Subsidiary Maximum Likelihood Iterative Decoding Based on Extrinsic Information

Fengfan Yang and Tho Le-Ngoc

Abstract: **This paper proposes a multimodal generalized Gaussian distribution (MGGD) to effectively model the varying statistical properties of the extrinsic information. A subsidiary maximum likelihood decoding (MLD) algorithm is subsequently developed to dynamically select the most suitable MGGD parameters to be used in the component maximum *a posteriori* (MAP) decoders at each decoding iteration to derive the more reliable metrics performance enhancement. Simulation results show that, for a wide range of block lengths, the proposed approach can enhance the overall turbo decoding performance for both parallel and serially concatenated codes in additive white Gaussian noise (AWGN), Rician, and Rayleigh fading channels.**

Index Terms: **Extrinsic information, iterative decoder, maximum a posteriori (MAP) decoder, multimodal generalized Gaussian distribution (MGGD), subsidiary maximum likelihood decoding (MLD).**

I. INTRODUCTION

Turbo coding [1] can provide an exceptional bit error rate (BER) close to the Shannon limit at very low signal-to-noise ratio (SNR) over an additive white Gaussian noise (AWGN) channel. It resorts to a creative strategy to prevent the worst cases of both lower weights inevitably resulted from each of the component recursive systematic convolutional (RSC) codes by using a well-designed interleaver. Therefore, the resultant weight of an overall turbo code is increased significantly. In general, a turbo decoder consists of two component MAP decoders using the maximum *a posteriori* (MAP) decoding algorithm or a variant. The soft information, known as the logarithmic likelihood ratio (LLR) of the *a posteriori* probability (APP), is generated from the component MAP decoder, and the extrinsic information of the computed LLR is exchanged between the component decoders. The effective ways applied to tackle the extrinsic information can be classified into two categories. One uses an approximate Gaussian model [1]. The other directly employs the extrinsic information from one component decoder to update the *a priori* probability for the other decoder [2]–[4]. The most obvious advantage of the latter scheme is the low computational complexity since the estimation of the mean and variance of the extrinsic information is avoided. In the case of infinite block length, the *a priori* information is statistically independent. Consequently, the MAP decoder provides the true APPs by applying the second scheme, and there is really no need for additional estimation of the extrinsic information distribution. It was reported in [3] that the second scheme outperforms the first one for a very long block length. However, in the finite-length case for short or medium blocks that are of great importance in practical applications, the *a priori* information is not truly independent, thus both schemes are approximate. Generally, such a problem also exists in many other iterative systems that adopt soft-input soft-output (SISO) APP module for the iterative decoding, e.g., serially concatenated codes, low-density parity-check (LDPC) codes, turbo equalization, joint turbo equalization, and multi-user detection.

This paper introduces an approach to model the statistical properties of the extrinsic information with a multimodal generalized Gaussian distribution (MGGD) and to use them in the derivation of the LLR for the component MAP decoders in order to enhance the turbo decoding performance for a wide range of block lengths. As will be shown in this paper, the statistical properties of the extrinsic information vary in different iterations and signal-to-noise ratios, and do not follow the Gaussian distribution. The MGGD with its parameters can describe more effectively such dynamic statistical properties of the extrinsic information. Taking into account the MGGD-based model of the extrinsic information, a subsidiary maximum likelihood decoding (MLD) is developed to determine the performance metrics and to dynamically select the most suitable MGGD parameters to be used in the component MAP decoders in each decoding iteration. The proposed approach is applicable to both parallel and serially concatenated codes.

The paper is organized as follows. After this introduction, in Section II, we briefly revisit the extrinsic information and its reinforcement effects for iterative decoding. Section III describes the introduced iterative decoding structure with subsidiary decoder, including the derivation of the MLD metrics for component MAP decoders based on the extrinsic information and channel outputs and the new selection criterion to dynamically choose the optimal MGGD of the extrinsic information for the component decoder by using the subsidiary MLD. Section IV presents the performance evaluation and comparison by simulations of both parallel and serially concatenated codes using the proposed approach over AWGN, Rician, and Rayleigh fading channels and for various block lengths.

II. EXTRINSIC INFORMATION AND MGGD DESCRIPTION

Fig. 1(a) shows the block diagram of the proposed turbo decoder using subsidiary MLD. Using the general turbo decoding structure, the proposed turbo decoder also consists of 2 component MAP Decoders I and II, which exchange their soft infor-

Manuscript received December 21, 2003; approved for publication by Jong-Seon No, Division I Editor, January 10, 2007.

F. Yang is with the Department of Electronic Engineering, Nanjing University of Aeronautics & Astronautics, Nanjing, 210016, P. R. China, email: yffee@nuaa.edu.cn.

T. Le-Ngoc is with the Department of Electrical & Computer Engineering, McGill University, Montreal, Quebec, Canada, H3A-2A7, email: tho@ece.mcgill.ca.

mation via the N-bit interleaver and deinterleaver. Based on a selected statistical model of the extrinsic information, the subsidiary MLD dynamically selects the most suitable values of the model parameters to be used by the component MAP decoders in deriving the performance metrics at each decoding iteration to enhance the overall turbo decoding performance.

In this section, we will examine the extrinsic information with focus on its reinforcement effects by use of the MGGD to describe more effectively its dynamic statistical properties.

A. Reinforcement Effects of the Extrinsic Information

As an illustrative example, we consider a rate-1/3 nonpunctured turbo code. Its component code is a rate-1/2 RSC code with the generator matrix $[1, (1+D^2)/(1+D+D^2)]$ as shown in Fig. 1(b). For a rate-1/2 binary systematic convolutional code of constraint length ν, the decoding of a block of N bits by MAP algorithm involves the iterative calculations of the following probability functions $a_k(s_k)$ and $b_k(s_k)$ and LLR soft output [1].

Forward recursion:

$$a_k(s_k) = \Pr\left\{s_{k-1}|R_1^{k-1}\right\} = \sum_{\forall s_{k-1}} \sum_{i=0}^{1} \frac{g_i(R_k, s_k, s_{k-1}) a_{k-1}(s_{k-1})}{\Lambda_k(R_k)}. \quad (1a)$$

Backward recursion:

$$b_k(s_k) = \frac{\Pr\left\{R_{k+1}^N | s_k = m\right\}}{\Pr\left\{R_{k+1}^N | R_1^k\right\}} = \sum_{\forall s_{k+1}} \sum_{i=0}^{1} \frac{g_i(R_{k+1}, s_{k+1}, s_k) b_{k+1}(s_{k+1})}{\Lambda_{k+1}(R_{k+1})}. \quad (1b)$$

LLR soft output:

$$L(d_k) = \log \frac{\sum_{\forall s_k} \sum_{\forall s_{k-1}} g_1(\bar{R}_k, s_{k-1}, s_k) a_{k-1}(s_{k-1}) b_k(s_k)}{\sum_{\forall s_k} \sum_{\forall s_{k-1}} g_0(\bar{R}_k, s_{k-1}, s_k) a_{k-1}(s_{k-1}) b_k(s_k)} \quad (1c)$$

where the probability function:

$$\Lambda_k(R_k) = \sum_{\forall s_k} \sum_{\forall s_{k-1}} \sum_{i=0}^{1} g_i(R_k, s_k, s_{k-1}) a_{k-1}(s_{k-1})$$

and $R_k = (r_k, r_k^{(e)})$ is the k-th received sample for $k = 1, 2, \cdots, N$. The channel output $r_k = (r_{k,s}, r_{k,r})$ is the demodulated sample, and the extrinsic information, $r_k^{(e)}$, is generated by the counterpart decoder. s_k is the state of the convolutional code at the k-th instant. The probability $g_i(R_k, s_k, s_{k-1})$, $i = 0, 1$, can be expressed as [1]

$$g_i(R_k, s_k, s_{k-1}) = q(d_k, s_k, s_{k-1}) \Pr\{r_{k,s}|d_k = i\} \times \Pr\{r_{k,r}|d_k = i, s_k, s_{k-1}\} \Pr\left\{r_k^{(e)}|d_k = i\right\} \quad (2a)$$

where

$$q(d_k, s_k, s_{k-1}) = \Pr\{d_k|s_k = m, s_{k-1} = m'\} \times \Pr\{s_k = m|s_{k-1} = m'\}. \quad (2b)$$

