

# Efficient Computation of EXIT Functions for Nonbinary Iterative Decoding

Jörg Kliewer, *Senior Member, IEEE*, Soon Xin Ng, *Member, IEEE*, and Lajos Hanzo, *Fellow, IEEE*

**Abstract**—The calculation of nonbinary extrinsic information transfer charts for the iterative decoding of concatenated index-based codes is addressed. We show that the extrinsic information at the output of a constituent *a posteriori* probability decoder can be calculated with very low complexity, where expensive histogram measurements are not required anymore. An example for turbo trellis-coded modulation demonstrates the capabilities of the proposed approach.

**Index Terms**—Extrinsic information transfer (EXIT) charts, nonbinary iterative decoding, turbo trellis-coded modulation (TTCM).

## I. INTRODUCTION

EXTRINSIC information transfer (EXIT) charts have recently emerged as a useful tool for analyzing the convergence properties of iterative decoding for concatenated coding schemes when the block length tends to infinity [1], [2]. As an advantage, this can be achieved *without* performing a costly decoder simulation. For example, the information provided by EXIT charts can be used to find good codes with guaranteed convergence for a given channel condition. Especially, capacity-approaching codes have been designed successfully by applying an EXIT-chart-based analysis [3], [4]. In [5], EXIT charts have been extended to the nonbinary (index-based) case, where serially concatenated systems with an inner trellis-coded modulation (TCM) and a space-time convolutional code, respectively, have been analyzed. Recently, in [6], this approach was applied to turbo TCM (TTCM) [7].

EXIT charts are composed from the input-output (or EXIT) characteristics for all constituent soft-input soft-output (SISO) decoders in terms of extrinsic mutual information. The standard method to calculate EXIT characteristics is to measure the histogram of the decoder's extrinsic soft output, followed by a numerical integration in order to determine the extrinsic mutual information. However, in the nonbinary case, the histogram computation becomes quite demanding, and even impossible for larger index lengths. This is due to the fact that the soft values,

Paper approved by A. H. Banihashemi, the Editor for Coding and Communication Theory of the IEEE Communications Society. Manuscript received June 22, 2005; revised February 28, 2006. This work was supported in part by the Leverhulme Trust, London, U.K., in part by EPSRC, Swindon, U.K., and in part by the EU under the auspices of the Phoenix and Newcom projects. This paper was presented in part at the 4th International Symposium on Turbo Codes and Related Topics, Munich, Germany, April 2006.

J. Kliewer was with the University of Southampton, Communications Research Group, Southampton, SO17 1BJ, U.K. He is now with the University of Notre Dame, Department of Electrical Engineering, Notre Dame, IN 45561 USA (e-mail: jkiewer@nd.edu).

S. X. Ng and L. Hanzo are with the University of Southampton, Communications Research Group, Southampton, SO17 1BJ, U.K. (e-mail: sxn@ecs.soton.ac.uk; lh@ecs.soton.ac.uk).

Digital Object Identifier 10.1109/TCOMM.2006.885050

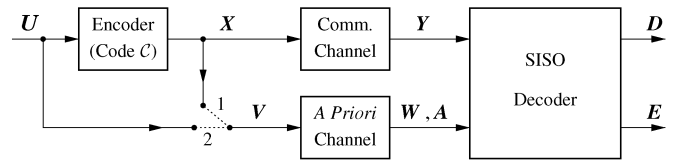


Fig. 1. Decoding model.

which are exchanged between the constituent decoders of an iterative decoding scheme, no longer consist of a single scalar value, as in the binary case, but now are represented by a vector of (logarithmic) probabilities for all elements of the index alphabet.

In this letter, we propose an efficient method for computing nonbinary EXIT charts from index-based *a posteriori* probabilities (APPs), which essentially represents a generalization of the approach presented in [8] for the binary case. The proposed technique is based on the fact that the index-based APPs generated at the output of a SISO decoder represent a sufficient statistic for the channel observations at its input. Recently, this fact was also employed in [9] for computing EXIT charts for nonbinary coset low-density parity-check (LDPC) codes. In contrast, the presented method can be employed for two or more nonbinary parallel or serially concatenated encoders and the corresponding iterative decoding schemes with constituent SISO decoders. As an advantage, the index-based extrinsic mutual information can be computed with considerably lower complexity, compared with using histogram-based approaches. The proposed method is well suited to analyzing nonbinary iterative decoding schemes with a large index alphabet, where an example will be given for a TTCM system employing a 32-quadrature amplitude modulation (QAM) signal set.

