Information Combining

Information Combining

Ingmar Land

Aalborg University
Denmark

Johannes Huber

Erlangen University
Germany

the essence of knowledge

Boston – Delft

Foundations and Trends® in Communications and Information Theory

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
USA
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

Library of Congress Control Number: 2006937415

The preferred citation for this publication is I. Land and J. Huber, Information Combining, Foundations and Trends® in Communications and Information Theory, vol 3, no 3, pp 227–330, November 2006

Printed on acid-free paper

ISBN: 1-933019-46-8

Foundations and Trends® in Communications and Information Theory
Volume 3 Issue 3, November 2006

Editorial Board

Editorial Scope

Foundations and Trends® in Communications and Information Theory will publish survey and tutorial articles in the following topics:

- Coded modulation
- Coding theory and practice
- Communication complexity
- Communication system design
- Cryptology and data security
- Data compression
- Data networks
- Demodulation and Equalization
- Denoising
- Detection and estimation
- Information theory and statistics
- Information theory and computer science
- Joint source/channel coding
- Modulation and signal design
- Multiuser detection
- Multiuser information theory
- Optical communication channels
- Pattern recognition and learning
- Quantization
- Quantum information processing
- Rate-distortion theory
- Shannon theory
- Signal processing for communications
- Source coding
- Storage and recording codes
- Speech and Image Compression
- Wireless Communications

Information for Librarians

Foundations and Trends® in Communications and Information Theory, November 2006, Volume 3, 4 issues. ISSN paper version 1567-2190. ISSN online version 1567-2328. Also available as a combined paper and online subscription.

Foundations and Trends® in
Communications and Information Theory
Vol. 3, No 3 (November 2006) 227–330

DOI: 10.1561/0100000013

Information Combining

Ingmar Land[1] and Johannes Huber[2]

[1] *Aalborg University, Denmark*
[2] *Erlangen University, Germany*

Abstract

Consider coded transmission over a binary-input symmetric memoryless channel. The channel decoder uses the noisy observations of the code symbols to reproduce the transmitted code symbols. Thus, it combines the information about individual code symbols to obtain an overall information about each code symbol, which may be the reproduced code symbol or its a-posteriori probability. This tutorial addresses the problem of "information combining" from an information-theory point of view: the decoder combines the mutual information between channel input symbols and channel output symbols (observations) to the mutual information between one transmitted symbol and all channel output symbols. The actual value of the combined information depends on the statistical structure of the channels. However, it can be upper and lower bounded for the assumed class of channels. This book first introduces the concept of mutual information profiles and revisits the well-known Jensen's inequality. Using these tools, the bounds on information combining are derived for single parity-check codes and for repetition codes. The application of the bounds is illustrated in four examples: information processing characteristics of coding schemes, including extrinsic information transfer (EXIT) functions; design of multiple turbo codes; bounds for the decoding threshold of low-density parity-check codes; EXIT function of the accumulator.

Contents

1

Introduction

In digital communications, the transmitter adds redundancy to the data to be transmitted, and the receiver exploits this redundancy to perform error correction. In this book, we restrict ourselves to binary linear channel codes and transmission over memoryless communication channels. The transmitter can thus be identified with the channel encoder and the receiver with the channel decoder. Because of the assumed channel model, the receiver obtains one noisy observation for each code symbol.

Each of these observations carries information about the corresponding code symbol at the channel input, of course. In addition to that, due to the code constraints that couple the code symbols, each observation also carries information about other code symbols. To exploit the redundancy in the code, the decoder combines all available information to estimate the value of each code symbol. In this chapter, the focus will be on optimal combining, i.e. combining such that all information about individual code symbols is retained.

This process of information combining can also be seen from an information theory point of view when the asymptotic case of codes of infinite length[1] is considered. For each code symbol, there is a mutual information between the code symbol and the noisy observation. These values of mutual information are "combined" to obtain a value of the mutual information between a code symbol (or an information symbol) and *all* observations. The decoder is thus interpreted as a processor for mutual information. This is done in the information processing characteristic (IPC) method [1, 2, 3].