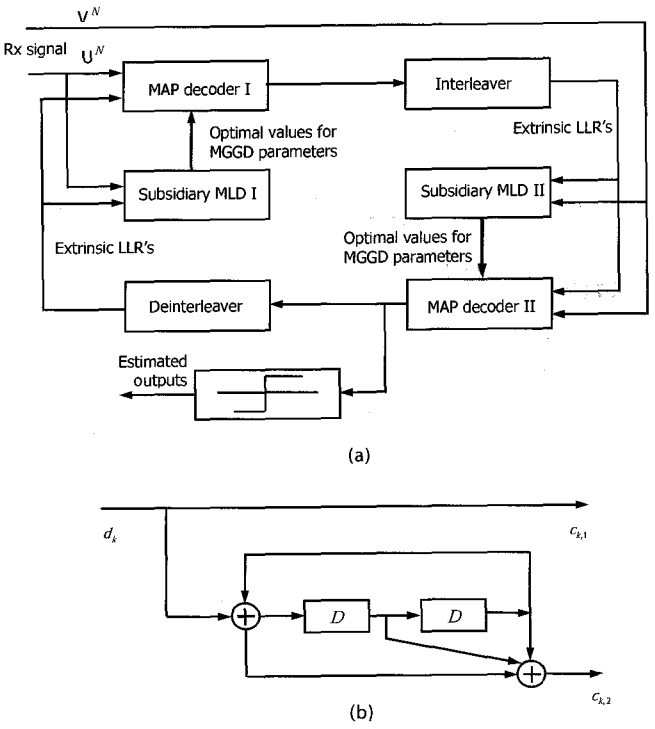

Fig. 1. Block diagram of a turbo decoder using subsidiary MLD: (a) Turbo decoder using subsidiary MLD, (b) example of a rate-1/2 RSC encoder with generator matrix $[1, (1+D^2)/1+D+D^2)]$.

The soft LLR and extrinsic information are obtained from (1c) with $\bar{R}_k = R_k$ and $\bar{R}_k = r_k$, respectively.

The function $a_{k-1}(s_{k-1})$ in (1a) is the probability of the state $s_{k-1} \in \boldsymbol{\pi} = \{\pi_0, \pi_1, \cdots, \pi_{2^\nu - 1}\}$ in the $(k-1)$-th instant conditioned on $R_1^{k-1} = (R_1, R_2, \cdots, R_{k-1})$. If $a_{k-1}(s_{k-1})$ is available, the conditional probability of the extrinsic information provided by its counterpart decoder, $\Pr\left\{r_k^{(e)}|d_k = 1\right\}$ in (2a) can implicitly exert selective and strengthened effects for the probability $g_i(R_k, s_k, s_{k-1})$. This leads to a significant increase in $a_k(s_k)$ for the desired state s_k in the k-th instant. The computation of $a_k(s_k)$, $k = 1, 2, \cdots, N$, is iteratively carried out by the forward recursion in (1a). For high SNRs and large number of decoding iterations, the following equality holds for a particular state $\pi_p \in \boldsymbol{\pi}$ in high probability, partially due to the selective contributions of the extrinsic information,

$$a_k(\pi_p) = \max_{0 \le j \le 2^\nu - 1} \{a_k(s_k)|s_k = \pi_j\}. \quad (3)$$

From the probabilistic viewpoint, the above equality implies that the favorite state $s_k = \pi_p$ should be more likely along the surviving path than any other states $(s_k \neq \pi_p)$ through the trellis diagram using the MAP algorithm. More effective reinforcement from the extrinsic information provides more selective impacts on the favorite state. Consequently, $b_k(s_k)$ with available $a_k(s_k)$ can be iteratively evaluated by using the backward recursion with $k = N-1, N-2, \cdots, 1$ in (1b). The current favorite state with respect to $b_k(s_k)$ should be also more likely along the correct path. Finally, the soft LLRs and extrinsic information (generated by the component decoder) computed for each information bit by (1c) should be more reliable thanks to the extrinsic information involved in the computations of $a_k(s_k)$ and $b_k(s_k)$.

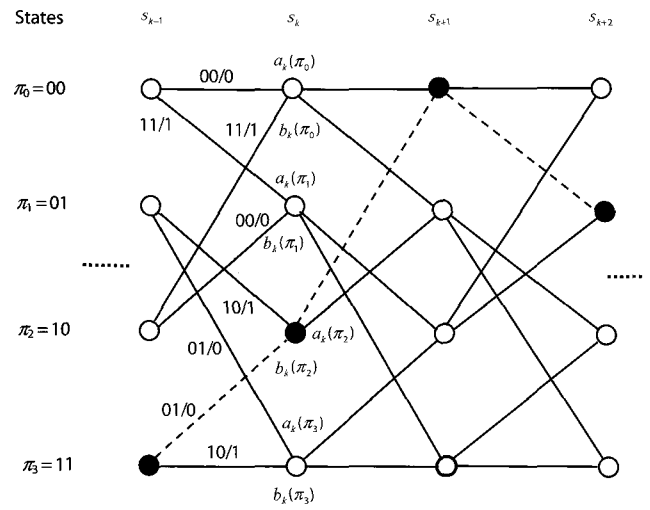

Fig. 2. RSC code trellis diagram with the particular transition between the favorite states related to information bits 011 (denoted by the dotted lines).

As an illustrative example, consider a block of 10,000 information bits, a 4-state, rate-1/3 non-punctured turbo code with component RSC codes and 100×100 block-interleaver shown in Fig. 1. The three consecutive transmitted bits at instants k, $(k+1)$, and $(k+2)$ are 0, 1, and 1, respectively with $k = 4999$. They form a valid path segment as indicated in Fig. 2. Using the MAP algorithm with Gaussian distributed extrinsic information [1] at a SNR of 0 dB, $a_j(m)$ and $b_j(m)$ for $j = k$, $(k+1)$, and $(k+2)$ are computed and given in Table 1 for the first component decoder at different iterations. The simulation results indicate that the largest value of $a_k(m)$ corresponds to $m = 10_2$ (in binary). Notice that the relative difference between $a_k(m = 10_2)$ and the others, $a_k(m \neq 10_2)$, becomes much larger as the number of iterations increases due to the reinforcement effects from the extrinsic information from the second decoder in the last iteration. Thus, $m = 10_2$ is the favorite state for $a_k(m)$ in the k-th epoch. The similar situations can be observed for $a_{k+1}(m)$ and $a_{k+2}(m)$ with the favorite states $m = 00_2$ and 01_2 in the $(k+1)$-th and $(k+2)$-th epochs, respectively. Accordingly, $b_k(m)$, $b_{k+1}(m)$, and $b_{k+2}(m)$ are closer to 1 for the favorite states. In other words, the reinforcement effects from the extrinsic information enables $a_k(m)$ and $b_k(m)$ to selectively choose the favorite state along the correct path with higher probability at each instant. This supports the validity of (3).

B. Multimodal Generalized Gaussian Distribution for the Extrinsic Information

The above discussions illustrate the important role of a reasonable statistical distribution model for extrinsic information to enhance the MAP decoding performance. More specifically, as indicated by the last term of right-hand side of (2a), this distribution varies from iteration to iteration (see also Fig. 11 of [1]). In general, *infinite-length* or *very long* turbo codes with a sufficiently large number of decoding iterations have the BER performance curves with an *abrupt* improvement at a certain signal-to-noise *threshold*, while *finite-length* turbo codes with an insufficiently large number of decoding iterations have a more *gentle* performance-improvement region covering a *wider* signal-to-noise range. Extrinsic information in performance-improvement region experiences a critical transition from low to high reliability and can have a *bimodal* distribution as shown in [5]. In this paper, we propose an MGGD family to dynamically characterize the extrinsic information distribution. At first, we consider the probability density function (pdf) of the generalized Gaussian distribution (GGD) defined in [5],

$$f(z;\alpha,\sigma,\mu) = \frac{\alpha\eta(\alpha,\sigma)}{2\Gamma(1/\alpha)}\exp\left(-[\eta(\alpha,\sigma)|z-\mu|]^{\alpha}\right) \quad (4a)$$

for $-\infty \leq z \leq \infty$, where μ and σ are the mean and standard deviation, respectively; $\Gamma(u) = \int_0^{+\infty} x^{u-1}e^{-x}\mathrm{d}x$ is a Gamma function and $\eta(\alpha,\sigma) = [\Gamma(3/\alpha)/\Gamma(1/\alpha)]^{1/2}/\sigma$ with a positive parameter α that characterizes the distribution shape. In particular, it becomes Laplacian, Gaussian, and uniform distribution for $\alpha = 1$, 2, and ∞, respectively. We then introduce the MGGD with its pdf as a linear combination of M GGD pdf's, i.e.,

$$f_M(z;\boldsymbol{\alpha},\boldsymbol{\sigma},\boldsymbol{\mu},\boldsymbol{\omega}) = \sum_{j=1}^{M}\frac{\omega_j\alpha_j\eta(\alpha_j,\sigma_j)}{2\Gamma(1/\alpha_j)}\exp\left(-[\eta(\alpha_j,\sigma_j)|z-\mu_j|]^{\alpha_j}\right) \quad (4b)$$

for $-\infty \leq z \leq \infty$, where M is the number of modes (e.g., $M = 2$ for bimodal distribution), $\boldsymbol{\alpha} = [\alpha_j, j = 1,2,\cdots,M]$, $\boldsymbol{\sigma} = [\sigma_j, j = 1,2,\cdots,M]$, $\boldsymbol{\mu} = [\mu_j, j = 1,2,\cdots,M]$, and $\boldsymbol{\omega} = [\omega_j, j = 1,2,\cdots,M]$, where ω_j's are the coefficients such that $\int_{-\infty}^{+\infty} f_M(z)\mathrm{d}z = 1$.

