Bounds on mutual information for simple codes using information combining

Land, Ingmar; Huettinger, S.; Hoeher, P.A.; Huber, J.

Document Version
Publisher's PDF, also known as Version of record

Link to publication from Aalborg University

Bounds on Mutual Information for Simple Codes Using Information Combining

Ingmar Land, Simon Huettinger, Peter A. Hoeher, and Johannes Huber

Abstract— For coded transmission over a memoryless channel, two kinds of mutual information are considered: the mutual information between a code symbol and its noisy observation and the overall mutual information between encoder input and decoder output. The overall mutual information is interpreted as a combination of the mutual informations associated with the individual code symbols. Thus, exploiting code constraints in the decoding procedure is interpreted as combining mutual informations. For single parity check codes and repetition codes, we present bounds on the overall mutual information, which are based only on the mutual informations associated with the individual code symbols. Using these mutual information bounds, we compute bounds on extrinsic information transfer (EXIT) functions and bounds on information processing characteristics (IPC) for these codes.

Index Terms— Mutual information, information combining, extrinsic information, concatenated codes, iterative decoding, extrinsic information transfer (EXIT) chart, information processing characteristic (IPC).

I. INTRODUCTION

In digital communications, the information on channel input symbols (code symbols) gained by observing the corresponding channel output symbols is expressed by mutual information [1]. This concept was originally introduced for communication channels, but it may also be applied to the super-channel formed by encoder, communication channel, and decoder. Then, mutual information is the information gained about the encoder inputs given the decoder outputs.

Assuming memoryless channels, information on each code symbol (the channel input) is provided by its noisy observation (the channel output), independently of the observations of the other code symbols. When redundant channel coding is applied, the decoder combines the channel outputs so as to take the code constraints into account. From an information theory point of view, the decoder can be interpreted as processing mutual information [2]. According to this interpretation, the mapping from mutual information per channel-use to mutual information between encoder input and decoder output characterizes the behavior and the capability of a decoder, or

Ingmar Land and Peter A. Hoeher are with the Information and Coding Theory Lab, Faculty of Engineering, University of Kiel, Kaiserstraße 2, D-24143 Kiel, Germany. (Email: il@tf.uni-kiel.de, ph@tf.uni-kiel.de; WWW: www.tf.uni-kiel.de/ict.)

Simon Huettinger was with the Institute for Information Transmission, University Erlangen-Nürnberg, Cauerstraße 7/LIT D-91058 Erlangen, Germany. He is now with Siemens AG, Germany. (Email: simon.huettinger@siemens.com.)

Johannes Huber is with the Institute for Information Transmission, University Erlangen-Nürnberg, Cauerstraße 7/LIT, D-91058 Erlangen, Germany. (Email: huber@lnt.de; WWW: www.lnt.de/lit.)

indeed, the behavior of a whole coding scheme comprising both encoder and decoder.

In this paper, we restrict ourselves to two recently introduced concepts, which are both based on this information theoretic interpretation: the information processing characteristic of coding schemes and the extrinsic information transfer chart for iterative decoding.

The information processing characteristic (IPC) plots the end-to-end mutual information of a coding scheme versus the mutual information of the communication channel [3], [4]. Thus, the focus is on the performance of the overall coding scheme. Using mutual information, a common scale can be used for all communication channels, and soft-values computed by the decoder can be taken into account.

The extrinsic information transfer (EXIT) chart method analyzes the behavior of constituent decoders involved in iterative decoding [5], [6]. Each constituent decoder accepts *a priori* values and computes extrinsic values. (Commonly, probabilities or log-likelihood ratios are used.) The mutual information between the encoder input or output (depending on the role of the constituent decoder within the iterative decoding structure) and the *a priori* values is called *a priori* information; the mutual information between the encoder input or output (same as above) and the extrinsic values is called extrinsic information. The input-output behavior of a constituent decoder is characterized by the EXIT function plotting the extrinsic information versus the *a priori* information.

Though IPCs and EXIT functions plot input mutual information versus output mutual information, they rely on the input distributions, or equivalently, on the corresponding channel models. One may ask whether it is possible to compute such functions which are valid for a whole class of channel models. In [7], this problem has been addressed for the first time.

Consider the system made of two binary-input symmetric memoryless channels (BISMCs) having the *same* uniformly distributed input; let I_1 and I_2 be their mutual informations. The overall mutual information I of the channel between the inputs of the channels and the outputs of both channels is then a combination of I_1 and I_2. In [7], tight bounds on I are presented which are based only on I_1 and I_2 and which are valid for all BISMCs. This concept of computing I using I_1 and I_2 is called *information combining*. (Even though the mutual information of a BISMC for uniformly distributed inputs is equal to the channel capacity, we prefer the term "mutual information" to emphasize that our focus is on "information" processing.)

The present paper generalizes from two channels with the same input, as discussed in [7], [8], to multiple channels with

inputs that are subject to code constraints. We consider two kinds of constraints: (a) The inputs are required to fulfill a parity check equation; (b) the inputs are required to be equal. The first case corresponds to a single parity check code, and the second case corresponds to a repetition code. We restrict ourselves to memoryless channels, so all code symbols may also be transmitted over the same channel. Bounds on the IPCs and on the EXIT functions are presented for both codes.

The paper is organized as follows: In Section II, the decoding model and the notation are introduced. In Section III, definitions and properties are given for binary-input symmetric memoryless channels. In Section IV, bounds on the combined information are derived for single parity check codes and for repetition codes. These results are applied in Section V to give bounds on EXIT functions and on IPCs for these codes. Finally, conclusions are drawn in Section VI.

II. DECODING MODEL AND NOTATION

Throughout this paper, random variables are denoted by uppercase letters, and realizations are denoted by lowercase letters. Vectors are written in boldface, and for subvectors of a vector $\boldsymbol{a} = [a_1, \ldots, a_J]$, we adopt the short-hand notation

$$\begin{aligned} \boldsymbol{a}_{[i,j]} &:= [a_i, a_{i+1}, \ldots, a_{j-1}, a_j], \\ \boldsymbol{a}_{\backslash i} &:= [a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_J]. \end{aligned}$$

Binary symbols, also called bits, are defined over $\mathbb{B} := \{-1, +1\}$ or, equivalently, over $\mathbb{F}_2 := \{0, 1\}$. For a binary symbol $a \in \mathbb{B}$, we write its equivalent representation over $\mathbb{F}_2$ as $\bar{a}$, and we use the convention

$$\begin{aligned} \bar{a} = 0 &\qquad \Leftrightarrow \qquad a = +1, \\ \bar{a} = 1 &\qquad \Leftrightarrow \qquad a = -1. \end{aligned}$$

The representation over $\mathbb{B}$ is more convenient when the focus is on the symmetry of a channel, whereas the representation over $\mathbb{F}_2$ is more common when dealing with the constraints of a linear code.

Consider a *systematic binary linear code* of length N with equiprobable code words $\boldsymbol{x} = [x_0, x_1, \ldots, x_{N-1}] \in \mathbb{B}^N$. The code bits are transmitted over independent *binary-input symmetric memoryless channels* (BISMCs), denoted by $X_i \to Y_i$, offering the mutual information $I(X_i; Y_i)$, $i = 0, 1, \ldots, N-1$. Channels of this kind are discussed in more detail in the following section.

The generality of this model is illustrated in the following two points:

- When the code is used as a constituent code of a concatenated coding scheme including an iterative decoder, the channel $X_i \to Y_i$ may be the communication channel. It may also be the virtual channel between code bit X_i and the soft estimate, provided by another constituent decoder and interpreted as *a priori* value; this channel is called *a priori* channel.

- The assumption of systematic codes is not a restriction, because the channels associated with the systematic bits may be interpreted as *a priori* channels. Rather, this represents a straightforward method to include *a priori* information on systematic bits in the decoding model.

This will be exploited in Section V, where EXIT functions are discussed.

The mutual information between two random variables X and Y is defined as the expected value

$$I(X; Y) := \mathrm{E}\left\{ \frac{p_{X|Y}(x|y)}{p_X(x)} \right\},$$

see [9]. The functions $p_X(x)$ and $p_{X|Y}(x|y)$ denote the probability distribution of X and the conditional probability distribution of X given Y, respectively; to be precise, they denote probability mass functions for discrete X and probability density functions for continuous X. Similarly, the conditional mutual information between X and Y given a third random variable J is defined as the expected value

$$I(X; Y|J) := \mathrm{E}\left\{ \frac{p_{X|Y,J}(x|y,j)}{p_{X|J}(x|j)} \right\},$$

where $p_{X|J}(x|j)$ and $p_{X|Y,J}(x|y,j)$ again denote the respective probability mass functions or probability density functions.

Three kinds of mutual information may be associated with each code bit X_i [10]: the intrinsic information, the extrinsic information and the complete information. The *intrinsic information* on code bit X_i is defined as

$$I_{\mathsf{int},i} := I(X_i; Y_i) \tag{1}$$

and provides the mutual information between X_i and its noisy observation. The intrinsic information is equal to the mutual information of the channel over which the code bit is transmitted. The *extrinsic information* on code bit X_i is defined as

$$I_{\mathsf{ext},i} := I(X_i; \boldsymbol{Y}_{\backslash i}) \tag{2}$$

and provides the mutual information between X_i and the noisy observations of all other code bits. This kind of mutual information follows the definition of extrinsic probabilities or extrinsic log-likelihood ratios used in iterative decoding, e.g., [11], [12]. Finally, the *complete information* on code bit X_i is defined as

$$I_{\mathsf{comp},i} := I(X_i; \boldsymbol{Y}) \tag{3}$$

and provides the mutual information between X_i and the observations of all code bits. The complete information may be formed by combining the intrinsic information and the extrinsic information [10].

III. PROPERTIES OF BINARY-INPUT SYMMETRIC MEMORYLESS CHANNELS

All channels in this paper are assumed to be binary-input symmetric memoryless channels (BISMCs). Examples of BISMCs are the binary symmetric channel (BSC), the binary erasure channel (BEC), and the binary-input additive white Gaussian noise (AWGN) channel. Some properties of BISMCs are discussed in this section.