Some classes of channel codes, e.g., low-density parity-check (LDPC) codes [4, 5], are iteratively decoded: two constituent decoders exchange extrinsic values, called messages, until they agree on a certain estimated codeword, the maximum number of iterations is reached, or another stopping criterion is fulfilled. (The term "extrinsic" will be introduced later.) These constituent decoders can also be interpreted as processors for mutual information, in this case of extrinsic mutual information. This is done in the extrinsic information transfer (EXIT) chart method [6, 7]

The mutual information resulting from the combining operation can be computed exactly if exact models of the channels between the code symbols and the observations (or messages) are assumed to be known, as in the IPC method and the EXIT chart method. Thus, the combined mutual information depends on the "input" mutual information and the channel models. These models (e.g. the Gaussian noise model), however, do not apply exactly.

This chapter addresses a generalization of these ideas. The channels are only assumed to be symmetric and memoryless. Thus, the exact value of the combined mutual information cannot be determined, but an upper and a lower bound can be given. This is referred to as *bounds on information combining* [8, 9]. These bounds depend then only on the values of the "input" mutual information but not on the specific channel model. This basic problem is interesting from a pure information-theory point of view. The results can, however, also be used to analyze coding schemes and iterative decoders; they can even be used to design codes for the whole class of memoryless symmetric channels [10, 11, 12, 13, 14,

[1] To be precise, ensembles of codes are considered and the code length tends to infinity.

15]. A closer look at these references as well as at references to similar or extended combining concepts are provided at the end of this chapter.

This book gives an introduction to the principles of information combining. The concept is described, the bounds for repetition codes and for single parity-check codes are proved, and some applications are provided. As we focus on the basic principles, we consider a binary symmetric source, binary linear channel codes, and binary-input symmetric memoryless channels.

Throughout this book, we use the following notation. Upper-case letters denote random variables, and lower-case letters denote realizations. Vectors and matrices are both written in boldface. The meaning of boldface upper-case letters becomes clear from the context.

1.1 Combining of Probabilities

To achieve very closely the information-theoretic performance bounds of digital communication systems, joint processing of information over long blocks of symbols is necessary. Within such blocks, information has to be combined in some sense, e.g., parity symbols are generated in a channel encoder by forming check sums over distinct subsets of the information symbols, which are fed into the encoder. For a linear block code $\mathcal{C}$ with length N of symbols taken from the binary field $\mathbb{F}_2 = \{0,1\}$, these check sums are specified by the rows of a $(N-K) \times N$ parity check matrix $\boldsymbol{H}$, where K denotes the number of dimensions of the linear subspace in $\mathbb{F}_2^N$ forming the code.

Consider a binary codeword $\boldsymbol{X} = (X_0, X_1, \ldots, X_{N-1})$ of length N that is generated from K binary information symbols; the information symbols are assumed to be independent and uniformly distributed. Each code symbol $X_i \in \{0,1\}$ is transmitted over a binary-input communication channel, which we assume to be symmetric, time-invariant, memoryless, and without feedback throughout this book. This binary input symmetric memoryless channel (BISMC) maps the input symbols X_i into output symbols Y_i taken from an M-ary set $\mathcal{Y} = \{0,1,\ldots,M-1\}$ in a random way according to the transition probabilities $\Pr(Y = j|X = x)$, see Fig. 1.1. A channel is said to be

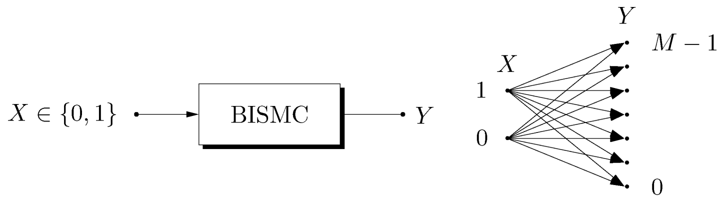

Fig. 1.1 Characterization of a BISMC by means of transition probabilities.

symmetric if it can be decomposed into strongly symmetric subchannels [16]; this is addressed in detail in Section 2.2.