III. ITERATIVE DECODING WITH SUBSIDIARY ML DECODER

In this section, we will first develop the subsidiary MLD algorithm and derive the corresponding metrics for each component MAP decoder. Next, we will describe in detail the proposed turbo decoding algorithm using subsidiary ML decoder.

A. Subsidiary MLD Algorithm with Extrinsic Information

Binary modulation scheme is assumed in the following discussion for the sake of simplicity. Consider a coded bit stream $\boldsymbol{c} = (\boldsymbol{c}_1, \boldsymbol{c}_2, \cdots)$, corresponding to an input information bit stream $\boldsymbol{d} = (\boldsymbol{d}_1, \boldsymbol{d}_2, \cdots)$, from a rate-$m_0/m_1$ binary convolutional encoder, where $\boldsymbol{d}_p = (d_{p,1}, d_{p,2}, \cdots, d_{p,m_0})$ and $\boldsymbol{c}_p = (c_{p,1}, c_{p,2}, \cdots, c_{p,m_1})$, $p = 1, 2, \cdots$. Both information bit, d_{p,j_1}, $j_1 = 1, 2, \cdots, m_0$, and coded bit, $c_{p,j}$, $j = 1, 2, \cdots, m_1$, are either $+1$ or -1.

The received signal can be denoted by $\boldsymbol{r} = (\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3, \cdots)$, where $\boldsymbol{r}_p = (r_{p,1}, r_{p,2}, \cdots, r_{p,m_1})$, and $r_{p,j} = \gamma_{p,j}c_{p,j} + n_{p,j}$. The vector $\boldsymbol{n} = (\boldsymbol{n}_1, \boldsymbol{r}_2, \cdots)$ with $\boldsymbol{n}_p = (n_{p,1}, n_{p,2}, \cdots, n_{p,m_1})$ and $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \boldsymbol{\gamma}_3, \cdots)$ with $\boldsymbol{\gamma}_p = (\gamma_{p,1}, \gamma_{p,2}, \cdots, \gamma_{p,m_1})$ represent the additive white Gaussian noise and channel fading, respectively.

Given the received $\boldsymbol{r}$ and the fading $\boldsymbol{\gamma}$, the optimal decoder selects the message $\boldsymbol{c}$ that maximizes the *a posteriori* probability $\Pr(\boldsymbol{c}|\boldsymbol{r},\boldsymbol{\gamma}) = \Pr(\boldsymbol{r}|\boldsymbol{c},\boldsymbol{\gamma})\Pr(\boldsymbol{c})/\Pr(\boldsymbol{r}|\boldsymbol{\gamma})$. $\Pr(\boldsymbol{r}|\boldsymbol{\gamma})$ is irrelevant to the trellis path $\boldsymbol{c}$. Furthermore, for equally probable transmission, $\Pr(\boldsymbol{c})$ is the same for all $\boldsymbol{c}$, and hence, maximizing the

Table 1. Values of $a_j(m)$ and $b_j(m)$ in three consecutive iterations.

j	m	$a_j(m)$			$b_j(m)$		
		Iteration 1	Iteration 2	Iteration 3	Iteration 4	Iteration 5	Iteration 6
k	00	1.170748×10^{-2}	1.734525×10^{-3}	1.483962×10^{-4}	9.404543×10^{-1}	3.302947×10^{-2}	8.592076×10^{-6}
	01	7.949190×10^{-3}	3.428098×10^{-4}	1.876200×10^{-5}	0.9600574	1.553554×10^{-3}	3.402420×10^{-5}
	10	**0.5939825**	**0.9951085**	**0.9996746**	**1.234800**	**1.004857**	**1.000326**
	11	0.3863608	2.814118×10^{-3}	1.582537×10^{-4}	0.6416505	1.970121×10^{-4}	7.183486×10^{-6}
$k+1$	00	**0.6767174**	**0.9978183**	**0.9998264**	**1.042049**	**1.002125**	**1.000174**
	01	5.115052×10^{-2}	1.820043×10^{-3}	1.535004×10^{-4}	0.7680731	3.285863×10^{-2}	3.507411×10^{-6}
	10	0.2177554	4.103730×10^{-5}	2.137436×10^{-6}	0.7916686	1.363674×10^{-2}	2.996263×10^{-4}
	11	5.437671×10^{-2}	3.205985×10^{-4}	1.796991×10^{-5}	1.529151	1.644967×10^{-3}	6.314683×10^{-5}
$k+2$	00	0.2532995	1.914310×10^{-4}	2.849346×10^{-4}	0.8115615	1.346133×10^{-2}	1.669763×10^{-5}
	01	**0.5749062**	**0.9996502**	**0.9997021**	**1.168876**	**1.000287**	**1.000298**
	10	8.833496×10^{-2}	1.314713×10^{-4}	1.108063×10^{-5}	0.9573521	0.4569463	5.029321×10^{-7}
	11	8.345930×10^{-2}	2.686217×10^{-5}	1.894146×10^{-6}	0.4537527	9.536437×10^{-3}	8.803777×10^{-4}

a posteriori probability $\Pr(\boldsymbol{c}|\boldsymbol{r},\boldsymbol{\gamma})$ is equivalent to maximizing the probability $\Pr(\boldsymbol{r}|\boldsymbol{c},\boldsymbol{\gamma})$, i.e.,

$$\max_{\forall \boldsymbol{c}\in\boldsymbol{C}} \Pr(\boldsymbol{c}|\boldsymbol{r},\boldsymbol{\gamma}) \equiv \max_{\forall \boldsymbol{c}\in\boldsymbol{C}} \Pr(\boldsymbol{r}|\boldsymbol{c},\boldsymbol{\gamma}). \tag{5}$$

Consider a rate-1/3 non-punctured turbo code with two identical rate-1/2 RSC encoders. One component decoder gets the extrinsic information, denoted by a vector $\boldsymbol{r}_e = (r_1^{(e)}, r_2^{(e)}, r_3^{(e)}, \cdots)$, from its counterpart decoder and the involved channel outputs. Hence, the metric derivation for a MLD should take into account the extrinsic information that is assumed to follow the MGGD in (4b). For a valid path (coded bit stream) $\boldsymbol{c}$ based on (5), the metric concerning the channel output $\boldsymbol{r}$ and extrinsic information $\boldsymbol{r}_e$ can be computed as follows

$$\begin{aligned}&\zeta(\boldsymbol{r},\boldsymbol{r}_e|\boldsymbol{c},\boldsymbol{\gamma},\boldsymbol{\alpha},\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\omega})\\ &\quad= -\log\left[\Pr(\boldsymbol{r},\boldsymbol{r}_e|\boldsymbol{c},\boldsymbol{\gamma},\boldsymbol{\alpha},\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\omega})\right]\\ &\quad= \zeta(\boldsymbol{r}|\boldsymbol{c},\boldsymbol{\gamma}) + \zeta(\boldsymbol{r}_e|\boldsymbol{c},\boldsymbol{\gamma},\boldsymbol{\alpha},\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\omega})\end{aligned} \tag{6}$$

where the overall metric can be decomposed into two independent sub-metrics in relation to the channel output (1st term) and the extrinsic information (2nd term), conditioned on the path $\boldsymbol{c}$ and the fading vector $\boldsymbol{\gamma}$ known by the decoder. The first sub-metric can be further expressed as

$$\begin{aligned}\zeta(\boldsymbol{r}|\boldsymbol{c},\boldsymbol{\gamma}) &= -\log\left[\Pr(\boldsymbol{r}|\boldsymbol{c},\boldsymbol{\gamma})\right]\\ &= \sum_p \sum_{j=1}^{m_1} \left[\log(\sqrt{2\pi}\sigma_G) + \frac{|r_{p,j} - \gamma_{p,j}c_{p,j}|^2}{2\sigma_G^2}\right]\end{aligned} \tag{7a}$$

where σ_G is the standard deviation of additive Gaussian noise. For a rate-1/3 turbo code, m_1 is, respectively, 2 and 1 for the 1st and 2nd component decoders (of the turbo decoder). Based on (4b), the second sub-metric corresponding to the MGGD extrinsic information is

$$\begin{aligned}\zeta(\boldsymbol{r}_e|\boldsymbol{c},\boldsymbol{\gamma},&\boldsymbol{\alpha},\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\omega}) = -\log\left[\Pr(\boldsymbol{r}_e|\boldsymbol{c},\boldsymbol{\gamma},\boldsymbol{\alpha},\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\omega})\right]\\ &= -\sum_p \left\{ \log\left[\sum_{j=1}^{M} \frac{\omega_j\alpha_j\eta(\alpha_j,\sigma_j)}{2\Gamma(1/\alpha_j)}\right.\right.\\ &\quad\left.\left.\times \exp\left(-\left[\eta(\alpha_j,\sigma_j)|r_p^{(e)} - \mu_j|\right]^{\alpha_j}\right)\right]\right\}.\end{aligned} \tag{7b}$$

