On the Achievable Extrinsic Information of Inner Decoders in Serial Concatenation

Jörg Kliewer, Axel Huebner, and Daniel J. Costello, Jr.
Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556, U.S.A.
e-mail: {jkliewer,huebner.7,costello.2}@nd.edu

***Abstract*— In this paper we address the extrinsic information transfer functions of inner decoders for a serially concatenated coding scheme. For the case of an AWGN channel, we give a universal proof for the fact that only inner encoders yielding an infinite output weight for a weight-one input sequence, such as recursive convolutional encoders, lead to perfect extrinsic information at the output of the corresponding SISO decoder. As an example we consider bit-interleaved coded modulation with iterative demapping (BICM-ID) and insert an additional recursive precoder prior to the mapping operation. Simulation results show that the proposed system does not suffer from an error floor and thus significantly outperforms BICM-ID systems that solely use mappings as inner encodings, even when they are optimized.**

I. Introduction

It has been shown in [1], for a serially concatenated coding scheme, that recursive inner encoders are preferable since they always lead to an interleaver gain, in contrast to nonrecursive or block encoders. A similar observation can be made by considering extrinsic information transfer functions [2], [3]. If the sequence of information bits can be described as an independent and identically distributed (iid) random process, only recursive inner encoders can achieve an average mutual information of one between the information bits and the extrinsic soft-values if perfect *a priori* information is applied. This fact has been observed by many researchers and, in the case of two constituent encoders, is a necessary condition for the absence of an error floor.

As an example we consider bit-interleaved coded modulation with iterative demapping (BICM-ID) [4]. BICM [5–7] represents a serial concatenation of an outer encoder, an interleaver, and a mapper or modulator which maps groups of bits to complex waveforms prior to transmission over the channel. Since the inner encoder is a nonrecursive one, perfect extrinsic information can never be achieved, leading to an error floor even for moderately distorted channels. The achievable average extrinsic information with perfect *a priori* information for the demodulator depends on the mapping, where it has been shown, e.g. in [8], that Gray mapping is unsuitable in order to obtain a significant gain from iterations. However, improved mappings have been given in [7], [9], [10]. In [11], a general mapping optimization algorithm is proposed on the basis of maximizing the achievable extrinsic information at the SISO demapper output, or, alternatively, minimizing the pairwise error probability. However, generally, a noticeable error floor is observed with all mappings, even when they are optimized.

As a main result of this paper we present a formal proof that only inner encoders producing an infinite output weight for a weight-one input sequence, such as recursive inner encoders, lead to perfect extrinsic information at the output of the corresponding SISO decoder. For the sake of simplicity we restrict ourselves to AWGN channels. In contrast to [12], where only rate-1 convolutional codes are considered via an analysis of their state-space representation, our proof uses information theoretic considerations and universally holds for arbitrary code rates and inner encoders (not just convolutional ones). As an example, our findings are then applied to BICM-ID where, in order to eliminate the error floor associated with soft demapping, we propose to add an accumulator as a recursive precoder prior to the mapper. At the decoder, the demapper is then replaced by a SISO decoder working on the joint accumulator / mapper trellis. Simulation results for Gray-mapped 16-QAM show the performance gain of the proposed precoded BICM-ID scheme compared to those approaches that solely employ optimized mappings without precoding. A related approach has been discussed in [13], where decoding is carried out by iterating between a turbo decoder and the demapper. While this approach theoretically also leads to perfect overall mutual information after the iterative decoder has converged in practice this depends on the error floor of the outer turbo code. Another drawback compared to our approach is the larger decoding complexity.

II. System model for inner encoding

Fig. 1 depicts the underlying system model for inner encoding in a serial concatenation. A binary sequence $\boldsymbol{V} = [V_1, \ldots, V_\ell, \ldots, V_N]$ of length N, with random variables (RVs) V_ℓ and realizations $v_\ell \in \{0, 1\}$, is applied to a rate-N/Q inner encoder $\mathcal{C}_i$. The resulting codeword $\boldsymbol{U} = [U_1, \ldots, U_\ell, \ldots, U_Q]$ is then transmitted over the BPSK-modulated AWGN communication channel. At the channel output the sequence $\boldsymbol{Y} = [Y_1, \ldots, Y_\ell, \ldots, Y_Q]$ consisting of soft-bit realizations $y_\ell \in \mathbb{R}$ with channel probability density function (pdf)

$$p(y_\ell|v_\ell) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left(-(y_\ell - (1 - 2\,v_\ell))^2/(2\sigma_n^2)\right), \quad (1)$$

is observed. Herein, $\sigma_n^2 = N_0/(2E_s)$ represents the normalized noise variance on the AWGN channel, E_s the

This work was partly supported by the German Research Foundation (DFG) under grant KL 1080/3-1, NSF grants CCR02-05310 and EEC-0203366, and NASA grant NNG05GH73G.

