Iterative Source-Channel Decoding with Reduced Error Floors

Laurent Schmalen and Peter Vary, *Fellow, IEEE*

***Abstract*—Audio-visual source encoders for digital mobile communications extract parameters that – due to delay and complexity constraints – exhibit some residual redundancy. This residual redundancy can be exploited by iterative source-channel decoding (ISCD) to improve the robustness against transmission noise by performing soft parameter detection as part of the decoding process. Systems employing ISCD at the receiving end often exhibit an observable error floor. While this error floor can be tolerated in some cases, it is often desirable to perfectly reconstruct the source codec parameters. In this paper, we explain the reasons for the error floor and propose two solutions for realizing ISCD systems with optimized error floor performance while maintaining the desired near-capacity waterfall behavior. All approaches aim at optimizing the distance properties of the (redundant) mapping of bit patterns to the source codec parameters. In some cases, especially if small quantizer code books are employed, good mappings cannot be found – in this case, the novel multi-dimensional bit mapping allows to reduce the error floor after decoding.**

***Index Terms*—Iterative Source-Channel Decoding (ISCD), Soft Decision Source Decoding (SDSD), Fixed-length codes**

I. Introduction

SHANNON's source-channel separation theorem [1] states that if the minimum achievable source coding rate of a given source is below the channel capacity, the source can be reliably transmitted over the channel by performing appropriate encoding and decoding operations: therein, source and channel coding can be separated. The original theorem from Shannon holds for stationary and ergodic sources and channels and for asymptotically large block lengths. Shannon already showed the sub-optimality of the separation theorem for multi-user systems in [2].

But also in more "traditional" cases, the constraints of the original separation theorem [1] are almost never fulfilled. In state-of-the-art audio, image and video coding standards, delay and computational complexity restrictions imply the use of short block lengths. Some terms of residual redundancy typically remain in the source codec output parameters (e.g., scale factors or predictor coefficients) due to these constraints and can be exploited at the receiver of a noisy communication system in order to increase the error robustness, e.g., [3], [4].

For these reasons, the joint study of source and channel coding needs to be considered for realizing transmission systems with good performance. Several techniques for joint source-channel coding have been introduced. The variety of these approaches can be divided into mainly two classes. In the first class, channel properties (like bit error rates) are utilized throughout the design of robust source encoding schemes, e.g., pseudo Gray coding [5], channel optimized vector quantization [6], source optimized channel codes [7], or the selection of the optimum rate distribution between source and channel coding [8]. In the second class, source statistics (capturing the natural residual source redundancy) are exploited throughout channel (de)coding. Famous examples include source controlled channel decoding [9], and its application to speech coding [10], [11].

Iterative Source-Channel Decoding (ISCD) [12]–[15] can be considered as an advancement of source controlled channel decoding [9], where in each decoding step the source statistics are iteratively refined in a Turbo-like process. The *a priori* knowledge on the residual redundancy, e.g., non-uniform probability distribution or auto-correlation, is utilized by a derivative of a *Soft Decision Source Decoder* (SDSD) [4] which exchanges extrinsic reliabilities with a channel decoder. In the literature, ISCD has been applied to systems employing *Variable Length Codes* (VLCs) and *Fixed Length Codes* (FLCs). In this paper we put the focus on FLCs, which are used in wireless systems like GSM, UMTS, and LTE and in recent speech and audio codecs, designed for VoIP amongst others, like G.729.1 and G.718. In fact, FLCs are a special case of VLCs and all results given in this paper can directly be generalized to ISCD for VLCs. An introduction to ISCD techniques for VLCs can be found in [16], [17]. The application of ISCD to discrete-valued sources (which can also be source codec parameters) is highlighted in [18].

The key element of the ISCD approach is to exploit residual redundancy which is contained in the output parameters of the source encoder. As mentioned above, this is mostly due to delay and computational complexity constraints. However, a source encoder can also deliberately keep redundancy in the signal in order to increase the error robustness. If this redundancy is removed, for example by differential encoding, the decoder is prone to error propagation.

EXtrinsic Information Transfer (EXIT) charts [19], [20] are a powerful analysis tool for an easy comparison of different system setups and for assessing the convergence performance of ISCD. Many optimizations and performance improvements of ISCD have been obtained by using EXIT charts, e.g., [15], [21]–[23]. However, ISCD systems optimized using EXIT charts show in some cases an error floor, which means that a fraction of the parameters are still reconstructed erroneously,

Laurent Schmalen has been with the RWTH Aachen University, Institute of Communication Systems and Data Processing (IND). He is *now* with Alcatel-Lucent Bell Labs, Stuttgart, Germany (e-mail: Laurent.Schmalen@alcatel-lucent.com).

Peter Vary is with the RWTH Aachen University, Institute of Communication Systems and Data Processing (IND), Germany (e-mail: vary@ind.rwth-aachen.de).

Digital Object Identifier 10.1109/JSTSP.2011.2167216

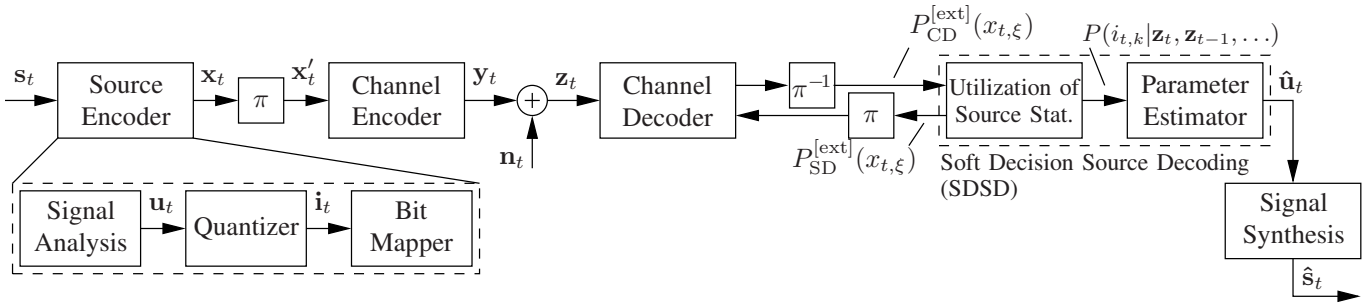

Fig. 1. Transmitter and receiver (base band model) of a transmission system utilizing ISCD.

even in good channel conditions. The error floor is mainly influenced by the distance properties of the bit mapping in the source encoding stage. In several source codecs, the exact reconstruction of the parameters is indispensable for guaranteeing reconstruction of the source signal. The optimization of ISCD schemes improving error floor and waterfall by means of EXIT charts and interleaver gain analysis is discussed in [24] in the context of VLCs with low redundancy. The case of FLCs is also examined in [24].

The goal of this paper is to reduce the error floor of ISCD in the context of FLCs by several means. First, it is shown how the selection of the bit mapping influences the error floor. We further show how the resulting loss in the waterfall region can be compensated by the use of a low complexity inner irregular code. In certain cases, especially if small quantizer code books are employed, the use of such an optimized bit mapping is not always possible. In these cases, an elegant transmitter modification can be applied. The resulting innovative *Multi-Dimensional Bit Mappings* (MDBMs) permit to further lower the error floor.

This paper is structured as follows: Section II introduces the considered system model and the notation. It is shown in Sec. III how the error floor can be considerably reduced by a proper selection of the bit mapping while maintaining a good waterfall performance. Section IV introduces the concept of MDBMs and shows how they permit to lower the error floor if small quantizer code books are employed. The equations for SDSD are given in Appendix A, while Appendix B shows how to compute *a priori* probabilities in the case of MDBMs.

II. System Model

Figure 1 depicts the base band system model considered throughout this paper. At time instant t, a source encoder extracts a frame of N_I source codec parameters $\mathbf{u}_t = (u_{t,1}, \ldots, u_{t,N_I})$ out of the audio-visual source signal $\mathbf{s}_t$. The source codec parameters can for example be gain factors, filter coefficients, pitch values, the residual signal, or transform coefficients of speech and audio codecs. These parameters are assumed to be either inter-frame correlated with correlation coefficient $\rho = \mathsf{E}\{U_{t,k}U_{t-1,k}\}/\mathsf{E}\{U^2_{t,k}\}$ or intra-frame correlated with correlation coefficient $\delta = \mathsf{E}\{U_{t,k}U_{t,k-1}\}/\mathsf{E}\{U^2_{t,k}\}$. Note that the upper-case letter $U_{t,k}$ denotes the random variable describing the source codec parameter $u_{t,k}$. The frame $\mathbf{u}_t$ is quantized using a Q-level scalar quantizer $\mathcal{Q}$ which maps the input parameter $u_{t,k}$ to a quantizer index $i_{t,k}$ denoting the selected entry of the quantizer code book $\mathbb{V} = \{\bar{v}^{(1)}, \ldots, \bar{v}^{(Q)}\} \subset \mathbb{R}$. In this paper, we only consider scalar quantization in order to demonstrate the concept, which can easily be extended to incorporate vector quantization. All quantization indices within a frame are grouped to a vector $\mathbf{i}_t = (i_{t,1}, \ldots, i_{t,N_I})$. To each quantizer index $i_{t,k}$ selected at time instant t and position k, a unique bit pattern $\mathbf{b}_{t,k} \in \mathbb{B} \doteq \{\bar{\mathbf{b}}^{(1)}, \ldots, \bar{\mathbf{b}}^{(Q)}\} \subseteq \mathbb{F}_2^M$ of M bits is assigned according to the (possibly redundant) bit mapping function $\mathcal{B}$ (with $\mathbb{F}_2 \doteq \{0, 1\}$). The single bits of the bit pattern $\mathbf{b}_{t,k}$ are denoted by $b_{t,k,\mu} \in \mathbb{F}_2$, with $\mu \in \{1, \ldots, M\} \subset \mathbb{N}_1$ denoting the μth entry of $\mathbf{b}_{t,k}$. The single bits of the possible bit patterns $\bar{\mathbf{b}}^{(q)} = (\bar{b}^{(q)}_1, \ldots, \bar{b}^{(q)}_\mu, \ldots, \bar{b}^{(q)}_M)$ are denoted by $\bar{b}^{(q)}_\mu$. If $M > \log_2 Q$, the bit mapping is called redundant, as it introduces artificial redundancy: more bits than actually necessary are spent to represent a quantizer index.