From (6), (7a), and (7b), the metric concerning the received signal $\boldsymbol{r}$ and extrinsic information $\boldsymbol{r}_e$ can be expressed as

$$\begin{aligned}&\zeta(\boldsymbol{r},\boldsymbol{r}_e|\boldsymbol{c},\boldsymbol{\gamma},\boldsymbol{\alpha},\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\omega})\\ &= \sum_p \left\{ \sum_{j=1}^{m_1} \left[\log(\sqrt{2\pi}\sigma_G) + \frac{|r_{p,j} - \gamma_{p,j}c_{p,j}|^2}{2\sigma_G^2}\right]\right.\\ &\quad - \log\left[\sum_{j=1}^{M} \frac{\omega_j\alpha_j\eta(\alpha_j,\sigma_j)}{2\Gamma(1/\alpha_j)}\right.\\ &\quad\left.\left.\times \exp\left(-\left[\eta(\alpha_j,\sigma_j)|r_p^{(e)} - \mu_j|\right]^{\alpha_j}\right)\right]\right\}.\end{aligned} \tag{8a}$$

For $M = 1$, the metric (8a) can be simplified to

$$\begin{aligned}&\zeta(\boldsymbol{r},\boldsymbol{r}_e|\boldsymbol{c},\boldsymbol{\gamma},\boldsymbol{\alpha},\boldsymbol{\mu},\boldsymbol{\sigma})\\ &= \sum_p \left[\sum_{j=1}^{m_1}(r_{p,j} - \gamma_{p,j}c_{p,j})^2 + |r_p^{(e)} - d_p\mu|^\alpha \frac{2\sigma_G^2}{\sigma_e^\alpha}\right.\\ &\quad\left.\times \left[\frac{\Gamma(3/\alpha)}{\Gamma(1/\alpha)}\right]^{\alpha/2} + \sigma_G^2 \log\left[\frac{\Gamma^3(1/\alpha)}{\alpha^2\Gamma(3/\alpha)}\right]\right]\end{aligned} \tag{8b}$$

where d_p is the p-th information bit corresponding to extrinsic information $r_p^{(e)}$. The mean, μ, and standard deviation, σ, of the extrinsic information can be approximately estimated by using (48) in [1].

Unlike the conventional MLD algorithm, in addition to the received signal $\boldsymbol{r}$ and path $\boldsymbol{c}$, the metric in (8a) also depends on the standard deviation σ_G of the Gaussian noise (i.e., hence signal-to-noise ratio), and the MGGD parameters, $\boldsymbol{\alpha}$, $\boldsymbol{\sigma}$, and $\boldsymbol{\mu}$, of the extrinsic information. For a given set of $\boldsymbol{r}$, $\boldsymbol{r}_e$, and $\boldsymbol{\gamma}$, the subsidiary MLD algorithm selects an optimum combination of path $\boldsymbol{c}^* \in \boldsymbol{C}$ and corresponding MGGD parameters $\boldsymbol{\alpha}^*$, $\boldsymbol{\sigma}^*$, and $\boldsymbol{\mu}^*$, to achieve the minimum metric, i.e.,

$$\begin{aligned}&\zeta(\boldsymbol{r},\boldsymbol{r}_e|\boldsymbol{c}^*,\boldsymbol{\gamma},\boldsymbol{\alpha}^*,\boldsymbol{\mu}^*,\boldsymbol{\sigma}^*,\boldsymbol{\omega}^*)\\ &\quad= \min_{\forall \boldsymbol{c}\in\boldsymbol{C},\boldsymbol{\alpha},\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\omega}} \left\{\zeta(\boldsymbol{r},\boldsymbol{r}_e|\boldsymbol{c},\boldsymbol{\gamma},\boldsymbol{\alpha},\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\omega})\right\}.\end{aligned} \tag{9}$$

The above selected surviving path $\boldsymbol{c}^*$ and MGGD parameters $\boldsymbol{\alpha}^*$, $\boldsymbol{\sigma}^*$, and $\boldsymbol{\mu}^*$ jointly maximize the *a posteriori* probability

in (5), which differs greatly from the conventional MLD (or Viterbi) algorithm.

B. Dynamic Selection of Optimal MGGD Extrinsic Information

This section introduces a decision strategy to choose the most appropriate conditional probability of the extrinsic information generated from the component decoders based on the obtained metric of the subsidiary MLD.

In the 1st iteration, since the extrinsic information is not available, the MAP Decoder I only uses the received signal to compute the LLR soft-outputs. These LLR soft-outputs are interleaved and then used as extrinsic information for the MAP Decoder II and its subsidiary MLD II as shown in Fig. 1(a).

The subsidiary MLD II then determines the best estimated values of the MGGD parameters for the extrinsic information. Subsequently, it applies (8a) to compute the metric of an arbitrary path through the trellis corresponding to the received channel output signal $\boldsymbol{r}$, channel fading vector $\boldsymbol{\gamma}$, and extrinsic information $\boldsymbol{r}_e$. Note that each metric is uniquely related to a combination of a path and specified values of the MGGD parameters. The subsidiary MLD II eventually selects the optimum combination of the surviving path $\boldsymbol{c}^*$ and corresponding MGGD parameters $\boldsymbol{\alpha}^*$, $\boldsymbol{\sigma}^*$, and $\boldsymbol{\mu}^*$ based on the maximum likelihood criterion in (9) and passes the selected MGGD parameters $\boldsymbol{\alpha}^*$, $\boldsymbol{\sigma}^*$, and $\boldsymbol{\mu}^*$ to the MAP Decoder II.

The Decoder II performs the MAP algorithm with its input received signal, extrinsic LLR's, and optimum MGGD parameter values (derived from the subsidiary MLD II), to produce its LLR soft-outputs. These LLR soft-outputs are de-interleaved and then passed to the MAP Decoder I and Subsidiary MLD I as updated extrinsic information for further decoding. This completes the 1st decoding iteration. Note that the LLR soft-outputs of the MAP Decoder II can also be used to produce the decoded outputs.

In the 2nd and subsequent iterations, the MAP Decoder I and its Subsidiary MLD I have both received signal and updated extrinsic information. The Subsidiary MLD I computes path metrics and provides the optimum MGGD parameters to the MAP Decoder I that, in turn, performs the MAP algorithm with its input received signal, updated extrinsic information (from MAP Decoder II), and optimum MGGD parameter values (derived from the subsidiary MLD I), to produce its LLR soft-outputs. With the updated extrinsic information (as the interleaved version of these LLR soft-outputs from the MAP Decoder I) and its received signal, the MAP Decoder II and its subsidiary MLD II will repeat the same component decoding process as in the previous iteration.

The roles of the two decoding algorithms differ greatly in the component decoder. The subsidiary MLD is also distinct from the well-known SOVA that can yield the soft LLR's to allow the decoding iteratively. The MAP algorithm with the optimal MGGD parameter values for extrinsic information is thus regarded as a principal decoding.

For each component decoder, the additional complexity compared with the conventional MAP algorithm increases with the number of modes, M, and MGGD parameters. Thus, we should keep M as small as possible to reduce the complexity provided that it can fully characterize the dynamic extrinsic information distribution over the interested SNR range and number of decoding iterations. It can be verified from (4a) and (4b) that when μ_j's are sufficiently close to each other, a single-mode ($M = 1$) GGD with appropriate selection of 3 parameters, α, μ, σ, can provide a reasonably good approximation of a given MGGD. The procedure to select an appropriate MGGD model and to estimate the values of its parameters is as follows.