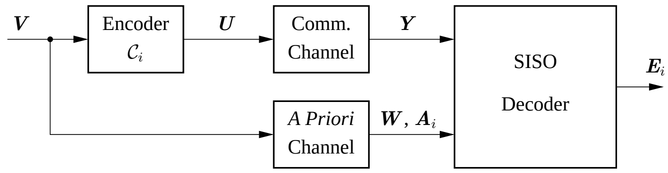

Fig. 1. System model for decoding the inner code in a serial concatenation [14].

transmitted energy per code bit, and N_0 the single-sided noise power spectral density. The *a priori* AWGN channel models the *a priori* information used at the inner constituent decoder in an iterative decoding scheme. At the output we observe the soft-bit sequence $\boldsymbol{W} = [W_1, \ldots, W_\ell, \ldots, W_N]$, where $w_\ell \in \mathbb{R}$ represents a soft-bit realization. The sequence

$$\boldsymbol{A}_i = [A_{i,1}, \ldots, A_{i,\ell}, \ldots, A_{i,N}], \quad \text{with} \quad A_{i,\ell} := L_a^{(i)}(V_\ell), \tag{2}$$

contains the corresponding *a priori* log-likelihood ratios (LLRs)

$$L_a^{(i)}(V_\ell) = \ln\left(\frac{p(w_\ell|V_\ell = 0)}{p(w_\ell|V_\ell = 1)}\right) \text{ for } \quad \ell \in \{1, 2, \ldots, N\}, \tag{3}$$

where the *a priori* channel pdf $p(w_\ell|v_\ell)$ is analogous to (1). The SISO decoder employs the outputs of both the communication and the *a priori* channels for computing extrinsic *a posteriori* probabilities (APPs) $P(V_\ell = v_\ell|\boldsymbol{y}, \boldsymbol{a}_{i,\backslash\ell})$, where $\boldsymbol{a}_{i,\backslash\ell} = [a_{i,1}, \ldots, a_{i,\ell-1}, a_{i,\ell+1}, \ldots, a_{i,N}]$ describes realizations of *a priori* LLRs with the ℓ-th element $a_{i,\ell}$ excluded. The sequence $\boldsymbol{E}_i$ contains the extrinsic LLRs and is defined as

$$\boldsymbol{E}_i = [E_{i,1}, \ldots, E_{i,\ell}, \ldots, E_{i,N}], \quad \text{with} \quad E_{i,\ell} := L_e^{(i)}(V_\ell), \tag{4}$$

and

$$L_e^{(i)}(V_\ell) = \ln\left(\frac{P(V_\ell = 0|\boldsymbol{y}, \boldsymbol{a}_{i,\backslash\ell})}{P(V_\ell = 1|\boldsymbol{y}, \boldsymbol{a}_{i,\backslash\ell})}\right) \text{ for } \quad \ell \in \{1, 2, \ldots, N\}. \tag{5}$$

Denoting the mutual information between two RVs X and Y as $I(X;Y)$, we define the quantities

$$I_{A_i} := \frac{1}{N}\sum_{\ell=1}^{N} I(V_\ell; A_{i,\ell}), \quad 0 \le I_{A_i} \le 1, \tag{6}$$

$$I_{E_i} := \frac{1}{N}\sum_{\ell=1}^{N} I(V_\ell; E_{i,\ell}), \quad 0 \le I_{E_i} \le 1. \tag{7}$$

Herein, I_{A_i} is defined as the average *a priori* information and denotes the average capacity of the *a priori* channel, whereas I_{E_i} refers to the average extrinsic information. By defining a continuous mapping between *a priori* and extrinsic information as $I_{E_i} = T_i(I_{A_i}, E_s/N_0)$, we obtain the inner extrinsic information transfer (EXIT) function or characteristic $T_i(\cdot)$ for the SISO decoder. Here, $T_i(\cdot)$ depends on both the capacity of the *a priori* and the communication channel, where the latter is characterized by the channel parameter E_s/N_0. As in (6), (7), both I_{A_o} and I_{E_o} can also be determined for the outer decoder. The mapping between I_{A_o} and I_{E_o} is then given as $I_{E_o} = T_o(I_{A_o})$, where $T_o(\cdot)$ represents the EXIT function of the outer channel decoder. An EXIT chart can now be obtained by plotting the transfer functions $T_{i/o}$ for both constituent decoders on a single diagram, where the axes must be swapped for one of the constituent decoders. For further details about EXIT charts we refer to the corresponding literature (e.g., [2], [3], [8], [14]).