Let $X \to Y$ denote a BISMC with input $X \in \mathbb{X} := \mathbb{B}$ and output $Y \in \mathbb{Y} \subseteq \mathbb{R}$, where $\mathbb{X}$ and $\mathbb{Y}$ denote the input and the output alphabet of the channel, respectively. The input is assumed to be uniformly distributed if not stated

otherwise. The transition probabilities of the channel are given by $p_{Y|X}(y|x)$, denoting the probability density function for continuous output alphabets and denoting the probability mass function for discrete output alphabets. In this paper, we call a binary-input channel symmetric[1] if

$$p_{Y|X}(y|x) = p_{Y|X}(-y|-x) \qquad (4)$$

for all $x \in \mathbb{X}$ and for all $y \in \mathbb{Y}$. The mutual information of the channel is denoted by

$$I := I(X; Y).$$

As mentioned above, two examples of BISMCs are the BSC and the BEC; they are depicted in Fig. 1 and Fig. 2. The BSC is defined by the crossover probability ϵ, and the BEC is defined by the erasure probability δ. The output alphabets are chosen such that the symmetry condition (4) is fulfilled; the output value $Y = 0$ corresponds to an erasure. Let

$$h(\xi) := -\xi \log_2 \xi - (1 - \xi) \log_2(1 - \xi)$$

denote the binary entropy function, where $\xi \in [0, 1]$; let further $h^{-1}(\eta)$ denote the inverse of $h(\xi)$ for $\xi \in [0, 1/2]$, where $\eta \in [0, 1]$. Then, the mutual information of the BSC is given by

$$I = 1 - h(\epsilon), \qquad (5)$$

and the mutual information of the BEC is given by
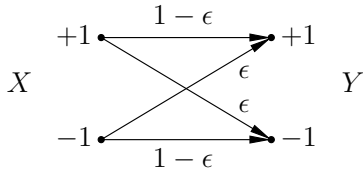
$$I = 1 - \delta. \qquad (6)$$

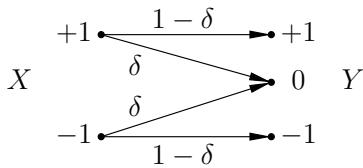Fig. 1.   Binary symmetric channel (BSC) with crossover probability ϵ.

Fig. 2.   Binary erasure channel (BEC) with erasure probability δ.

In the following, we show that every BISMC can be separated into strongly symmetric subchannels that are BSCs, due to symmetry and binary input. For this separation, we use the absolute value of the channel output.

[1] This definition of symmetry implies a small, but not significant loss of generality. For example, an AWGN channel with nonzero-mean noise is not symmetric according to our definition, although it can be regarded as symmetric, of course. The derivations based on the given definition of symmetry can easily be extended to a more general definition, which may follow the well-known one for channels with discrete output alphabets, see [9].

Let the random variable J be defined as the absolute value of Y,

$$J := |Y|,$$

where $J \in \mathbb{J} := \{y \in \mathbb{Y} : y \geq 0\}$. The elements of the output alphabet $\mathbb{Y}$ are now partitioned into subsets

$$\mathbb{Y}(j) := \begin{cases} \{+j, -j\} & \text{for } j \in \mathbb{J}\backslash\{0\}, \\ \{0\} & \text{for } j = 0. \end{cases}$$

With these definitions, J indicates which output set $\mathbb{Y}(j)$ the output symbol Y belongs to.

Due to the introduced partitioning, the random variable J separates the symmetric channel $X \to Y$ into strongly symmetric subchannels denoted by $X \to Y|J = j$. Therefore, J is called the subchannel indicator.

The subchannels for $J > 0$ are BSCs. The subchannel for $J = 0$ is a BEC with erasure probability 1. For convenience, we interpret also this channel as a BSC, namely as a BSC with crossover probability $1/2$. This interpretation does not restrict generality, because the properties of this channel are not changed. However, the following derivations become easier, since now *all* subchannels are BSCs.

The probability distribution of the subchannels is given by

$$q(j) := p_J(j),$$

$j \in \mathbb{J}$, where $p_J(j)$ denotes the probability density function for continuous output alphabets and the probability mass function for discrete output alphabets. The conditional crossover probabilities $\epsilon(j)$ of the subchannels are defined as

$$\epsilon(j) := \begin{cases} p_{Y|X,J}(-j|+1, j) & \text{for } j \in \mathbb{J}\backslash\{0\}, \\ \frac{1}{2} & \text{for } j = 0. \end{cases}$$

As all subchannels are BSCs, the mutual information of the subchannel specified by $J = j$ is given by

$$I(j) := I(X; Y|J = j) = 1 - h(\epsilon(j)). \qquad (7)$$

Using the above definitions, the mutual information of the BISMC can be written as the expected value of the mutual information of its subchannels:

$$\begin{aligned} I &= I(X;Y) = I(X;Y,J) \\ &= I(X;J) + I(X;Y|J) \\ &= \mathop{\mathrm{E}}_{j \in \mathbb{J}}\{I(X;Y|J = j)\} \\ &= \mathop{\mathrm{E}}_{j \in \mathbb{J}}\{I(j)\}. \end{aligned} \qquad (8)$$

In the first line, we used the fact that J is a function of Y. In the second line, the chain rule for mutual information [9] has been applied. As X and J are statistically independent, we have $I(X; J) = 0$.

This concept of separating of BISMCs into BSCs can easily be generalized to include channels with vector-valued outputs. Let $X \to \boldsymbol{Y}$ denote a BISMC with input $X \in \mathbb{X} := \mathbb{B}$ and output $\boldsymbol{Y} = [Y_1, Y_2, \ldots, Y_n]$, with $Y_i \in \mathbb{Y}_i \subseteq \mathbb{R}$, $i = 1, 2, \ldots, n$. The set $\mathbb{X}$ denotes the input alphabet and $\mathbb{Y} := \mathbb{Y}_1 \times \mathbb{Y}_2 \times \cdots \times \mathbb{Y}_n$ denotes the output alphabet of the channel. The input is assumed to be uniformly distributed if not stated otherwise. The transition probabilities of the

channel are given by $p_{\boldsymbol{Y}|X}(\boldsymbol{y}|x)$, denoting the probability density function for continuous output alphabets and denoting the probability mass function for discrete output alphabets. The channel is assumed to be symmetric, i.e.,

$$p_{\boldsymbol{Y}|X}(\boldsymbol{y}|x) = p_{\boldsymbol{Y}|X}(-\boldsymbol{y}|-x)$$

for all $x \in \mathbb{X}$ and for all $\boldsymbol{y} \in \mathbb{Y}$ (cf. Footnote 1).

The definition of an appropriate subchannel indicator for the separation into BSCs is slightly more complicated than before. Using $\mathbb{J}_1 := \{y \in \mathbb{Y}_1 : y \geq 0\}$ and

$$\mathbb{J} := \mathbb{J}_1 \times \mathbb{Y}_2 \times \mathbb{Y}_3 \times \cdots \times \mathbb{Y}_n,$$

we define the subchannel indicator $\boldsymbol{J} \in \mathbb{J}$ as

$$\boldsymbol{J} := \boldsymbol{j} \quad \text{for } \boldsymbol{y} \in \{\boldsymbol{j}, -\boldsymbol{j}\}.$$

The resulting subchannels $X \to \boldsymbol{Y}|\boldsymbol{J} = \boldsymbol{j}$ are BSCs for $\boldsymbol{J} \neq \boldsymbol{0}$. The subchannel for $\boldsymbol{J} = \boldsymbol{0}$ is a BEC with erasure probability 1, and it is again interpreted as a BSC with crossover probability $1/2$.

The probability distribution $q(\boldsymbol{j})$, the crossover probability $\epsilon(\boldsymbol{j})$, and the mutual information $I(\boldsymbol{j})$ for each subchannel follow the above definitions for a BISMC with scalar output. As before, the mutual information of the BISMC, $I := I(X; \boldsymbol{Y})$, can be written as the expected value of the mutual information of its subchannels:

$$I = \underset{\boldsymbol{j} \in \mathbb{J}}{\mathrm{E}}\{I(\boldsymbol{j})\}. \tag{9}$$

The described separation of a BISMC into binary symmetric subchannels is utilized in the following sections.

IV. Bounds on Mutual Information

Consider the decoding model introduced in Section II. Since code bits are subject to the code constraints, the extrinsic information on a particular code bit is a combination of the intrinsic informations on the other code bits. Similarly, the complete information on a code bit is a combination of the intrinsic information and the extrinsic information on this code bit. (Notice that intrinsic, extrinsic, and complete information are mutual informations.)

If certain models are assumed for the channels over which the code bits are transmitted, the extrinsic information and the complete information can be computed precisely. This is done in the EXIT chart method and in the IPC method. Often, the binary-input AWGN channel, the binary erasure channel, or the binary symmetric channel are used as channel models. On the other hand, if only the intrinsic information on each code bit is known, and not the underlying channel model, only bounds on the extrinsic and on the complete information can be given.

This section deals with bounds on the extrinsic information for single parity check codes and for repetition codes and with bounds on the complete information. The bounds correspond to the cases where the individual channels are BSCs or BECs.

The main results are stated in Theorem 1, Theorem 2, and Theorem 3. These theorems generalize the work presented in [7] and [13]. Nevertheless, the proofs are based on the same concepts, where the separation of BISMCs into BSCs

plays a central role. Motivated by the initial work in [7], [8], the authors of [14] produced similar results, but used different methods of proof. They also determined decoding thresholds for low-density parity-check codes.

A. Extrinsic Information for the Single Parity Check Code

Consider a single parity check (SPC) code of length N, which is defined by the constraint

$$\bar{X}_0 \oplus \bar{X}_1 \oplus \cdots \oplus \bar{X}_{N-1} = 0 \tag{10}$$

on the code bits $\bar{X}_i \in \mathbb{F}_2$. (Note that the code bits may be equivalently represented by $X_i \in \mathbb{B}$, cf. Section II.) The code constraint and the transmission channels $X_i \to Y_i$ are shown in Fig. 3 for $N = 4$. In the following, we consider only the extrinsic information on code bit X_0. Due to the symmetric structure of the code, the expressions for the other code bits are similar. In the following, we discuss first the case where the channels for the code bits are all BECs, and then the case where the channels are all BSCs.
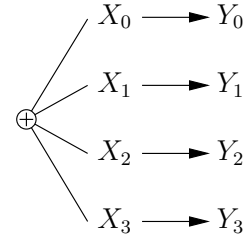

Fig. 3.   Single parity check code of length $N = 4$.

If the channels are all BECs, the value of code bit X_0 can be recovered with certainty if no erasure has occurred, i.e., if $Y_i \neq 0$ for $i = 1, 2, \ldots, N - 1$. (Note that an erasure corresponds to $Y_i = 0$.) This happens with probability $(1 - \delta_1)(1 - \delta_2) \cdots (1 - \delta_{N-1})$. If we have one or more erasures, no extrinsic information on code bit X_0 is available. Using (6) and the above probability, it can easily be seen that

$$I_{\text{ext},0}^{\text{BEC}} = I_{\text{int},1} I_{\text{int},2} \cdots I_{\text{int},N-1}. \tag{11}$$

For the case where the channels are all BSCs, it is useful to introduce the following function.

Definition 1
Let $\xi_1, \xi_2, \ldots, \xi_n \in [0, 1]$, $n \geq 1$. We define the binary information function for serial concatenation for $n = 1$ as

$$f_1^{\text{ser}}(\xi_1) := \xi_1,$$

for $n = 2$ as

$$f_2^{\text{ser}}(\xi_1, \xi_2) := 1 - h\big((1 - \epsilon_1)\epsilon_2 + \epsilon_1(1 - \epsilon_2)\big),$$

and for $n > 2$ as

$$f_n^{\text{ser}}(\xi_1, \xi_2, \ldots, \xi_n) := f_2^{\text{ser}}\big(\xi_1, f_{n-1}^{\text{ser}}(\xi_2, \xi_3, \ldots, \xi_n)\big),$$

where $\epsilon_i := h^{-1}(1 - \xi_i)$ for $i = 1, 2, \ldots, n$.