## II. NOTATION AND DECODING MODEL

Random variables (RVs) are denoted with upper-case letters, and their corresponding realizations with lower-case letters. Sequences of RVs and realizations are indicated by boldface italic letters (as “ $\mathbf{U}$ ” or “ $\mathbf{u}$ ”). Furthermore, boldface roman letters denote vectors (as “ $\mathbf{E}$ ” or “ $\mathbf{e}$ ”). The time instant is denoted with “ $k$ ,” and a bit-index with “ $\ell$ .”

The decoding model used in this paper is depicted in Fig. 1 [2]. For the iterative decoding of an outer code in a serial concatenated scheme, the connection marked with “1” is active in Fig. 1, and the communication channel is not present or completely corrupted. All other scenarios for iterative decoding (inner serially concatenated code and parallel concatenation) have a constituent decoder modeled by Fig. 1 with the switch in position “2.” In Fig. 1, the stationary input index sequence  $\mathbf{U} = [U_1, U_2, \dots, U_k, \dots]$  consists of RVs  $U_k$ , where the corresponding realizations  $u_k$  have an index length of  $M$  bits, and are taken from a finite alphabet  $\mathcal{U} = \{0, 1, \dots, 2^M - 1\}$ . For each symbol, a channel code  $\mathcal{C}$  with rate  $R = M/N$

generates an  $N$ -bit index  $X_k$ , which leads to the sequence  $\mathbf{X} = [X_1, X_2, \dots, X_k, \dots]$ . The data is then transmitted over the communication channel where the sequence  $\mathbf{Y}$  is observed at the output.

The *a priori* channel (or extrinsic channel, as it is denoted in [2]) models the *a priori* information used at the constituent decoders. The input sequence  $\mathbf{V} = [V_1, V_2, \dots, V_k, \dots]$  has realizations  $v_k$  from the alphabet  $\mathcal{V} = \{0, 1, \dots, 2^K - 1\}$ , where  $K = N$  if, in Fig. 1, the switch is in position "1" (i.e., in the case of an outer decoder for serial concatenation), and  $K = M$ , otherwise. For the noisy observation  $\mathbf{W}$  at the output of the *a priori* channel, the same considerations hold as for the output  $\mathbf{Y}$  of the communication channel. Furthermore, the vector sequence  $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_k, \dots]$  contains the *a priori* information in form of conditional probability density functions (pdfs) with the vector-valued RVs  $\mathbf{A}_k$  having the realizations

$$\mathbf{a}_k = [p(w_k|V_k=0), p(w_k|V_k=1), \dots, p(w_k|V_k=2^K-1)]. \quad (1)$$

The SISO decoder then employs both the output of the communication and the *a priori* channel for computing both the *a posteriori* information  $\mathbf{D} = [\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_k, \dots]$  and extrinsic information  $\mathbf{E} = [\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_k, \dots]$  [10]. The latter is then used as *a priori* information for the other constituent decoder. The sequence  $\mathbf{D}$  comprises the RVs  $\mathbf{D}_k$  with the realizations

$$\mathbf{d}_k = [P(U_k = 0|\mathbf{y}, \mathbf{w}), P(U_k = 1|\mathbf{y}, \mathbf{w}), \dots, P(U_k = 2^M - 1|\mathbf{y}, \mathbf{w})] \quad (2)$$

where  $P(U_k = u_k|\mathbf{y}, \mathbf{w})$  represents the APPs for the source hypothesis  $U_k = u_k$ . The extrinsic sequence  $\mathbf{E}$  is described by the RVs  $\mathbf{E}_k$  with

$$\mathbf{e}_k = [P(V_k = 0|\mathbf{y}, \mathbf{w}_{[k]}), P(V_k = 1|\mathbf{y}, \mathbf{w}_{[k]}), \dots, P(V_k = 2^K - 1|\mathbf{y}, \mathbf{w}_{[k]})] \quad (3)$$

where the notation  $\mathbf{w}_{[k]}$  denotes that the entry corresponding to the time instant  $k$  is excluded from the sequence  $\mathbf{w}$ .