From the measured pdf (histogram) of a given extrinsic information, we first identify the number of distinct modes, M, and μ_i's (as the maxima), $i = 1, \cdots, M$, and estimate the weighting coefficients ω_i's, $i = 1, \cdots, M$. Second, for each distinct mode indicated by μ_i, we estimate the best values of σ_i and α_i to fit the measured extrinsic information histogram. Finally, the values of $\boldsymbol{\omega}$ are readjusted to formulate the MGGD model as a linear combination of M GGD pdf's in (4b) with the estimated $\boldsymbol{\alpha}$, $\boldsymbol{\sigma}$, and $\boldsymbol{\mu}$.

In many cases of *long* turbo codes with a sufficiently large number of decoding iterations, their error performance curves exhibit an *abrupt* improvement at a certain signal-to-noise ratio (SNR) *threshold*, SNR_{TH}, that approximately divides the error performance into two regions: Unusable high error rate region for $\text{SNR} < \text{SNR}_{\text{TH}}$, and very low error rate floor for $\text{SNR} > \text{SNR}_{\text{TH}}$. Statistical properties of extrinsic information in these two regions can be approximately represented by a single-modal ($M = 1$) GGD. For short turbo codes with relatively low number of decoding iterations, the error performance curves exhibit a more *gentle* performance-improvement region, covering to a *wider* SNR range. As extrinsic information in the performance-improvement region experiences a critical transition from low to high reliability, it can have a *bimodal* distribution as shown in [6]. In such cases, MGGD with $M = 2$ is more appropriate. In other words, in practice, a MGGD model (4b) with $M = 2$ could be quite adequate for a wide SNR range and both long and short turbo codes with different numbers of decoding iterations

IV. SIMULATION RESULTS

This section presents the simulation results on the performance of an iterative decoder over AWGN, Rician and Rayleigh fading channels using three schemes: (i) Gaussian model [1] for extrinsic information initially proposed by C. Berrou, denoted by "*Gaussian*," (ii) proposed subsidiary MLD and MGGD model for the extrinsic information, denoted by "*Subsidiary*," and (iii) direct use of the extrinsic information generated by one component decoder to update the *a priori* probabilities for the other decoder [3], [4], denoted by "*directly update (DU)*." For the MGGD of the extrinsic information, we consider the finite set $\Omega = \{\alpha | \alpha = (8 + i)/16, i = 0, 1, 2, \cdots, 56\}$ with 57 discrete values of α.

A. Parallel Concatenated Convolutional Codes in AWGN and Rayleigh Fading Channels

We first investigate the cases of medium and long block lengths over AWGN, Rician, and Rayleigh fading channels. Two identical component 8-state RSC codes with generator ma-

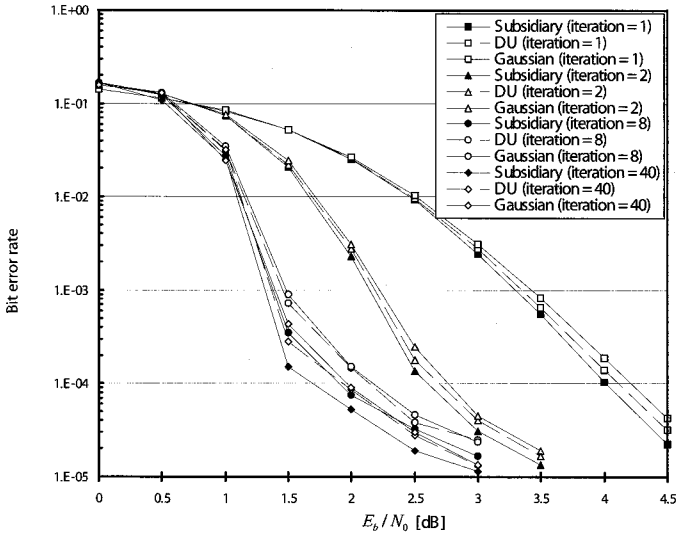

Fig. 3. Performance of an 8-state, rate-1/2 punctured turbo code with a medium block length of 1,600 bits over an AWGN channel using three different schemes for extrinsic information.

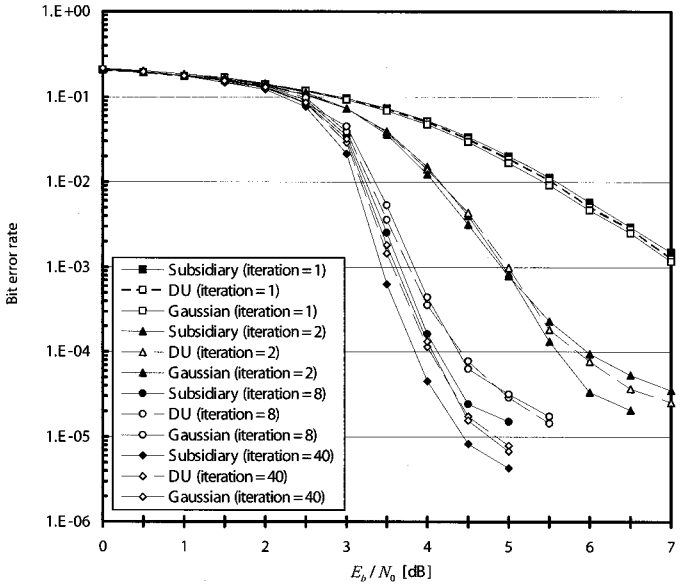

Fig. 4. Performance of an 8-state rate-1/2 punctured turbo code with medium block length of 1,600 bits over a Rayleigh fading channel using three different schemes for the extrinsic information.

trix $[1, (1+D^2+D^3)/(1+D+D^3)]$ are employed to construct a rate-1/2 punctured turbo code.

A.1 Medium Block and Various Numbers of Decoding Iterations

Figs. 3–5 present the BER performance of the iterative turbo decoder with a medium block of 1,600 bits using the three schemes over AWGN, Rician and Rayleigh fading channels with 1, 2, 8, and 40 decoding iterations. An S-random interleaver [2] with $S = 28$ and block length of 1,600 bits is used. In general, the error performance curves show that the performance-improvement region is gentle in the cases of 1 or 2 decoding iterations and becomes more and more abrupt as the number of decoding iterations increases. For the proposed decoding

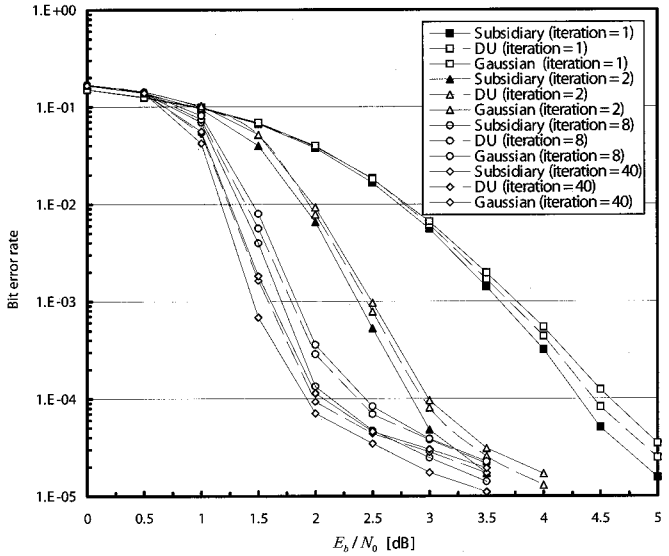

Fig. 5. Performance of an 8-state rate-1/2 punctured turbo code with medium block length of 1,600 bits over a Rician fading channel with $K = 5$ dB using three different strategies for the extrinsic information.

scheme *Subsidiary*, we used MGGD to represent the statistical properties of extrinsic information in different simulation cases and SNRs. We observed that $M = 2$ is needed for the cases of 40 decoding iterations with SNRs corresponding to the rather abrupt performance-improvement regions, i.e., $E_b/N_0 = 1$ dB an AWGN channel (Fig. 3), $E_b/N_0 = 3.5$ dB in a Rayleigh fading channel (Fig. 4), and $E_b/N_0 = 1.5$ dB in a Rician fading with $K = 5$ dB (Fig. 5). In all other cases and SNRs of Figs. 3–5, only $M = 1$ is required. The results confirm that the extrinsic information experiences critical transition from low to high reliability, and can have a *bimodal* distribution at the SNRs where an abrupt performance improvement likely happens. In a more gentle performance-improvement region, its two modes could be very close to a value and a single-modal GGD can adequately represents its statistical properties.