We have included the case $n = 1$, so that the following formulas can be written in a more compact form.

The function $f_n^{\text{ser}}(.)$ describes the mutual information of the channel formed by a serial concatenation of n independent BSCs. The input of the first channel is assumed to be uniformly distributed. With $I_1, I_2, \ldots, I_n$ denoting the mutual information of each BSC, respectively, the mutual information between the input of the first BSC and the output of the last BSC is given by $f_n^{\text{ser}}(I_1, I_2, \ldots, I_n)$. This is explained in more detail in Appendix I. The function defined above is now used to express the extrinsic information.

Using the chain rule of mutual information [9], the extrinsic information on code bit X_0 can be written as

$$
\begin{aligned}
I(X_0; \boldsymbol{Y}_{[1,N-1]}) &= \\
&= I(X_0; \boldsymbol{Y}_{[1,N-2]}) + I(X_0; Y_{N-1} | \boldsymbol{Y}_{[1,N-2]}). \quad (12)
\end{aligned}
$$

The first term is equal to zero, i.e., $I(X_0; \boldsymbol{Y}_{[1,N-2]}) = 0$, as X_0 and $\boldsymbol{Y}_{[1,N-2]}$ are independent if neither X_{N-1} nor Y_{N-1} are known.

To determine the second term, we use the representation of the code bits and the channel outputs over $\mathbb{F}_2$, written as $\bar{X}_i$ and $\bar{Y}_i$ (cf. Section II). (Note that this does not change mutual information, of course.) Let binary random variables $Z_i \in \mathbb{F}_2$, $i = 0, 1, \ldots, N-1$, be defined as

$$
\begin{aligned}
Z_0 &:= \bar{X}_0, \\
Z_1 &:= Z_0 \oplus \bar{X}_1, \\
Z_2 &:= Z_1 \oplus \bar{X}_2, \\
&\cdots \\
Z_{N-2} &:= Z_{N-3} \oplus \bar{X}_{N-2}, \\
Z_{N-2} &= \bar{X}_{N-1}, \\
Z_{N-1} &:= \bar{Y}_{N-1}.
\end{aligned}
$$

Notice that the penultimate line is not a definition, but an equality which results from the previous definitions and the parity check equation (10). It follows from the definitions that all Z_i are uniformly distributed and that

$$
I(\bar{X}_0; \bar{Y}_{N-1} | \bar{\boldsymbol{Y}}_{[1,N-2]}) = I(Z_0; Z_{N-1} | \bar{\boldsymbol{Y}}_{[1,N-2]}). \quad (13)
$$

For the time being, assume $\bar{\boldsymbol{Y}}_{[1,N-2]} = \bar{\boldsymbol{y}}_{[1,N-2]}$, where $\bar{\boldsymbol{y}}_{[1,N-2]} \in \mathbb{F}_2^{N-2}$ denotes an arbitrary but fixed realization of $\bar{\boldsymbol{Y}}_{[1,N-2]}$. Then, the random variables Z_i form a Markov chain,

$$
Z_0 \to Z_1 \to Z_2 \to \cdots \to Z_{N-3} \to Z_{N-2} \to Z_{N-1},
$$

where each pair $Z_i \to Z_{i+1}$, $i = 0, 1, \ldots, N-2$ can be interpreted as a BSC. The mutual information of each BSC is as follows:

- $Z_0 \to Z_1$: The code bit $\bar{X}_1$ represents the error bit of this BSC. The crossover probability of the channel $\bar{X}_1 \to \bar{Y}_1$ and that of the channel $\bar{Y}_1 \to \bar{X}_1$ are equal due to the uniform distribution of $\bar{X}_1$. Thus, we have for the crossover probability ϵ_1 of the channel $Z_0 \to Z_1$:

$$
\begin{aligned}
\epsilon_1 \in \left\{ p_{\bar{X}_1 | \bar{Y}_1}(1 | \bar{y}') : \bar{y}' \in \mathbb{F}_2 \right\} &= \\
&= \left\{ h^{-1}(1 - I_{\text{int},1}), 1 - h^{-1}(1 - I_{\text{int},1}) \right\}.
\end{aligned}
$$

The mutual information of this channel is then given by $I(Z_0; Z_1) = 1 - h(\epsilon_1) = I_{\text{int},1}$, which is independent of $\bar{y}_1$.

- $Z_i \to Z_{i+1}$, $i = 1, 2, \ldots, N-3$: Similar to $Z_0 \to Z_1$, the mutual information is given by $I(Z_i; Z_{i+1}) = I_{\text{int},i}$.

- $Z_{N-2} \to Z_{N-1}$: This channel is identical to the channel $\bar{X}_{N-1} \to \bar{Y}_{N-1}$, and thus its mutual information is given by $I(Z_{N-2}; Z_{N-1}) = I_{\text{int},N-1}$.

Note that the mutual information of each BSC is independent of $\bar{\boldsymbol{y}}_{[1,N-2]}$.

We have a serial concatenation of BSCs and know the mutual information of each one, and thus we can apply the binary information function for serial concatenation according to Definition 1:

$$
\begin{aligned}
I(Z_0; Z_{N-1} | \bar{Y}_{[1,N-2]} = \bar{y}_{[1,N-2]}) &= \\
&= f_{N-1}^{\text{ser}}(I_{\text{int},1}, I_{\text{int},2}, \ldots, I_{\text{int},N-1}).
\end{aligned}
$$

Since this mutual information is independent of $\bar{\boldsymbol{y}}_{[1,N-2]}$, we have

$$
\begin{aligned}
I(Z_0; Z_{N-1} | \bar{Y}_{[1,N-2]}) &= \\
&= \mathop{\mathrm{E}}_{\bar{\boldsymbol{y}}_{[1,N-2]} \in \mathbb{F}_2^{N-2}} \left\{ I(Z_0; Z_{N-1} | \bar{Y}_{[1,N-2]} = \bar{y}_{[1,N-2]}) \right\} \\
&= \mathop{\mathrm{E}}_{\bar{\boldsymbol{y}}_{[1,N-2]} \in \mathbb{F}_2^{N-2}} \left\{ f_{N-1}^{\text{ser}}(I_{\text{int},1}, I_{\text{int},2}, \ldots, I_{\text{int},N-1}) \right\} \\
&= f_{N-1}^{\text{ser}}(I_{\text{int},1}, I_{\text{int},2}, \ldots, I_{\text{int},N-1}). \quad (14)
\end{aligned}
$$

Using (14) in (13), we obtain

$$
I(\bar{X}_0; \bar{Y}_{N-1} | \bar{\boldsymbol{Y}}_{[1,N-2]}) = f_{N-1}^{\text{ser}}(I_{\text{int},1}, I_{\text{int},2}, \ldots, I_{\text{int},N-1}).
$$

After substituting this result into (12), we obtain the extrinsic information on code bit X_0 for the case where all channels are BSCs:

$$
I_{\text{ext},0}^{\text{BSC}} = f_{N-1}^{\text{ser}}(I_{\text{int},1}, I_{\text{int},2}, \ldots, I_{\text{int},N-1}). \quad (15)
$$

The two cases discussed (where either the channels are all BECs or all BSCs) represent bounds on the extrinsic information for a code bit, when only the intrinsic informations for code bits are known, and not the underlying channel models. This is proved by the following theorem.

Theorem 1 (Extrinsic Information for SPC Code)
Let $X_0, X_1, \ldots, X_{N-1} \in \mathbb{B}$ denote the code bits of a single parity check code of length N. Let $X_i \to Y_i$, $i = 1, 2, \ldots, N-1$, denote $N-1$ independent BISMCs having mutual information $I(X_i; Y_i)$. Let the intrinsic information on code bit X_i be defined by $I_{\text{int},i} := I(X_i; Y_i)$, $i = 1, 2, \ldots, N-1$, and let the extrinsic information on code bit X_0 by defined by $I_{\text{ext},0} := I(X_0; \boldsymbol{Y}_{\setminus 0})$. Then, the following tight bounds hold:

$$
\begin{aligned}
I_{\text{ext},0} &\geq I_{\text{int},1} I_{\text{int},2} \cdots I_{\text{int},N-1}, \\
I_{\text{ext},0} &\leq f_{N-1}^{\text{ser}}(I_{\text{int},1}, I_{\text{int},2}, \ldots, I_{\text{int},N-1}).
\end{aligned}
$$

The lower bound is achieved if the channels are all BECs, and the upper bound is achieved if the channels are all BSCs.

To prove this theorem, we use the following lemma.

Lemma 1

The binary information function for serial concatenation (Definition 1) has the following two properties:

(a) $f_n^{\text{ser}}(\xi_1, \xi_2, \ldots, \xi_n)$ *is convex-$\cap$ in each ξ_i, $i = 1, 2, \ldots, n$;*

(b) $f_n^{\text{ser}}(\xi_1, \xi_2, \ldots, \xi_n)$ *is lower-bounded by the product of its arguments:*

$$f_n^{\text{ser}}(\xi_1, \xi_2, \ldots, \xi_n) \geq \xi_1 \xi_2 \cdots \xi_n.$$

The proof follows from Lemma 2 in [7] or Lemma 2 in [15] by induction.

Theorem 1 can now be proved as follows.

Proof: The extrinsic information does not change if it is written conditioned on the subchannel indicators $\boldsymbol{J}_{\backslash 0}$,

$$
\begin{aligned}
I_{\text{ext},0} &= I(X_0; \boldsymbol{Y}_{\backslash 0}) = I(X_0; \boldsymbol{Y}_{\backslash 0}, \boldsymbol{J}_{\backslash 0}) \\
&= I(X_0; \boldsymbol{J}_{\backslash 0}) + I(X_0; \boldsymbol{Y}_{\backslash 0} | \boldsymbol{J}_{\backslash 0}) \\
&= I(X_0; \boldsymbol{Y}_{\backslash 0} | \boldsymbol{J}_{\backslash 0}),
\end{aligned}
$$

since X_0 is independent of $\boldsymbol{J}_{\backslash 0}$ (cf. (8) and corresponding comments). Thus, we can write

$$
\begin{aligned}
I_{\text{ext},0} &= I(X_0; \boldsymbol{Y}_{[1,N-1]} | \boldsymbol{J}_{[1,N-1]}) = \\
&= \underset{\boldsymbol{j}_{[1,N-1]}}{\mathrm{E}} \left\{ I(X_0; \boldsymbol{Y}_{[1,N-1]} | \boldsymbol{J}_{[1,N-1]} = \boldsymbol{j}_{[1,N-1]}) \right\} = \\
&= \underset{\boldsymbol{j}_{[1,N-1]}}{\mathrm{E}} \left\{ f_{N-1}^{\text{ser}} \big(I_{\text{int},1}(j_1), \ldots, I_{\text{int},N-1}(j_{N-1}) \big) \right\}.
\end{aligned}
$$
(16)

