf{j,k}}, 1 \le j \le r, 1 \le k \le t$ are i.i.d. random variables, the conditional density of $\tilde{\mathbf{L}}$ is the rt-fold convolution of the conditional density of $\tilde{\mathbf{L}}_{\mathbf{j,k}}$. The conditional density of $\tilde{\mathbf{L}}_{\mathbf{j,k}}$ can be derived as (same result in Eq. (42) of [6])

$$P_L(l|b) = \frac{\hat{\sigma}^2}{2\sqrt{1+2\hat{\sigma}^2}}\exp\left\{\frac{bl - \sqrt{1+2\hat{\sigma}^2}|l|}{2}\right\}, \tag{24}$$

where $\hat{\sigma}^2 = \frac{\sigma^2 t}{\epsilon}$. Conditioned on $\mathbf{b} = b$, the mean and variance of $\tilde{\mathbf{L}}_{\mathbf{j,k}}$ are $\frac{2b}{\hat{\sigma}^2} = \frac{2\epsilon b}{\sigma^2 t}$ and $\frac{4}{\hat{\sigma}^2} + \frac{4}{\hat{\sigma}^4} = \frac{4\epsilon}{\sigma^2 t} + \frac{4\epsilon^2}{\sigma^4 t^2}$ respectively.

Next we consider the asymptotic conditional density of extrinsic information when the number of antennas becomes large. By the Central Limit Theorem (Theorem 4.7-1 [15]), conditioned on $\mathbf{b} = b$, the distribution of

$$\frac{\sum_{j=1}^{r}\sum_{k=1}^{t}(\tilde{\mathbf{L}}_{j,k} - \frac{2\epsilon b}{\sigma^2 t})}{\sqrt{rt(\frac{4\epsilon}{\sigma^2 t} + \frac{4\epsilon^2}{\sigma^4 t^2})}} \tag{25}$$

converges to the standard normal distribution as rt increases to infinity. Therefore, if rt is large, we are motivated to approximate the conditional density of extrinsic information $\tilde{\mathbf{L}} = \sum_{j=1}^{r}\sum_{k=1}^{t}\tilde{\mathbf{L}}_{j,k}$ by a Gaussian density

$$\mathcal{N}(\frac{2r\epsilon}{\sigma^2}b, \frac{4r\epsilon}{\sigma^2} + \frac{4r\epsilon^2}{\sigma^4 t}). \tag{26}$$

Furthermore, we have

$$\frac{4r\epsilon}{\sigma^2} + \frac{4r\epsilon^2}{\sigma^4 t} = \frac{4r\epsilon}{\sigma^2}(1 + \frac{\epsilon}{\sigma^2 t}). \tag{27}$$

If $\frac{\epsilon}{\sigma^2 t}$ is small, the density in Eq. (26) can be further approximated by

$$\mathcal{N}(\frac{2r\epsilon}{\sigma^2}b, \frac{4r\epsilon}{\sigma^2}), \tag{28}$$

which fits in the model of the density in Eq. (1).

We conclude that when the priors are perfect ($I_a = 1$), it is reasonable to approximate the conditional density of extrinsic information by $\mathcal{N}(\frac{2r\epsilon}{\sigma^2}b, \frac{4r\epsilon}{\sigma^2})$ if the following conditions hold

- $r \cdot t$ is large;
- $\frac{\epsilon}{\sigma^2 t}$ is small;

D. Density of Extrinsic Information when $I_a < 1$

When the *a priori* information is not perfect, i.e., $I_a < 1$, it is practically impossible to determine the density of extrinsic information. However, if the dispersion matrices of the BL-LD code satisfy

$$\frac{A_i A_j^\dagger + A_j A_i^\dagger}{\sigma^2} \to 0, \forall i \neq j, \tag{29}$$

where "$\to$" means "be close to", motivated by Theorem 1 and its proof, we conjecture that the conditional density of extrinsic information at $I_a < 1$ be close to the conditional density at $I_a = 1$ (for detailed argument, refer to [16]). Furthermore, if the conditions summarized at the end of last subsection are satisfied, it is reasonable to use the type of density $\mathcal{N}(\frac{\sigma_L^2}{2}b, \sigma_L^2)$ to approximate the conditional density of extrinsic information at arbitrary $I_a < 1$. As for how close to 0 is close enough, we do not have an answer yet.

