Advances in
Radio Science
Open Access Proceedings

Improved fault tolerance of Turbo decoding based on optimized index assignments

J. Geldmacher and J. Götze

TU Dortmund University, Information Processing Lab, Otto-Hahn-Str. 4, 44227 Dortmund, Germany

Correspondence to: J. Geldmacher (jan.geldmacher@ieee.org)

Abstract. This paper investigates the impact of an error-prone buffer memory on a channel decoder as employed in modern digital communication systems. On one hand this work is motivated by the fact that energy efficient decoder implementations may not only be achieved by optimizations on algorithmic level, but also by chip-level modifications. One of such modifications is so called aggressive voltage scaling of buffer memories, which, while achieving reduced power consumption, also injects errors into the likelihood values used during the decoding process. On the other hand, it has been recognized that the ongoing increase of integration density with smaller structures makes integrated circuits more sensitive to process variations during manufacturing, and to voltage and temperature variations. This may lead to a paradigm shift from 100 %-reliable operation to fault tolerant signal processing. Both reasons are the motivation to discuss the required co-design of algorithms and underlying circuits. For an error-prone receive buffer of a Turbo decoder the influence of quantizer design and index assignment on the error resilience of the decoding algorithm is discussed. It is shown that a suitable design of both enables a compensation of hardware induced bits errors with rates up to 1 % without increasing the computational complexity of the decoder.

1 Introduction

Two impact factors for the power consumption of a signal processing circuit can be identified: Power consumption of the logic (roughly related to the algorithmic complexity) and power consumption of the involved buffer memory. The latter is related to the number of read-write operations and their reduction has become a key factor when designing energy efficient algorithms. Especially in state-of-the-art baseband

signal processing, where large blocks are often processed in an iterative fashion, memory access is a major power consuming part. At the same time, system on chips (SoCs) are dominated in terms of area by the embedded memory.

Given that embedded memories are already highly structured and highly optimized devices, there is less potential for energy optimization if hard quality constraints are reinforced. However, if relaxing quality requirements in a controlled fashion, significant energy reduction can be achieved: In aggressive voltage scaling (AVS) (Hegde and Shanbhag, 2001; Djahromi et al., 2007; Makhzan et al., 2007), the supply voltage of an embedded circuit is deliberately reduced below the required threshold. This leads to substantial reduction of power consumption, but also to unreliable operation of the circuit. But while the latter would introduce processing errors and serious performance degradation of the whole system if circuit logic was affected, it can be tolerated up to a certain level for some applications in case of embedded memories. This is for example true for baseband signal processing systems, like channel decoders, which are actually designed to deal with error-prone data and can thus be extended to deal with deliberately introduced errors as well. As an example, Hussien et al. (2010) propose an error-resilient Viterbi decoder architecture, where power savings of 15 % to 20 % are achieved while the bit error rate (BER) performance degradation is insignificant.

Besides deliberately introduced circuit failures, the increasing integration density of ICs also leads to an increased sensitivity against variations during the manufacturing process and harsh operating conditions (radiation, temperature) (Baumann, 2005; Borkar, 2005; Ghosh and Roy, 2010; Mitra et al., 2011). These so called single event upsets have been traditionally tackled on circuit level and there is vast amount of research in this field, c.f. e.g. Mitra et al. (2005);

Yoshimoto et al. (2012). However, it has also been recognized (Shanbhag et al., 2010; Mitra et al., 2010; Karakonstantis et al., 2012; Kleeberger et al., 2013), that the development of fault tolerant (error resilient) systems cannot be dealt with in a single perspective, but rather that a cross-layer view is required: A co-design of embedded circuits and signal-processing algorithms will be necessary to efficiently exploit potentials of approaches like AVS and to obtain relatively reliable systems based on unreliable underlying components. This approach constitutes a "paradigm shift from 100 %-reliable operation to fault tolerant signal processing" (Novak et al., 2010).