III. Conditions for perfect inner extrinsic information

In the following we give a formal proof for the fact that an inner encoder producing an infinite output weight for a weight-one input sequence (e.g., a recursive convolutional encoder) is sufficient for achieving $I_{E_i}(I_{A_i} = 1\,\text{bit}) = 1\,\text{bit}$.

The average extrinsic information I_{E_i} at the output of a general inner SISO decoder in a serially concatenated scheme can be expressed as

$$I_{E_i} = \frac{1}{N}\sum_{\ell=1}^{N} I(V_\ell; E_{i,\ell}) = 1 - \frac{1}{N}\sum_{\ell=1}^{N} H(V_\ell|E_{i,\ell}), \tag{8}$$

where $H(V_\ell|E_{i,\ell})$ represents the conditional entropy of V_ℓ given $E_{i,\ell}$. In (8) we assume, without loss of generality, that $H(V_\ell) = 1\,\text{bit}$ for all $\ell \in \{1, 2, \ldots, N\}$.

In [14], [15] it is shown that, for a SISO decoder emitting *a posteriori* probabilities (APPs), the mutual information $I(V_\ell; E_{i,\ell})$ can also be written as

$$\begin{aligned} I(V_\ell; E_{i,\ell}) &= I(V_\ell; \boldsymbol{A}_{i,\backslash\ell}, \boldsymbol{Y}) \\ &= H(V_\ell) - H(V_\ell|\boldsymbol{A}_{i,\backslash\ell}, \boldsymbol{Y}). \end{aligned} \tag{9}$$

We define the following information bit sequence realizations $\boldsymbol{v}_{\backslash\ell} := [v_1, \ldots, v_{\ell-1}, v_{\ell+1}, \ldots, v_N]$ and $\boldsymbol{v}_{\ell\to q} := [v_1, \ldots, v_{\ell-1}, q, v_{\ell+1}, \ldots, v_N]$ with $q \in \{0, 1\}$. The conditional entropy $H(V_\ell|\boldsymbol{A}_{i,\backslash\ell}, \boldsymbol{Y})$ in (9) may now be further expanded as

$$\begin{aligned} H(V_\ell|\boldsymbol{A}_{i,\backslash\ell}, \boldsymbol{Y})\Big|_{I_A=1} &= H(V_\ell|\,\boldsymbol{V}_{\backslash\ell}, \boldsymbol{Y}) \\ &= \int_{\boldsymbol{y}} \sum_{\boldsymbol{v}_{\backslash\ell}} p(\boldsymbol{y}, \boldsymbol{v}_{\backslash\ell}) \sum_{q=0}^{1} P(V_\ell = q|\boldsymbol{v}_{\backslash\ell}, \boldsymbol{y}) \cdot \\ &\qquad \log_2(P(V_\ell = q|\boldsymbol{v}_{\backslash\ell}, \boldsymbol{y}))\, d\boldsymbol{y}. \end{aligned} \tag{10}$$

Lemma 1. *Let the all-zero information sequence* $\boldsymbol{v}^{(0)}$ *of length* N *be applied to a recursive convolutional encoder* $\mathcal{C}$ *of rate* $R = N/Q$*. Let the resulting BPSK-modulated codeword* $\boldsymbol{c}_0^{(0)} = 1 - 2\,\mathcal{C}\{\boldsymbol{v}^{(0)}\}$ *be transmitted over an AWGN channel with noise variance* σ_n^2*, leading to the channel observation* $\boldsymbol{y}$ *with error sequence* $\boldsymbol{e}_0 = \boldsymbol{y} - \boldsymbol{c}_0^{(0)}$*. Furthermore, let* $\boldsymbol{e}_1 = \boldsymbol{y} - \boldsymbol{c}_1^{(0)} = \boldsymbol{y} - (1 - 2\,\mathcal{C}\{\boldsymbol{v}_{\ell'\to 1}^{(0)}\})$ *for the same observation* $\boldsymbol{y}$ *denote the resulting error sequence when the weight-one information sequence* $\boldsymbol{v}_{\ell'\to 1}^{(0)}$ *is transmitted. All information bits except the one at bit position* ℓ'*,* $\ell' \in \{1, 2, \ldots, N\}$*, are assumed to be perfectly known at the decoder. Then, for*