The results indicate that the introduced Scheme *Subsidiary* outperforms Schemes *DU* and *Gaussian* for all the considered numbers of iterations except in the 1st iteration for the Rayleigh fading channel, where Scheme *Gaussian* offers a slightly better performance over Schemes *DU* and *Subsidiary*. For the AWGN channel, as shown in Fig. 3, Scheme *Subsidiary* offers performance gains of about 0.25 dB and 0.30 dB at a BER of 10^{-4} the 8th and 40th iterations as compared to Scheme *DU*. Both Schemes DU and Gaussian have almost the same BER performance in all iterations. In a Rayleigh fading channel (Fig. 4), Scheme *Subsidiary* offers a higher performance gain as compared to Schemes *DU* and *Gaussian* in the 2nd, 8th, and 40th iterations. For instance, at a BER of 10^{-4} in the 2nd, Scheme *Subsidiary* provides a gain of 0.25 dB and 0.30 dB as compared to Schemes *DU* and *Gaussian*, respectively. For the same BER in the 8th iteration, Scheme *Subsidiary* outperforms Schemes *DU* and *Gaussian* by 0.30 dB. The gain offered by Scheme *Subsidiary* over Schemes *DU* and *Gaussian* at a BER of 10^{-5} is increased to 0.4 dB in the 40th iteration. Fig. 5 demonstrates the similar results over a Rician fading channel with a direct-to-

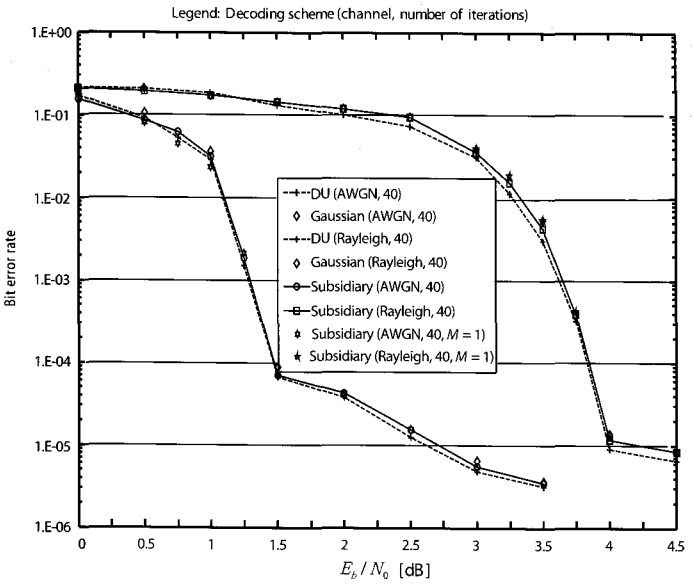

Fig. 6. Performance of an 8-state rate-1/2 punctured turbo code with long block length of 22,500 bits in AWGN and Rayleigh fading channels using three different strategies for the extrinsic information.

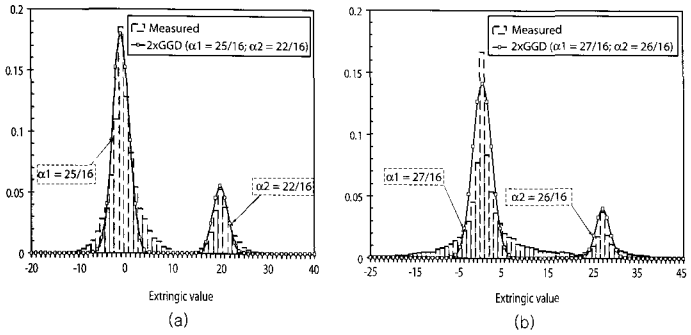

Fig. 7. Histograms and approximated MGGD pdf's of the extrinsic information at SNR in the BER waterfall regions (8-state, rate-1/2 punctured turbo code of block length of 22,500 bits, and 40 iterations): (a) Subsidiary with $M = 2, \omega_1 = 0.81, \omega_2 = 0.19, \sigma_1^2 = 3.85, \sigma_2^2 = 2.5$ (AWGN, SNR = 1.25 dB), (b) subsidiary with $M = 2, \omega_1 = 0.79, \omega_2 = 0.21, \sigma_1^2 = 4.88, \sigma_2^2 = 3.3$ (Rayleigh, SNR = 3.75 dB).

diffuse power ratio of $K = 5$ dB.

The results also show that the performance gain by increasing the number of iterations becomes much smaller from the 8th to the 40th iteration as compared to that from the 1st to the 2nd iteration and from the 2nd to the 8th iteration. This indicates that performance obtained after 40 or more iterations, approaches its limit.

A.2 Long Block and Large Number of Decoding Iterations

To examine the asymptotical behavior of the three schemes with a very long block and a sufficiently large number of iterations, we consider a 150×150 block-interleaver with 40 decoding iterations. Simulation runs for different cases were performed over an E_b/N_0 range from 0 dB to 4.5 dB with a normal step of 0.5 dB. In the BER waterfall regions, i.e., 0.5 dB~1.5 dB for AWGN and 3 dB~4 dB for Rayleigh, an E_b/N_0 step of 0.25 dB is used to obtain more detailed results. The simulation results in Fig. 6 over both AWGN and Rayleigh fading channels confirm that Scheme *DU* is an optimal approach [3] for a very long block and a sufficiently large number of iterations. The performance of Scheme *Subsidiary* closely approaches that of Scheme *DU*, and is lightly better than that of Scheme *Gaussian* over all SNRs in both channels. Scheme *Subsidiary* uses the MGGD and we notice that $M = 2$ is only needed in the BER waterfall regions, i.e., 0.5 dB~1.5 dB and 3 dB~4 dB for AWGN and Rayleigh fading channels, respectively. Fig. 7 shows the example plots of the histograms and MGGD pdf of the extrinsic information, at the central points of the steepest waterfall regions, i.e., $E_b/N_0 = 1.25$ dB and $E_b/N_0 = 3.75$ dB, for AWGN and Rayleigh fading channels, respectively. For other SNRs outside of the BER waterfall regions, only $M = 1$ is required for MGGD used in Scheme *Subsidiary*. The simulation results again confirm the bimodal properties of the extrinsic information in the abrupt BER waterfall regions. However, three schemes have a comparable performance. This indicates that the bimodal characteristics of the extrinsic information do not have strong influence on the BER performance in the case of a very long block and a sufficiently large number of iterations. By simulation, we observe that Scheme *Subsidiary* with single-modal GGD model can also provide a performance very close to that of Scheme *DU* as shown in Fig. 6.

B. Serially Concatenated Convolutional Codes in AWGN Channels

We also examined the performance of serially concatenated convolutional code (SCCC) using three schemes. The SCCC is a rate-1/3 code by concatenating a 4-state rate-1/2 recursive convolutional code as an outer code and a 4-state rate-2/3 recursive convolutional code as an inner code [7]. The generator matrices for the outer and inner code are G_1 and G_2, respectively, where

$$G_1 = \left[1 \quad \frac{1+D^2}{1+D+D^2}\right], \quad G_2 = \begin{bmatrix} 1 & 0 & \frac{1+D^2}{1+D+D^2} \\ 0 & 1 & \frac{1+D}{1+D+D^2} \end{bmatrix} \tag{10}$$

Fig. 8 shows the simulation results with interleaving depths of 40×40 (medium) and 150×150 (long) with 3, 20, and 40 iterations. All the error performance curves have a rather gentle performance-improvement region and, for all cases and SNRs, only $M = 1$ is needed. This indicates that in a rather gentle performance-improvement region, the single-modal GGD can adequately model the extrinsic information in different SNRs.

The simulation results indicates that Scheme *Subsidiary* outperforms Scheme *DU* and *Gaussian* for medium block and has a similar performance as Scheme DU for long block in both the 20st and 40st decoding iterations. In general, performance results obtained with 20 and 40 iterations are not much different, i.e., they approach the limit.

C. Statistical Characteristics of Extrinsic Information

To provide an insight of the statistical characteristics of the extrinsic information, we consider a 4-state, rate-1/3, nonpunctured turbo code with the component RSC shown in Fig. 2, and a 40×40 (short) block-interleaver.

We first investigate the statistical characteristics of the extrinsic information at low SNRs outside the BER waterfall regions.

Table 2. Distributions of α (in %) in various cases.