In context of baseband signal processing and channel decoding, a few authors have addressed a co-design with erroneous circuits under different conditions: The inherent fault-tolerance of communication systems is exploited by Djahromi et al. (2007), where the authors identify a duality between communication channel errors and hardware induced errors. Based on this observation they propose to adapt the supply voltage depending on the current working condition of the system; by reducing the supply voltage in case of good reception conditions they achieve power savings of around 45 % in a WCDMA receiver The resilience of Viterbi and MaxLog decoding against timing errors and memory errors introduced by voltage overscaling is treated in Liu et al. (2009), where power savings of about 44 % and 38 % are reported for Viterbi and MaxLog decoders, respectively. The authors of Abdallah and Shanbhag (2009) also deal with Viterbi decoding and the influence of timing errors due to process variations and voltage scaling. They investigate several low level methods to improve error resilience and achieve significant power savings of up to 71 % with a small tolerated loss of coding gain. A more abstract treatment of error-resilient Viterbi decoding using an erroneous receive buffer is provided by Hussien et al. (2010), where the authors employ a statistical model of the combined communication channel and hardware noise to derive a suitable branch metric. They also show that the branch metric computation can be kept simple in case of Two's complement representation of the quantization symbols and report a reduction of power consumption in the order of 15 % to 20 % with small loss of coding gain. The authors extend their approach in Hussien et al. (2011) to LDPC and Turbo decoders considering the receive (ARQ) buffer; in Khairy et al. (2012) they also study its application to MIMO detection. For an LDPC decoder Alles et al. (2007); May et al. (2008) propose a reliability-aware design, which improves the protection against timing and signal errors by using simple error correction and detection techniques. Low and high level approaches to improve error-resilience of a Turbo decoder are investigated by Brehm et al. (2012). On algorithm level, the authors propose an increase of iteration number to facilitate resilience against timing errors and soft errors. The same authors also study the resilience of a MIMO-BICM system with iterative receiver and hardware errors that manifest as transient bit errors in various

components (Gimmler-Dumont et al., 2012) and conclude that for low error rates, performance of the original system can almost be retained by employing additional iterations.

Considering the problem of improving resilience of channel decoding algorithms against memory defects under more general, algorithmic perspective, it is interesting, that only the works by Kurdahi, Eltawil, et al. (Hussien et al., 2010, 2011) consider modification and adaption of the decoding algorithm itself. Furthermore, the strong relation of the given problem to source channel coding and robust quantizer design has only been pointed out in Novak et al. (2010); Roth et al. (2012), while its significance is long known in those fields. Novak et al. (2010) consider a MIMO BICM system with unreliable storage of the received log-likelihood ratio (LLR) values. The authors investigate the influence of (redundant) index assignments and simple forward error correction codes for the indices in terms of achievable rate and conclude that in case of non-redundant indices the selection of a robust index assignment is crucial, while in case of redundant labels the decision between simple FEC coding and customized index assignment depends on the SNR operation state of the underlying system. The authors address the first point more extensively in Roth et al. (2012) and stress the importance of application specific index assignments. They discuss the cases of repetition coding and convolutional coding and derive optimized index assignments through exhaustive search. For a Turbo decoder with unreliable LLR buffer, an index assignment optimization strategy based on the EXIT characteristics of the decoder is described in Geldmacher and Götze (2013). The authors show that the resulting assignments provide improved error-resilience without increasing decoding complexity.

In this article, the influence of optimized index assignment and quantizer design on a Turbo decoder with unreliable receive buffer memory is studied. This scenario of communication channel with quantized output and successive unreliable buffer memory is modeled as a cascade of two discrete memoryless channels (DMCs); the index assignment is regarded as a way to connect both DMCs. It is shown that joint optimization of index assignment and quantizer can significantly improve error-resilience of the decoder, without increasing its computational complexity.

2   Problem description

2.1   System model

The model shown in Fig. 1 is employed to discuss the problem of an unreliable buffer memory. A source generates Gaussian distributed values $\tilde{r}$ from the original equiprobable and independent symbols $x \in \{\pm 1\}$,

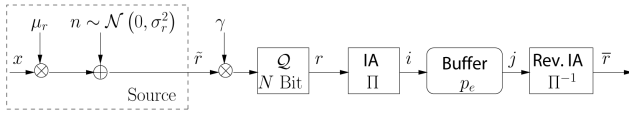$$\tilde{r} = \mu_r x + n, \quad \text{where } n \sim \mathcal{N}(0, \sigma_r^2). \tag{1}$$

Figure 1. Model of a Gaussian source with successive quantizer and buffer memory.

Figure 2. Abstract representation of the AWGN channel with successive quantizer and unreliable buffer memory as DMC cascade.