If the a-priori probability $\Pr(X_i = 0)$ and the channel transition probabilities $\Pr(Y_i = y_i|X_i = x_i)$ are known, a-posteriori probabilities

$$p_i := \Pr(X_i = 0|Y_i = y_i)$$
$$= \frac{\Pr(X_i = 0)\Pr(Y_i = y_i|X_i = 0)}{\Pr(X_i = 0)\Pr(Y_i = y_i|X_i = 0) + (1 - \Pr(X_i = 0))\Pr(Y_i = y_i|X_i = 1)} \tag{1.1}$$

are available after observing $Y_i = y_i$ for each individual code symbol. Usually, the vector $\boldsymbol{p} = (p_0, \ldots, p_{N-1})$ of these probabilities after transmission, but before decoding, is referred to as the *intrinsic probabilities for the code symbols* obtained from the communication channel [17].

Without any restriction of generality, we specify a probability on a binary variable $X \in \{0,1\}$ with respect to the value 0, i.e., $\Pr(X = 0|.)$ throughout the book. Of course, probability ratios $\Pr(X = 0|.)/(1 - \Pr(X = 0|.))$ or their logarithms, the so-called L-value $\ln\big(\Pr(X = 0|.)/(1 - \Pr(X = 0|.))\big)$ are synonymous to this notation, but in contrast to the mainstream in technical literature in the field of communications, we think that for theoretical derivations pure probabilities are more convenient than other types of probability specifications: A lot of nonlinear functions can be avoided, some equations are much more evident and easier to handle, and many readers may be more familiar with the language of basic probability theory than with specialized notation popular only in the coding and communications communities. Of course, for implementation of a decoder in hard- or software, probability ratios or, more pronounced, L-value notation may offer a lot of advantages. But the intentions of this tutorial book are quite different; here, the development and understanding of the basic theory is the main focus.

Seen from a general point of view, values of information for individual symbols have to be combined in some way for exploiting the constraints within a sequence of symbols. *Information combining* happens in source encoding for extraction of redundancy from a source sequence or in channel decoding for improvement of data reliability. But there are many further fields where data processing essentially is some sort of information combining. To illustrate what we mean by information combining, we use the example of decoding a linear block code. Without loss of generality, the processing for code symbol X_0 will be further addressed in this example.

In a linear code, each parity check equation (e.g., Qth row of the parity check matrix $\boldsymbol{H}$) that includes X_0 provides further information on the code symbol X_0 by means of the other symbols X_{i_l} due to the check constraint

$$X_0 = X_{i_1} \oplus X_{i_2} \oplus X_{i_3} \oplus \cdots \oplus X_{i_L}. \tag{1.2}$$

Based on the intrinsic probabilities $p_i = \Pr(X_i = 0|y_i)$ of the residual symbols in a check sum, the *extrinsic probability* of code symbol X_0,

$$P_{\mathsf{ext},0} = \Pr(X_0 = 0|y_{i_1}, y_{i_2}, \ldots, y_{i_L}), \tag{1.3}$$

is computed. This probability on a code symbol is called *extrinsic* because it is calculated using only the channel outputs corresponding to the other code symbols but not the channel output corresponding to the symbol itself (see e.g. [18]).

In the case of a memoryless channel, the extrinsic probability $P_{\mathsf{ext},0}$ results in

$$P_{\mathsf{ext},0} = \frac{1}{2}\prod_{l=1}^{L}(2p_{i_l} - 1) + \frac{1}{2}. \tag{1.4}$$

(Remember that the codewords are assumed to be equiprobable.) This famous equation [4] can easily be derived from the case where only three symbols are involved ($X_0 = 0$ if both symbols X_1 and X_2 are 0 or 1)

$$\begin{aligned} P_{\mathsf{ext}} &= \Pr(X_1 \oplus X_2 = 0 | y_1, y_2) \\ &= p_1 p_2 + (1 - p_1)(1 - p_2) \\ &= \frac{1}{2}(2p_1 - 1)(2p_2 - 1) + \frac{1}{2} \end{aligned} \tag{1.5}$$

and by induction from $L - 1$ to L. Notice that (1.4) also corresponds to the probability of observation of symbol 0 at the output of a chain (*series*) of L binary symmetric channels (BSCs) with crossover probabilities $\epsilon_i = 1 - p_i$ when symbol 0 is fed to its input, see Fig. 1.2. Therefore, we refer to (1.4) as the basic formula for *serial combining* of information.