The argument in the second line corresponds to the case where the channels are all BSCs, due to the conditions. Therefore, this expression can be replaced by the function $f_{N-1}^{\text{ser}}(\ldots)$ according to Definition 1

Next, the two properties of the function $f_{N-1}^{\text{ser}}(\ldots)$ given in Lemma 1 are exploited. Using Lemma 1(b) in (16), we obtain

$$
\begin{aligned}
\underset{\boldsymbol{j}_{[1,N-1]}}{\mathrm{E}} &\left\{ f_{N-1}^{\text{ser}} \big(I_{\text{int},1}(j_1), \ldots, I_{\text{int},N-1}(j_{N-1}) \big) \right\} \geq \\
&\geq \underset{\boldsymbol{j}_{[1,N-1]}}{\mathrm{E}} \left\{ I_{\text{int},1}(j_1) \cdots I_{\text{int},N-1}(j_{N-1}) \right\} = \\
&\qquad\qquad = I_{\text{int},1} I_{\text{int},2} \cdots I_{\text{int},N-1},
\end{aligned}
$$

where in the last line, (8) was used. On the other hand, due to Lemma 1(a), Jensen's inequality (e.g., [9]) can be applied in (16), and we obtain

$$
\begin{aligned}
\underset{\boldsymbol{j}_{[1,N-1]}}{\mathrm{E}} &\left\{ f_{N-1}^{\text{ser}} \big(I_{\text{int},1}(j_1), \ldots, I_{\text{int},N-1}(j_{N-1}) \big) \right\} \leq \\
&\leq f_{N-1}^{\text{ser}} \Big(\underset{j_1}{\mathrm{E}} \{ I_{\text{int},1}(j_1) \}, \ldots \\
&\qquad \ldots, \underset{j_{N-1}}{\mathrm{E}} \{ I_{\text{int},N-1}(j_{N-1}) \} \Big) = \\
&= f_{N-1}^{\text{ser}} (I_{\text{int},1}, I_{\text{int},2}, \ldots, I_{\text{int},N-1}).
\end{aligned}
$$

Thus, we have the two bounds.

According to (11) and (15), the lower bound and the upper bound are actually achieved when the channels are all BECs or all BSCs, respectively. ∎

B. Extrinsic Information for the Repetition Code

Consider a repetition code of length N, which is defined by the constraint

$$X_0 = X_1 = \cdots = X_{N-1} \tag{17}$$

on the code bits X_i. This code constraint and the transmission channels $X_i \rightarrow Y_i$ are shown in Fig. 4 for $N = 4$. In the following, we consider only the extrinsic information on code bit X_0. Due to the symmetric structure of the code, the expressions for the other code bits are similar. In the following, we discuss first the case where the channels are all BECs, and then the case where the channels are all BSCs.
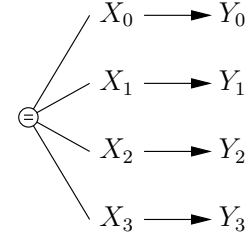

Fig. 4. Repetition code of length $N = 4$.

If the channels are all BECs, the value of code bit X_0 can be recovered with certainty provided that not all channel outputs are erasures. (Note that an erasure corresponds to $Y_i = 0$.) Otherwise, no extrinsic information on code bit X_0 is available. Using (6) and the probabilities of these events, it can easily be seen that

$$I_{\text{ext},0}^{\text{BEC}} = 1 - (1 - I_{\text{int},1})(1 - I_{\text{int},2}) \cdots (1 - I_{\text{int},N-1}). \tag{18}$$

For the case where the channels are all BSCs, we introduce the following function.

Definition 2

Let $\xi_1, \xi_2, \ldots, \xi_n \in [0,1]$, $n \geq 1$. Let further $\boldsymbol{r} = [r_1, r_2, \ldots, r_n]$, $r_i \in \mathbb{B}$. We define the binary information function for parallel concatenation as

$$f_n^{\text{par}}(\xi_1, \xi_2, \ldots, \xi_n) := -\sum_{\boldsymbol{r} \in \mathbb{B}^n} \psi(\boldsymbol{r}) \log_2 \psi(\boldsymbol{r}) - \sum_{i=1}^{n} (1 - \xi_i)$$

with

$$\psi(\boldsymbol{r}) := \frac{1}{2} \Big(\prod_{i=1}^{n} \varphi_i(r_i) + \prod_{i=1}^{n} \big(1 - \varphi_i(r_i) \big) \Big),$$

and

$$\varphi_i(r_i) := \begin{cases} \epsilon_i & \text{for } r_i = +1, \\ 1 - \epsilon_i & \text{for } r_i = -1, \end{cases}$$

where $\epsilon_i := h^{-1}(1 - \xi_i)$ for $i = 1, 2, \ldots, n$.

Similar to Definition 1, we have included the case $n = 1$, so that the following formulas can be written in a more compact form.

This function describes the mutual information of a channel formed by a parallel concatenation of n independent BSCs, i.e., BSCs having the same input. This input is assumed to be

uniformly distributed. With $I_1, I_2, \ldots, I_n$ denoting the mutual information of each BSC, respectively, the mutual information between the (common) input and the vector of all channel outputs is given by $f_n^{\mathsf{par}}(I_1, I_2, \ldots, I_n)$. Appendix II provides further details.

As we are interested in the extrinsic information on code bit X_0, the given scenario can be interpreted as the transmission of code bit X_0 over $N-1$ BSCs (corresponding to code bits $X_1, X_2, \ldots, X_{N-1}$). Thus, the extrinsic information on code bit X_0 can be written as

$$I_{\mathsf{ext},0}^{\mathsf{BSC}} = f_{N-1}^{\mathsf{par}}(I_{\mathsf{int},1}, I_{\mathsf{int},2}, \ldots, I_{\mathsf{int},N-1}). \tag{19}$$

The two cases considered above represent the bounds on the extrinsic information on the code bit when only the intrinsic information on code bits is known, and not the underlying channel models. This is proved by the following theorem.

Theorem 2 (Extrinsic Information for Rep. Code)
Let $X_0, X_1, \ldots, X_{N-1} \in \mathbb{B}$ denote the code bits of a repetition code of length N. Let $X_i \to Y_i$, $i = 1, 2, \ldots, N-1$, denote $N-1$ independent BISMCs having mutual information $I(X_i; Y_i)$. Let the intrinsic information on code bit X_i be defined by $I_{\mathsf{int},i} := I(X_i; Y_i)$, $i = 1, 2, \ldots, N-1$, and let the extrinsic information on code bit X_0 by defined by $I_{\mathsf{ext},0} := I(X_0; \boldsymbol{Y}_{\backslash 0})$. Then, the following tight bounds hold:

$$\begin{aligned} I_{\mathsf{ext},0} &\geq f_{N-1}^{\mathsf{par}}(I_{\mathsf{int},1}, I_{\mathsf{int},2}, \ldots, I_{\mathsf{int},N-1}), \\ I_{\mathsf{ext},0} &\leq 1 - (1 - I_{\mathsf{int},1})(1 - I_{\mathsf{int},2}) \cdots (1 - I_{\mathsf{int},N-1}). \end{aligned}$$

The lower bound is achieved if the channels are all BSCs, and the upper bound is achieved if the channels are all BECs.

Note that BSCs achieve the lower bound for the repetition code, but the upper bound for the single parity check code. For BECs, the reverse holds.

To prove this theorem, we use the following lemma.

Lemma 2
The binary information function for parallel concatenation of two channels (Definition 2) has the following two properties:

(a) $f_2^{\mathsf{par}}(\xi_1, \xi_2)$ *is convex-$\cup$ in ξ_1 and ξ_2;*
(b) $f_2^{\mathsf{par}}(\xi_1, \xi_2)$ *is upper-bounded as*

$$f_2^{\mathsf{par}}(\xi_1, \xi_2) \leq \xi_1 + \xi_2 - \xi_1 \xi_2.$$

The proof follows immediately by observing that

$$f_2^{\mathsf{par}}(\xi_1, \xi_2) = \xi_1 + \xi_2 - f^{\mathsf{ser}}(\xi_1, \xi_2)$$

and Lemma 1.

These two properties are now exploited in the proof of Theorem 2.

Proof: Let a random variable $Z \in \mathbb{B}$ be defined as $Z := X_0$. We use Z to write the constraint for the repetition code given in (17) as

$$\begin{aligned} X_0 &= X_1 = Z, \\ Z &= X_2 = X_3 = \cdots = X_{N-1}. \end{aligned}$$

The first line is now interpreted as the constraint of a repetition code of length 3, where the code bits are X_0, X_1, and Z. The

channel $Z \to \boldsymbol{Y}_{[2,N-1]}$ can be considered as a channel with vector-valued output, as described in Section III. It can easily be seen that this channel is a BISMC. Let

$$I_Z := I(Z; \boldsymbol{Y}_{[2,N-1]})$$

denote the mutual information of this channel, and let $\boldsymbol{J}_Z \in \mathbb{J}_Z$ denote its subchannel indicator.

Following similar ideas as in the proof of Theorem 1, we can write the extrinsic information using the binary information function given in Definition 2:

$$\begin{aligned} I_{\mathsf{ext},0} &= I(X_0; \boldsymbol{Y}_{[1,N-1]}) = I(X_0; \boldsymbol{Y}_{[1,N-1]} | J_1, \boldsymbol{J}_Z) \\ &= \mathop{\mathrm{E}}_{\substack{j_1 \in \mathbb{J}_1 \\ \boldsymbol{j}_Z \in \mathbb{J}_Z}} \left\{ I(X_0; \boldsymbol{Y}_{[1,N-1]} | J_1 = j_1, \boldsymbol{J}_Z = \boldsymbol{j}_Z) \right\} \\ &= \mathop{\mathrm{E}}_{\substack{j_1 \in \mathbb{J}_1 \\ \boldsymbol{j}_Z \in \mathbb{J}_Z}} \left\{ f_2^{\mathsf{par}}(I_{\mathsf{int},1}(j_1), I_Z(\boldsymbol{j}_Z)) \right\} \tag{20} \end{aligned}$$

First, we prove the lower bound. Consider the expectation with respect to j_1. Using Lemma 2(a) and Jensen's inequality, we obtain

$$\mathop{\mathrm{E}}_{j_1 \in \mathbb{J}_1} \left\{ f_2^{\mathsf{par}}(I_{\mathsf{int},1}(j_1), I_Z(\boldsymbol{j}_Z)) \right\} \geq f_2^{\mathsf{par}}(I_{\mathsf{int},1}, I_Z(\boldsymbol{j}_Z)),$$

where equality holds if $|\mathbb{J}_1| = 1$, i.e., if $X_1 \to Y_1$ is a BSC. Since the expectation with respect to $\boldsymbol{j}_Z$ is independent of the channel $X_1 \to Y_1$, the extrinsic information is minimal with respect to this channel if it is a BSC. This holds for all other channels in a similar way, and thus the extrinsic information is minimal if the channels are all BSCs. Together with (19), this proves the lower bound.