Next, we consider a BL-LD code (an LD code in Eq. (36) of [11], with QPSK modulation)

$$\begin{aligned}X(:,1) &= \begin{pmatrix} b_1 + b_3 + \frac{i(b_6+b_7)}{\sqrt{2}} + ib_8 \\ \frac{b_2-b_4-ib_5}{\sqrt{2}} + \frac{-ib_6+ib_7}{2} \\ 0 \end{pmatrix}, \\ X(:,2) &= \begin{pmatrix} \frac{-b_2+b_4-ib_5}{\sqrt{2}} + \frac{-ib_6+ib_7}{2} \\ b_1 - \frac{ib_6+ib_7}{\sqrt{2}} \\ -\frac{b_2+b_4-ib_5}{\sqrt{2}} + \frac{-ib_6+ib_7}{2} \end{pmatrix}, \\ X(:,3) &= \begin{pmatrix} 0 \\ \frac{b_2+b_4+ib_5}{\sqrt{2}} + \frac{-ib_6+ib_7}{2} \\ b_1 - b_3 + \frac{ib_6+ib_7}{\sqrt{2}} - ib_8 \end{pmatrix}, \\ X(:,4) &= \begin{pmatrix} \frac{b_2-b_4+ib_5}{\sqrt{2}} + \frac{ib_6-ib_7}{2} \\ -b_3 + ib_8 \\ -\frac{b_2+b_4-ib_5}{\sqrt{2}} + \frac{-ib_6+ib_7}{2} \end{pmatrix}.\end{aligned} \tag{30}$$

We simulate the conditional density of extrinsic information with $r = 3$ and $E_b/N_o = 0$ dB. In the computation of bit energy E_b, we assume a rate $\frac{1}{2}$ outer code is used.

In Fig. 2, the *a priori* mutual information I_a is equal to one. Extrinsic information data are generated according to Eq. (18). The solid curve corresponds to the histogram of the extrinsic information conditioned on $\mathbf{b} = 1$, while the dashed

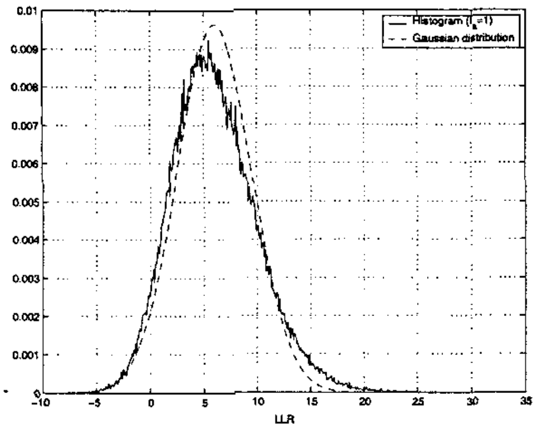

Fig. 2. The histogram of extrinsic information when *a priori* mutual information I_a is 1. E_b/N_o is 0 dB. The dashed curve is the Gaussian approximation.

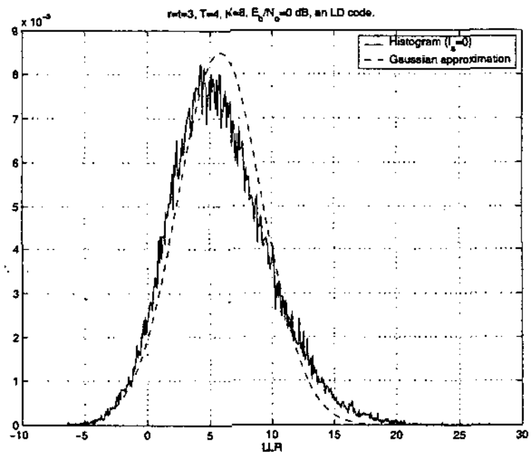

Fig. 3. The histogram of extrinsic information when *a priori* mutual information I_a is 0. E_b/N_o is 0 dB. The dashed curve is the Gaussian approximation.