Intrinsic and several extrinsic probabilities on a certain code symbol X are independent as long as the exploited check equations do not contain further code symbols in common and the channel is memoryless, as a memoryless channel acts independently on each of the code symbols. The task, to merge intrinsic and extrinsic probabilities on one symbol into a combined information is equivalent to the situation when a binary code symbol is transmitted over L parallel and independent channels or to the application of a repetition code of rate $1/L$ and transmission of the code symbols over a memoryless channel, see Fig. 1.3.

Thus, the second basic operation of information combining in channel decoding is to merge different, independent messages on individual code symbols and referring to Fig. 1.3, we denominate this operation as *parallel information combining*. Without loss of generality, a uniform a-priori distribution of X can be assumed because one of those "channels" may also be used to specify an a-priori probability on the variable X: a-priori knowledge is nothing else but a further independent source of extrinsic information. Basic probability calculation yields for two

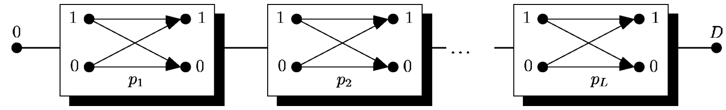

Fig. 1.2 Interpretation of Eq. (1.4) by a chain of BSCs with crossover probabilities $\epsilon_i = 1 - p_i$: serial information combining.

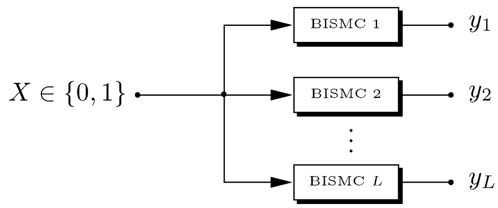

Fig. 1.3 Parallel information combining.

parallel channels (uniform a-priori distribution, cf. Fig. 1.3, too)

$$\begin{aligned}\Pr(X=0|y_1,y_2) &= \frac{p_1p_2}{p_1p_2+(1-p_1)(1-p_2)} \\ &=: p_1 \otimes p_2\,. \end{aligned} \tag{1.6}$$

In the same way, the corresponding result for L parallel channels is obtained:

$$\begin{aligned}\Pr(X=0|y_1,y_2,\ldots,y_L) &= p_1 \otimes p_2 \otimes \cdots \otimes p_L\,, \\ &= \frac{\prod_{l=1}^{L} p_l}{\prod_{l=1}^{L} p_l + \prod_{l=1}^{L}(1-p_l)}. \end{aligned} \tag{1.7}$$

Equation (1.6) is one of the reasons why probability ratios or L-values are very popular in this context: Combining independent a-posteriori probabilities on a binary symbol corresponds to the product of probability ratios or the sum of L-values, respectively. The binary operation "$\otimes$" induces an Abelian group $\mathbb{G} = \{\otimes, [0,1]\}$ onto the set $[0,1]$ of probabilities and by calculating the L-values, i.e., by the function $L : [0,1] \to \mathbb{R} : \ln(x/(1-x))$, an isomorphic mapping of the group $\mathbb{G}$ to $\{+, \mathbb{R}\}$ is established [19]. (Notice that for the basic combining equation (1.4) for check equations (serial information combining), such a nice accordance to L-values does not exist. Unfortunately, the corresponding formulas are rather involved when L-values are used, see (4.3).)

1.2 Combining of Mutual Information

The parallel and serial combination of probabilities on binary variables, i.e., Equations (1.4) and (1.6), are the basic operations for (iterative) soft-decision decoding of linear binary codes. They also form the two key operations for iterative decoding of LDPC codes (details for LDPC codes are provided in Section 6.3). Therefore, we intend to analyze these basic information combining operations in a more general context, looking rather on averages than on individual channel actions and observations as it is usually done in information theory.