To prove the upper bound, we determine first the extrinsic information for the case where the channel $X_1 \to Y_1$ is a BEC. If $Y_1 \neq 0$, the information on code bit X_0 is equal to one. If $Y_1 = 0$ (corresponding to an erasure), the information on code bit X_0 is equal to I_Z. The first event happens with probability $1 - \delta_1 = I_{\mathsf{int},1}$, and the second event happens with probability $\delta_1 = 1 - I_{\mathsf{int},1}$, according to (6). Thus, the extrinsic information can be written as

$$\begin{aligned} I_{\mathsf{ext},0} &= I_{\mathsf{int},1} \cdot 1 + (1 - I_{\mathsf{int},1}) \cdot I_Z \\ &= I_{\mathsf{int},1} + I_Z - I_{\mathsf{int},1} I_Z. \end{aligned}$$

On the other hand, the extrinsic information can be upper-bounded using Lemma 2(b) in (20):

$$\begin{aligned} I_{\mathsf{ext},0} &\leq \mathop{\mathrm{E}}_{\substack{j_1 \in \mathbb{J}_1 \\ \boldsymbol{j}_Z \in \mathbb{J}_Z}} \left\{ I_{\mathsf{int},1}(j_1) + I_Z(\boldsymbol{j}_Z) - I_{\mathsf{int},1}(j_1) \cdot I_Z(\boldsymbol{j}_Z) \right\} = \\ &= I_{\mathsf{int},1} + I_Z - I_{\mathsf{int},1} \cdot I_Z. \end{aligned}$$

When comparing the last two relations, we see that the upper bound with respect to channel $X_1 \to Y_1$ is achieved when this channel is a BEC. Since the same reasoning can be applied to the other channels, the extrinsic information is maximal if all channels are BECs. Together with (18), this proves the upper bound. ∎

C. Complete Information by Combining Intrinsic and Extrinsic Information

In the previous two sections, we have discussed bounds on the extrinsic information. This section deals with bounds on the complete information, which is a combination of the intrinsic information and the extrinsic information.

Consider a code bit X_0 on which we have intrinsic information $I_{int,0}$ and extrinsic information $I_{ext,0}$. Furthermore, consider a length 3 repetition code with code bits X_0', X_1', and X_2', where the intrinsic information on code bits X_1' and X_2' is equal to $I_{int,0}$ and $I_{ext,0}$, respectively. It can easily be seen that the extrinsic information on code bit X_0' is equal to the complete information on code bit X_0.

Using this interpretation, we can immediately apply Theorem 2 to bound the complete information on code bit X_0:

Theorem 3 (Complete Information)
Let $X_0, X_1, \ldots, X_{N-1} \in \mathbb{B}$ denote the code bits of a linear code of length N. Let $X_i \to Y_i$, $i = 1, 2, \ldots, N-1$, denote $N-1$ independent BISMCs. Let the intrinsic information on code bit X_0 be defined by $I_{int,0} := I(X_0; Y_0)$, let the extrinsic information on code bit X_0 be defined by $I_{ext,0} := I(X_0; \boldsymbol{Y}_{\setminus 0})$, and let the complete information on code bit X_0 be defined by $I_{comp,0} := I(X_0; \boldsymbol{Y})$. Then, the following bounds hold:

$$\begin{aligned} I_{comp,0} &\geq f_2^{par}(I_{int,0}, I_{ext,0}), \\ I_{comp,0} &\leq 1 - (1 - I_{int,0})(1 - I_{ext,0}). \end{aligned}$$

The lower bound is achieved if the intrinsic channel $X_0 \to Y_0$ and the extrinsic channel $X_0 \to \boldsymbol{Y}_{\setminus 0}$ are BSCs. The upper bound is achieved if the intrinsic and the extrinsic channel are BECs.

V. APPLICATIONS

In this section, we consider the extrinsic information transfer (EXIT) functions [5], [6] and the information processing characteristics (IPCs) [2] for single parity check codes and repetition codes. The EXIT function and the IPC both describe properties of a coding scheme, including encoder and decoder, using the mutual information as a measure. In the following section, we shortly revise these two concepts.

A. EXIT Function and Information Processing Characteristic

The EXIT function describes the behavior of a constituent decoder involved in iterative decoding. This function is defined in the following, using the notation introduced in Section II. We assume that the observations Y_i of code bits X_i are either outputs of the communication channel, called channel values, or soft values (commonly, probabilities or LLRs) provided by other constituent decoders, called *a priori* values. (If both a channel value and an *a priori* value is available for a code bit, we may (conceptually) extend the code with a replica of this bit and associate the channel value with the original code bit and the *a priori* value with the replica.)

Let $\mathbb{I} := \{0, 1, \ldots, N-1\}$ denote the index set of all code bits. Let further $\mathbb{I}_{ch} \subseteq \mathbb{I}$ denote the index set of all code bits for which channel values are available, and let $\mathbb{I}_{apri} \subseteq \mathbb{I}$ denote the index set of all code bits for which *a priori* values are available. Last, let $\mathbb{I}_{syst}$ denote the index set of the systematic code bits, and let $K := |\mathbb{I}_{syst}|$ denote its cardinality. Thus, the code rate is given by $R := K/N$.

The channel between a code bit and its channel value is called the *communication channel*, and it is assumed to be the same for all code bits in $\mathbb{I}_{ch}$. The mutual information of the communication channel is called the *channel information*, and it is denoted by

$$I_{ch} := I(X_i; Y_i), \tag{21}$$

for $i \in \mathbb{I}_{ch}$.

The virtual channel between a code bit and its *a priori* value, called the *a priori channel*, is assumed to be the same for all code bits in $\mathbb{I}_{apri}$. The mutual information of the *a priori* channel is called the *a priori information*, and it is denoted by

$$I_{apri} := I(X_i; Y_i), \tag{22}$$

for $i \in \mathbb{I}_{apri}$.

Assume an optimal symbol-by-symbol decoder computing the extrinsic log-likelihood ratio (LLR) w_i for code bit X_i as

$$w_i := \ln \frac{\Pr(X_i = +1 | \boldsymbol{Y}_{\setminus i} = \boldsymbol{y}_{\setminus i})}{\Pr(X_i = -1 | \boldsymbol{Y}_{\setminus i} = \boldsymbol{y}_{\setminus i})},$$

$i \in \mathbb{I}_{apri}$. Notice that extrinsic values are usually computed only for code bits, for which *a priori* values are available; this *a priori* value may represent no knowledge (corresponding to an *a priori* LLR that is equal to zero) before the first iteration. It can be shown that

$$I(X_i; W_i) = I(X_i; \boldsymbol{Y}_{\setminus i}), \tag{23}$$

which corresponds to the optimality of the decoder with respect to maximal symbol-wise mutual information [16].

The virtual channel between a code bit and its extrinsic LLR is called the *extrinsic channel*. It can easily be seen that it is a BISMC. The *extrinsic information* is the average mutual information of the extrinsic channels corresponding to code bits in $\mathbb{I}_{apri}$,

$$\begin{aligned} I_{ext} &:= \frac{1}{|\mathbb{I}_{apri}|} \sum_{i \in \mathbb{I}_{apri}} I(X_i; W_i) \\ &= \frac{1}{|\mathbb{I}_{apri}|} \sum_{i \in \mathbb{I}_{apri}} I_{ext,i}, \end{aligned} \tag{24}$$

where (2) and (23) were used. The *EXIT function* plots the extrinsic information versus the *a priori* information. If channel information (in the sense of mutual information, as introduced above) is available, it is used as the parameter.

The IPC describes the behavior of the whole coding scheme. This is discussed in the sequel, using the notation introduced in Section II. We assume that channel information (again in the sense of mutual information, as introduced above) is available for all code bits. The IPC plots the mutual information between the inputs of the encoder and the outputs of the decoder versus the channel information. In the following, we consider the IPC for optimal symbol-by-symbol decoding and the IPC for optimal word-decoding (optimal with respect to maximal mutual information).

Consider first an optimal symbol-by-symbol decoder computing *a posteriori* LLRs for each systematic code bit X_i,

$$l_i := \ln \frac{\Pr(X_i = +1 | \boldsymbol{Y} = \boldsymbol{y})}{\Pr(X_i = -1 | \boldsymbol{Y} = \boldsymbol{y})},$$

for $i \in \mathbb{I}_{\mathsf{syst}}$. Similar to (23), it can be shown [16] that

$$I(X_i; L_i) = I(X_i; \boldsymbol{Y}). \tag{25}$$

It can easily be seen that the virtual channel between a systematic code bit and its *a posteriori* LLR is a BISMC. The *complete symbol-wise information* is the average mutual information between a code bit and its *a posteriori* LLR,

$$
\begin{aligned}
I_{\mathsf{comp}} &:= \frac{1}{K} \sum_{i \in \mathbb{I}_{\mathsf{syst}}} I(X_i; L_i) \\
&= \frac{1}{K} \sum_{i \in \mathbb{I}_{\mathsf{syst}}} I_{\mathsf{comp},i},
\end{aligned} \tag{26}
$$

where (3) and (25) were used. The IPC for optimal symbol-by-symbol decoding is defined as the function mapping the channel information to the complete symbol-wise information.

Consider now an optimal word-decoder computing the *a posteriori* probability of each code word. The resulting mutual information between encoder input and decoder output is called the *complete word-wise information per systematic bit*, and it is denoted by $I_{W,\mathsf{comp}}$. Since it is equal to the mutual information between a code word and the vector of outputs of the communication channel, we can define it as

$$I_{W,\mathsf{comp}} := \frac{1}{K} I(\boldsymbol{X}; \boldsymbol{Y}). \tag{27}$$

The IPC for optimal word-decoding is defined as the function mapping the channel information to the complete word-wise information per systematic bit.

The IPC of a coding scheme of code rate R is upper-bounded by $\min\{I_{\mathsf{ch}}/R, 1\}$, as shown in [3]. This corresponds to the IPC of an *ideal coding scheme* defined as a coding scheme leading to the minimum error rate for a communication channel with given channel information I_{ch}.

EXIT functions and IPCs are valid for certain models for the communication channel and for the (virtual) *a priori* channel. They may be computed using the efficient and convenient method proposed in [16]. If only the channel information and the *a priori* information are given, and not the underlying channel models, we can still give bounds on those functions. This is done for the single parity check code and for the repetition code in the sequel. (Notice again that *a priori* information, extrinsic information, and complete information are mutual informations.)

B. Single Parity Check Code

Consider a single parity check code as defined in Section IV-A. For this code, we present first bounds on the EXIT function and then bounds on its IPCs.

Initially, we assume that *a priori* information is available for all code bits and there is no channel information for any code bit, i.e., $|\mathbb{I}_{\mathsf{apri}}| = N$ and $|\mathbb{I}_{\mathsf{ch}}| = 0$. This corresponds to the decoding operation for a check node in the iterative decoder for

a low-density parity-check code, see [17]–[19]. It also applies in the case where a single parity check code is used as an outer code in a serially concatenated coding scheme.