The source abstractly models for example an AWGN communication channel, with a certain gain μ_r and a noise variance σ_r^2, or it may be a so called *a priori* channel representing a SISO component in an iterative processing system, e.g. a MAP decoder. Given the continuous, Gaussian distributed values $\tilde{r}$, the N-Bit quantizer assigns discrete reconstruction values r from a finite set $\mathcal{Q}$ of size $|\mathcal{Q}| = 2^N$ in a suitable fashion. In the following the quantizer is designed using uniform and fixed reconstruction values from a set

$$\mathcal{Q} = \{-2^{d-1} + k2^{-f} : 0 \le k < 2^N\}, \tag{2}$$

where d and f are decimal and fractional bits and $d + f = N$. This approach is simple and might not include the optimum quantizer design with respect to the mean-squared-error (MSE) or the average mutual information (MI), but it seems practically relevant and enables us to discuss effects of the memory channel disturbance. To account for changing channel conditions (μ_r and σ_r^2), the scaling factor γ can be used as an adaptive gain control and scale the quantizer input to the quantization range. In some use cases, for example if the source models a constituent decoder in a Turbo decoding framework, the scaler γ might, however, be fixed for practical reasons. Given the set $\mathcal{Q}$, the probability mass function (PMF) $P(r|x)$ of the quantized values r conditioned on x can be found by integrating over the individual quantization intervals.

The quantization scheme coincides with the Two's complement representation, if the assignment from quantization index k to the index i, whose binary representation is stored in the buffer, is selected as $i = k \oplus 2^{d-1}$ with $\oplus$ denoting the XOR operation and $i \in \{0, \ldots, 2^{N-1}\}$. In general, however, any mapping from quantization values to indices may be used. We refer to this bijective mapping as the index assignment and denote a specific one-to-one relation between k and i by Π, $i = \Pi(k)$, and its reverse as Π^{-1}. The binary representation of the integer i is written to the buffer memory. As in Novak et al. (2010); Khajeh et al. (2010), spatially independent and uniformly distributed errors on the memory cells are assumed, such that the memory channel can be seen as a binary symmetric channel (BSC) with input and output the binary representations of i and j, respectively. It is referred to as the memory channel in the following. Given the bit error probability p_e, the probability of reading an index j under the condition that i has been written to the memory is given as
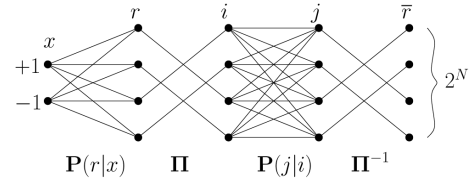
$$P(j|i) = p_e^{d_H(i,j)} (1 - p_e)^{N - d_H(i,j)}, \tag{3}$$

where $d_H(i, j)$ is the Hamming distance between i and j. Given the fact that the index assignment is a bijective mapping, then for the PMF of the reconstructed quantization symbols $\bar{r} \in \mathcal{Q}$ conditioned on x it holds that

$$P(\bar{r}|x) = P(\Pi^{-1}(j)|x). \tag{4}$$

2.2 Characteristics of memory channel

The transmission matrix (Cover and Thomas, 1991, p. 189) of the AWGN channel with quantized outputs and binary input is a (2×2^N)-matrix and it can be written as

$$\mathbf{P}(r|x) = \begin{bmatrix} P(q_0|-1) & \ldots & P(q_{2^N-1}|-1) \\ P(q_0|+1) & \ldots & P(q_{2^N-1}|+1) \end{bmatrix}, \tag{5}$$

with $q_k \in \mathcal{Q}$ the reconstruction values of the quantizer. The $(2^N \times 2^N)$-matrix of the memory channel is given as

$$\mathbf{P}(j|i) = \begin{bmatrix} P(0|0) & \ldots & P(2^N - 1|0) \\ \vdots & \ddots & \vdots \\ P(0|2^N - 1) & \ldots & P(2^N - 1|2^N - 1) \end{bmatrix}. \tag{6}$$

To obtain the transmission matrix $\mathbf{P}(\bar{r}|x)$ of the concatenation of both channels, the index assignment has also to be considered. As it is a mapping of quantizer output to memory channel input, it may be represented by a matrix $\mathbf{\Pi}$, that permutes the columns of $\mathbf{P}(r|x)$ in the required order. The reverse permutation has to be applied to the columns of $\mathbf{P}(j|i)$ using the transpose of $\mathbf{\Pi}$. Then we have the transmission matrix as

$$\mathbf{P}(\bar{r}|x) = \mathbf{P}(r|x) \mathbf{\Pi} \mathbf{P}(j|i) \mathbf{\Pi}^T, \tag{7}$$

where $\mathbf{\Pi}$ is for example given as

$$\mathbf{\Pi} = \begin{bmatrix} & \mathbf{I}_{2^{N-1}} \\ \mathbf{I}_{2^{N-1}} & \end{bmatrix} \text{ for Two's complement,} \tag{8}$$

or as $\mathbf{\Pi} = \mathbf{I}_{2^N}$ for natural binary coding (NBC).