One of the key concepts in iterative decoding is the use of extrinsic probabilities (or extrinsic L-values). Correspondingly, the basic problem that we will address in the following sections is to find tight bounds on the mutual information $I(X_0;Y_1,\dots,Y_{L-1})$ for the serial and parallel combination of information solely based on the mutual information $I(X_i;Y_i)$ provided by the channels for transmission of the individual symbols. Notice that this is an *extrinsic mutual information* (e.g., [6]) with respect to X_0 as it is the mutual information between the code symbol X_0 and the observations of only other code symbols; the direct observation of X_0 is omitted.

An introductory example is serial or parallel information combining for binary erasure channels (BECs) with erasure probabilities γ_i and capacities $I_i = 1 - \gamma_i$, cf. Fig. 1.4, which really is the simplest one.

The combination of L received symbols in a check equation leads to an erasure if at least one of the transmitted symbols is erased; otherwise, we get a surely correct extrinsic information. Therefore, the erasure probability of the combined channel reads $\gamma = 1 - \prod_{i=1}^{L}(1-\gamma_i)$, which is equivalent to the formula

$$I = \prod_{i=1}^{L} I_i \tag{1.8}$$

for serial information combining.

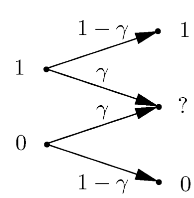

Fig. 1.4 Binary erasure channel (BEC) with erasure probability γ. The erasure is denoted by "?".

Transmission of binary symbol over L parallel BECs yields perfect knowledge at the receiver side if at least one of these channels does not deliver an erasure. Thus, the erasure probability of L parallel BECs is $\gamma = \prod_{i=1}^{L} \gamma_i$, and the overall mutual information (or capacity) reads

$$I = 1 - \prod_{i=1}^{L} (1 - I_i) . \tag{1.9}$$

Unfortunately, such explicit solutions do not exist in general, but we are able to derive rather tight bounds on information combining, if the individual binary input symmetric channels are only specified by their mutual information (or capacity).

1.3 Outline and Related Work

The bounds on information combining will enable us to analyze various properties of coding schemes and iterative decoding procedures in a very general way. "Mutual information" has proven to be a very useful and relevant measure to characterize a channel by a single parameter only. Correspondingly, applying it leads to easy tools to derive fairly tight performance bounds or to optimize coding schemes (e.g., the design of multiple turbo codes, see Section 6.2).

For that purpose, we will recapitulate the basic properties of BISMCs in Chapter 2 and define a new tool to fully specify channels of that type, called the mutual information profile (MIP) of a BISMC. In Chapter 3, Jensen's well-known formula is revisited and extended to a pair of inequalities, i.e., to a lower and an upper bound on the expectation of a real random variable after processing by a convex function; we will identify the probability density functions (pdfs) for real random variables for which those bounds are tight, irrespective of the actual convex function.

Equipped with these prearrangements, the central theorems of this book, i.e., bounds on mutual information for serial and parallel combination of information on binary variables, are derived in a straightforward way in Chapters 4 and 5. Chapter 6 is dedicated to examples and applications of information combining: information processing characteristic of coding schemes, design of multiple turbo codes, and bounds on EXIT functions and bounds on thresholds for convergence of iterative decoding of LDPC codes, and EXIT functions for RA codes.

The problem of information combining for parallel channels has been addressed in [2, 20] for the first time. Here, the so-called information processing characteristic (IPC) for a coding scheme has been introduced, too, cf. Section 6.1. In [21] an example has been given on how to use an IPC and information combining for a coarse estimation of bit error probability (BEP) and BEP-curves for concatenated coding schemes. In [22, 23] the analysis and optimization of multiple turbo codes by means of information combining was proposed. Surprisingly, tight bounds on the combined extrinsic information from several constituent codes in a so-called extended serial setup decoder leads to an analysis of the iterative decoding process, which is as simple as EXIT charts for the concatenation of only two constituent codes.

A more rigorous mathematical background to information combining has been introduced in [8] by finding the proof that there are simple tight bounds on parallel information combining for the case of two channels. Initiated by that, the results were generalized and applied to code design by two groups. In [12,10,9,13], the proofs are explicitly based on the decomposition of symmetric channels into binary symmetric subchannels and the concept of mutual information profiles, which may give a more intuitive access to this subject. Furthermore, these authors address only the basic case of binary symmetric sources and channels without memory, and the optimization with respect to all channels involved. In [14, 15], the proofs are based on [24], which is a generalization of Mrs Gerber's lemma [25], and thus have a more abstract character. These authors also address the question of a uniform source with memory and the role of the symmetry of the source. Furthermore, they show that the optimization can also be done with respect to the individual channels involved.