Then, the extrinsic information on code bit X_0 can be bounded according to Theorem 1. Since the extrinsic information according to (24) is equal to the extrinsic information on code bit X_0, we have the bounds

$$
\begin{aligned}
I_{\mathsf{ext}} &\geq (I_{\mathsf{apri}})^{N-1}, \\
I_{\mathsf{ext}} &\leq f_{N-1}^{\mathsf{ser}}(I_{\mathsf{apri}}, I_{\mathsf{apri}}, \ldots, I_{\mathsf{apri}}).
\end{aligned}
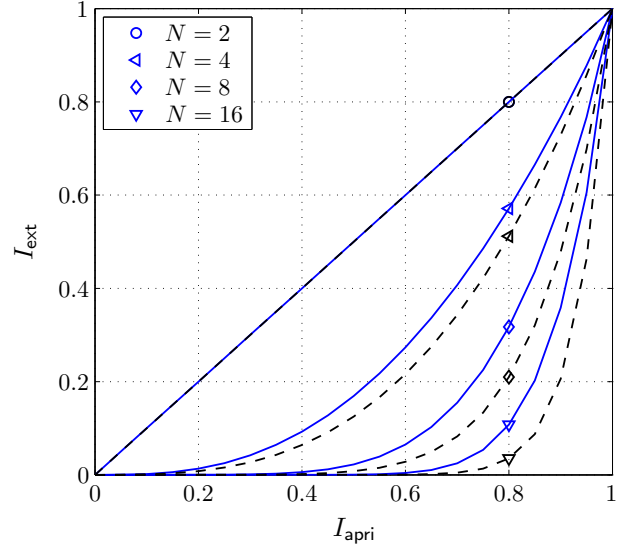$$

These bounds are illustrated in Fig. 5.

Fig. 5. Bounds on the EXIT functions for single parity check codes of several code lengths N. (Upper bounds: solid line; lower bounds: dashed line. Mutual information is given in bit/use.)

Assume now that channel information on code bit X_{N-1} and *a priori* information on all other code bits are available, i.e., $|\mathbb{I}_{\mathsf{apri}}| = N - 1$ and $|\mathbb{I}_{\mathsf{ch}}| = 1$. This is the case if single parity check codes are used as inner codes in serially concatenated coding schemes, and only the parity bits are transmitted over the communication channel. Using Theorem 1, we obtain the bounds

$$
\begin{aligned}
I_{\mathsf{ext}} &\geq I_{\mathsf{ch}} \cdot (I_{\mathsf{apri}})^{N-2}, \\
I_{\mathsf{ext}} &\leq f_{N-1}^{\mathsf{ser}}(I_{\mathsf{ch}}, I_{\mathsf{apri}}, \ldots, I_{\mathsf{apri}}).
\end{aligned}
$$

These bounds are illustrated in Fig. 6. Obviously, the extrinsic information cannot become larger than the channel information, even if the *a priori* information is equal to 1. This makes this code unattractive as the inner code of a serially concatenated coding scheme, because iterative decoding can never achieve mutual information of 1, i.e., be without errors.

Consider now the IPC, assuming that we have channel information on all code bits. The complete information on code bit X_0 is a combination of the intrinsic information $I_{\mathsf{int},0} = I_{\mathsf{ch}}$ and the extrinsic information $I_{\mathsf{ext},0}$, and it can be bounded according to Theorem 3. The extrinsic information on code bit X_0 is bounded as

$$I_{\mathsf{ch}}^{N-1} \leq I_{\mathsf{ext},0} \leq f_{N-1}^{\mathsf{ser}}(I_{\mathsf{ch}}, \ldots, I_{\mathsf{ch}}),$$
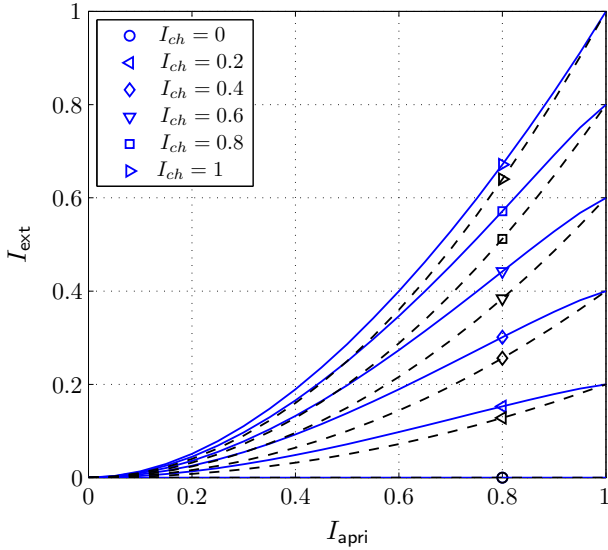
Fig. 6. Bounds on the EXIT functions for single parity check codes of length $N = 4$ for several values of channel information I_{ch}. (Upper bounds: solid line; lower bounds: dashed line. Mutual information is given in bit/use.)

Fig. 7. Bounds on the IPC for optimal symbol-by-symbol decoding for single parity check codes of several code lengths N. (Upper bounds: solid line; lower bounds: dashed line; ideal coding scheme: dash-dotted line. Mutual information is given in bit/use.)

according to Theorem 1. Using the lower bound on the extrinsic information in the lower bound on the complete information, we obtain

$$I_{\text{comp},0} \geq f_2^{\text{par}}(I_{\text{ch}}, I_{\text{ch}}^{N-1}). \tag{28}$$

Similarly, using the upper bound on the extrinsic information in the upper bound on the complete information, we obtain

$$I_{\text{comp},0} \leq 1 - (1 - I_{\text{ch}})(1 - f_{N-1}^{\text{ser}}(I_{\text{ch}}, I_{\text{ch}}, \ldots, I_{\text{ch}})). \tag{29}$$

In contrast to the bounds on the extrinsic information, the above bounds on the complete information are not tight, since we "mixed" channel models. The extrinsic information is minimal if the channels are BECs; in this case, the resulting "extrinsic channel" is also a BEC. On the other hand, the formula for the lower bound of the combined information holds with equality if both channels are BSCs. Due to this contradiction, the lower bound on the complete information is not tight. In a similar way, we can argue for the upper bound.

It can easily be seen that the complete information is the same for each systematic bit. Therefore, bounds on the complete symbol-wise information, defined in (26), are given by

$$
\begin{aligned}
I_{\text{comp}} &\geq f_2^{\text{par}}(I_{\text{ch}}, I_{\text{ch}}^{N-1}), \\
I_{\text{comp}} &\leq 1 - (1 - I_{\text{ch}})(1 - f_{N-1}^{\text{ser}}(I_{\text{ch}}, I_{\text{ch}}, \ldots, I_{\text{ch}})).
\end{aligned} \tag{30}
$$

Thus, we have bounds on the IPC for optimal symbol-by-symbol decoding.

These bounds are plotted in Fig. 7. For each code length N, a large gap between the upper bound and the bound given by the ideal coding scheme can be observed. This gap becomes smaller if word-decoding is applied. We now consider this case.
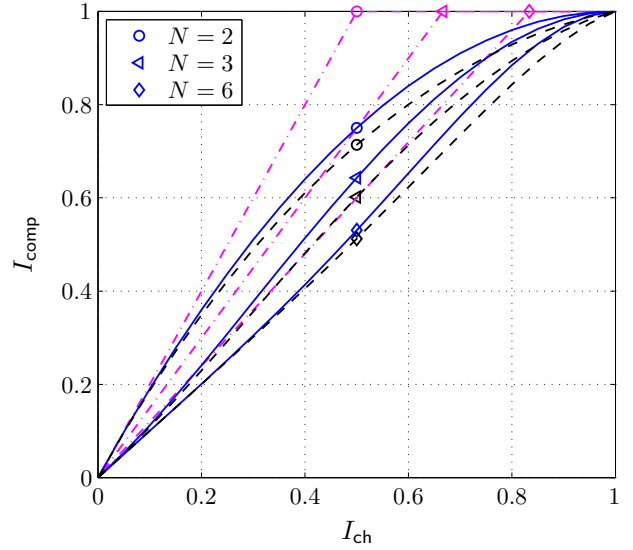
We can apply the chain rule for mutual information and write the complete word-wise information as

$$
\begin{aligned}
I(\boldsymbol{X}; \boldsymbol{Y}) &= I(X_0; \boldsymbol{Y}) + I(X_1; \boldsymbol{Y}|X_0) + \\
&\quad + I(X_2; \boldsymbol{Y}|\boldsymbol{X}_{[0,1]}) + \\
&\quad + I(X_3; \boldsymbol{Y}|\boldsymbol{X}_{[0,2]}) + \\
&\quad + \cdots + I(X_{N-2}; \boldsymbol{Y}|\boldsymbol{X}_{[0,N-3]}) \\
&= I(X_0; \boldsymbol{Y}) + I(X_1; \boldsymbol{Y}_{[1,N-1]}|X_0) + \\
&\quad + I(X_2; \boldsymbol{Y}_{[2,N-1]}|\boldsymbol{X}_{[0,1]}) + \\
&\quad + I(X_3; \boldsymbol{Y}_{[3,N-1]}|\boldsymbol{X}_{[0,2]}) + \\
&\quad + \cdots + I(X_{N-2}; \boldsymbol{Y}_{[N-2,N-1]}|\boldsymbol{X}_{[0,N-3]}).
\end{aligned} \tag{31}
$$

In the latter expression, we have omitted observations for given code bits, because they do not contribute to the mutual information.

The first term in this sum is identical to the complete symbol-wise information on code bit X_0, $I_{\text{comp},0}$. The second term corresponds to the complete symbol-wise information on a code bit for a single parity check code of length $N - 1$, because the code bit X_0 is known. Proceeding in a similar way, it can be seen that the term $I(X_i; \boldsymbol{Y}_{[i,N-1]}|\boldsymbol{X}_{[0,i-1]})$ corresponds to the complete symbol-wise information on a code bit for a single parity check code of length $N - i$, $i = 0, 1, \ldots, N - 3$. Thus, we can apply the bounds given in (28) and (29).

The last term corresponds to a single parity check code of length 2. Thus, $I(X_{N-2}; Y_{N-2})$ and $I(X_{N-1}; Y_{N-1})$ represent the intrinsic and the extrinsic information on code bit X_{N-2}, respectively. The complete information, $I(X_{N-2}; \boldsymbol{Y}_{[N-2,N-1]}|\boldsymbol{X}_{[0,N-3]})$, can then be bounded using

Theorem 3:

$$f_2^{\mathsf{par}}(I_{\mathsf{ch}}, I_{\mathsf{ch}}) \leq$$
$$\leq I(X_{N-2}; \boldsymbol{Y}_{[N-2,N-1]} | \boldsymbol{X}_{[0,N-3]}) \leq$$
$$\leq 1 - (1 - I_{\mathsf{ch}})(1 - I_{\mathsf{ch}}).$$

When applying these bounds in (31), we get bounds on the IPC for optimal word-decoding:

$$
\begin{aligned}
I_{W,\mathsf{comp}} &\geq \frac{1}{K} \sum_{i=0}^{N-2} f_2^{\mathsf{par}}(I_{\mathsf{ch}}, I_{\mathsf{ch}}^{N-1-i}), \\
I_{W,\mathsf{comp}} &\leq \frac{1}{K} \sum_{i=0}^{N-2} \Big(1 - (1 - I_{\mathsf{ch}}) \cdot \\
&\qquad\qquad \cdot (1 - f_{N-1-i}^{\mathsf{ser}}(I_{\mathsf{ch}}, \dots, I_{\mathsf{ch}})) \Big) = \\
&= 1 - (1 - I_{\mathsf{ch}}) \cdot \\
&\qquad \cdot \Big(1 - \frac{1}{K} \sum_{i=0}^{N-2} f_{N-1-i}^{\mathsf{ser}}(I_{\mathsf{ch}}, \dots, I_{\mathsf{ch}}) \Big).
\end{aligned}
\tag{32}
$$

(Note that $K = N - 1$.)