### III. INDEX-BASED EXIT CHARTS

EXIT charts visualize the input–output characteristics of the constituent SISO decoders in terms of a mutual information transfer between the sequence  $\mathbf{V}$  and *a priori* information  $\mathbf{A}$  at the input of the decoder, as well as between  $\mathbf{V}$  and  $\mathbf{E}$  at the output, respectively. Denoting the mutual information between two RVs  $X$  and  $Y$  as  $I(X; Y)$ , we may define for a given length  $Q$  of the sequence  $\mathbf{V}$  the quantities

$$I_A := \frac{1}{Q} \sum_{k=1}^Q I(V_k; \mathbf{A}_k), \quad I_E := \frac{1}{Q} \sum_{k=1}^Q I(V_k; \mathbf{E}_k). \quad (4)$$

Herein,  $I_A$  represents the average *a priori* information and  $I_E$  the average extrinsic information, respectively. Note that these quantities are defined as approximations for the sequence-wise mutual information per index  $1/Q \cdot I(\mathbf{V}; \mathbf{A})$  and  $1/Q \cdot I(\mathbf{V}; \mathbf{E})$ ,

respectively. The transfer characteristic (or function)  $T$  of the constituent decoder is given as  $I_E = T(I_A, \rho)$ , where  $\rho$  represents a quality parameter for the communication channel which could be the noise variance of an additive white Gaussian noise (AWGN) channel, for example. In the case of an outer decoder in a serially concatenated scheme,  $T$  does not depend on  $\rho$ . An EXIT chart can now be obtained by plotting the transfer characteristics  $T$  for both constituent decoders within a single diagram, where the axes must be swapped for one of the constituent decoders [1] (normally, the outer one for serial concatenation).

The mutual information  $I(V_k; \mathbf{A}_k)$  in (4) can be expressed as

$$I(V_k; \mathbf{A}_k) = \sum_{v_k=0}^{2^K-1} \int_{\mathbf{a}_k} p(\mathbf{a}_k|v_k) P(v_k) \log_2 \left( \frac{p(\mathbf{a}_k|v_k)}{p(\mathbf{a}_k)} \right) d\mathbf{a}_k \quad (5)$$

with

$$p(\mathbf{a}_k) = \sum_{v_k=0}^{2^K-1} p(\mathbf{a}_k|v_k) P(v_k) \quad (6)$$

and the *a priori* probabilities  $P(v_k)$  for the indices  $v_k$ . The  $2^K$ -dimensional integration in (5) can be evaluated numerically, where the pdf  $p(\mathbf{a}_k|v_k)$  in many cases can be obtained analytically (e.g., for the AWGN channel [1]). Likewise, we can write  $I(V_k; \mathbf{E}_k)$  in the same manner as in (5) and (6), where  $\mathbf{A}_k$  is replaced by  $\mathbf{E}_k$ . The standard approach to determine  $p(\mathbf{e}_k|v_k)$  is by computing a  $2^K$ -dimensional histogram [1], [5], [6] which, in practice, is exceedingly difficult to compute for  $K > 3$ . Furthermore, the evaluation of the  $2^K$ -dimensional integral in (5) becomes almost impossible for larger index lengths  $K$ . In fact, for a discretization of the integrals with  $L$  points corresponding to a certain interval of the logarithmic vector elements<sup>1</sup> in  $\mathbf{a}_k$  and  $\mathbf{e}_k$  (say, e.g.,  $[-20, 0]$ ) we need to sum over  $L^{2^K}$  discrete numbers. Even a calculation of the mutual information for  $K = 3$  bits and  $L = 100$  points (which should be sufficient, according to our experiments) involves a summation over  $10^{16}$  elements; for  $K = 4$ , this number rises to  $10^{32}$  elements. Therefore, the calculation of an index-based transfer function seems to be computationally feasible only for 2-bit or 3-bit indices without losing the computational advantage of an EXIT-chart based design, compared with a full decoder simulation. In the next section, we will present a more feasible way of computing index-based transfer characteristics.