From Eq. (7) it is already apparent that the concatenation of quantized AWGN channel and unreliable buffer memory can be seen as cascade of two DMCs. Figure 2 illustrates this cascade for the case of $N = 2$ Bit and Two's complement index assignment. It can be noted that the index assignment is just a way of connecting the first DMC to the second one.

To derive the transmission matrix it has intuitively been assumed, as illustrated in Fig. 2, that the output of the memory channel only depends on its input, such that

$$P(\overline{r}|r,x) = P(\overline{r}|r). \tag{9}$$

Now let R, $\overline{R}$ and X be the random variables representing r, $\overline{r}$ and x, respectively. Then Eq. (9) is justified by the observation that once R is known, $\overline{R}$ cannot reveal any additional information about X, because there is no connection between X and $\overline{R}$ other than the cascade of communication and memory channel. This means that the MI of X and $\overline{R}$ conditioned on R vanishes,

$$I(X;\overline{R}|R) = 0, \tag{10}$$

and thus $X, R, \overline{R}$ form a Markov chain $X \rightarrow R \rightarrow \overline{R}$. Now some interpretations can be derived: First the chain rule of mutual information (Cover and Thomas, 1991, p. 22) may be applied as

$$\begin{aligned} I(X;R,\overline{R}) &= I(X;R) + I(X;\overline{R}|R) \\ &= I(X;\overline{R}) + I(X;R|\overline{R}). \end{aligned} \tag{11}$$

Then using Eq. (10) yields

$$I(X;\overline{R}) = I(X;R) - I(X;R|\overline{R}). \tag{12}$$

Equation (12) shows that the MI of $\overline{R}$ and X can only be smaller than or equal to the MI of R and X, given non-negativity of the MI. Equality holds if $p_e = 0$, in which case $R = \overline{R}$ and consequently $I(X;R|\overline{R}) = 0$:

$$\begin{aligned} I(X;\overline{R}) &< I(X;R) \quad \text{for} \quad p_e > 0 \quad \text{and} \\ I(X;\overline{R}) &= I(X;R) \quad \text{for} \quad p_e = 0. \end{aligned} \tag{13}$$

Thus the memory channel output will be degraded if $p_e > 0$, while it matches the output of the communication channel for $p_e = 0$. The term $I(X;R|\overline{R})$ from Eq. (12), which denotes the MI loss at memory channel output, can be further rewritten,

$$\begin{aligned} I(X;R|\overline{R}) &= H(R|\overline{R}) - H(R|\overline{R},X) \\ &= \left(H(R|\overline{R}) - H(\overline{R}) \right) + H(R) - H(R|\overline{R},X) \\ &= -I(R;\overline{R}) + H(R) - H(R|\overline{R},X), \end{aligned} \tag{14}$$

and it can be seen that the degradation of MI depends on

- the error probability p_e of the memory channel, represented by the MI $I(R;\overline{R})$ of its in- and output,

- on the design of the quantizer, represented by the entropy $H(R)$ of its output, and

- the interaction of all components, including the index assignment, represented by $H(R|\overline{R},X)$.

Thus in order to optimize the MI $I(X;\overline{R})$ and with it the MI loss due to the memory channel, all system parameters have to be jointly taken into account. Therefore, the optimization should be carried out subject to the quantizer (represented by the scaler γ) and the index assignment Π. This approach is investigated in the following section for a Turbo decoding scenario.

3   Turbo decoder with unreliable receive buffer

3.1   System model

In the following, the transmission of BPSK-modulated Turbo coded data with rate R_c over a communications channel is modeled as shown in Fig. 3: Binary information symbols u are encoded using two parallel concatenated recursive systematic convolutional encoders, such that the coded symbols v consist of a systematic part and two parity parts from first and second encoder. BPSK modulation of v results in the equally probable modulation symbols $x = 2v - 1$. A gain factor μ_r may be applied, such that the symbols $\mu_r x$ are transmitted eventually.