The bit mapping can always be considered to be the composite function $\mathcal{B} = \check{\mathcal{B}} \circ \mathcal{B}^{[\mathrm{NB}]}$, i.e., $\mathcal{B}(i_{t,k}) = \left(\check{\mathcal{B}} \circ \mathcal{B}^{[\mathrm{NB}]}\right)(i_{t,k}) = \check{\mathcal{B}}\left(\mathcal{B}^{[\mathrm{NB}]}(i_{t,k})\right)$ with $\mathcal{B}^{[\mathrm{NB}]} : \mathbb{I} \to \mathbb{F}_2^{M^{[\mathrm{NB}]}}$ denoting the natural binary bit mapping. $\mathcal{B}^{[\mathrm{NB}]}$ generates the natural binary representation of the index $i_{t,k}$ of length $M^{[\mathrm{NB}]} \doteq \lceil \mathrm{ld}\, Q \rceil$ according to $\mathcal{B}^{[\mathrm{NB}]}(i_{t,k}) = \mathbf{b}^{[\mathrm{NB}]}_{t,k} = \left(b^{[\mathrm{NB}]}_{t,k,1}, \ldots, b^{[\mathrm{NB}]}_{t,k,M^{[\mathrm{NB}]}}\right)$ with $b^{[\mathrm{NB}]}_{t,k,\mu} = \left\lfloor \frac{i_{t,k}}{2^{M^{[\mathrm{NB}]}-\mu}} \right\rfloor \bmod 2$. The expression $\lceil \chi \rceil$ denotes the smallest integer greater than or equal to χ and $\lfloor \chi \rfloor$ denotes the largest integer less than or equal to χ. Natural binary (NB) simply signifies that the binary representation of the natural number describing the quantizer index is used. In what follows, the superscript "$^{[\mathrm{NB}]}$" always refers to the natural binary component of the bit mapping.

The second constituent position-dependent function of the mapping, $\check{\mathcal{B}}_k : \mathbb{F}_2^{M^{[\mathrm{NB}]}} \to \mathbb{B}_k$ can be regarded as being a (potentially non-linear) block code of rate $r_{\mathrm{BM}} = M^{[\mathrm{NB}]}/M$. Two prominent examples of non-redundant mapping functions $\check{\mathcal{B}}_k$ are the Gray mapping and the pseudo-Gray mapping [5]. If the bit mapping is generated using a linear block code, the function $\check{\mathcal{B}}_k$ can be described as a multiplication with a generator matrix $\mathbf{G}^{[\mathrm{BM}]}$, i.e., $\mathbf{b}_{t,k} = \mathbf{b}^{[\mathrm{NB}]}_{t,k} \cdot \mathbf{G}^{[\mathrm{BM}]}$, with $\dim \mathbf{G}^{[\mathrm{BM}]} = M^{[\mathrm{NB}]} \times M$. After the bit mapping, the N_I individual bit patterns $\mathbf{b}_{t,k}$ are grouped to a bit vector $\mathbf{x}_t \doteq (\mathbf{b}_{t,1}, \ldots, \mathbf{b}_{t,N_I}) = (x_{t,1}, \ldots, x_{t,\xi}, \ldots, x_{t,N_X})$. The size of the bit vector is $N_X \doteq N_I M$. The single bits $x_{t,\xi}$ of $\mathbf{x}_t$ are indexed by ξ. As the bit mapping is considered to be a code,

the rate of the bit mapping is defined by

$$r_{\mathrm{BM}} = \frac{N_I \cdot \mathrm{ld}\, Q}{N_X} = \frac{\mathrm{ld}\, Q}{M} \quad \overset{\text{if } \mathrm{ld}\, Q = M^{[\mathrm{NB}]}}{=} \quad \frac{M^{[\mathrm{NB}]}}{M}. \qquad (1)$$

Following the bit mapping, the bit vector $\mathbf{x}_t$ is permuted by a bijective interleaver function π which maps the bit vector $\mathbf{x}_t$ of length N_X to an (interleaved) bit vector $\mathbf{x}'_t$ of the same length. In this paper, we limit the interleaving to the present frame $\mathbf{x}_t$ in order not to violate a possible delay constraint required in audio or video communication.

After interleaving, a convolutional channel encoder of rate $r_{\mathrm{CC}} = N_X/N_E$ encodes $\mathbf{x}'_t$ to $\mathbf{y}_t = (y_{t,1}, \ldots, y_{t,\eta}, \ldots, y_{t,N_E})$ consisting of N_E bipolar bits $y_{t,\eta} \in \{\pm 1\}$. In Turbo-like systems designed for iterative decoding, the rate of the (inner) channel code can be $r_{\mathrm{CC}} = 1$ [20]. We restrict our considerations to convolutional codes, as, according to [20], [25], the inner component of a capacity-achieving serially concatenated system should be a *recursive* convolutional code of rate $r_{\mathrm{CC}} = 1$.

On the channel, the bipolar symbols of $\mathbf{y}_t$ (with symbol energy $E_{\mathrm{s}} = 1$) are subject to additive white Gaussian noise (AWGN) with known variance $\sigma_n^2 = N_0/2$. After transmitting the bipolar values over the channel, a vector of noisy values $\mathbf{z}_t = (z_{t,1}, \ldots, z_{t,N_E}) = \mathbf{y}_t + \mathbf{n}_t$ is received.

The aim of ISCD is to jointly exploit the channel-related knowledge, the artificial channel coding redundancy, the artificial redundancy possibly introduced by a redundant bit mapping as well as the natural residual source redundancy for approximating the *a posteriori* probabilities $P(i_{t,k}|\mathbf{z}_t, \mathbf{z}_{t-1}, \ldots)$ that are used to estimate the source codec parameters. For the attainment of this aim, a channel decoder and a *Soft Decision Source Decoder* (SDSD) iteratively exchange extrinsic information in a Turbo-like process [12], [14]. The SDSD equations taking into account either inter-frame (AK1-INTER) or intra-frame correlation (AK1-INTRA) are summarized in App. A.

After a fixed number Ω of receiver iterations, the estimated *a posteriori* probabilities are used to reconstruct the quantizer reproduction values $\hat{v}_{t,k}$ using MAP estimation by $\hat{v}_{t,k} = \mathcal{Q}^{-1}(\hat{i}_{t,k}) = \bar{v}^{(\hat{i}_{t,k})}$, with [4]

$$\hat{i}_{t,k} = \underset{\forall q \in \mathbb{I}}{\arg\max}\, P(I_{t,k} = q|\mathbf{z}_t, \mathbf{z}_{t-1}, \ldots). \qquad (2)$$

Finally, the estimated source parameter vector $\hat{\mathbf{u}}_t$ is obtained by concatenating all the estimated values $\hat{v}_{t,k}$, i.e., $\hat{\mathbf{u}}_t = (\hat{v}_{t,1} \cdots \hat{v}_{t,N_I}) = (\hat{u}_{t,1} \cdots \hat{u}_{t,\kappa} \cdots \hat{u}_{t,N_I})$. Using $\hat{\mathbf{u}}_t$, the signal synthesis stage of the source decoder can reconstruct the audio-visual source signal $\hat{\mathbf{s}}_t$. As we assume that the goal of the ISCD stage is to perfectly reconstruct the quantizer code book indices $\hat{i}_{t,k}$, we mostly utilize the *symbol error rate* (SER) $P(\hat{I} \neq I)$ as performance assessment measure.

III. Error Floor Reduction by Distance Optimized Bit Mappings and Irregular Channel Codes

The performance of the ISCD system with different soft decision source decoders shall be compared in simulation examples. In order to show reproducible results, the source codec parameters are modeled by a Gauss-Markov source generating either inter-frame correlated sets of parameters (correlation coefficient $\rho = 0.9$) or intra-frame correlated parameters (correlation coefficient $\delta = 0.9$). Such high correlation coefficients frequently occur in speech and audio codecs [4], [26]. The source emits $N_I = 250$ parameters per frame which are quantized by a $Q = 16$ level scalar *Lloyd-Max Quantizer* (LMQ). Bit patterns of $M = 8$ bits, $\forall k \in \{1, \ldots, N_I\}$ are assigned to the quantizer indices according to a repetition coded bit mapping. The generator matrix $\mathbf{G}^{[\mathrm{BM}]}$ generates a repetition code, i.e., $\mathbf{G}^{[\mathrm{BM}]} = (\mathbf{I}_4\ \mathbf{I}_4)$ with $\mathbf{I}_4$ denoting the 4×4 identity matrix. This repetition coded bit mapping has been found to result in a very good performance for $\rho = 0.9$ and AK1-INTER source decoding [27] and is in fact the optimal redundant bit mapping (in terms of decoding convergence) that is obtained by a systematic generator matrix $\mathbf{G}^{[\mathrm{BM}]}$ for the given setup (i.e., source properties, quantizer, and utilized channel code). After S-random interleaving with $\mathsf{S} = 15$ [28], a rate $r_{\mathrm{CC}} = 1$ recursive convolutional code with memory $J = 3$ (J denotes the number of memory elements of the shift register) and generator polynomial $\mathbb{G}^{[\mathrm{CC}]}(D) = 1/(1 + D + D^2 + D^3)$ is applied.

In the first publications on ISCD, usually non-redundant bit mappings were used together with rate $r_{\mathrm{CC}} < 1$ convolutional codes. However, it has been found in [21], [29], [30] that a redundant bit mapping can lead to significant improvements. Utilizing redundant bit mappings together with $r_{\mathrm{CC}} = 1$ codes complies with the design rules for serially concatenated codes given in [20], [31]: The inner code shall be a recursive convolutional code (see also [25]) of rate $r_{\mathrm{CC}} = 1$ and the outer code shall have a minimum Hamming distance ≥ 2. The repetition coded bit mapping realizes this constraint on the Hamming distance [27]. To the best of our knowledge, many publications focusing on the design of (redundant) bit mappings for ISCD aim at a good convergence behavior and often neglect the error floor performance. An error floor analysis in the VLC case (and also applied to FLCs) is given for example in [24].

The dashed lines in the simulation results shown in Fig. 2-a) show that this system setup leads to a relatively high error floor in terms of SER, although having a very good waterfall behavior. Note that the residual source correlation can be interpreted as an additional code of rate $r_{\mathrm{SC}}^{[\mathrm{Markov}]}$, such that the theoretical performance limit is shifted towards lower values of E_{s}/N_0 (or E_{b}/N_0 respectively) [32]. The error floor is due to the properties of the (serially concatenated) ISCD transmitter: It is mainly characterized by the interleaver size N_X and the *Input Output Weight Enumerating Function* (IOWEF) of the component codes [31] and thus also closely related to the minimum Hamming distance $d_{\min}$ of the codes.