The present book is mainly based on [13] and the slides to [26] where the material was presented in a way that emphasizes the tutorial aspect. This is the main focus of this book as well. Therefore, we will follow the approaches of the first research group mentioned above.

Even though the present book focuses on pure combining of mutual information, references to similar or extended concepts should be given in the following. Mutual information is probably one the most successfully applied parameter of a memoryless channel. However, such a channel can also be characterized by other parameters, of course, like the expectation of the conditional bit probabilities (expected "soft-bit"), the Bhattacharyya noise parameter, the mean-square error (MSE), see [27, 28, 29, 30, 31]. Instead of using only one parameter to describe a channel, two such parameters may be used, as considered in [28, 32]. Since more parameters may characterize a channel more precisely than a single parameter, the resulting bounds may be tighter.

References

[1] J. Huber and S. Huettinger, "Information Processing and Combining in Channel Coding," in *Int. Symp. on Turbo Codes & Rel. Topics*, Brest, France, September 2003, pp. 95–102.

[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, October 2001.

[3] S. Huettinger, "Analysis and Design of Power-Efficient Coding Schemes," Ph.D. dissertation, University Erlangen-Nürnberg, Germany, 2004.

[4] R. Gallager, "Low-Density Parity-Check Codes," *IEEE Trans. Inform. Theory*, vol. 8, no. 1, pp. 21–28, January 1962.

[5] D. J. MacKay, "Good Error-Correcting Codes Based on Very Sparse Matrices," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 399–431, March 1999.

[6] S. ten Brink, "Convergence Behavior of Iteratively Decoded Parallel Concatenated Codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, October 2001.

[7] S. ten Brink, G. Kramer, and A. Ashikhmin, "Design of Low-Density Parity-Check Codes for Modulation and Detection," *IEEE Trans. Inform. Theory*, vol. 52, no. 14, pp. 670–678, April 2004.

[8] 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, September 2003, pp. 39–42.

[9] Ingmar Land, Simon Huettinger, Peter A. Hoeher, and Johannes Huber, "Bounds on Information Combining," *IEEE Trans. Inform. Theory*, vol. 51, no. 2, pp. 612–619, February 2005.

[10] I. Land, S. Huettinger, P. A. Hoeher, and J. Huber, "Bounds on Mutual Information for Simple Codes Using Information Combining," *Ann. Télécommun.*, vol. 60, no. 1/2, pp. 184–214, January/February 2005.

[11] I. Land, P. A. Hoeher, and J. Huber, "Analytical Derivation of EXIT Charts for Simple Block Codes and for LDPC Codes Using Information Combining," in *Proc. European Signal Processing Conf. (EUSIPCO)*, Vienna, Austria, September 2004.

[12] ——, "Bounds on Information Combining for Parity-Check Equations," in *Proc. Int. Zurich Seminar on Communications (IZS)*, Zurich, Switzerland, February 2004, pp. 68–71.

[13] I. Land, "Reliability Information in Channel Decoding – Practical Aspects and Information Theoretical Bounds," Ph.D. dissertation, University of Kiel, Germany, 2005. [Online]. Available: http://e-diss.uni-kiel.de/diss_1414/

[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, October 2003.

[15] ——, "Extremes of Information Combining," *IEEE Trans. Inform. Theory*, vol. 51, no. 4, pp. 1313–1325, April 2005.

[16] R. G. Gallager, *Information Theory and Reliable Communication.* John Wiley & Sons, 1968.

[17] 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.

[18] C. Berrou and A. Glavieux, "Near Optimum Error Correcting Coding and Decoding: Turbo-Codes," *IEEE Trans. Commun.*, vol. 44, no. 10, pp. 1261–1271, October 1996.