These bounds are plotted in Fig. 8. When comparing the results to Fig. 7, we see that the gap to the bound given by the ideal coding scheme is now relatively small as long as the channel information is smaller than about half the code rate. On the other hand, the upper bound (solid line) is slightly above the IPC of the ideal coding scheme. This shows that this bound is not tight.
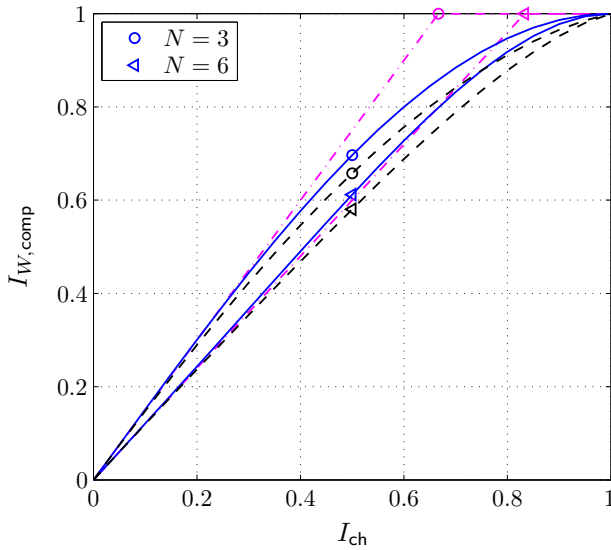

Fig. 8. Bounds on the IPC for optimal word decoding for single parity check codes of several code lengths N. (Upper bounds: solid line; lower bounds: dashed line; ideal coding scheme: dash-dotted line. Mutual information is given in bit/use.)

C. Repetition Code

Consider a repetition code as defined in Section IV-B. For this code, we present first bounds on the EXIT function and then bounds on the IPCs.

Initially, we assume that *a priori* information is available for all code bits and there is no channel information for any code bit, i.e., $|\mathbb{I}_{\mathsf{apri}}| = N$ and $|\mathbb{I}_{\mathsf{ch}}| = 0$. This is the case if repetition codes are used as outer codes in serially concatenated coding schemes, e.g., in repeat accumulate codes [20], [21] or DRS codes [22], [23]. Then, the extrinsic information on code bit X_0 can be bounded according to Theorem 2. Since the extrinsic information according to (24) is equal to the extrinsic information on code bit X_0, we have the bounds

$$
\begin{aligned}
I_{\mathsf{ext}} &\geq f_{N-1}^{\mathsf{par}}(I_{\mathsf{apri}}, I_{\mathsf{apri}}, \dots, I_{\mathsf{apri}}), \\
I_{\mathsf{ext}} &\leq 1 - (1 - I_{\mathsf{apri}})^{N-1}.
\end{aligned}
$$

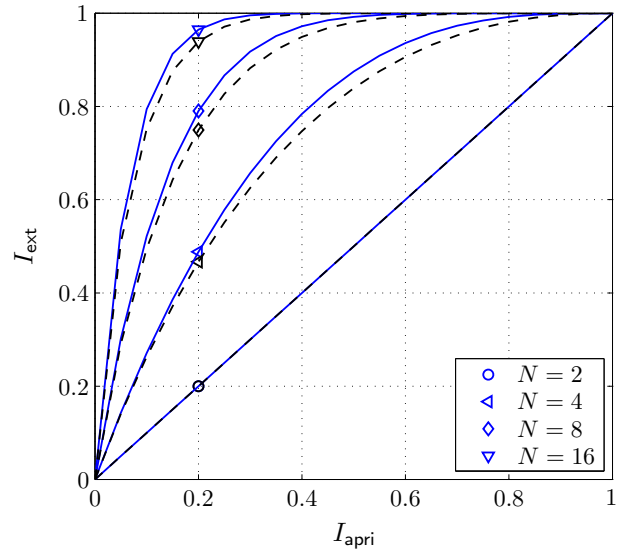These bounds are depicted in Fig. 9 for several code lengths.

Fig. 9. Bounds on the EXIT functions for repetition codes of several code lengths N. (Upper bounds: solid line; lower bounds: dashed line. Mutual information is given in bit/use.)

Assume now that the channel information on code bit X_{N-1} and *a priori* information for all other code bits are available, i.e., $|\mathbb{I}_{\mathsf{apri}}| = N - 1$ and $|\mathbb{I}_{\mathsf{ch}}| = 1$. This corresponds to the decoding operation for a variable node in the iterative decoder for a low-density parity-check code. It is also the case for repetition codes used in systematic repeat accumulate codes, see, e.g., [21]. Using Theorem 2, we obtain the bounds

$$
\begin{aligned}
I_{\mathsf{ext}} &\geq f_{N-1}^{\mathsf{par}}(I_{\mathsf{ch}}, I_{\mathsf{apri}}, \dots, I_{\mathsf{apri}}), \\
I_{\mathsf{ext}} &\leq 1 - (1 - I_{\mathsf{ch}})(1 - I_{\mathsf{apri}})^{N-2}.
\end{aligned}
$$

These bounds are depicted in Fig. 10. In direct contrast to the curves for the single parity check codes in Fig. 6, these curves start with $I_{\mathsf{ext}} = I_{\mathsf{ch}}$ (at $I_{\mathsf{apri}} = 0$) and end with $I_{\mathsf{ext}} = 1$ (at $I_{\mathsf{apri}} = 1$) for increasing *a priori* information. The latter makes these codes particularly suitable for iterative decoding.

Consider now the IPC, assuming that we have channel information on all code bits. It is sufficient to discuss the complete information on code bit X_0, as it is identical to both the complete symbol-wise information and the complete word information.

Computing the complete information on code bit X_0 is equivalent to computing the extrinsic information on a code
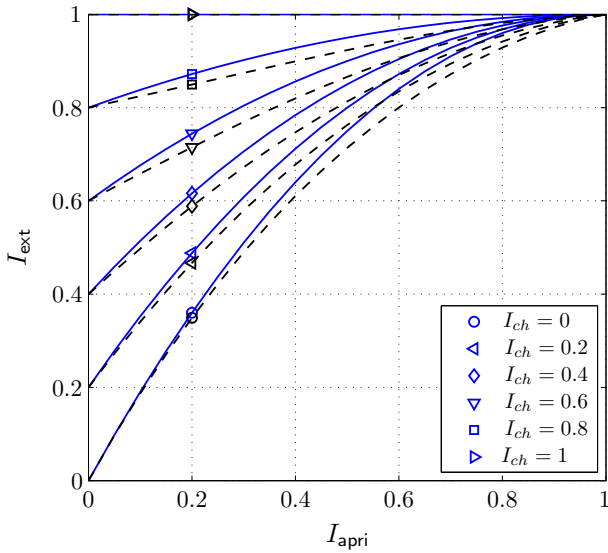
Fig. 10. Bounds on the EXIT functions for repetition codes of length $N = 4$ for several values of channel information I_{ch}. (Upper bounds: solid line; lower bounds: dashed line. Mutual information is given in bit/use.)
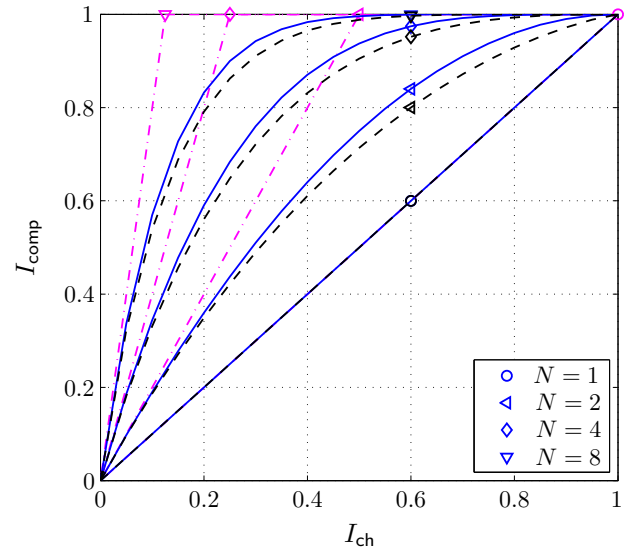

Fig. 11. Bounds on the IPC for repetition codes of several code lengths N. (Upper bounds: solid line; lower bounds: dashed line; ideal coding scheme: dash-dotted line. Mutual information is given in bit/use.)

bit of a repetition code of length $N + 1$, where the intrinsic information of all other code bits is equal to the channel information. Thus, we can apply Theorem 2 to obtain bounds on the complete symbol-wise information:

$$
\begin{aligned}
I_{\text{comp}} &\geq f_N^{\text{par}}(I_{\text{ch}}, \ldots, I_{\text{ch}}), \\
I_{\text{comp}} &\leq 1 - (1 - I_{\text{ch}})^N.
\end{aligned}
$$

Due to $I_{W,\text{comp}} = I_{\text{comp}}$, the same bounds hold for the complete word-wise information. Thus we have bounds on the IPC for optimal decoding.

These bounds are plotted in Fig. 11. As was the case for the bounds for the single parity check code in Fig. 7, we observe a large gap between the upper bound and the IPC of the ideal coding scheme, unless the channel information is very small or very large.

VI. Conclusions

In this paper, we have presented bounds on functions which map mutual information per channel-use to mutual information between encoder input and decoder output.

Based on the separability of binary-input symmetric memoryless channels (BISMC) into BSCs, we have proved bounds on the extrinsic information for single parity check codes and for repetition codes and bounds on the combination of extrinsic and intrinsic information. These bounds are achieved when the channels are BSCs or BECs.

Using these results, we have bounded EXIT functions and information processing characteristics (IPCs) for single parity check codes and for repetition codes. As opposed to the original EXIT functions and IPCs, these bounds are not only valid for specific channel models, but for all BISMCs.

The present paper extends the principle of information combining presented in [2], and also it extends the bounding of combined information presented in [7]. For future research, two aspects may be of special interest: first, the generalization

to nonbinary channels, and second, the generalization to more complicated codes.