### IV. EFFICIENT COMPUTATION OF INDEX-BASED EXIT TRANSFER FUNCTIONS

Starting point of the derivation of the proposed approach is the following theorem, which relates the extrinsic information  $\mathbf{E}_k$  at time instant  $k$  to the observed sequences  $\mathbf{Y}$  and  $\mathbf{W}$  at the input of the SISO decoder in Fig. 1.

*Theorem 1:* Let the entries of the vector  $\mathbf{E}_k$  represent extrinsic APPs for the realizations  $v_k \in \mathcal{V}$ . Then, the vector  $\mathbf{E}_k$

<sup>1</sup>Instead of using the pdfs/probability mass functions in both  $\mathbf{a}_k$  and  $\mathbf{e}_k$  directly, the logarithm of these quantities is employed for numerical reasons. However, this does not change the value of  $I(V_k; \mathbf{A}_k)$  and  $I(V_k; \mathbf{E}_k)$ .

contains the same amount of information as the sequences  $\mathbf{Y}$  and  $\mathbf{W}_{[k]}$ . In other words, we have

$$I(V_k; \mathbf{E}_k) = I(V_k; \mathbf{Y}, \mathbf{W}_{[k]}).$$

In order to prove *Theorem 1*, we need the following lemma.

*Lemma 1:* Let  $\mathbf{e}_k$  be defined as in (3). Then, for all  $v_k \in \mathcal{V}$ ,  $P(v_k|\mathbf{e}_k) = P(v_k|\mathbf{y}, \mathbf{w}_{[k]})$ .

*Proof:* The proof can be carried out as a modification of the binary case in [8], where in contrast to [8], the following proof holds for an arbitrary *a priori* distribution of the indices  $V_k$ . The conditional pdf of  $\mathbf{e}_k$  with respect to the index  $v_k$  can be expressed with the conditional pdf  $p(\mathbf{y}, \mathbf{w}_{[k]}|V_k = v_k)$  as

$$p(\mathbf{e}_k|V_k = v_k) = \int_{\mathbf{y}, \mathbf{w}_{[k]}:\mathbf{e}_k} p(\mathbf{y}, \mathbf{w}_{[k]}|V_k = v_k) d\mathbf{y}d\mathbf{w}_{[k]} \quad (7)$$

where the notation “ $\mathbf{y}, \mathbf{w}_{[k]} : \mathbf{e}_k$ ” means that for a given  $\mathbf{e}_k$ , the integration is carried out over the whole set  $\{\mathbf{y}, \mathbf{w}_{[k]} : [P(V_k = 0|\mathbf{y}, \mathbf{w}_{[k]})], P(V_k = 1|\mathbf{y}, \mathbf{w}_{[k]}), \dots, P(V_k = 2^K - 1|\mathbf{y}, \mathbf{w}_{[k]})] \stackrel{\perp}{=} \mathbf{e}_k\}$ . From (7), by applying the Bayes theorem and by using the definition  $e_{k,\lambda} := P(V_k = \lambda|\mathbf{y}, \mathbf{w}_{[k]})$ ,  $\lambda \in \mathcal{V}$ , for the  $\lambda$ th entry of the vector  $\mathbf{e}_k$ , we obtain

$$\begin{aligned} p(\mathbf{e}_k|V_k = \lambda) &= \frac{e_{k,\lambda}}{P(V_k = \lambda)} \int_{\mathbf{y}, \mathbf{w}_{[k]}:\mathbf{e}_k} p(\mathbf{y}, \mathbf{w}_{[k]}) d\mathbf{y}d\mathbf{w}_{[k]} \\ &= \frac{e_{k,\lambda}}{P(V_k = \lambda)} p(\mathbf{e}_k). \end{aligned} \quad (8)$$

Herein, a similar relation as in (7) is employed for  $p(\mathbf{y}, \mathbf{w}_{[k]})$  without the dependence on  $V_k$ . From (8), we obtain  $P(V_k = \lambda|\mathbf{e}_k) = e_{k,\lambda} = P(V_k = \lambda|\mathbf{y}, \mathbf{w}_{[k]})$ , which completes the proof with  $\lambda = v_k$ . ■