curve corresponds to a Gaussian distribution $\mathcal{N}(\mu, 2\mu)$ where μ is equal to the mean of extrinsic information conditioned on $\mathrm{b} = 1$. The Gaussian distribution approximates the histogram quite well, which is predictable since $r \cdot t = 9$ and $\frac{\epsilon}{\sigma^2 t}$ is low. Fig. 3 shows the histogram and Gaussian approximation when the *a priori* mutual information I_a is equal to zero. In this case, extrinsic information data are generated through SISO decoding. It can be found that the approximation is still quite accurate since the condition in Eq. (29) is approximately satisfied. Since $0 \leq I_a \leq 1$, the histogram of extrinsic information would also be well approximated by a Gaussian density of type $\mathcal{N}(\mu, 2\mu)$ if $0 < I_a < 1$.

We will like to point out that through simulations, we observed that even if the conditional density of extrinsic information does not look like the Gaussian density, sometimes the EXIT chart technique still works well. It demonstrates the robustness of this technique.

III. Conclusion

We have studied the conditional density of extrinsic information at the output of space time demodulator. We have found that when (1) the product of the number of transmit and receive antennas $r \cdot t$ is large; (2) $\frac{\epsilon}{\sigma^2 t}$ is small, i.e., the SNR is low or the number of transmit antennas t is large; (3) the modulation schemes satisfy $\frac{A_i A_j^\dagger + A_j A_i^\dagger}{\sigma^2} \to 0, \forall i \neq j$, the conditional density of extrinsic information is reasonably approximated by a Gaussian density $\mathcal{N}(\frac{\sigma_L^2}{2} b, \sigma_L^2)$ when the *a priori* information is also Gaussian distributed.

References

[1] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel." submitted to IEEE Transactions on Communications, July 2001.

[2] A. van Zelst, R. van Nee, and G. Awater, "Turbo-BLAST and its performance," in *Proc. VTC*, May 2001.

[3] H. Su and E. Geraniotis, "Space-time turbo codes with full antenna diversity," *IEEE Transactions on Communications*, vol. 49, pp. 47–59, Jan. 2001.

[4] Y. Liu, M. P. Fitz, and O. Y. Takeshita, "Full rate space-time turbo codes," *IEEE Journal on Selected Areas in Communications*, vol. 19, pp. 969–980, May 2001.

[5] S. ten Brink, G. Kramer, and A. Ashikhmin, "Design of low-density parity-check codes for multi-antenna modulation and detection." submitted to IEEE Transactions on Communications, June 2002.

[6] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Transactions on Communications*, vol. 49, pp. 1727–1737, Oct 2001.

[7] S. ten Brink and B. M. Hochwald, "Detection thresholds of iterative MIMO processing," in *Proc. ISIT*, p. 22, 2002.

[8] N. Wiberg, *Codes and decoding on general graphs*. PhD thesis, Linköping Univ., Sweden, 1996.

[9] H. El Gamal and A. R. Hammons, Jr., "Analyzing the turbo decoder using the Gaussian approximation," *IEEE Transactions on Information Theory*, vol. 47, pp. 671–686, Feb. 2001.

[10] S. -Y. Chung, T. J. Richardson, and R. L. Urbanke, "Analysis of sum-product decoding of low-density parity-check codes using a Gaussian approximation," *IEEE Transactions on Information Theory*, vol. 47, pp. 657–670, Feb. 2001.

[11] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Transactions on Information Theory*, vol. 48, pp. 1804–1824, July 2002.

[12] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory*, vol. 45, pp. 1456–1467, July 1999.

[13] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 1451–1458, October 1998.

[14] Y. Jiang, R. Koetter, and A. Singer, "On the separability of demodulation and decoding for communications over multiple-antenna block fading channels," *IEEE Transactions on Information Theory*, vol. 49, pp. 2709–2713, October 2003.

[15] H. Stark and J. W. Woods, *Probability, Random Processes, and Estimation Theory for Engineers*. Prentice-Hall, 1994.

[16] Y. Jiang, R. Koetter, and A. Singer, "Some properties of iterative MIMO processing." In preparation.