One approach to lower the error floor is to utilize a different outer code with better distance properties (as the repetition code only exhibits $d_{\min} = 2$). For the given setup $M^{[\mathrm{NB}]} = 4$, $r_{\mathrm{BM}} = \frac{1}{2}$, i.e., $M = 8$, there exists only a single linear block code of maximum Hamming distance $d_{\min} = 4$ with generator matrix $\mathbf{G}^{[\mathrm{BM}]} = (\mathbf{I}_4\ \ \mathbf{1}_{4\times 4} - \mathbf{I}_4)$ with $\mathbf{1}_{4\times 4}$ denoting the 4×4 all-1 matrix [33]. A similar matrix (with several systematic repetitions) has also been used in [34].

Figure 2 shows the comparison between both mappings. By increasing the distance by a factor of 2, the error floor can be

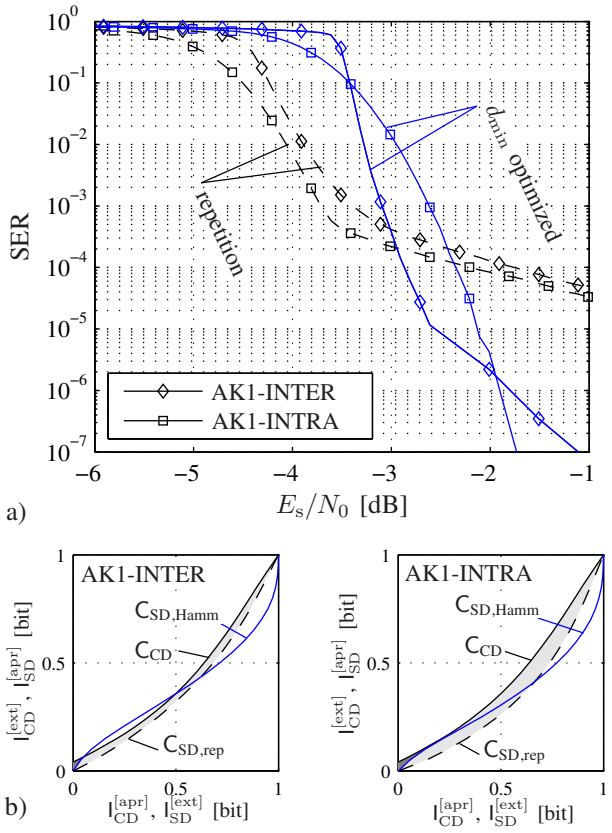

Fig. 2. Comparison of the effect of different block code based redundant bit mappings with different source decoders in ISCD with $\rho = 0.9$ or $\delta = 0.9$, LMQ scalar quantization, $N_I = 250$, $r_{\text{BM}} = \frac{1}{2}$ bit mapping, 8-state conv. code ($\mathbb{G}^{[\text{CC}]}(D) = 1/(1 + D + D^2 + D^3)$), $\Omega = 25$.
a) Symbol error rate.
b) EXIT charts at $E_s/N_0 = -4\,\text{dB}$.

significantly lowered. In fact, a union bound analysis according to [31] confirms that the error floor could be decreased by more than two decades (the SER is decreased by a factor larger than 220) and that the error floor in the case of the optimized bit mapping has a steeper slope. This analysis has been conducted by comparing the error rate union bounds of the corresponding serially concatenated channel codes consisting of a block code as outer code and the convolutional code as inner code, assuming a uniform interleaver [31]. Details on the error floor and an asymptotic interleaver gain analysis in the context of ISCD with VLCs and FLCs can be found in [24]. The simulation results depicted in Fig. 2 confirm the expected results: The error floor is considerably decreased and the slope of the floor is higher, leading to considerable gains if low SERs are targeted. Astonishingly, we have not found an error floor in the AK1-INTRA case for SERs $< 10^{-7}$ in our simulations. Although a floor seems to manifest itself at $E_s/N_0 \approx -2\,\text{dB}$, the SER curve quickly drops afterwards. The simulation has been carried out using a total number of $2 \cdot 10^8$ frames consisting of $N_I = 250$ indices per E_s/N_0 value.

The drawback of this distance-optimized bit mapping is that the waterfall has been shifted by 0.7 - 1 dB towards higher values of E_s/N_0. This can be explained by the EXIT charts

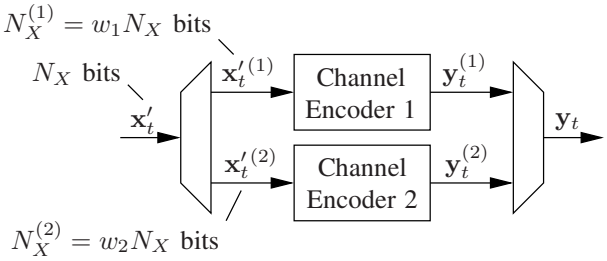

Fig. 3. Base band model for the transmitter with an irregular channel code.

depicted in Fig. 2-b) (for the AK1-INTER and AK1-INTRA cases). While the characteristic of the repetition code $C_{\text{SD,rep}}$ and the channel decoder characteristic C_{CD} have well matching shapes, the Hamming distance optimized code exhibits a characteristic $C_{\text{SD,Hamm}}$ with a distinct curvature.

One way to overcome this problem is to search for a different rate $r_{\text{CC}} = 1$ inner recursive convolutional code. The concept of *irregular bit mappings*, introduced in [22], [35], is not applied as its usage would imply a lower $d_{\min}$. In the irregular bit mapping case, high rate mappings have to be used together with low rate mappings. The high rate mappings naturally have a lower $d_{\min}$ and therefore lower the effective distance. If a bit mapping with $d_{\min} = 4$ should be retained, a different inner code has to be used. An elegant way to design a very well matching inner code is to apply the concept of irregular codes [36] to the inner code [23], [37].

The irregular inner codes presented in [23] are based on *Randomly Punctured Recursive Systematic Convolutional* (RPRSC) codes. In this example, we show how to use simple non-punctured codes of rate $r_{\text{CC}} = 1$. The codes shall be selected such that the overall decoding complexity of the channel decoder is reduced. Maunder *et al.* also employ non-punctured rate $r_{\text{CC}} = 1$ codes in [37]. Due to the curved shape of the SDSD EXIT characteristic, it can be challenging to find well matching rate-1 recursive convolutional codes.

In order to demonstrate that a system with irregular codes does not need to be complex, we show a very simple system employing an inner irregular code consisting of $M_{\mathcal{E}} = 2$ convolutional codes. This encoder is depicted in the base band model of Fig. 3. The interleaved bit vector $\mathbf{x}'_t$ is partitioned into $\mathbf{x}'^{(1)}_t$ and $\mathbf{x}'^{(2)}_t$ according to $\mathbf{x}'_t = (\mathbf{x}'^{(1)}_t, \mathbf{x}'^{(2)}_t)$. The irregular inner encoder then encodes the sub-vector $\mathbf{x}'^{(1)}_t$ consisting of $N_X^{(1)} \doteq w_1 N_X$ bits with the first code and the sub-vector $\mathbf{x}'^{(2)}_t$ consisting of $N_X^{(2)} \doteq w_2 N_X$ bits with the second code. The goal of the inner code optimization is to find a weight vector $\mathbf{w} = (w_1, w_2)^T$. In this example, the first convolutional code is a rate $r_{\text{CC}} = 1$ *Recursive Non-Systematic Convolutional* (RNSC) code with memory $J = 2$ and octal generator $\mathbb{G}^{[\text{CC}]}(D) = (1 + D + D^2)/(1 + D)$, while the second RNSC code of rate $r_{\text{CC}} = 1$ has $J = 1$ and $\mathbb{G}^{[\text{CC}]}(D) = 1/(1 + D)$ (accumulator). Both codes can be represented by a trellis diagram with either 4 or 2 states, respectively. This means that the (channel decoding) complexity is more than halved compared to the $J = 3$ code previously used. In the case of two component codes, the

TABLE I
OPTIMUM WEIGHTS $\mathbf{w}_{\text{OPT}}$ FOR THE PROPOSED IRREGULAR CODE WITH $M_{\mathcal{E}} = 2$ COMPONENT CODES WITH $\mathbb{G}_1^{[\text{CC}]}(D) = (1 + D + D^2)/(1 + D)$ AND $\mathbb{G}_2^{[\text{CC}]}(D) = 1/(1 + D)$ FOR BOTH SDSD ALGORITHMS (CORRELATION $\rho = 0.9$ OR $\delta = 0.9$).

SDSD Algorithm	Optimization E_s/N_0	w_1	w_2
AK1-INTER	-4.2 dB	0.574	0.426
AK1-INTRA	-4.9 dB	0.489	0.511

optimization problem can be formulated as [23], [36]

$$\mathbf{w}_{\text{opt}} = \arg\min_{\mathbf{w}} \left\| \mathbf{C}_{\text{CD}} \cdot \mathbf{w} - \mathbf{c}_{\text{SD,inv}} \right\|_2 \tag{3}$$

subject to

$$\mathbf{C}_{\text{CD}} \cdot \mathbf{w} > \mathbf{c}_{\text{SD,inv}} + \mathbf{o}\,, \tag{4}$$

$$w_1 + w_2 = 1, \tag{5}$$

$$0 \le w_j \le 1, \qquad \forall j \in \{1, \ldots, M_{\mathcal{E}} = 2\}\,. \tag{6}$$

The matrix $\mathbf{C}_{\text{CD}}$ contains Ξ sample points of both channel decoder characteristics and $\mathbf{c}_{\text{SD,inv}}$ contains Ξ sample points of the inverse SDSD characteristic $\mathsf{C}_{\text{SD}}^{-1}$. The offset vector $\mathbf{o}$ can control the width of the decoding tunnel. Constraint (4) ensures an open decoding tunnel while constraints (5) and (6) guarantee the validity of the weights $\mathbf{w} = (w_1 \;\; w_2)^T$. As all component codes are of identical rate $r_{\text{CC}} = 1$, no additional rate constraint needs to be considered. For a detailed description of irregular inner codes, the reader is referred to [23].