*Proof of Theorem 1:* The mutual information  $I(V_k; \mathbf{Y}, \mathbf{W}_{[k]})$  can be expanded as

$$\begin{aligned} I(V_k; \mathbf{Y}, \mathbf{W}_{[k]}) &= \sum_{v_k=0}^{2^K-1} \int \int p(v_k, \mathbf{y}, \mathbf{w}_{[k]}) \log_2 \left( \frac{P(v_k|\mathbf{y}, \mathbf{w}_{[k]})}{P(v_k)} \right) d\mathbf{y}d\mathbf{w}_{[k]} \\ &= \sum_{v_k=0}^{2^K-1} \int \int \int p(v_k, \mathbf{y}, \mathbf{w}_{[k]}, \mathbf{e}_k) \log_2 \left( \frac{P(v_k|\mathbf{e}_k)}{P(v_k)} \right) d\mathbf{e}_k d\mathbf{y}d\mathbf{w}_{[k]} \\ &= \sum_{v_k=0}^{2^K-1} \int_{\mathbf{e}_k} p(v_k, \mathbf{e}_k) \log_2 \left( \frac{P(v_k|\mathbf{e}_k)}{P(v_k)} \right) d\mathbf{e}_k \\ &= I(V_k; \mathbf{E}_k) \end{aligned} \quad (9)$$

where (9) uses *Lemma 1* and the definition of a marginal pdf. ■

Note that *Theorem 1* represents an index-based formulation of the well-known fact that the APPs at the output of the SISO decoder represent a sufficient statistic of the received sequences  $\mathbf{y}$  and  $\mathbf{w}_{[k]}$ . In the following, we show that *Theorem 1* can be used to compute the average extrinsic information from the extrinsic APPs available at the SISO decoder output.

*Theorem 2:* Let the extrinsic APP  $P(v_\kappa|\mathbf{y}, \mathbf{w}_{[\kappa]})$  represent an ergodic RV for all  $v_\kappa \in \mathcal{V}$ , and let  $\mathbf{V}$  be a stationary index sequence. Then, the average extrinsic information is given by

$$I_E = H(V_1) + \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{\kappa=1}^L \sum_{v_\kappa=0}^{2^K-1} P(v_\kappa|\mathbf{y}, \mathbf{w}_{[\kappa]}) \cdot \log_2 (P(v_\kappa|\mathbf{y}, \mathbf{w}_{[\kappa]})).$$

*Proof:* Combining *Theorem 1* and (4), we can write, by using the assumed stationary of  $\mathbf{V}$

$$\begin{aligned} I_E &= \frac{1}{Q} \sum_{k=1}^Q I(V_k; \mathbf{Y}, \mathbf{W}_{[k]}) \\ &= H(V_1) - \underbrace{\frac{1}{Q} \sum_{k=1}^Q H(V_k|\mathbf{Y}, \mathbf{W}_{[k]})}_{=:A}. \end{aligned} \quad (10)$$

The sum over the conditional entropies denoted with  $A$  in (10) can be further expanded as

$$\begin{aligned} A &= -\frac{1}{Q} \sum_{k=1}^Q \sum_{v_k=0}^{2^K-1} \int \int p(v_k|\mathbf{y}, \mathbf{w}_{[k]}) \log_2 (P(v_k|\mathbf{y}, \mathbf{w}_{[k]})) d\mathbf{y}d\mathbf{w}_{[k]} \\ &= -E \left\{ \frac{1}{Q} \sum_{k=1}^Q \sum_{v_k=0}^{2^K-1} P(v_k|\mathbf{y}, \mathbf{w}_{[k]}) \log_2 (P(v_k|\mathbf{y}, \mathbf{w}_{[k]})) \right\}. \end{aligned} \quad (11)$$

Using the assumed ergodicity for the extrinsic APPs and defining  $L := BQ$ , where  $B$  denotes the number of length- $Q$  blocks to be averaged, proves the theorem for  $B \rightarrow \infty$ . ■

The result from *Theorem 2* can now be used for finite  $L$  to approximate the average extrinsic information  $I_E$  by simply time-averaging a function of the extrinsic APPs  $P(v_\kappa|\mathbf{y}, \mathbf{w}_{[\kappa]})$  over a block of  $L$  source symbols  $v_\kappa$ . In order to achieve a small approximation error variance, the block length  $L$  must be reasonably large. Furthermore, the entropy  $H(V_1)$  can be easily determined from the *a priori* index distributions  $P(v_k)$ . Note that *Theorem 2* may also be used to calculate the average *a priori* information  $I_A$ .