$Q \to \infty$, $N \to \infty$, and fixed R

$$P(v_\ell^{(0)} = 0 | \boldsymbol{v}_{\backslash \ell'}^{(0)}, \boldsymbol{y}) = \begin{cases} 1 & \text{for } |\boldsymbol{e}_0| < |\boldsymbol{e}_1|, \\ \frac{1}{2} & \text{for } |\boldsymbol{e}_0| = |\boldsymbol{e}_1|, \\ 0 & \text{for } |\boldsymbol{e}_0| > |\boldsymbol{e}_1|, \end{cases}$$

$$P(v_\ell^{(0)} = 1 | \boldsymbol{v}_{\backslash \ell'}^{(0)}, \boldsymbol{y}) = \begin{cases} 0 & \text{for } |\boldsymbol{e}_0| < |\boldsymbol{e}_1|, \\ \frac{1}{2} & \text{for } |\boldsymbol{e}_0| = |\boldsymbol{e}_1|, \\ 1 & \text{for } |\boldsymbol{e}_0| > |\boldsymbol{e}_1|. \end{cases}$$

where $|\boldsymbol{e}| = \sqrt{\sum_{\ell=1}^{Q} e_\ell^2}$.

Proof. Clearly, a weight-one information sequence $\boldsymbol{v}_{\ell \to 1}^{(0)}$ results in an infinite weight at the output of a recursive encoder if the blocklength tends to infinity. When the encoded all-zero information sequence $\mathcal{C}\{\boldsymbol{v}^{(0)}\}$ is transmitted over the BPSK-modulated AWGN channel, $\boldsymbol{y} = \boldsymbol{c}_0^{(0)} + \boldsymbol{n}$ is observed at the channel output. Herein, $\boldsymbol{n} = [n_1, \ldots, n_\ell, \ldots, n_Q]$, where n_ℓ, $\ell \in \{1, 2, \ldots, Q\}$, represents a zero-mean Gaussian noise sample realization with variance σ_n^2. Additionally, achieving the same observed sequence $\boldsymbol{y}$ with the BPSK-modulated codeword $\boldsymbol{c}_1^{(0)}$ leads to the error sequences

$$\begin{aligned} \boldsymbol{e}_0 &= \boldsymbol{n}, \\ \boldsymbol{e}_1 &= \boldsymbol{y} - \boldsymbol{c}_1^{(0)} = \boldsymbol{c}_0^{(0)} - \boldsymbol{c}_1^{(0)} + \boldsymbol{n} = \boldsymbol{d} + \boldsymbol{n} \end{aligned} \tag{11}$$

with $\boldsymbol{d} = [d_1, \ldots, d_\ell, \ldots, d_Q] := \boldsymbol{c}_0^{(0)} - \boldsymbol{c}_1^{(0)}$.

The conditional probabilities $P(V_\ell^{(0)} = q | \boldsymbol{v}_{\backslash \ell}^{(0)}, \boldsymbol{y})$, $q \in \{0, 1\}$, can be described by using Bayes' theorem and assuming equiprobable information sequences as follows:

$$P(V_\ell^{(0)} = q | \boldsymbol{v}_{\backslash \ell}^{(0)}, \boldsymbol{y}) = \frac{p(\boldsymbol{y} | \boldsymbol{v}_{\ell \to q}^{(0)})}{p(\boldsymbol{y} | \boldsymbol{v}^{(0)}) + p(\boldsymbol{y} | \boldsymbol{v}_{\ell \to 1}^{(0)})}. \tag{12}$$

For an AWGN channel the corresponding pdfs in (12) are given by

$$p(\boldsymbol{y} | \boldsymbol{v}^{(0)}) = \frac{1}{\sqrt{2\pi}\sigma_n} \prod_{\ell=1}^{Q} \exp\left(-\frac{n_\ell^2}{2\sigma_n^2}\right) \tag{13}$$

and

$$p(\boldsymbol{y} | \boldsymbol{v}_{\ell \to 1}^{(0)}) = \frac{1}{\sqrt{2\pi}\sigma_n} \prod_{\ell=1}^{Q} \exp\left(-\frac{(d_\ell + n_\ell)^2}{2\sigma_n^2}\right). \tag{14}$$

Combining (12), (13), and (14) yields the expression

$$P(V_\ell^{(0)} = 0 | \boldsymbol{v}_{\backslash \ell}^{(0)}, \boldsymbol{y}) = \frac{1}{1 + \exp\left(\dfrac{|\boldsymbol{e}_0|^2 - |\boldsymbol{e}_1|^2}{2\,\sigma_n^2}\right)}. \tag{15}$$

We now consider the following three cases:

(i) $|\boldsymbol{e}_0| < |\boldsymbol{e}_1|$: The exponential term in (15) can be rewritten as

$$\exp\left(\frac{|\boldsymbol{e}_0|^2 - |\boldsymbol{e}_1|^2}{2\,\sigma_n^2}\right) = \underbrace{\exp\left(-\frac{1}{\sigma_n^2}\sum_{\ell=1}^{Q} d_\ell\, n_\ell\right)}_{=:\alpha(Q)} \underbrace{\exp\left(-\frac{1}{2\,\sigma_n^2}\sum_{\ell=1}^{Q} d_\ell^2\right)}_{=:\beta(Q)}. \tag{16}$$