The optimization yields optimum weights $\mathbf{w}_{\text{opt}}$ which are summarized in Table I for the four distinct SDSD algorithms. Note that the optimization is performed at different channel qualities, depending on the algorithm used. Figure 4 depicts the simulation results. Significant gains are observed compared to the system employing the Hamming distance optimized bit mapping and the regular $J = 3$ code. The very good waterfall performance of the reference system with repetition code according to [27] is almost obtained, however, the error floor is significantly reduced. Furthermore, the algorithmic complexity of the channel decoder is also reduced, due to the smaller number of trellis states. Note that although the error floor is reduced compared to the first system employing the bit mapping with the repetition code, the error floor region starts at higher SER values compared to the system with optimized bit mapping and the regular $J = 3$ convolutional code (SER of $\approx 10^{-4}$ compared to $\approx 10^{-5}$ for the AK1-INTER case). However, due to the earlier convergence, this defect is compensated for, such that in the AK1-INTER case, a better SER performance is obtained in all observed channel conditions with the system employing the irregular inner code. In the AK1-INTRA case, an error floor can now be observed, contrary to the case with the optimized bit mapping and the regular convolutional code. However, the performance is still better in this case for an SER $\ge 2 \cdot 10^{-7}$. The slightly different error floor behavior may be due to the worse distance properties of the irregular convolutional code compared to the regular $J = 3$ code. A detailed analysis based on the concept of uniform interleaving, taking into account the IOWEF of outer and inner code, could be performed [24], [31] in order to get more insights. Such an analysis is however not within the scope of this paper.

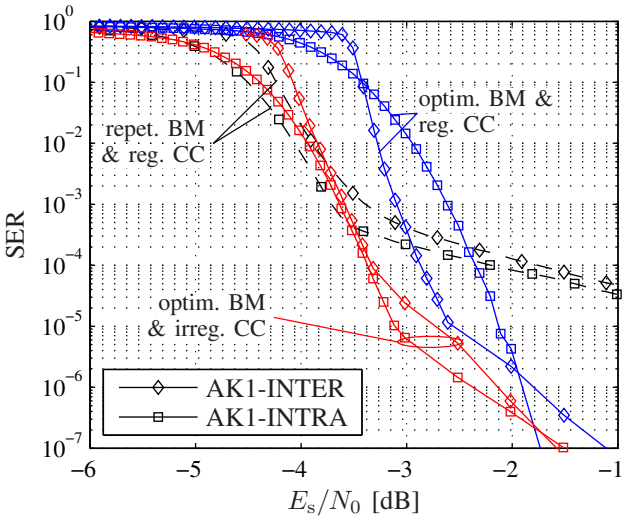

Fig. 4. Comparison of a regular channel code ($J = 3$ and $\mathbb{G}^{[\text{CC}]}(D) = 1/(1+D+D^2+D^3)$) and an irregular code built from $M_{\mathcal{E}} = 2$ convolutional codes, $\Omega = 25$, MAP estimation, other simulation settings as in Fig. 2.
Dashed black (– – –): repetition code and regular channel code
Solid blue (——): Hamming distance optimized, regular convolutional code
Solid red (——): Hamming distance optimized, irregular convolutional code.

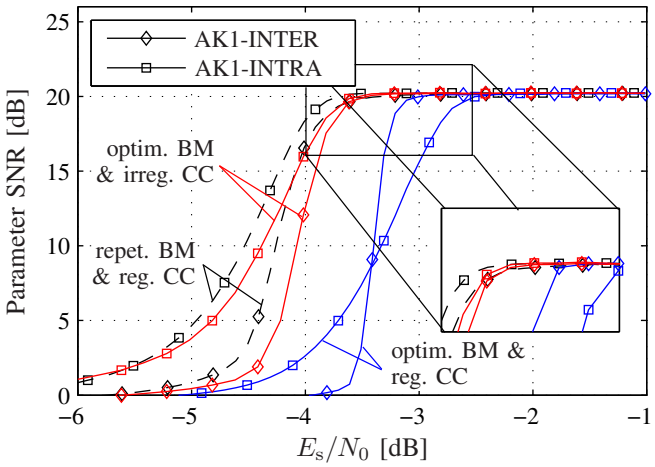

Fig. 5. Reconstruction SNR performance for the setup utilized in Fig. 4
Dashed black (– – –): repetition code and regular channel code
Solid blue (——): Hamming distance optimized, regular convolutional code
Solid red (——): Hamming distance optimized, irregular convolutional code.
8-state conv. code ($\mathbb{G}^{[\text{CC}]}(D) = 1/(1 + D + D^2 + D^3)$), $\Omega = 25$.

Figure 5 additionally shows the reconstruction parameter SNR $\mathsf{E}\{U^2\}/\mathsf{E}\{(U - \hat{U})^2\}$, i.e., the SNR between original and reconstructed source codec parameters, for all three setups depicted in Fig. 4. The same behavior as in Fig. 4 is observed. The maximum attainable SNR is given by the quantizer and amounts to 20.22 dB for the given settings. Note that there is no straightforward relation between SER and reconstruction SNR, if MAP estimation according to (2) is performed. A single symbol error can either have no influence at all on the parameter SNR (if the original parameter had a value corresponding to the quantizer decision boundary and the symbol error leads to the quantizer reproduction level with

the same distortion) or a very high influence (if the distortion between original and erroneously reconstructed parameter is maximized). If the symbol error rate is small enough, the reconstruction SNR difference is however almost not noticeable. Our presented approach with bit mappings based on linear block codes does not aim at maximizing the reconstruction SNR in the error floor. Such an optimization should map adjacent quantizer code book entries to code words with low Hamming distance, as these are more likely in error, taking also into account the (possibly unequal) parameter distribution and the parameter correlation. Such an optimization is not within the scope of this paper, where we aim at minimizing the symbol error rate, which is necessary if, for instance, source codecs are employed that require a certain residual SER or even the error-free reconstruction of certain parameters. Furthermore, the utilized concept based on block coded bit mappings permits one to implement an easy stopping criterion [22]. Approaches optimizing the bit mapping with respect to the reconstruction SNR are presented, e.g., in [38], [39]. Further note that if a maximization of the reconstruction SNR is targeted, the MAP estimation of the parameters according to (2) should be replaced by an MMSE estimation, maximizing the reconstruction SNR. The MMSE estimation rule is given by [4]

$$\hat{v}_{t,k} = \sum_{q=1}^{Q} \bar{v}^{(q)} \cdot P(I_{t,k} = q | \mathbf{z}_t, \mathbf{z}_{t-1}, \ldots) \,. \tag{7}$$

In some applications, the SDSD EXIT characteristics may vary from frame to frame due to the instationarity of the source, such that a re-optimization of the channel code is required on a frame-by-frame basis. In [36], an efficient realization of the optimization has been presented. However, even the simple basic approach of [36] is sufficient, if only two codes are used. The optimum vector $\mathbf{w}'$ fulfilling constraint (5) is found by [36]

$$\mathbf{w}' = \mathbf{w}_0 - \frac{\mathbf{1}_{1\times 2}\mathbf{w}_0 - 1}{\mathbf{1}_{1\times 2}\left(\mathbf{C}_{\text{CD}}^T\mathbf{C}_{\text{CD}}\right)^{-1}\mathbf{1}_{2\times 1}} \left(\mathbf{C}_{\text{CD}}^T\mathbf{C}_{\text{CD}}\right)^{-1}\mathbf{1}_{2\times 1} \tag{8}$$

with

$$\mathbf{w}_0 = \left(\mathbf{C}_{\text{CD}}^T\mathbf{C}_{\text{CD}}\right)^{-1}\mathbf{C}_{\text{CD}}^T\mathbf{c}_{\text{SD,inv}} \tag{9}$$

denoting the optimum solution of the unconstrained problem (3). $\mathbf{1}_{1\times 2}$ denotes the two-dimensional row vector containing all ones. Note that only matrices of maximum size 2×2 need to be inverted in (8) and (9) thus leading to a very low computational complexity of the optimization.

Using $\mathbf{w}'$, which only fulfills the constraint (5), the optimum solution can be found by a steepest descent approach [40] by exploiting the fact that $\frac{\partial}{\partial \mathbf{w}}\|\mathbf{C}_{\text{CD}}\mathbf{w} - \mathbf{c}_{\text{SD,inv}}\|_2 = 2(\mathbf{C}_{\text{CD}}\mathbf{w} - \mathbf{c}_{\text{SD,inv}})$ [36]. The steepest descent approach is an iterative algorithm, where in each iteration the gradient (multiplied with a small enough step size) is subtracted from $\mathbf{w}'$. In the present case, the gradient $\frac{\partial}{\partial \mathbf{w}}\|\mathbf{C}_{\text{CD}}\mathbf{w} - \mathbf{c}_{\text{SD,inv}}\|_2$ as well as the gradients corresponding to the (convex) constraint (6) need to be considered. The resulting weights are subsequently used as $\mathbf{w}_0$ in (8) to compute a new $\mathbf{w}'$ for the next iteration. The optimum vector $\mathbf{w}_{\text{opt}}$ is either obtained after a fixed number of iterations or if the frame-to-frame variation of $\|\mathbf{C}_{\text{CD}}\mathbf{w} - \mathbf{c}_{\text{SD,inv}}\|_2$ is sufficiently small.

IV. Error Floor Reduction by Multi-Dimensional Bit Mappings

Increasing the Hamming distance of the redundant bit mapping is not always possible, especially if the number of quantization levels Q is small. The Hamming distance of the bit mapping is upper bounded by the Singleton bound with $d_{\min} \leq M - \operatorname{ld} Q + 1$ (see, e.g., [41]). If $\operatorname{ld} Q = M^{[\text{NB}]}$, then the Singleton bound can also be expressed as $d_{\min} \leq M - M^{[\text{NB}]} + 1 = 1 + M^{[\text{NB}]}(\frac{1}{r_{\text{BM}}} - 1)$. For example, if $Q = 4$ ($M^{[\text{NB}]} = 2$) and $r_{\text{BM}} = 1/2$, $d_{\min}$ is bounded by $d_{\min} \leq 3$. However, linear block codes over $\mathbb{F}_2$ of size 2×4 only achieve $d_{\min} \leq 2$.

In order to achieve higher distances $d_{\min}$ without modifying the rate r_{BM} and Q, a different approach has to be considered. Following the Singleton bound with constant r_{BM}, increasing $d_{\min}$ results in larger codes. Unfortunately, the size of the code is limited by Q. In order to be able to utilize larger codes, *Multi-Dimensional Bit Mappings* (MDBMs), which group several quantizer indices to one *super index* and utilize a larger code as redundant bit mapping, can be used. This larger code allows a larger Hamming distance $d_{\min}$ and thus results in a lower error floor.