## V. EXAMPLE

As an example, we consider a parallel concatenated TCM scheme employing two identical memory-4 rate-4/5 recursive systematic convolutional codes, with feedback ( $g_r$ ) and feedforward ( $g_1, \dots, g_4$ ) generator polynomials of  $(g_r, g_1, g_2, g_3, g_4)_8 = (23, 2, 4, 10, 20)_8$ , which have been found experimentally. For the signal set, the 2-D 32-QAM from [11] is used, and both the communication and the *a priori* channel in Fig. 1 represent AWGN channels. Note that the index length of  $M = K = 4$  bits does not permit a histogram-based computation method for the EXIT characteristics of the constituent decoders. In order to determine the extrinsic information  $\mathbf{E}$ , the conditional probabilities of the systematic indices can be derived from the index-based channel pdf via a marginal probability and subsequent subtraction from the APPs in the logarithmic domain at each TCM SISO decoder output

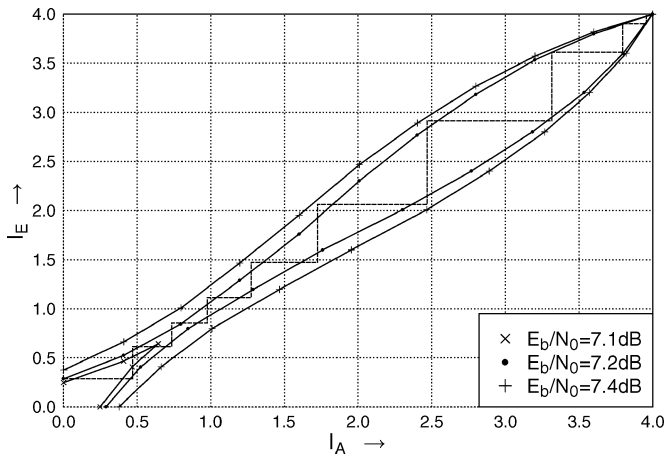


Fig. 2. EXIT chart for TTCM and average decoding trajectory (block length 50 000 symbols, signal set 32-QAM,  $M = 4$ , memory-4 rate-4/5 recursive systematic convolutional code).

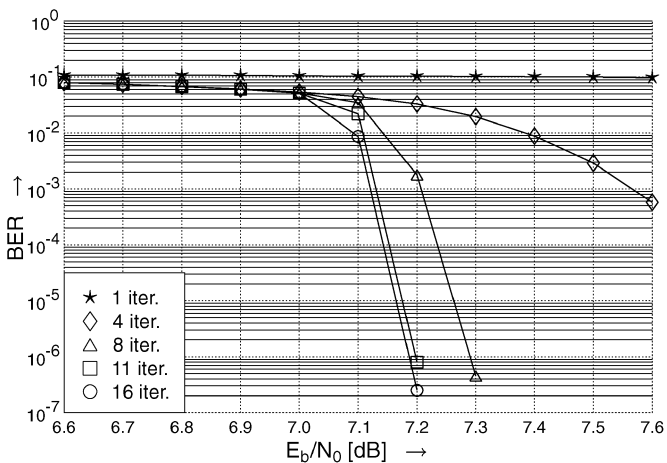


Fig. 3. BER versus  $E_b/N_0$  on the communication channel for the 32-QAM TTCM system (blocklength 50 000 symbols, memory-4 rate-4/5 recursive systematic convolutional code).