Herein, $\alpha(Q)$ does not necessarily tend to zero. However, due to the fact that $d_\ell \in \{0, 2\}$ and $\lim_{Q\to\infty} w(\boldsymbol{d}/2) \to \infty$, where $w(\cdot)$ denotes the Hamming weight, it is clear that $\lim_{Q\to\infty} \beta(Q) = 0$ in (16). But since $|\boldsymbol{e}_0| < |\boldsymbol{e}_1|$, the overall product $\alpha(Q)\,\beta(Q)$ in (16) tends to zero as well. Thus, from (15) we obtain for $N \to \infty$, leading to $Q \to \infty$ with $Q = N/R$ for fixed R,

$$\lim_{Q\to\infty} P(v_\ell^{(0)} = 0 | \boldsymbol{v}_{\backslash \ell'}^{(0)}, \boldsymbol{y}) = 1,$$
$$\lim_{Q\to\infty} P(v_\ell^{(0)} = 1 | \boldsymbol{v}_{\backslash \ell'}^{(0)}, \boldsymbol{y}) = 0.$$

(ii) $|\boldsymbol{e}_0| > |\boldsymbol{e}_1|$: We exchange $\boldsymbol{v}^{(0)}$ and $\boldsymbol{v}_{\ell' \to 1}^{(0)}$ in the above derivation, where now

$$\boldsymbol{e}_1 = \boldsymbol{y} - \boldsymbol{c}_1^{(0)} = \boldsymbol{n}, \quad \boldsymbol{e}_0 = -\boldsymbol{d} + \boldsymbol{n}.$$

Analogously to the previous case, it can be shown from (15) that

$$\lim_{Q\to\infty} P(v_\ell^{(0)} = 0 | \boldsymbol{v}_{\backslash \ell'}^{(0)}, \boldsymbol{y}) = 0,$$
$$\lim_{Q\to\infty} P(v_\ell^{(0)} = 1 | \boldsymbol{v}_{\backslash \ell'}^{(0)}, \boldsymbol{y}) = 1.$$

(iii) $|\boldsymbol{e}_0| = |\boldsymbol{e}_1|$: Here, we obtain

$$\lim_{Q\to\infty} P(v_\ell^{(0)} = 0 | \boldsymbol{v}_{\backslash \ell'}^{(0)}, \boldsymbol{y}) = \frac{1}{2},$$
$$\lim_{Q\to\infty} P(v_\ell^{(0)} = 1 | \boldsymbol{v}_{\backslash \ell'}^{(0)}, \boldsymbol{y}) = \frac{1}{2}.$$

Note that for a given $\boldsymbol{v}^{(0)}$ and $\boldsymbol{v}_{\ell' \to 1}^{(0)}$, this case occurs only for specific observations $\boldsymbol{y}$ lying on a $(Q-1)$-dimensional hyperplane characterized by $|\boldsymbol{e}_0| = |\boldsymbol{e}_1|$.

□

We can now state the main result of this section in the following theorem.

Theorem 1. *Only an inner recursive convolutional encoder $\mathcal{C}_i$ in a serially concatenated scheme achieves $I_E(I_A = 1\,bit) = 1\,bit$ for an information blocklength tending to infinity and transmission over a BPSK-modulated AWGN communication channel.*

Proof. Let $\boldsymbol{c}_0 := 1 - 2\,\mathcal{C}_i\{\boldsymbol{v}_{\ell \to 0}\}$ and $\boldsymbol{c}_1 := 1 - 2\,\mathcal{C}_i\{\boldsymbol{v}_{\ell \to 1}\}$, respectively. Assume that we transmit the BPSK codeword $\boldsymbol{c}_0$ over an AWGN channel, where the sequence $\boldsymbol{y}$ is observed at the channel output. Due to the linearity of the encoder we may then express the error sequences $\boldsymbol{e}_{0/1}$ required for receiving $\boldsymbol{y}$ when the codeword $\boldsymbol{c}_{0/1}$ is transmitted as

$$\boldsymbol{e}_0 = \boldsymbol{n}, \quad \boldsymbol{e}_1 = \boldsymbol{c}_0 - \boldsymbol{c}_1 + \boldsymbol{n} = \boldsymbol{c}_0^{(0)} - \boldsymbol{c}_1^{(0)} + \boldsymbol{n},$$

which are the same as in (11). Using the result from Lemma 1, we conclude that either