The MDBM approach is similar to the concept of *multi-dimensional mappings*, originally considered for trellis-coded modulation [42] and later applied among others to *Bit-Interleaved Coded Modulation with Iterative Decoding* (BICM-ID) [43], [44]. The application of MDBMs is straightforward at the transmitter side: Ψ consecutive quantizer indices $i_{t,k}, \ldots, i_{t,k+\Psi-1}$ are grouped to a super index $i^{\star}_{t,k'}$, $k' = \lceil k/\Psi \rceil$ with

$$i^{\star}_{t,k'} = 1 + \sum_{j=1}^{\Psi} \left(i_{t,(k'-1)\Psi+j} - 1\right) Q^{\Psi-j} \tag{10}$$

and $i^{\star}_{t,k'} \in \mathbb{I}^{\star} \doteq \{1, \ldots, Q^{\Psi}\}$. A frame then contains a total number of $N_I^{\star} \doteq \lceil N_I/\Psi \rceil$ super indices. To the super index $i^{\star}_{t,k'}$, the natural binary representation $\mathbf{b}^{[\text{NB}]\star}_{t,k'}$ consisting of $M^{[\text{NB}]\star} \doteq \Psi \cdot M^{[\text{NB}]}$ bits is assigned, followed by the (possibly) redundant part of the bit mapping. After bit mapping, $N_I^{\star}$ super bit patterns $\mathbf{b}^{\star}_{t,k'}$, $k' \in \{1, \ldots, N_I^{\star}\}$ of $M^{\star}$ bits result. As in the one-dimensional case, $\mathbf{b}^{\star}_{t,k'} \in \mathbb{B}^{\star} \subseteq \mathbb{F}_2^{M^{\star}}$, with $\mathbb{B}^{\star} \doteq \{\bar{\mathbf{b}}^{(1)\star}, \ldots, \bar{\mathbf{b}}^{(Q^{\Psi})\star}\}$. Because the rate r_{BM} of the bit mapping is constrained to a fixed value, the condition $N_I M = N_I^{\star} M^{\star}$ has to hold.

Thus, choosing MDBMs with dimensions $\Psi > 1$ permits the use of a longer code, and thus to (theoretically) increase the Hamming distance of the outer component code in the serially concatenated system [41].

Finally note that using MDBMs together with scalar quantizers ($\Pi = 1$) can also be interpreted as using $\Pi = \Psi$ dimensional vector quantization together with (one-dimensional) bit mappings. The (separable) code book of this resulting vector quantizer is composed of the original scalar code book in each dimension. However, as it is often not possible to change the

quantizer (as the source codec is given and fixed), MDBMs have to be used for achieving higher distances $d_{\min}$ of the outer code.

A. Receiver Modifications

In order to be able to decode MDBMs, the SDSD has to be modified accordingly. The different kinds of decoding algorithms have to be considered separately, however, the first steps are common to all four algorithms. The factors $\theta_{t,k}(q)$ and $\theta_{t,k}^{[\text{ext}]\setminus m}(q)$ given by (15) and (16) have to be replaced by $\theta_{t,k'}^{\star}(q^{\star})$ and $\theta_{t,k'}^{[\text{ext}]\setminus m,\star}(q^{\star})$, according to

$$\theta_{t,k'}^{\star}(q^{\star}) \doteq \prod_{\mu=1}^{M^{\star}} P_{\text{CD}}^{[\text{ext}]}\left(B_{t,k',\mu}^{\star} = \bar{b}_{k',\mu}^{(q^{\star})\star}\right) \tag{11}$$

$$\theta_{t,k'}^{[\text{ext}]\setminus\chi,\star}(q^{\star}) \doteq \prod_{\substack{\mu=1\\ \mu\neq\chi}}^{M^{\star}} P_{\text{CD}}^{[\text{ext}]}\left(B_{t,k',\mu}^{\star} = \bar{b}_{k',\mu}^{(q^{\star})\star}\right). \tag{12}$$

Equation (11) has to be evaluated for each distinct $q^{\star} \in \{1,\ldots,Q^{\Psi}\}$. Equation (12) has to be evaluated for each distinct pair of $(q^{\star},\chi)$, with $\chi \in \{1,\ldots,M^{\star}\}$.

Unfortunately, the usage of MDBMs increases the complexity of SDSD. If $N_I^{\star} = \lceil N_I/\Psi\rceil = N_I/\Psi$, a total number of $N_I^{\star}Q^{\Psi}(M^{\star})^2 = N_I\Psi Q^{\Psi}M^2$ multiplications are required for computing all $\theta_{t,k}^{\star}(q^{\star})$ and $\theta_{t,k}^{[\text{ext}]\setminus\chi,\star}(q^{\star})$ of a frame. The standard SDSD only requires $N_I Q M^2$ multiplications. This corresponds to an increase of the number of multiplications per frame by a factor of $\Psi Q^{\Psi-1}$ common to all MDBM decoders. As the complexity increase grows exponentially with Ψ, the selection of a small dimension Ψ is suggested in order to keep the receiver moderate in its complexity. Note that the aforementioned (and the following) complexity considerations are only an indication of the complexity increase. A real implementation would be performed in the logarithmic domain, thus requiring additions instead of multiplications for instance. The re-utilization of intermediate results in, e.g., (11) and (12) also permits to slightly reduce the number of operations.

In what follows, the different flavors of the MDBM-SDSD are explained in detail. The corresponding MDBM-SDSD variables are also annotated by a superscript * for a better differentiation.

AK1-INTER: In the case of inter-frame correlation, the different indices within a frame are assumed to be statistically independent. This allows to write the conditional inter-frame *a priori* probabilities of the super indices as

$$P(I_{t,k'}^{\star} = q^{\star} | I_{t-1,k'}^{\star} = \tilde{q}^{\star}) = \prod_{j=1}^{\Psi} P(I_{t,\Psi(k'-1)+j} = q_j | I_{t-1,\Psi(k'-1)+j} = \tilde{q}_j) \tag{13}$$

with $q^{\star} = 1 + \sum_{j=1}^{\Psi}(q_j - 1)Q^{\Psi-j}$ and $\tilde{q}^{\star} = 1 + \sum_{j=1}^{\Psi}(\tilde{q}_j - 1)Q^{\Psi-j}$. In this case, the SDSD complexity is also increased. The factors $A_{t,k}^{\star}(q^{\star})$ (intermediate result of the temporal forward recursion) require $Q^{2\Psi}$ MAC operations while the computation of the temporal forward recursion $\alpha_{t,k}^{\star}(q^{\star})$ requires Q^{Ψ} multiplications per super index. However, the operations only need to be carried out $N_I^{\star} = N_I/\Psi$ times per frame. Thus, in the forward recursion, the number of MAC operations increases by a factor of $\frac{1}{\Psi}Q^{2(\Psi-1)}$ while the number of multiplications increases by a factor of $\frac{1}{\Psi}Q^{(\Psi-1)}$. The number of MAC operations per bit required for computing the extrinsic information increases from Q to Q^{Ψ}.

AK1-INTRA: In contrast to the AK0 and AK1-INTER cases, the different indices within a frame are not independent anymore if intra-frame correlation is exploited. In this case, it is shown in App. B that the *a priori* probabilities on super index level can be expressed as

$$P(I_{t,k'}^{\star} = q^{\star} | I_{t,k'-1}^{\star} = \tilde{q}^{\star}) = P(I_{t,\Psi(k'-1)+1} = q_1 | I_{t,\Psi(k'-2)+\Psi} = \tilde{q}_{\Psi}) \cdot \times \prod_{j=2}^{\Psi} P(I_{t,\Psi(k'-1)+j} = q_j | I_{t,\Psi(k'-1)+j-1} = q_{j-1}) \tag{14}$$

with $q^{\star} = 1 + \sum_{j=1}^{\Psi}(q_j - 1)Q^{\Psi-j}$ and $\tilde{q}^{\star} = 1 + \sum_{j=1}^{\Psi}(\tilde{q}_j - 1)Q^{\Psi-j}$.

The complexity increase in this case is again comparable to the AK1-INTER case as the forward recursion is basically identical and there is an additional backward recursion whose complexity scales by the same factor as the complexity of the forward recursion (see AK1-INTER case). For the computation of extrinsic information, Q^{Ψ} multiplications (instead of Q) are necessary in addition to the Q^{Ψ} MACs.

The implication of the complexity considerations beforehand is that the computational complexity of MDBM roughly increases by $Q^{\Psi-1}$. For this reason, the dimension Ψ (with $\Psi \geq 2$) should be kept as small as possible. Already with $\Psi = 2$, a considerable decrease of the error floor due to the higher $d_{\min}$ (in the case of a well-chosen redundant bit mapping) can be observed, as will be shown below by a simulation example.

Note that although the complexity is increased by MDBMs, effective methods to reduce the overall SDSD complexity exist, such as a reduction of the state transitions [45] or an efficient implementation based on a fast matrix search technique that allows to implement the algorithm using complexity of order $O(Q)$ (or $O(Q^{\Psi})$ in the case of MDBMs) if the residual source correlation fulfills certain requirements [46].

B. Simulation Example

The capabilities of MDBMs shall be demonstrated by a simulation example. The source setup of Section III is used with $Q = 4$ level scalar LMQ. The overall coding rate of the bit mapping shall be $r_{\text{BM}} = 1/2$. In the one-dimensional case, the number of linear block codes with Hamming distance $d_{\min} \geq 2$ is limited due to the small dimension of the generator matrix (here 2×4). For the example, we use either the repetition code (i.e., $\mathbf{G}^{[\text{BM}]} = (\mathbf{I}_2\ \mathbf{I}_2)$, $\forall k \in \{1,\ldots,N_I\}$) or the multiple parity check code (i.e., $\mathbf{G}^{[\text{BM}]} = (\mathbf{I}_2\ \mathbf{1}_{2\times 2})$, $\forall k \in \{1,\ldots,N_I\}$). The minimum distance of these codes is $d_{\min} = 2$. In fact, for the given settings, only systematic linear codes with $d_{\min} = 2$ can be found.