Appendix I
Capacity of Serially Concatenated Binary Symmetric Channels

Consider n binary symmetric channels (BSCs) $X_i \rightarrow Y_i$, $X, Y_i \in \mathbb{B}$, $i = 1, 2, \ldots, n$, which are serially concatenated such that $Y_i = X_{i+1}$ for $i = 1, 2, \ldots, n - 1$. The input of the first channel is assumed to be uniformly distributed. The mutual information of each individual channel is denoted by $I_i := I(X_i; Y_i)$, $i = 1, 2, \ldots, n$. The end-to-end mutual information between the input of the first and the output of the last channel is denoted by $I := I(X_1; Y_n)$.

It is now shown that the end-to-end mutual information is given by the binary information function for serial concatenation according to Definition 1, i.e.,

$$
I = f_n^{\text{ser}}(I_1, I_2, \ldots, I_n). \tag{33}
$$

As the serially concatenated channel $X_1 \rightarrow Y_n$ is symmetric, the mutual information I is equal to the channel capacity.

We start with the case $n = 2$. It can easily be seen that the serially concatenated channel $X_1 \rightarrow Y_2$ is also a BSC. Let ϵ_{12} denote its crossover probability. We have an error on this channel if an error occurs either on the first or on the second channel. With $\epsilon_1 = h^{-1}(1 - I_1)$ and $\epsilon_2 = h^{-1}(1 - I_2)$ denoting the crossover probabilities of the two individual channels, we can compute the crossover probability of the serially concatenated channel as

$$
\epsilon_{12} = (1 - \epsilon_1)\epsilon_2 + \epsilon_1(1 - \epsilon_2).
$$

Thus, its mutual information is given by

$$
\begin{aligned}
I(X_1; Y_2) &= 1 - h(\epsilon_{12}) \\
&= 1 - h\Big((1 - \epsilon_1)\epsilon_2 + \epsilon_1(1 - \epsilon_2)\Big),
\end{aligned}
$$

and we have the proof of (33) for $n = 2$.

The general case can easily be shown by induction.

Appendix II
Capacity of Parallel Concatenated Binary Symmetric Channels

Consider n binary symmetric channels (BSCs) $X \rightarrow Y_i$, $X, Y_i \in \mathbb{B}$, $i = 1, 2, \ldots, n$, that have the same input X. Following the accepted practice for parallel concatenated codes, see, e.g., [24], we call these channels parallel concatenated. The input is assumed to be uniformly distributed. The mutual information of each channel is denoted by $I_i := I(X; Y_i)$, $i = 1, 2, \ldots, n$. The vector of channel outputs is written as $\boldsymbol{Y} := [Y_1, Y_2, \ldots, Y_n]$. The overall mutual information between the input and the vector of channel outputs is denoted by $I := I(X; \boldsymbol{Y})$.

It is now shown that the overall mutual information is given by the binary information function for parallel concatenation according to Definition 2, i.e.,

$$I = f_n^{\mathsf{par}}(I_1, I_2, \ldots, I_n). \tag{34}$$

As the parallel concatenated channel $X \rightarrow \boldsymbol{Y}$ is symmetric, the mutual information I is equal to the channel capacity.

To start with, we write the overall mutual information as

$$I = I(X; \boldsymbol{Y}) = H(\boldsymbol{Y}) - H(\boldsymbol{Y}|X). \tag{35}$$

The first term can be computed using the joint probabilities of the channel outputs,

$$\begin{aligned} H(\boldsymbol{Y}) &= \mathrm{E}\{-\log_2 p_{\boldsymbol{Y}}(\boldsymbol{y})\} \\ &= -\sum_{\boldsymbol{y} \in \mathbb{B}^n} p_{\boldsymbol{Y}}(\boldsymbol{y}) \cdot \log_2 p_{\boldsymbol{Y}}(\boldsymbol{y}) \end{aligned}$$

with

$$\begin{aligned} p_{\boldsymbol{Y}}(\boldsymbol{y}) &= \sum_{x \in \mathbb{B}} p_{X,\boldsymbol{Y}}(x, \boldsymbol{y}) \\ &= \sum_{x \in \mathbb{B}} p_X(x) \cdot p_{\boldsymbol{Y}|X}(\boldsymbol{y}|x) \\ &= \sum_{x \in \mathbb{B}} p_X(x) \cdot \prod_{i=1}^{n} p_{Y_i|X}(y_i|x). \end{aligned}$$

In the last line, we used the conditional independence of the channel outputs for a given channel input. Due to the uniform input distribution, we have $p_X(x) = 1/2$. The transition probabilities of each channel can be expressed using its mutual information:

$$p_{Y_i|X}(y_i|x) \in \{\epsilon_i, 1 - \epsilon_i\}$$

with

$$\epsilon_i := h^{-1}(1 - I_i),$$

$i = 1, 2, \ldots, n$. Thus, the joint probability of a vector of channel outputs y can be obtained according to

$$p_{\boldsymbol{Y}}(\boldsymbol{y}) = \frac{1}{2}\left(\prod_{i=1}^{n} \varphi_i(y_i) + \prod_{i=1}^{n}\bigl(1 - \varphi_i(y_i)\bigr)\right),$$

with

$$\varphi_i(y_i) := \begin{cases} \epsilon_i & \text{for } y_i = +1, \\ 1 - \epsilon_i & \text{for } y_i = -1. \end{cases}$$

The second term in (35) can be written as

$$H(\boldsymbol{Y}|X) = \sum_{i=1}^{n} H(Y_i|X) = = \sum_{i=1}^{n}(1 - I_i),$$

where again, the conditional independence of the channel outputs for a given channel input was used.

By substituting the above equations into (35), we have the proof of (34).

References

[1] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, July and Oct. 1948.

[2] S. Huettinger, J. Huber, R. Johannesson, and R. Fischer, "Information processing in soft-output decoding," in *Proc. Allerton Conf. on Communications, Control, and Computing*, Monticello, Illinois, USA, Oct. 2001.

[3] S. Huettinger, J. Huber, R. Fischer, and R. Johannesson, "Soft-output-decoding: Some aspects from information theory," in *Proc. Int. ITG Conf. on Source and Channel Coding*, Berlin, Germany, Jan. 2002, pp. 81–90.

[4] J. Huber and S. Huettinger, "Information processing and combining in channel coding," in *Proc. Int. Symp. on Turbo Codes & Rel. Topics*, Brest, France, Sept. 2003, pp. 95–102.

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

[6] ——, "Code characteristic matching for iterative decoding of serially concatenated codes," *Ann. Télécommun.*, vol. 56, no. 7-8, pp. 394–408, 2001.

[7] I. Land, S. Huettinger, P. A. Hoeher, and J. Huber, "Bounds on information combining," in *Proc. Int. Symp. on Turbo Codes & Rel. Topics*, Brest, France, Sept. 2003, pp. 39–42.

[8] S. Huettinger and J. Huber, "Performance estimation for concatenated coding schemes," in *Proc. IEEE Inform. Theory Workshop*, Paris, France, Mar./Apr. 2003, pp. 123–126.

[9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, 1991.

[10] S. Huettinger and J. Huber, "Extrinsic and intrinsic information in systematic coding," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Lausanne, Switzerland, June 2002, p. 116.

[11] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," *IEEE Trans. Commun.*, vol. 44, no. 10, pp. 1261–1271, Oct. 1996.

[12] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Trans. Inform. Theory*, vol. 42, no. 2, pp. 429–445, Mar. 1996.

[13] I. Land, P. A. Hoeher, and J. Huber, "Bounds on information combining for parity-check equations," in *Proc. Int. Zurich Seminar on Communications (IZS)*, Zurich, Switzerland, Feb. 2004.

[14] I. Sutskover, S. Shamai (Shitz), and J. Ziv, "Extremes of information combining," in *Proc. Allerton Conf. on Communications, Control, and Computing*, Monticello, Illinois, USA, Oct. 2003.

[15] A. D. Wyner and J. Ziv, "A theorem on the entropy of certain binary sequences and applications: Part I," *IEEE Trans. Inform. Theory*, vol. 19, no. 6, pp. 769–772, Nov. 1973.

[16] I. Land, P. A. Hoeher, and S. Gligorević, "Computation of symbol-wise mutual information in transmission systems with LogAPP decoders and application to EXIT charts," in *Proc. Int. ITG Conf. on Source and Channel Coding*, Erlangen, Germany, Jan. 2004.

[17] R. Gallager, "Low-density parity-check codes," *IEEE Trans. Inform. Theory*, vol. 8, no. 1, pp. 21–28, Jan. 1962.

[18] D. J. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 399–431, Mar. 1999.

[19] T. J. Richardson and R. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001.

[20] D. Divsalar, H. Jin, and R. J. McEliece, "Coding theorems for 'turbo-like' codes," in *Proc. Allerton Conf. on Communications, Control, and Computing*, Monticello, Illinois, USA, Sept. 1998, pp. 201–210.

[21] S. ten Brink and G. Kramer, "Design of repeat-accumulate codes for iterative detection and decoding," *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2764–2772, Nov. 2003.

[22] S. ten Brink, "Rate one-half code for approaching the Shannon limit by 0.1 dB," *Electron. Letters*, vol. 36, no. 15, pp. 1293–1294, July 2000.

[23] S. Huettinger, S. ten Brink, and J. Huber, "Turbo-code representation of RA-codes and DRS-codes for reduced decoding complexity," in *Proc. Conf. Inform. Sciences and Systems (CISS)*, The Johns Hopkins University, Baltimore, MD, USA, Mar. 2001, pp. 118–123.

[24] S. Benedetto and G. Montorsi, "Unveiling turbo codes: Some results on parallel concatenated coding schemes," *IEEE Trans. Inform. Theory*, vol. 42, no. 2, pp. 409–428, Mar. 1996.

Ingmar Land Ingmar Land studied electrical engineering at the Universities of Ulm and Erlangen-Nürnberg, Germany. After receiving his Dipl.-Ing. degree in 1999, he joined the University of Kiel, Germany, as a research and teaching assistant. His research topics are information and coding theory with focus on iterative decoding.

Simon Huettinger Simon Huettinger received the Dr.-Ing. Degree from the University of Erlangen-Nürnberg in June 2003. From 2000 to 2003 he was with the Institute for Information Transmission of the University of Erlangen-Nürnberg. Presently, he is with Siemens AG Corporate Technology.

Peter A. Hoeher Peter A. Hoeher received the Dipl.-Eng. and Dr.-Eng. degrees from the Technical University of Aachen, Germany, and the University of Kaiserslautern, Germany, in 1986 and 1990. From 1986 to 1998, he has been with the German Aerospace Research Establishment (DLR) in Oberpfaffenhofen. In 1992, he was on leave at AT&T Bell Laboratories in Murray Hill, NJ. Since 1998, he is a Professor at the University of Kiel, Germany.

Johannes Huber Johannes Huber is Professor for communication engineering at the Universität Erlangen-Nürnberg, Germany since 1991. His research interests are information and coding theory, modulation schemes, algorithms for signal detection and adaptive equalization, multiple-input multiple-output (MIMO) channels, concatenated coding together with iterative decoding.