$$\begin{aligned} &\lim_{Q\to\infty} P(v_\ell^{(0)} = 0 | \boldsymbol{v}_{\setminus \ell'}^{(0)}, \boldsymbol{y}) = 1, \\ &\lim_{Q\to\infty} P(v_\ell^{(0)} = 1 | \boldsymbol{v}_{\setminus \ell'}^{(0)}, \boldsymbol{y}) = 0, \quad \text{or} \\ &\lim_{Q\to\infty} P(v_\ell^{(0)} = 0 | \boldsymbol{v}_{\setminus \ell'}^{(0)}, \boldsymbol{y}) = 0, \\ &\lim_{Q\to\infty} P(v_\ell^{(0)} = 1 | \boldsymbol{v}_{\setminus \ell'}^{(0)}, \boldsymbol{y}) = 1. \end{aligned} \tag{17}$$

for all information sequence realizations $\boldsymbol{v} \in \mathcal{V}$ and any communication channel observation $\boldsymbol{y}$. Inserting (17) into (10) and considering that $\lim_{x\to 0} x \log x = 0$ proves the theorem. (Note that the case $|\boldsymbol{e}_0| = |\boldsymbol{e}_1|$ in Lemma 1 need not be considered in the integration of (10) since the corresponding $\boldsymbol{y}$ are located on a singular hyperplane with respect to the function to be integrated.) □

It is clear from the proof of Theorem 1 that the condition $I_{E_i}(I_{A_i} = 1\,\text{bit}) = 1\,\text{bit}$ cannot be satisfied for inner coding schemes, where an information sequence weight of one generates a finite codeword weight. This holds for all block codes and non-recursive convolutional encoders, as well as for modulators or ISI channels. Hence, Theorem 1 provides an alternative explanation of the well-known fact that non-recursive convolutional encoders should not be used as inner encoders in a serially concatenated coding scheme [1].

IV. Example: BICM with recursive precoding

As an example we consider a BICM-ID scheme with recursive precoding, where the corresponding encoder is shown in Fig. 2. The binary input sequence $\boldsymbol{B}$ is applied to a rate-R_o outer encoder $\mathcal{C}_o$ generating the sequence $\boldsymbol{C}$. After bit-interleaving using the permutation function π, we obtain $\boldsymbol{V} = [V_1, \ldots, V_\ell, \ldots, V_N] = \pi(\boldsymbol{C})$. In order to be able to achieve an average extrinsic information of $I_{E_i} = 1\,\text{bit}$, we insert an accumulator as a recursive precoder with generator polynomial $g(D) = 1/(1+D)$ before the mapping operation. The resulting sequence is denoted $\boldsymbol{V}'$. Alternatively, we may decompose $\boldsymbol{V}'$ into M-bit symbols, resulting in the symbol vector $\boldsymbol{I}' = [I'_1, \ldots, I'_k, \ldots, I'_L]$, with $N = ML$, where k denotes the symbol-time instant. Each I'_k has a realization $i'_k \in \mathcal{I}$ in the finite alphabet $\mathcal{I} = \{0, 1, \ldots, 2^M - 1\}$. The i'_k are then mapped to signal points u_k in the complex plane using the mapping function $\mathcal{M} : \mathcal{I} \to \mathbb{C}$, where u_k is a realization of U_k and $\boldsymbol{U} = [U_1, \ldots, U_k, \ldots, U_L]$. Then $\boldsymbol{U}$ is transmitted over an AWGN channel with the parameter E_s/N_0. The received sequence $\boldsymbol{Y}$ at the output of the communication channel is then applied to the iterative decoder, which contains both an outer SISO decoder for encoder $\mathcal{C}_o$ and an inner symbol-by-symbol SISO decoder working on the joint trellis for the accumulator / 2^M-ary mapper as constituent decoders. The trellis for the inner decoder has 2^M states, and

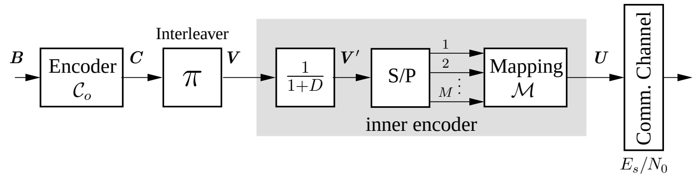

Fig. 2. BICM encoder structure with a recursive precoder.

each state s_{k-1} at time instant $k-1$ has transitions to all states s_k at time instant k, where 2^{M-1} transitions per state always correspond to both the same input symbol $i_k \in \mathcal{I}$ and output symbol $i'_k \in \mathcal{I}$. An example is given for QPSK with $M = 2$ in Fig. 3, where the state transitions are labeled i_k/i'_k to indicate the input and output symbols. Decoding is carried out via a symbol-based BCJR algorithm [16] which computes the APPs $P(I_k = i_k|\boldsymbol{y}, P_a(I_k = i_k|\boldsymbol{y}))$ given the *a priori* conditional probabilities $P_a(I_k = i_k|\boldsymbol{y})$.