In order to increase the possible Hamming distance of the bit mapping, a multi-dimensional bit mapping with $\Psi = 2$

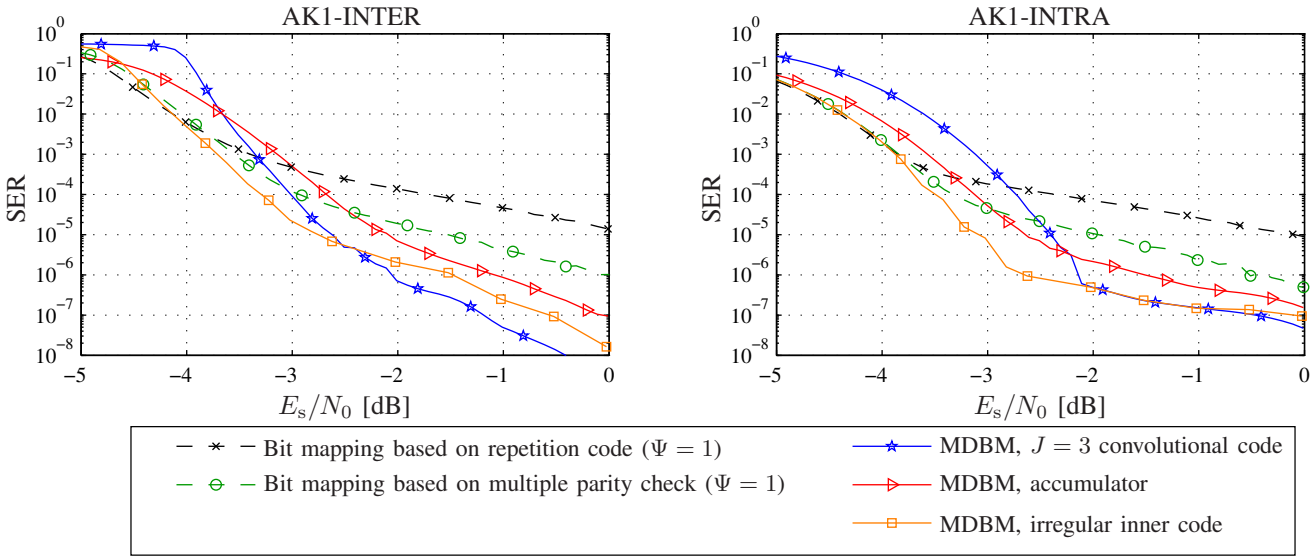

Fig. 6. Comparison of different source decoders with and without MDBMs with $\rho = 0.9$, or $\delta = 0.9$, scalar LMQ, $Q = 4$, $N_I = 250$, $r_{\text{BM}} = \frac{1}{2}$ bit mapping, $J = 3$, $r_{\text{CC}} = 1$ convolutional code ($\mathbb{G}^{[\text{CC}]}(D) = 1/(1 + D + D^2 + D^3)$) or $J = 1$, $r_{\text{CC}} = 1$ convolutional code ($\mathbb{G}^{[\text{CC}]}(D) = 1/(1 + D)$), or $r_{\text{CC}} = 1$ irregular convolutional code according to Tab. II, $\Omega = 25$ iterations, MAP estimation.

TABLE II
OPTIMUM WEIGHTS $\mathbf{w}_{\text{OPT}}$ FOR THE PROPOSED IRREGULAR CODE WITH $M_{\mathcal{E}} = 2$ COMPONENT CODES WITH $\mathbb{G}_1^{[\text{CC}]}(D) = (1 + D + D^2)/(1 + D)$ AND $\mathbb{G}_2^{[\text{CC}]}(D) = 1/(1 + D)$ FOR BOTH MDBM-SDSD ALGORITHMS (CORRELATION $\rho = 0.9$ OR $\delta = 0.9$).

SDSD Algorithm	Optimization E_s/N_0	w_1	w_2
AK1-INTER	-4.75 dB	0.639	0.361
AK1-INTRA	-5.7 dB	0.581	0.419

is employed. The utilized generator matrix is the same as used in Section III, i.e., $\mathbf{G}^{[\text{BM}]\star} = (\mathbf{I}_4 \quad \mathbf{1}_{4\times 4} - \mathbf{I}_4)$. Again, a block consists of $N_I = 250$ quantizer indices ($N_I/\Psi = 125$ super indices). The utilized convolutional codes are the $J = 3$, $r_{\text{CC}} = 1$ RNSC code with $\mathbb{G}^{[\text{CC}]}(D) = 1/(1 + D + D^2 + D^3)$ and additionally in the MDBM case, a low-complexity $r_{\text{CC}} = 1$ RNSC code with $J = 1$ and $\mathbb{G}^{[\text{CC}]}(D) = 1/(1 + D)$ (accumulator). Furthermore, in order to optimize the waterfall behavior of the MDBM setup, we have utilized the same irregular inner code as in Sec. III, optimized using the EXIT chart of the respective MDBM SDSDs. The resulting weighting factors are summarized in Tab. II.

Simulation results are given in Fig. 6. It can be observed that the systems employing the one-dimensional bit mapping suffer from a high error floor, which also depends on the generator matrix used for the bit mapping. The repetition coded bit mapping suffers from the highest error floor. This is mainly due to the fact that the repetition code leads to two code words with Hamming weight two, while the multiple parity check code generates a single code word with Hamming weight two. If the dimension of the bit mapping is increased to $\Psi = 2$ and a bit mapping generator matrix with $d_{\min} = 4$ is employed, the error floor can be considerably reduced. However, as in Section III, a shift of the waterfall region towards better channel qualities is observed. This shift can be compensated by utilizing different channel codes, for instance the irregular inner code presented in Sec. III. Especially in the AK1-INTRA case, the MDBM system with the irregular inner code outperforms all other setups, while in the AK1-INTER case, a slightly elevated error floor compared to the setup with MDBMs and the $J = 3$ regular inner code, is observed.

In this Section, we have shown that the utilization of multi-dimensional bit mappings permits to the effectively lower the error-floor of ISCD, especially if small quantizer code books are employed. The resulting shift of the waterfall performance can be efficiently combated by irregular inner codes.

V. CONCLUSIONS

As several known *Iterative Source-Channel Decoding* (ISCD) systems showing good waterfall performance have the drawback of an observable error floor, several alternative methods for reducing this error floor have been studied in this paper. It has been demonstrated that the error floor can be substantially lowered by a proper selection of the bit mapping. In addition, a carefully designed low-complexity inner irregular convolutional code can overcome the degraded waterfall performance of such bit mappings. Due to its simplicity, this code can also be efficiently employed in the scenarios where a frame-by-frame adaptation is necessary.

If small quantizer code books are employed, a bit mapping leading to low error floors is often not available. In this case, we have proposed *Multi-Dimensional Bit Mappings* (MDBMs), grouping several consecutive quantizer indices to multi-dimensional super indices, thus increasing the search space for beneficial (in terms of a low error floor) bit mappings. The improved error floor performance of ISCD with MDBMs has been confirmed by simulation examples and it has been shown that the resulting shift of the waterfall region can be compensated by the use of an irregular inner code.

In practical applications, usually quantizers with large code books are utilized besides quantizers with small code books.

In this case, the application of both presented methods permits to effectively lower the overall decoding error floor.

APPENDIX A
SOFT DECISION SOURCE DECODING EQUATIONS

For a derivation of the SDSD equations in the probability domain, we refer to the literate, e.g., [12], [14], [15].

The first steps are common for both SDSD algorithms. The channel decoder computes extrinsic information for all M bits of all N_I bit patterns. In what follows, this extrinsic information of the μth bit of the kth parameter in the frame at time instant t is denoted by $P_{\mathrm{CD}}^{[\mathrm{ext}]}(b_{t,k,\mu})$. Note that in an actual implementation of ISCD, L-values [47] are usually used instead of probabilities for numerical reasons. The SDSD equations in the logarithmic domain using L-values are given, e.g., in [48].

Under the assumption of a memoryless channel, a set of reliabilities

$$\theta_{t,k}(q) \doteq \prod_{\mu=1}^{M} P_{\mathrm{CD}}^{[\mathrm{ext}]}\left(B_{t,k,\mu} = \bar{b}_{k,\mu}^{(q)}\right) \tag{15}$$

is determined for each different quantization index $q \in \mathbb{I}$ at each position k within a frame, resulting in a total number of $N_I Q$ different expressions for each frame at time instant t. Note that $\bar{b}_{k,\mu}^{(q)}$ denotes the μth entry of the bit vector $\bar{\mathbf{b}}_k^{(q)}$, i.e., the μth bit of the pattern. We further define the extrinsic channel-related reliabilities considering all bits of a bit pattern with the exception of the χth bit

$$\theta_{t,k}^{[\mathrm{ext}]\setminus\chi}(q) \doteq \prod_{\substack{\mu=1 \\ \mu\neq\chi}}^{M} P_{\mathrm{CD}}^{[\mathrm{ext}]}\left(B_{t,k,\mu} = \bar{b}_{k,\mu}^{(q)}\right). \tag{16}$$

Using the expressions in (15) and (16), the *a posteriori* probabilities and the extrinsic output of the SDSD can be computed, depending on the available source statistics.

A. Exploitation of Inter-frame Correlation (AK1-INTER)

If *a priori* information describing the dependency of quantization indices between consecutive frames is available, the temporal correlation ρ is exploited. In this case, the N_I indices are modeled by N_I stationary Markov processes $I_{t,k}$ of first order, and the *a priori* knowledge is given by the probabilities $P(I_{t,k} = i_{t,k} | I_{t-1,k} = i_{t-1,k})$.

The *a posteriori* probabilities are approximated by [4]

$$P(I_{t,k} = q | \mathbf{z}_t, \mathbf{z}_{t-1}, \ldots) = \alpha_{t,k}(q) \qquad \forall q \in \mathbb{I}, \tag{17}$$

with $\alpha_{t,k}(q)$ being evaluated in the inter-frame forward recursion

$$\begin{aligned}\alpha_{t,k}(q) &= \frac{1}{\mathrm{K}_1} \cdot \theta_{t,k}(q) \cdot \sum_{\tilde{q}=1}^{Q} P(I_{t,k}=q | I_{t-1,k}=\tilde{q}) \cdot \alpha_{t-1,k}(\tilde{q}) \\ &\doteq \frac{1}{\mathrm{K}_1} \cdot \theta_{t,k}(q) \cdot A_{t,k}(q), \end{aligned} \tag{18}$$

and K_1 being a normalization constant ensuring that $\sum_{q=1}^{Q} \alpha_{t,k}(q) = 1$. The values $\alpha_{0,k}$ (the transmission is assumed to start at time instant $t = 1$) are initialized by

$$\alpha_{0,k}(q) = P(I_{1,k} = q) \qquad \forall q \in \mathbb{I}. \tag{19}$$

As the values $\alpha_{t,k}(q)$ are reused in the subsequent frame, they need to be stored, resulting in memory requirements of $N_I \cdot Q$ values.

The extrinsic probabilities which are fed back to the channel decoder for use in the subsequent iteration are given by (for $\ell \in \mathbb{F}_2$) [14]

$$P_{\mathrm{SD}}^{[\mathrm{ext}]}(b_{t,k,\mu} = \ell) = \frac{1}{\mathrm{K}_2} \sum_{\substack{q=1 \\ \bar{b}_{k,\mu}^{(q)} = \ell}}^{Q} \theta_{t,k}^{[\mathrm{ext}]\setminus\mu}(q) \cdot A_{t,k}(q) \tag{20}$$

with $A_{t,k}(q) = \sum_{\tilde{q}=1}^{Q} P(I_{t,k}=q | I_{t,k-1}=\tilde{q}) \alpha_{t,k-1}(\tilde{q})$.