[19] Johannes Huber, "Grundlagen der Wahrscheinlichkeitsrechnung für iterative Decodierverfahren," *e&i: Elektrotechnik und Informationstechnik*, vol. 119, no. 11, pp. 386–394, November 2002.

[20] 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, January 2002, pp. 81–90.

[21] S. Huettinger and J. Huber, "Performance Estimation for Concatenated Coding Schemes," in *Proc. IEEE Inform. Theory Workshop*, Paris, France, March/April 2003, pp. 123–126.

[22] ——, "Design of Multiple–Turbo–Codes with Transfer Characteristics of Component Codes," in *Proc. Conf. Inform. Sciences and Systems (CISS)*, Priniceton University, Princeton, NJ, USA, March 2002.

[23] ——, "Analysis and Design of Power Efficient Coding Schemes with Parallel Concatenated Convolutional Codes," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1251–1258, July 2006.

[24] N. Chayat and S. Shamai (Shitz), "Expansion of an Entropy Property for Binary Input Memoryless Symmetric Channels," *IEEE Trans. Inform. Theory*, vol. 35, no. 5, pp. 1077–1079, September 1989.

[25] 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, November 1973.

[26] J. Huber, I. Land, and P. A. Hoeher, "Information Combining: Models, Bounds and Applications — A Tutorial (Invited Talk)," in *Proc. IEEE Int. Symp. Inform. Theory and Its Applications (ISITA)*, Parma, Italy, October 2004.

[27] D. Burshtein and G. Miller, "Bounds on the Performance of Belief Propagation Decoding," *IEEE Trans. Inform. Theory*, vol. 48, no. 1, pp. 112–122, January 2002.

[28] C. Wang, S. Kulkarni, and H. Poor, "On Finite-Dimensional Bounds for LDPC-like Codes with Iterative Decoding," in *Proc. IEEE Int. Symp. Inform. Theory and its Applications (ISITA)*, Parma, Italy, October 2004.

[29] H. Jin, A. Khandekar, and R. J. McEliece, "Irregular Repeat-Accumulate Codes," in *Proc. Int. Symp. on Turbo Codes & Rel. Topics*, Brest, France, September 2000, pp. 1–8.

[30] K. Bhattad and K. Narayanan, "An MSE Based Transfer Chart to Analyze Iterative Decoding Schemes," in *Proc. Allerton Conf. on Communications, Control, and Computing*, Allerton House, Monticello, IL, USA, 2004.

[31] M. Tuechler, S. ten Brink, and J. Hagenauer, "Measures for Tracing Convergence of Iterative Decoding Algorithms," in *Proc. Int. ITG Conf. on Source and Channel Coding*, Berlin, Germany, January 2002, pp. 53–60.

[32] I. Sutskover, S. Shamai (Shitz), and J. Ziv, "Constrained Information Combining: Theory and Applications for LDPC Coded Systems," submitted to IEEE Trans. Inform. Theory.

[33] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* John Wiley & Sons, 1991.

[34] M. Hellman and J. Raviv, "Probability of Error, Equivocation, and the Chernoff Bound," *IEEE Trans. Inform. Theory*, vol. 16, no. 4, pp. 368–372, July 1970.

[35] 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, March 1996.

[36] J. Huber, personal communication, July 2004.

[37] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate," *IEEE Trans. Inform. Theory*, pp. 284–287, March 1974.

[38] P. Robertson, E. Villebrun, and P. Hoeher, "A Comparison of Optimal and Sub-Optimal MAP Decoding Algorithms Operating in the Log-Domain," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Seattle, USA, 1995, pp. 1009–1013.

[39] P. Robertson, P. Hoeher, and E. Villebrun, "Optimal and Sub-Optimal Maximum a Posteriori Algorithms Suitable for Turbo Decoding," *Europ. Trans. Telecommun.*, vol. 8, no. 2, pp. 119–125, March/April 1997.

[40] F. Jelinek, "A Fast Sequential Decoding Algorithm Using a Stack," *IBM J. Res. and Develop.*, vol. 13, pp. 675–685, February 1969.

[41] J. B. Anderson and S. Mohan, "Sequential Coding Algorithms: A Survey and Cost Analysis," *IEEE Trans. Inform. Theory*, vol. 32, pp. 169–176, February 1984.