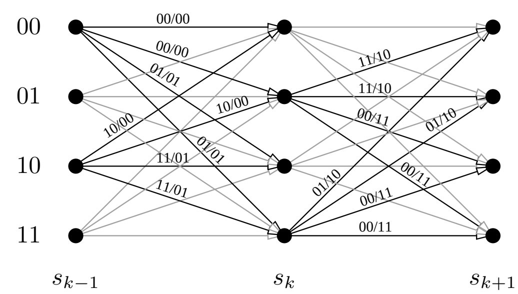

Fig. 3. Trellis representation for the joint accumulator / 2^M-ary mapper for QPSK with $M = 2$. The labels on the state transitions denote "(input symbol) / (output symbol)".

Numerical results are obtained for an inner accumulator combined with Gray-mapped 16-QAM, where the outer encoder is a memory-1 recursive systematic convolutional (RSC) encoder with feedforward (g_1) and feedback (g_r) polynomials $(g_r, g_1) = (3, 2)_8$ in octal form and a code rate of $R_o = 1/2$. This system has an effective throughput of $\eta = R_o M = 2\,\text{bit/s/Hz}$. Fig. 4 shows the corresponding EXIT chart and some "snapshot" decoding trajectories for $E_s/N_0 = 7.5\,\text{dB}$. We observe that the EXIT characteristic for the combined inner accumulator / mapper achieves $I_{E_i} = 1\,\text{bit}$ for perfect *a priori* information, as guaranteed by Theorem 1.

The bit error rate (BER) versus E_b/N_0 curve, with $E_b = E_s/\eta$, is shown in Fig. 5. Besides the proposed system, two BICM-ID systems employing optimized mapping without recursive precoding are also shown for comparison. The M16a mapping is proposed in [11], where it was obtained by an optimization approach minimizing the difference $1 - I_{E_i}(I_{A_i} = 1\,\text{bit})$. In [7] the MSP mapping was suggested as a good trade-off between large I_{E_i} for both zero and perfect *a priori* information. For both the proposed system and the M16a and MSP mappings, the $(g_r, g_1) = (3, 2)_8$ outer RSC encoder is used. The BER was averaged over 1000 simulated transmissions for a blocklength of 20,000 information bits, and 20 iterations were allowed for the iterative decoder. It

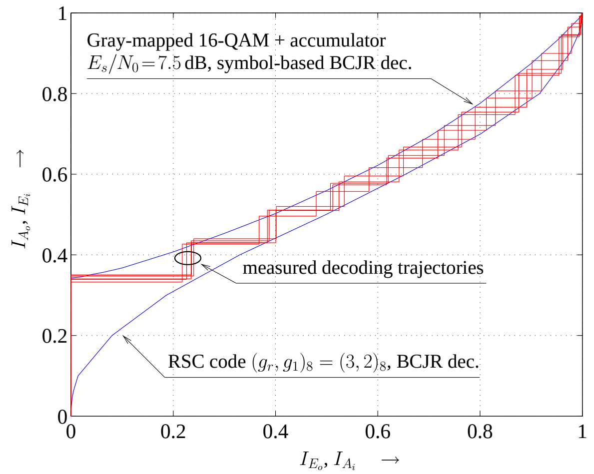

Fig. 4. EXIT chart and "snapshot" decoding trajectories for the accumulator / 16-QAM Gray mapper with a memory-1 $(g_r, g_1) = (3,2)_8$ RSC outer encoder.

can be seen from Fig. 5 that, in contrast to both the MSP and M16a mappings, no error floor is observed for the joint accumulator / Gray mapper, where all simulated transmissions were observed as error-free for $E_b/N_0 > 4$ dB. This confirms the fact that $I_{E_i}(I_{A_i} = 1\,\text{bit}) = 1\,\text{bit}$ is a necessary condition for the absence of an error floor in this case. In particular, at a BER of 10^{-6}, the joint accumulator / Gray mapper outperforms the M16a mapping by over 1 dB. This shows the advantage of the proposed scheme compared to using optimized mappings without recursive precoding.