	Range of α									
	[8/16, 15/16)	[15/16, 22/16)	[22/16, 29/16)	[29/16, 31/16)	**2**	[33/16, 36/16)	[36/16, 43/16)	[43/16, 50/16)	[50/16, 57/16)	[57/16, 64/16]
AWGN, 3rd iteration, SNR = −2 dB, DEC1	0.0	0.0	63.0	14.7	**5.2**	16.6	0.5	0.0	0.0	0.0
AWGN, 3rd iteration, SNR = −2 dB, DEC2	0.0	0.0	28.0	26.4	**8.8**	33.6	3.2	0.0	0.0	0.0
AWGN, 3rd iteration, SNR = 1 dB, DEC1	0.0	0.0	6.9	12.1	**4.2**	15.3	19.9	13.8	12.7	15.1
AWGN, 3rd iteration, SNR = 1 dB, DEC2	0.0	0.0	13.1	16.2	**4.8**	14.5	17.4	10.3	12.0	11.7
AWGN, 12th iteration, SNR = 1 dB, DEC1	0.0	11.0	59.4	0.72	**0.22**	1.06	0.5	0.7	0.6	25.8
AWGN, 12th iteration, SNR = 1 dB, DEC2	0.0	2.1	66.6	1.52	**0.44**	1.54	0.5	0.7	1.0	25.6
Rayleigh, 3rd iteration, SNR = 0 dB, DEC1	0.0	0.0	17.0	30.3	**7.8**	29.1	13.9	1.9	0.0	0.0
Rayleigh, 3rd iteration, SNR = 0 dB, DEC2	0.0	0.0	52.8	19.1	**5.0**	20.7	1.50	0.0	0.0	0.0
Rayleigh, 3rd iteration, SNR = 3 dB, DEC1	0.0	0.0	10.4	21.6	**0.4**	21.7	22.3	8.6	7.1	1.9
Rayleigh, 3rd iteration, SNR = 3 dB, DEC2	0.0	0.0	27.1	24.1	**7.2**	23.5	11.5	3.9	1.7	1.0
Rayleigh, 12th iteration, SNR = 3 dB, DEC1	0.0	0.0	41.7	22.3	**7.1**	24.8	1.7	0.5	0.2	1.7
Rayleigh, 12th iteration, SNR = 3 dB, DEC2	0.0	0.0	58.1	17.2	**4.0**	16.8	0.8	0.1	0.4	1.7

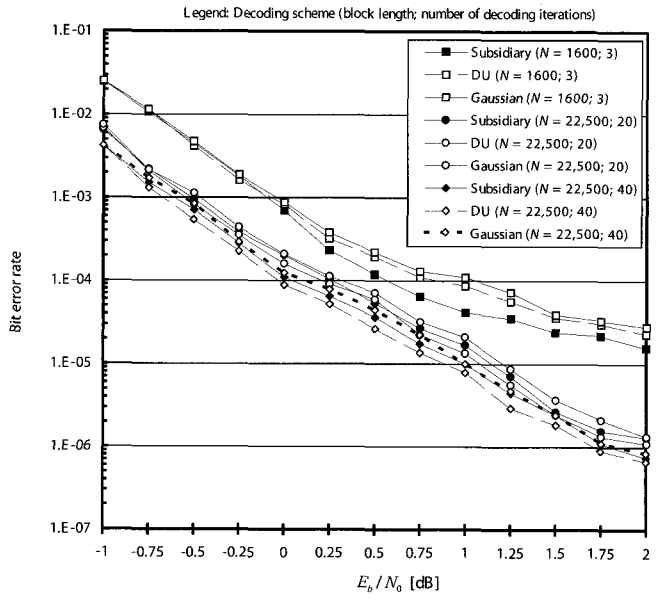

Fig. 8. Performance of an 8-state rate-1/3 SCCC in an AWGN channel using three different strategies for the extrinsic information with different block lengths and numbers of iterations.

Table 2 summarizes the statistical distributions of the selected parameter α (over a range from 0.5 to 4.0) of the GGD model of the extrinsic information sent to the two component decoders for different SNRs based on the simulation results of 10^6 Monte-Carlo tests. At low SNR of -2 dB and the 3rd iteration, 63.0% of the selected values of α fall in the interval of $[22/16, 29/16]$ and only 5.2% for exactly chosen at 2 (Gaussian) for the first decoder. For the second decoder, only 8.8% of its selected values at 2. At SNR of 1 dB for the same iteration, the selected values of α spread over a large range $[22/16, 64/16]$ with only 4.2% and 4.8% of its selected values at 2 for the the 1st and 2nd decoders, respectively. For SNR of 1 dB and the 12th iteration in an AWGN channel, the distribution of the selected values of α is quite different from that in the 3rd iteration with 59.4% (for $1st$ decoder) and 66.6% (for 2nd decoder) in $[22/16, 29/16]$, and the percentage for α at 2 is negligibly small for both decoders. The statistical distributions of the selected values of α for two decoders from the 3rd and 12th iterations are closer for SNR of 3 dB in a Rayleigh fading channel. The percentage for α at 2 is also small for both decoders at 3rd and 12th iterations and SNR of 0 dB and 3 dB in the case of Rayleigh fading channel.

Fig. 9 shows the measured histograms and their corresponding approximated pdf's of the extrinsic information sent to the first decoder in the 3rd decoding iteration for 3 schemes: *Gaussian*, *Subsidiary*, and *DU*, over AWGN (with SNR of -2 dB) and Rayleigh fading (with SNR of 0 dB) channels. It shows that the approximated GGD pdf's closely follow the corresponding measured histograms of the extrinsic information obtained by Scheme *Subsidiary* in both AWGN (Fig. 9(b) with $\alpha = 24/16$) and Rayleigh (Fig. 9(e) with $\alpha = 29/16$) channels while Scheme *Gaussian* provides much poorer approximated pdf's as shown in Figs. 9(a) and 9(d). Furthermore, the approximated GGD pdf's obtained by Scheme *Subsidiary* in Fig. 9(b) and Fig. 9(e) are also close to the measured histograms of the extrinsic information obtained by Scheme *DU* in Fig. 9(c) and Fig. 9(f), respectively.

Fig. 10 shows the *bimodal* distribution of the extrinsic information for the first decoder at the 20th iteration in a waterfall region over AWGN (with SNR of 0.4 dB) and Rayleigh fading (with SNR of 1.0 dB) channels as discussed in [6], and the cor-

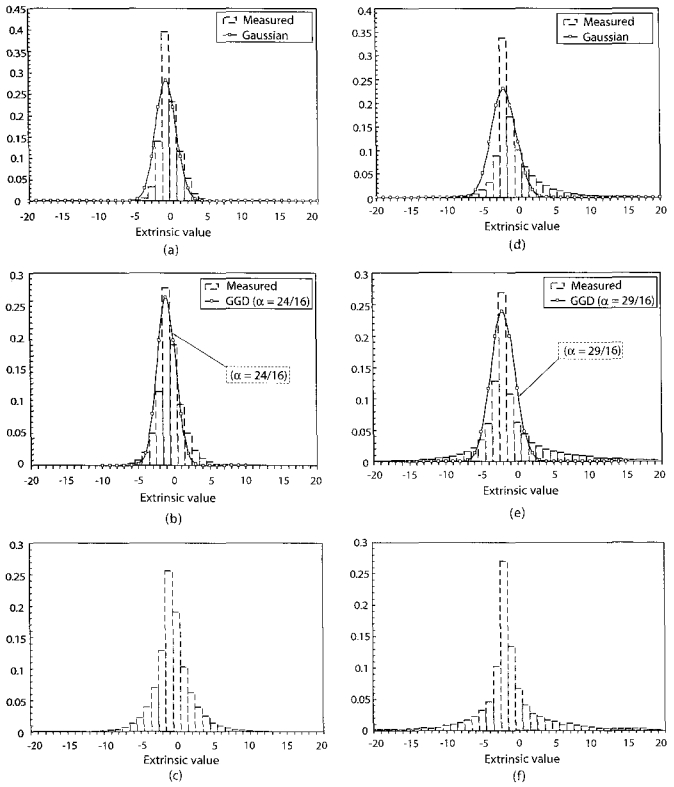

Fig. 9. Histograms and approximated pdf's of the extrinsic information (4-state, rate-1/3 non-punctured turbo code with block length of 1,600 bits at the 3rd iteration, and SNR outside the BER waterfall regions): (a) Gaussian (AWGN, SNR $= -2$ dB, $\sigma^2 = 2.02$), (b) Subsidiary with $M = 1$ (AWGN, SNR $= -2$ dB, $\sigma^2 = 1.77$), (c) DU (AWGN, SNR $= -2$ dB), (d) Gaussian (Rayleigh, SNR $= 0$ dB, $\sigma^2 = 3.0$), (e) Subsidiary with $M = 1$ (Rayleigh, SNR $= 0$ dB, $\sigma^2 = 2.76$), (f) DU (Rayleigh, SNR $= 0$ dB).