[6]. However, since all coded bits in a TTCM symbol are transmitted together over the channel as an inseparable symbol, and hence, are jointly affected by the channel noise, the interdependence of the systematic and parity bits of each TTCM symbol may be increased. On the other hand, computing  $\mathbf{E}$  requires that the RVs for the systematic and parity parts of the channel observation are independent [6]. Thus, only an approximation of the true extrinsic information may be obtained. Fig. 2 shows the resultant EXIT chart computed using *Theorem 2* for a blocklength of  $L = 50\,000$  4-bit indices, where the *a priori* information  $I_A$  is calculated by assuming independent bits for the index  $V_k$ . For  $E_b/N_0 = 7.2$  dB, the average decoding trajectory obtained from real simulations shows convergence. Fig. 3 visualizes the bit-error rate (BER) versus the  $E_b/N_0$  on the communication channel. As we can see, an infinitesimally low BER is obtained for  $E_b/N_0 > 7.2$  dB. This is consistent with the decoding trajectory recorded for  $E_b/N_0 = 7.2$  dB in Fig. 2, where convergence is achieved after 11 iterations. However, we experimentally determined the true convergence

threshold of the 32-QAM TTCM scheme, which is  $E_b/N_0 = 7.1$  dB. Hence, there is a 0.1 dB mismatch associated with the EXIT chart prediction, which may be mainly due to the limited validity of having independent systematic and parity bits.<sup>2</sup>

## VI. CONCLUSIONS

We have derived an efficient technique for computing non-binary EXIT charts from the extrinsic APPs for the information symbols available at the output of the constituent SISO decoders. The proposed method exploits the fact that these APPs represent a sufficient statistic for both the received sequence at the output of the communication channel and the *a priori* information, which has been shown to hold true also in the index-based case. Instead of computing multidimensional histograms as in the conventional approach, we simply need to average over a function of the extrinsic APPs for a long block of source data. The advantage of the proposed approach especially becomes prominent for index lengths larger than two or three bits, where histogram-based EXIT characteristics become exceedingly difficult to compute. The proposed technique can be used for a large variety of index-based serial and parallel concatenated systems with constituent APP decoders.

## ACKNOWLEDGMENT

The authors would like to thank I. Land for fruitful discussions and O. R. Alamri for contributing parts of the simulation software.

## REFERENCES

- [1] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.
- [2] A. Ashikhmin, G. Kramer, and S. ten Brink, "Extrinsic information transfer functions: Model and erasure channel properties," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2657–2673, Nov. 2004.
- [3] S. ten Brink, "Rate one-half code for approaching the Shannon limit by 0.1 dB," *Electron. Lett.*, vol. 36, no. 15, pp. 1293–1294, Jul. 2000.
- [4] M. Tüchler and J. Hagenauer, "EXIT charts and irregular codes," in *Proc. 36th Annu. Conf. Inf. Sci. Syst.*, Princeton, NJ, Mar. 2002, CD-ROM.
- [5] A. Grant, "Convergence of non-binary iterative decoding," in *Proc. IEEE Global Telecommun. Conf.*, San Antonio, TX, Nov. 2001, pp. 1058–1062.
- [6] H. Chen and A. Haimovich, "EXIT charts for turbo trellis-coded modulation," *IEEE Commun. Lett.*, vol. 8, no. 11, pp. 668–670, Nov. 2004.
- [7] P. Robertson and T. Wörz, "Bandwidth-efficient turbo trellis-coded modulation using punctured component codes," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 2, pp. 206–218, Feb. 1998.
- [8] I. Land, P. Hoeher, and S. Gligorević, "Computation of symbol-wise mutual information in transmission systems with log APP decoders and application to EXIT charts," in *Proc. Int. ITG Conf. Source Channel Coding*, Erlangen, Germany, Jan. 2004, pp. 195–202.
- [9] A. Bennatan and D. Burshtein, "Design and analysis of nonbinary LDPC codes for arbitrary discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 549–583, Feb. 2006.
- [10] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 429–445, Mar. 1996.
- [11] G. Ungerböck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 1, pp. 55–67, Jan. 1981.

<sup>2</sup>We observed, even for a blocklength of 5000 information symbols, that the EXIT curves computed from the proposed method perfectly match those computed using the histogram-based method, in the context of the eight-state 8-PSK TTCM scheme of [6, Fig. 3].