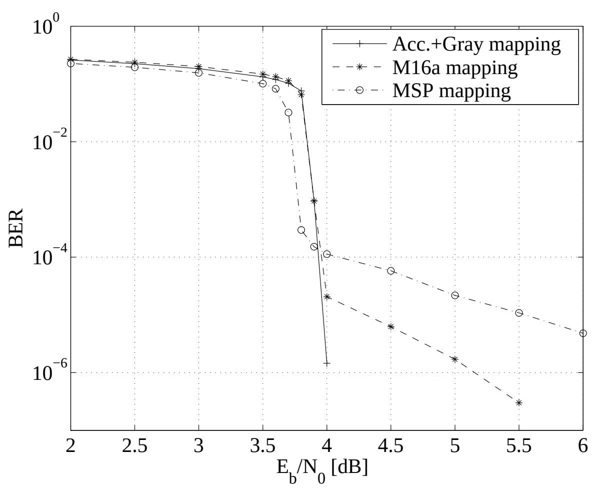

Fig. 5. BER versus E_b/N_0 for the proposed inner joint accumulator / 16-QAM Gray mapper and the M16a [11] and MSP [7] mappings for the 16-QAM constellation (outer $(g_r, g_1) = (3,2)_8$ RSC encoder, 1000 simulated transmissions, blocklength 20,000 information bits).

V. Conclusions

We have considered inner encoding in a serially concatenated coding scheme. For the corresponding decoder we have given a universal proof for the fact that an encoder that produces an infinite output weight for a weight-one input sequence (e.g., a recursive convolutional encoder) leads to perfect extrinsic information when perfect *a priori* information is available. As an example, we considered a BICM-ID scheme in which the SISO demapper does not achieve perfect extrinsic information even when optimized symbol mappings for the constellation points of the modulation are used. This leads to a noticeable error floor in the BER performance curve. We proposed adding an accumulator as a recursive precoder before the mapper along with Gray mapping, and at the receiver the demapper in the iterative decoder is replaced by a symbol-by-symbol SISO decoder on the joint accumulator / mapper trellis. Simulation results for a large number of channel realizations have not shown any error floor for the proposed scheme, in contrast to using optimized mappings without precoding. Finally, we note that the proposed scheme may further be optimized by choosing mappings other than Gray mapping.

References

[1] 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, no. 3, pp. 909–926, May 1998.

[2] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Comm.*, vol. 49, no. 10, pp. 1727–1737, October 2001.

[3] J. Huber and S. Huettinger, "Information processing and combining in channel coding," in *Proc. 3rd International Symposium on Turbo Codes & Related Topics*, Brest, France, Sept. 2003, pp. 95–102.

[4] S. ten Brink, S. Speidel, and R.-H. Yan, "Iterative demapping and decoding for multilevel modulation," in *Proc. IEEE Global Telecommun. Conf.*, Sydney, Australia, Nov. 1998, vol. 1, pp. 579–584.

[5] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," *IEEE Trans. Comm.*, vol. 40, no. 5, pp. 873–884, May 1992.

[6] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

[7] A. Chindapol and J. A. Ritcey, "Design, analysis, and performance evaluation for BICM-ID with square QAM constellations in Rayleigh fading channels," *IEEE J. Sel. Areas in Commun.*, vol. 19, no. 5, pp. 944–957, May 2001.

[8] S. ten Brink, "Convergence of iterative decoding," *Electron. Lett.*, vol. 35, no. 10, pp. 806–808, May 1999.

[9] X. Li, A. Chindapol, and J.A. Ritcey, "Bit-interleaved coded modulation with iterative decoding and 8PSK signaling," *IEEE Trans. Comm.*, vol. 50, no. 8, pp. 1250–1257, Aug. 2002.

[10] J. Tan and G. L. Stüber, "Analysis and design of symbol mappers for iteratively decoded BICM," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 662–672, Mar. 2005.

[11] F. Schreckenbach, N. Goertz, J. Hagenauer, and G. Bauch, "Optimization of symbol mappings for bit-interleaved coded modulation with iterative decoding," *IEEE Commun. Lett.*, vol. 7, no. 12, pp. 593–595, Dec. 2003.

[12] M. Tuechler, "Design of serially concatenated codes depending on the block length," *IEEE Trans. Comm.*, vol. 52, no. 2, pp. 209–218, Feb. 2004.

[13] F. Schreckenbach and G. Bauch, "Irregular signal constellations, mappings and precoder," in *Proc. International Symposium on Information Theory and its Applications*, Parma, Italy, Oct. 2004, pp. 1332–1336.

[14] 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.

[15] I. Land, P. Hoeher, and S. Gligorević, "Computation of symbol-wise mutual information in transmission systems with logAPP decoders and application to EXIT charts," in *Proc. International ITG Conference on Source and Channel Coding (SCC)*, Erlangen, Germany, Jan. 2004, pp. 195–202.

[16] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory*, pp. 294–287, Mar. 1974.