[42] C. Kuhn and J. Hagenauer, “Iterative List-Sequential (LISS) Detector for Fading Multiple-Access Channels,” in *Proc. IEEE Globecom Conf.*, Dalles, Texas, USA, November/December 2004.

[43] S. ten Brink, “Design of Serially Concatenated Codes Based on Iterative Decoding Convergence,” in *Proc. Int. Symp. on Turbo Codes & Rel. Topics*, Brest, France, September 2000, pp. 319–322.

[44] A. Ashikhmin, G. Kramer, and S. ten Brink, “Extrinsic Information Transfer Functions: Model and Erasure Channel Properties,” *IEEE Trans. Inform. Theory*, 2004.

[45] Ingmar Land and Johannes Huber, “Information Processing in Ideal Coding Schemes with Code-Symbol Decoding,” in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Adelaide, Australia, September 2005.

[46] C. Berrou, A. Glavieux, and P. Thitimajshima, “Near Shannon Limit Error-Correcting Coding and Decoding: Turbo Codes,” *Proc. IEEE Int. Conf. Commun. (ICC)*, pp. 1064–1070, May 1993.

[47] S. ten Brink, “Code Characteristic Matching for Iterative Decoding of Serially Concatenated Codes,” *Ann. Télécommun.*, vol. 56, no. 7–8, pp. 394–408, 2001.

[48] F. Brännström, “Convergence Analysis and Design of Multiple Concatenated Codes,” Ph.D. dissertation, Chalmers University of Technology, Göteborg, Sweden, 2004.

[49] H.-A. Loeliger, “An Introduction to Factor Graphs,” *IEEE Signal Processing Mag.*, vol. 21, no. 1, pp. 28–41, January 2004.

[50] F. R. Kschischang, B. J. Frey, and H. A. Loeliger, “Factor Graphs and the Sum-Product Algorithm,” *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 498–519, February 2001.

[51] T. J. Richardson and R. L. Urbanke, “The Capacity of Low-Density Parity-Check Codes under Message-Passing Decoding,” *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 599–618, Feb 2001.

[52] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, “Design of Capacity-Approaching Irregular Low-Density Parity-Check Codes,” *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 619–637, February 2001.

[53] R. G. Gallager, “Low Density Parity Check Codes,” Ph.D. dissertation, Cambridge, MA, USA, 1963.

[54] D. J. MacKay and R. Neal, “Near Shannon Limit Performance of Low Density Parity Check Codes,” *Electron. Lett.*, vol. 33, no. 6, pp. 457–458, March 1997.

[55] S.-Y. Chung, T. J. Richardson, and R. L. Urbanke, “Analysis of Sum-Product Decoding of Low-Density Parity-Check Codes Using a Gaussian Approximation,” *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 657–670, February 2001.

[56] E. Sharon, A. Ashikhmin, and S. Litsyn, “EXIT Functions for Binary Input Memoryless Symmetric Channels,” *IEEE Trans. Inform. Theory*, vol. 54, no. 7, pp. 1207–1214, July 2006.

[57] D. Divsalar, H. Jin, and R. J. McEliece, “Coding Theorems for Turbo-Like Codes,” in *Proc. Allerton Conf. on Communications, Control, and Computing*, Allerton House, Monticello, IL, USA, September 1998, pp. 201–210.

[58] D. Divsalar, S. Dolinar, and F. Pollara, "Low Complexity Turbo-like Codes," in *Proc. Int. Symp. on Turbo Codes & Rel. Topics*, Brest, France, September 2000, pp. 73–80.

[59] 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, November 2003.

[60] I. Land, J. Sayir, and P. A. Hoeher, "Bounds on Information Combining for the Accumulator of Repeat-Accumulate Codes without Gaussian Assumption," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Chicago, USA, June/July 2004, p. 443.

[61] Y. Jiang, A. Ashikhmin, R. Koetter, and A. C. Singer, "Extremal Problems of Information Combining," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Adelaide, Australia, September 2005, pp. 1236–1240.

[62] 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, March 1996.

[63] S. Shamai (Shitz), personal communication, September 2003.