Note that in this work, we do not allow any additional delay and therefore only consider a single frame during decoding and no information from future frames is available.

B. Exploitation of Intra-frame Correlation (AK1-INTRA)

If *a priori* information describing the relation between quantization indices within one frame are available, the intra-frame correlation δ is exploited. In that case, the N_I indices in a frame emerge from a stationary Markov process of first order and the *a priori* knowledge is modeled by the probabilities $P(I_{t,k} = i_{t,k} | I_{t,k-1} = i_{t,k-1})$.

Unlike in the AK1-INTER case, the decoding algorithm can now exploit information from all neighboring positions, i.e., past and future positions $k-1$ and $k+1$, as a complete frame is processed in one pass of the SDSD. The equations for determining *a posteriori* probabilities and extrinsic information exploit a forward-backward recursion similar to the BCJR algorithm [15], [49].

The *a posteriori* probabilities $P(I_{t,k} = q | \mathbf{z}_t, \mathbf{z}_{t-1}, \ldots) = P(I_{t,k} = q | \mathbf{z}_t)$ (no inter-frame relations present) amount to

$$P(I_{t,k} = q | \mathbf{z}_t) = \frac{1}{\mathrm{K}_1} \cdot \alpha_{t,k}(q) \cdot \beta_{t,k}(q) \qquad \forall q \in \mathbb{I}, \tag{21}$$

with K_1 being a normalization constant ensuring $\sum_{q=1}^{Q} P(I_{t,k} = q | \mathbf{z}_t) = 1$ and $\alpha_{t,k}(q)$ and $\beta_{t,k}(q)$ being evaluated in the forward and backward recursions

$$\alpha_{t,k}(q) = \frac{\theta_{t,k}(q)}{\mathrm{K}_2} \sum_{\tilde{q}=1}^{Q} P(I_{t,k}=q | I_{t,k-1}=\tilde{q}) \alpha_{t,k-1}(\tilde{q}) \tag{22}$$

$$\doteq \frac{\theta_{t,k}(q)}{\mathrm{K}_2} \cdot A'_{t,k}(q) \tag{23}$$

$$\beta_{t,k-1}(q) = \frac{1}{\mathrm{K}_3} \sum_{\tilde{q}=1}^{Q} \theta_{t,k}(\tilde{q}) P(I_{t,k}=\tilde{q} | I_{t,k-1}=q) \beta_{t,k}(\tilde{q}) \tag{24}$$

with the initialization

$$\alpha_{t,0}(q) = P(I_{t,1} = q) \qquad \forall q \in \mathbb{I} \tag{25}$$

$$\beta_{t,N_I}(q) = \frac{1}{Q} \qquad \forall q \in \mathbb{I}. \tag{26}$$

The factors K_2 and K_3 in (22) and (24) are used to ensure that $\sum_{q=1}^{Q} \alpha_{t,k}(q) = 1$ and $\sum_{q=1}^{Q} \beta_{t,k}(q) = 1$, $\forall k \in \{1, \ldots, N_I\}$.

The extrinsic information is given by (for $\ell \in \mathbb{F}_2$)

$$P_{\mathrm{SD}}^{[\mathrm{ext}]}(b_{t,k,\mu} = \ell) = \frac{1}{\mathrm{K}_4} \sum_{\substack{q=1 \\ \bar{b}_{k,\mu}^{(q)} = \ell}}^{Q} \theta_{t,k}^{[\mathrm{ext}]\setminus\mu}(q) \beta_{t,k}(q) A'_{t,k}(q). \tag{27}$$

The factor $A'_{t,k}(q) = \sum_{\tilde{q}=1}^{Q} P(I_{t,k}=q|I_{t,k-1}=\tilde{q})\alpha_{t,k-1}(\tilde{q})$ is required both in (22) and (27) and can therefore be computed only once and then reused.

APPENDIX B
A PRIORI PROBABILITIES FOR AK1-INTRA MULTI-DIMENSIONAL SDSD

The probability $P(I^{\star}_{t,k'} = q^{\star}|I^{\star}_{t,k'-1} = \tilde{q}^{\star})$ is decomposed using Bayes' theorem and by exploiting the first order Markov property of the single indices, leading to

$$\begin{aligned}
&P(i^{\star}_{t,k'}|i^{\star}_{t,k'-1}) = \\
&P(i_{t,\Psi(k'-1)+1},...,i_{t,\Psi(k'-1)+\Psi}|i_{t,\Psi(k'-2)+1},...,i_{t,\Psi(k'-2)+\Psi}) = \\
&\frac{P(i_{t,\Psi(k'-1)+1},...,i_{t,\Psi(k'-1)+\Psi},i_{t,\Psi(k'-2)+1},...,i_{t,\Psi(k'-2)+\Psi})}{P(i_{t,\Psi(k'-2)+1},...,i_{t,\Psi(k'-2)+\Psi})} \\
&\quad = \frac{P(i_{t,\mathfrak{k}},...,i_{t,\mathfrak{k}-\Psi+1},i_{t,\mathfrak{k}-\Psi},...,i_{t,\mathfrak{k}-2\Psi+1})}{P(i_{t,\mathfrak{k}-\Psi},...,i_{t,\mathfrak{k}-2\Psi+1})}
\end{aligned}$$

with the index substitution $\mathfrak{k} = \Psi(k'-1)+\Psi$. By applying the chain rule of probability and by exploiting the Markov property of neighboring indices, we get

$$\begin{aligned}
&P(i^{\star}_{t,k'}|i^{\star}_{t,k'-1}) = \\
&\quad P(i_{t,\mathfrak{k}}|i_{t,\mathfrak{k}-1})\cdot P(i_{t,\mathfrak{k}-1}|i_{t,\mathfrak{k}-2})\cdots P(i_{t,\mathfrak{k}-\Psi+1}|i_{t,\mathfrak{k}-\Psi})\cdot \\
&\times \frac{P(i_{t,\mathfrak{k}-\Psi}|i_{t,\mathfrak{k}-\Psi-1})\cdots P(i_{t,\mathfrak{k}-2\Psi+2}|i_{t,\mathfrak{k}-2\Psi+1})\cdot P(i_{t,\mathfrak{k}-2\Psi+1})}{P(i_{t,\mathfrak{k}-\Psi}|i_{t,\mathfrak{k}-\Psi-1})\cdots P(i_{t,\mathfrak{k}-2\Psi+2}|i_{t,\mathfrak{k}-2\Psi+1})\cdot P(i_{t,\mathfrak{k}-2\Psi+1})} \\
&\quad = P(i_{t,\mathfrak{k}}|i_{t,\mathfrak{k}-1})\cdot P(i_{t,\mathfrak{k}-1}|i_{t,\mathfrak{k}-2})\cdots P(i_{t,\mathfrak{k}-\Psi+1}|i_{t,\mathfrak{k}-\Psi}) \\
&= P(i_{t,\Psi(k'-1)+\Psi}|i_{t,\Psi(k'-1)+\Psi-1})\cdots P(i_{t,\Psi(k'-1)+1}|i_{t,\Psi(k'-1)}) \\
&= P(i_{t,\Psi(k'-1)+\Psi}|i_{t,\Psi(k'-1)+\Psi-1})\cdots P(i_{t,\Psi(k'-1)+1}|i_{t,\Psi(k'-2)+\Psi}).
\end{aligned}$$

This means that the product of the $\Psi-1$ crossover probabilities between each of the Ψ indices contained in one super index has to be multiplied by the crossover probability between the last index of super index $i^{\star}_{t,k'-1}$ ($i_{t,\Psi(k'-2)+\Psi}$) and the first index of $i^{\star}_{t,k'}$ ($i_{t,\Psi(k'-1)+1}$). Rewriting this fact using the product notation and by distinguishing two neighboring super indices directly leads to

$$\begin{aligned}
&P(I^{\star}_{t,k'} = q^{\star}|I^{\star}_{t,k'-1} = \tilde{q}^{\star}) = \\
&\quad P(I_{t,\Psi(k'-1)+1}=q_1|I_{t,\Psi(k'-2)+\Psi}=\tilde{q}_{\Psi})\cdot \\
&\quad \times \prod_{j=2}^{\Psi} P(I_{t,\Psi(k'-1)+j}=q_j|I_{t,\Psi(k'-1)+j-1}=q_{j-1})
\end{aligned}$$

with

$$i^{\star}_{t,k'} = 1 + \sum_{j=1}^{\Psi}\left(i_{t,(k'-1)\Psi+j} - 1\right)Q^{\Psi-j}$$

and thus $q^{\star} = 1 + \sum_{j=1}^{\Psi}(q_j - 1)Q^{\Psi-j}$ and $\tilde{q}^{\star} = 1 + \sum_{j=1}^{\Psi}(\tilde{q}_j - 1)Q^{\Psi-j}$.

ACKNOWLEDGMENTS

This work was financed in part by the European Union project *FlexCode* under Grant FP6-2002-IST-C 020023-2 and in part by the German research cluster of excellence UMIC.

The authors would like to thank Dr. Marc Adrat and Dr. Thorsten Clevorn for many inspiring and fruitful discussions.

REFERENCES

[1] C. E. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, July and October 1948.

[2] C. E. Shannon, "Two-Way Communication Channels," *Berkeley Symposium on Mathematics, Statistics, and Probability*, vol. 1, 1961, pp. 611–644.

[3] K. Sayood and J. Borkenhagen, "Use of Residual Redundancy in the Design of Joint Source/Channel Coders," *IEEE Trans. Commun.*, vol. 39, no. 6, pp. 838–846, June 1991.

[4] T. Fingscheidt and P. Vary, "Softbit Speech Decoding: A New Approach to Error Concealment," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 3, pp. 240–251, Mar. 2001.

[5] K. Zeger and A. Gersho, "Pseudo-Gray Coding," *IEEE Trans. Commun.*, vol. 38, no. 12, pp. 2147–2158, Dec. 1990.

[6] N. Farvardin and V. Vaishampayan, "Optimal Quantizer Design for Noisy Channels: An Approach to Combined Source - Channel Coding," *IEEE Trans. Inf. Theory*, vol. 33, no. 6, pp. 827–838, Nov. 1987.