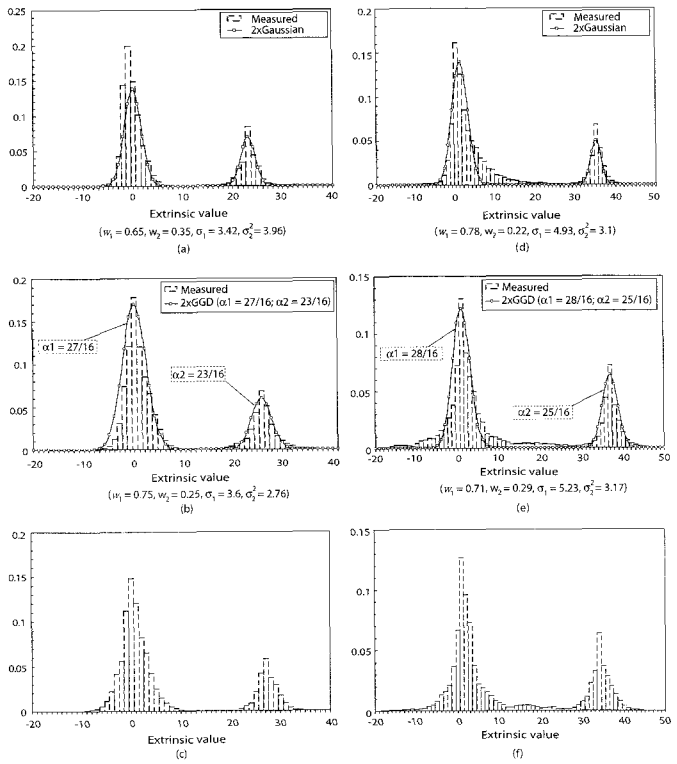

Fig. 10. Histograms and approximated pdf's of the extrinsic information (4-state, rate-1/3 non-punctured turbo code with block length of 1,600 bits at the 20th iteration, and SNR in the BER waterfall regions): (a) Gaussian (AWGN, SNR $= 0.4$ dB), (b) Subsidiary with $M = 2$ (AWGN, SNR $= 0.4$ dB), (c) DU (AWGN, SNR $= 0.4$ dB), (d) Gaussian (Rayleigh, SNR $= 1$ dB), (e) Subsidiary with $M = 2$ (Rayleigh, SNR $= 1$ dB), (f) DU (Rayleigh, SNR $= 1$ dB).

responding approximated *bimodal* pdf's for 3 schemes.

The approximated *bimodal* GGD pdf's for Scheme *Subsidiary* are very close to the measure histograms in both channels: AWGN (Fig. 10(b) with $\alpha_1 = 27/16$, $\omega_1 = 0.75$ and $\alpha_2 = 23/16$, $\omega_2 = 0.25$) and Rayleigh (Fig. 10(e) with $\alpha_1 = 28/16$, $\omega_1 = 0.71$ and $\alpha_2 = 25/16$, $\omega_2 = 0.29$). The value σ_i^2 estimated from each lobe is the variance of the pdf employed to match the lobe by the scaling factor ω_i. The approximated *bimodal* GGD pdf's obtained by Scheme *Subsidiary* in Fig. 10 (e) and Fig. 10(e) are also close to the measured *bimodal* histograms of the extrinsic information obtained by Scheme *DU* in Figs. 10(c) and 10(f), respectively.

For a fair comparison, in Figs. 10(a) and 10(d), we considered *bimodal Gaussian* pdf's by using (4b) but with $M = 2$ and $\alpha_i = 2$, $(i = 1, 2)$. The results in these figures show that Scheme *Gaussian* provides much poorer approximated pdf's for the extrinsic information than Scheme *Subsidiary* in the waterfall region.

V. CONCLUSION

Simulation results indicate that the statistical characteristics of the extrinsic information vary in different iterations and do not follow the exact Gaussian model. The MGGD with selectable parameters is shown to be more effective in dynamically characterizing the extrinsic information generated by the component MAP decoders of a turbo decoder. A component decoder can choose the most appropriate MGGD of the extrinsic information with the optimum parameters based on the proposed subsidiary MLD and selection criterion at each decoding iteration to enhance the overall turbo decoding performance. Simulation results show that for short and medium block lengths the turbo decoder using the proposed scheme outperforms the conventional schemes using direct approach and Gaussian model for the extrinsic information in AWGN, Rician, and Rayleigh fading channels and for both parallel and serially concatenated codes. For long blocks and large number of iterations, as the *a priori* information becomes statistically independent, the performance of the proposed scheme approaches that of the direct approach. In summary, the MGGD family can provide a very effective way to trace the extrinsic information that is dynamically varied with the SNRs and decoding iterations. The simulation results indicate that in an abrupt performance-improvement region $M = 2$ is needed while in most other cases, $M = 1$ is adequate. The performance improvement over all SNR ranges demonstrated by the simulations proves the effectiveness of the proposed approach using subsidiary MLD and GGD model for the extrinsic information. The simulation results indicate that in an abrupt performance-improvement region $M = 2$ is needed while in most other cases, $M = 1$ is adequate.

REFERENCES

[1] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo codes," *IEEE Trans. Commun.*, vol. 44, pp. 1261–1271, Oct. 1996.

[2] W. Feng, J. Yuan, and S. Vucetic, "A code-matched interleaver design for turbo codes," *IEEE Trans. Commun.*, vol. 50, pp 926–937, June 2002.

[3] G. Colavolpe, G. Ferrari, and R. Raheli, "Extrinsic information in iterative decoding: A unified view," *IEEE Trans. Commun.*, vol. 49, pp. 2088–2094, Dec. 2001.

[4] P. Robertson, "Illuminating the structure of parallel concatenated recursive systematic (TURBO) codes," in *Proc. IEEE Globecom'94*, San Francisco, CA, 1994, pp. 1298–1303.

[5] N. Farvardin and V. Vaishampayan, "Optimum quantizer design for noisy channels: An approach to combined source-channel coding," *IEEE Trans. Inf. Theory*, vol. 33, pp. 827–838, Nov. 1987.

[6] J. W. Lee and R. E. Blahut, "Bit error rate of finite length turbo codes," in *Proc. ICC 2003*, Anchorage, AK, 2003, pp. 2728–2732.

[7] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: Performance analysis, design, and iterative decoding," *IEEE Trans. Inf. Theory*, vol. 44, pp. 909–926, May 1998.

Fengfan Yang received the B.Sc., M.Sc., and Ph.D. degrees from Nanjing University of Aeronautics & Astronautics (NUAA), Northwestern Polytechnical University and Southeast University in 1990, 1993, and 1997 in P. R. China, respectively, all in electronic engineering. He has been with College of Information Science and Technology, Nanjing University of Aeronautics & Astronautics since May 1997. From October 1999 to May 2003, he was a research associate at Centre for Communication Systems Research, University of Surrey, UK, and Department of Electrical and Computer Engineering (ECE), McGill University, Canada. He currently spends one year sabbatical leave as a visiting scholar at ECE, McGill University. His major research interests are information theory and channel coding, especially at iteratively decodable codes, such as turbo codes and LDPC codes, and their applications for mobile and satellite communications.

Tho Le-Ngoc obtained his B.Eng. (with Distinction) in Electrical Engineering in 1976, his M.Eng. in Microprocessor Applications in 1978 from McGill University, Montreal, and his Ph.D. in Digital Communications 1983 from the University of Ottawa, Canada. During 1977–1982, he was with Spar Aerospace Limited involved in the development and design of satellite communications systems. During 1982–1985, he was an engineering manager of the Radio Group in the Department of Development Engineering of SRTelecom Inc., developed the new point-to-multipoint DA-TDMA/TDM Subscriber Radio System SR500. During 1985–2000, he was a professor the Department of Electrical and Computer Engineering of Concordia University. Since 2000, he has been with the Department of Electrical and Computer Engineering of McGill University. His research interest is in the area of broadband digital communications with a special emphasis on Modulation, Coding, and Multiple-Access Techniques. He is a senior member of the Ordre des Ingénieur du Quebec, a fellow of the Institute of Electrical and Electronics Engineers (IEEE), a fellow of the Engineering Institute of Canada (EIC), and a fellow of the Canadian Academy of Engineering (CAE). He is the recipient of the 2004 Canadian Award in Telecommunications Research, and recipient of the IEEE Canada Fessenden Award 2005.