[7] S. Heinen and P. Vary, "Source-Optimized Channel Coding for Digital Transmission Channels," *IEEE Trans. Commun.*, vol. 53, no. 4, pp. 592–600, Apr. 2005.

[8] B. Hochwald and K. Zeger, "Tradeoff Between Source and Channel Coding," *IEEE Trans. Inf. Theory*, vol. 43, no. 5, pp. 1412–1424, Sept. 1997.

[9] J. Hagenauer, "Source-Controlled Channel Decoding," *IEEE Trans. Commun.*, vol. 43, no. 9, pp. 2449–2457, Sept. 1995.

[10] F. Alajaji, N. Phamdo, and T. Fuja, "Channel Codes that Exploit the Residual Redundancy in CELP-encoded Speech," *IEEE Trans. Speech Audio Process.*, vol. 4, no. 5, pp. 325–336, Sept. 1996.

[11] T. Fazel and T. Fuja, "Robust Transmission of MELP-compressed Speech: An Illustrative Example of Joint Source-Channel Decoding," *IEEE Trans. Commun.*, vol. 51, no. 6, pp. 973–982, June 2003.

[12] N. Görtz, "Iterative Source-Channel Decoding Using Soft-In/Soft-Out Decoders," *IEEE International Symposium on Information Theory (ISIT)*, June 2000, p. 173.

[13] T. Hindelang, T. Fingscheidt, N. Seshadri, and R. V. Cox, "Combined Source/Channel (De-)Coding: Can A Priori Information Be Used Twice?" *IEEE International Conference on Communications (ICC)*, June 2000, pp. 1208–1212.

[14] M. Adrat, P. Vary, and J. Spittka, "Iterative Source-Channel Decoder Using Extrinsic Information from Softbit-Source Decoding," *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Salt Lake City, UT, USA, May 2001.

[15] M. Adrat and P. Vary, "Iterative Source-Channel Decoding: Improved System Design Using EXIT Charts," *EURASIP Journal on Applied Signal Processing (Special Issue: Turbo Processing)*, vol. 2005, pp. 928–941, May 2005.

[16] C. Guillemot and P. Siohan, "Joint Source-Channel Decoding of Variable-Length Codes with Soft Information: A Survey," *EURASIP Journal on Applied Signal Processing (Special Issue: Turbo Processing)*, vol. 2005, pp. 906–927, May 2005.

[17] J. Kliewer and R. Thobaben, "Iterative Joint Source-Channel Decoding of Variable-Length Codes Using Residual Source Redundancy," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 919–929, May 2005.

[18] X. Jaspar, C. Guillemot, and L. Vandendorpe, "Joint Source-Channel Turbo Techniques for Discrete-Valued Sources: From Theory to Practice," *Proc. IEEE*, vol. 95, no. 6, pp. 1345–1361, June 2007.

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

[20] A. Ashikhmin, G. Kramer, and S. ten Brink, "Extrinsic Information Transfer Functions: Model and Erasure Channel Properties," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2657–2673, Nov. 2004.

[21] J. Kliewer, N. Görtz, and A. Mertins, "Iterative Source-Channel Decoding With Markov Random Field Source Models," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3688–3701, Oct. 2006.

[22] L. Schmalen, P. Vary, T. Clevorn, and B. Schotsch, "Efficient Iterative Source-Channel Decoding Using Irregular Index Assignments," *International ITG Conference on Source and Channel Coding (SCC)*, Ulm, Germany, Jan. 2008.

[23] R. Thobaben, L. Schmalen, and P. Vary, "Joint Source-Channel Coding with Inner Irregular Codes," *IEEE International Symposium on Information Theory (ISIT)*, Toronto, Canada, July 2008, pp. 1153–1157.

[24] X. Jaspar and L. Vandendorpe, "Joint Source-Channel Codes Based on Irregular Turbo Codes and Variable Length Codes," *IEEE Trans. Commun.*, vol. 56, no. 11, pp. 1824–1835, Nov. 2008.

[25] J. Kliewer, A. Huebner, and D. J. Costello, Jr., "On the Achievable Extrinsic Information of Inner Decoders in Serial Concatenation," *IEEE International Symposium on Information Theory (ISIT)*, July 2006, pp. 2680–2684.
[26] S. Heinen and M. Adrat, "Optimal MMSE Estimation for Vector Sources with Spatially and Temporally Correlated Elements," *Advances in Digital Speech Transmission*, R. Martin, U. Heute, and C. Antweiler, Eds. John Wiley & Sons, Ltd., Jan. 2008, ch. 11, pp. 311–328.
[27] T. Clevorn, L. Schmalen, P. Vary, and M. Adrat, "On Redundant Index Assignments for Iterative Source-Channel Decoding," *IEEE Commun. Lett.*, vol. 12, no. 7, pp. 514–516, July 2008.
[28] D. Divsalar and F. Pollara, "Multiple Turbo Codes for Deep-Space Communications," *JPL TDA Progress Report*, vol. 42, no. 121, pp. 66–77, May 1995.
[29] T. Clevorn, M. Adrat, and P. Vary, "Turbo DeCodulation using Highly Redundant Index Assignments and Multi-Dimensional Mappings," *International Symposium on Turbo Codes & Related Topics and International ITG Conference on Source and Channel Coding (SCC)*, Munich, Germany, Apr. 2006.
[30] A. Q. Pham, L. L. Yang, and L. Hanzo, "Joint Optimization of Iterative Source and Channel Decoding Using Over-Complete Source Mapping," *IEEE Vehicular Technology Conference (VTC-Fall)*, Baltimore, MD, USA, Sept. 2007.
[31] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial Concatenation of Interleaved Codes: Performance Analysis, Design, and Iterative Decoding," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 909–926, May 1998.
[32] R. Thobaben and J. Kliewer, "Design Considerations for Iteratively-decoded Source-Channel Coding Schemes," *Allerton Annual Conference on Communication, Control and Computing*, Monticello, IL, USA, Sept. 2006.
[33] A. E. Brouwers, "Bounds on Linear Codes," *Handbook of Coding Theory*, V. S. Pless and W. C. Huffmann, Eds. Elsevier, 1998, pp. 265–461.
[34] Nasruminallah and L. Hanzo, "EXIT-Chart Optimized Short Block Codes for Iterative Joint Source and Channel Decoding in H.264 Video Telephony," *IEEE Trans. Veh. Technol.*, vol. 58, no. 8, pp. 4306–4315, Oct. 2009.
[35] R. G. Maunder, J. Wang, S. X. Ng, L.-L. Yang, and L. Hanzo, "On the Performance and Complexity of Irregular Variable Length Codes for Near-Capacity Joint Source and Channel Coding," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1338–1347, Apr. 2008.
[36] M. Tüchler and J. Hagenauer, "EXIT Charts of Irregular Codes," *Conference on Information Sciences and Systems (CISS)*, Princeton, NJ, USA, Mar. 2002.
[37] R. G. Maunder and L. Hanzo, "Near-Capacity Irregular Variable Length Coding and Irregular Unity Rate Coding," *IEEE Trans. Wireless Commun.*, vol. 8, no. 11, pp. 5500–5507, Nov. 2009.
[38] N. Görtz, "Optimization of Bit Mappings for Iterative Source-Channel Decoding," *International Symposium on Turbo Codes & Related Topics*, Brest, France, Sept. 2003.
[39] T. Clevorn, P. Vary, and M. Adrat, "Parameter SNR Optimized Index Assignments and Quantizers based on First Order A Priori Knowledge for Iterative Source-Channel Decoding," *Conference on Information Sciences and Systems (CISS)*, Princeton, NJ, USA, Mar. 2006.
[40] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization*. London: Academic Press, 1981.
[41] F. MacWilliams and N.Sloane, *The Theory of Error-Correcting Codes*. North Holland, 1977.
[42] L.-F. Wei, "Trellis Coded Modulation with Multidimensional Constellations," *IEEE Trans. Inf. Theory*, vol. 33, no. 4, pp. 483–501, July 1987.
[43] N. H. Tran and H. H. Nguyen, "Improving the Performance of QPSK BICM-ID by Mapping on the Hypercube," *IEEE Vehicular Technology Conference (VTC-Fall)*, Los Angeles, CA, USA, Sept. 2004.
[44] F. Simoens, H. Wymeersch, H. Bruneel, and M. Moeneclaey, "Multi-dimensional Mapping for Bit-Interleaved Coded Modulation with BPSK/QPSK Signaling," *IEEE Commun. Lett.*, vol. 9, no. 5, pp. 453–455, May 2005.
[45] L. Schmalen, P. Vary, M. Adrat, and T. Clevorn, "Complexity-Reduced Iterative Source-Channel Decoding by Conditional Quantization," *International Symposium on Turbo Codes & Related Topics*, Lausanne, Switzerland, Sept. 2008, pp. 327–332.
[46] X. Wu, S. Dumitrescu, and Z. Whang, "Monotonicity-Based Fast Algorithms for MAP Estimation of Markov Sequences Over Noisy Channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 7, pp. 1539–1544, July 2004.
[47] J. Hagenauer, E. Offer, and L. Papke, "Iterative Decoding of Binary Block and Convolutional Codes," *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 429–445, Mar. 1996.
[48] L. Schmalen, "Iterative Source-Channel Decoding: Design and Optimization for Heterogeneous Networks," Ph.D. dissertation, RWTH Aachen University, 2011.
[49] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate," *IEEE Trans. Inf. Theory*, vol. 20, no. 2, pp. 284–287, Mar. 1974.

Laurent Schmalen received the Dipl.-Ing. degree in information and communication technology and the Dr.-Ing. degree from RWTH Aachen University, Aachen, Germany, in 2005 and 2011, respectively. From 2005 to 2010, he was with the Institute of Communication Systems and Data Processing at RWTH Aachen University. In 2011, he joined Alcatel-Lucent Bell Labs. His research interests include channel coding, iterative receiver processing, modulation, joint source-channel coding, and optical fiber communications.

Peter Vary (M'85–SM'04–F'09) received the Dipl.-Ing. degree in electrical engineering from the University of Darmstadt, Darmstadt, Germany, in 1972 and the Dr.-Ing. degree from the University of Erlangen- Nuremberg, Erlangen, Germany, in 1978.

In 1980, he joined Philips Communication Industries (PKI), Nuremberg, Germany, where he became head of the Digital Signal Processing Group. Since 1988, he has been a Professor at RWTH Aachen University, Aachen, Germany, and head of the Institute of Communication Systems and Data Processing. His main research interests are speech coding, joint source-channel coding, error concealment, and speech enhancement including noise suppression, acoustic echo cancellation, and artificial wideband extension.