The communication channel is modeled as an AWGN channel, so that the received values $\tilde{r}$ can be written as

$$\tilde{r} = \mu_r x + n, \qquad \text{where } n \sim \mathcal{N}(0, \sigma_r^2), \tag{15}$$

where σ_r^2 denotes the noise power of the channel. The channel is characterized by its signal-to-noise ratio:

$$E_b/N_0|_{\text{dB}} = 10\log_{10}\frac{\mu_r^2}{2R_c\sigma_r^2}. \tag{16}$$

Before quantization, the received symbols are weighted with the channel reliability L_c and scaled to the quantizers dynamic range using γ. It is assumed that perfect channel state information is available to the receiver, such that the quantizer input will be the scaled LLR

$$\begin{aligned} \gamma \cdot \log\frac{p(\tilde{r}|v=1)}{p(\tilde{r}|v=0)} &= \gamma \cdot \log\frac{\exp\left(-(\tilde{r}-\mu_r)^2/2\sigma_r^2\right)}{\exp\left(-(\tilde{r}+\mu_r)^2/2\sigma_r^2\right)} \\ &= \gamma\frac{2\mu_r}{\sigma_r^2}\tilde{r} = \gamma L_c\tilde{r}. \end{aligned} \tag{17}$$

The quantizer assigns reconstruction values r from the set of Two's complement values as given in Eq. (2) with $f = 0$ and $N = d$. The scaling factor γ still allows to achieve a set of reconstruction values including any combination of d and f with $N = d + f$.

After quantization the index assignment Π assigns indices i to the quantized values and the corresponding binary representation is stored in the receive buffer. The buffer is characterized by its error probability p_e, such that the restored indices j may differ from the written indices i. Based on the corresponding restored symbols $\overline{r}$ the Turbo decoder computes an estimate $\hat{u}$ of the original information by iterating
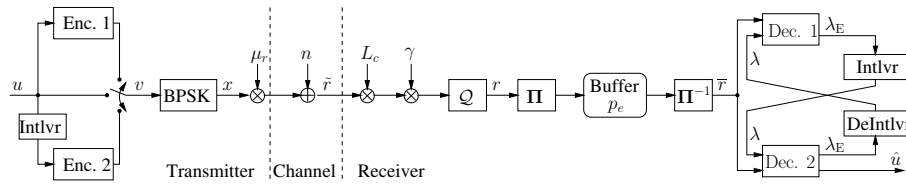
Figure 3. System model for Turbo coded transmission with unreliable receive and LLR buffer.

between the two constituent decoders. Both constituent decoders implement the BCJR algorithm (Bahl et al., 1974) in log-domain on the trellis of the encoder to produce extrinsic LLRs λ_E of the information symbols: More concisely, the first decoder employs the systematic part and the first parity part of $\overline{r}$, while the second decoder uses the interleaved systematic part and the second parity part of $\overline{r}$. The extrinsic LLRs λ_E generated by the first/second decoder are interleaved/de-interleaved and become the *a priori* LLRs λ of the second/first decoder in the next stage.

A rate $R_c = 1/3$ Turbo code with generator $\mathbf{G}(D) = \left[1, \frac{D^3+D+1}{D^3+D^2+1}\right]$ and length 2^{15} random interleaver is employed in the following. Both encoders are terminated and the resulting tails are transmitted along with the systematic and parity parts. The transmission gain is $\mu_r = 1$. In the following a conventional LogMAP based Turbo decoder is compared to a fault tolerant (FT) decoder. The number of iterations carried out by the decoder is fixed to 8. The FT decoder works like the conventional decoder, with the exception that its transition metric is based on the actual transition probability $P(\overline{r}|x)$ of the cascade of communication and memory channel (Geldmacher and Götze, 2012). It employs a look-up table (LUT) of size 2^N which holds the log domain representation of $P(\overline{r}|x)$. The values obtained by indexing this LUT using the indices j replace the received values in the transition metric of a conventional MAP algorithm. The LUT is constructed once per received frame by estimating the statistics of r, computing $P(\overline{r}|x)$ based on Eq. (7) and storing its logarithm.

3.2 Optimized index assignment

Following Eq. (13), the unreliable receive buffer, represented by a BSC with error probability p_e will cause a decrease of MI. The MI $I(X; \overline{R})$ is clearly a function of $P(\overline{r}|x)$ which in turn depends for a fixed E_b/N_0, p_e and $\mathcal{Q}$ on the scaling parameter γ and the index assignment Π. Thus the MI for a given pair (γ, Π) is denoted by $I_{(\gamma,\Pi)}(X; \overline{R})$. Consequently it is reasonable that for a given E_b/N_0, p_e and $\mathcal{Q}$, there is a combination (Π^*, γ^*) that maximizes $I_{(\gamma,\Pi)}(X; \overline{R})$:

$$\left(\Pi^*, \gamma^*\right) = \arg\max_{\Pi,\gamma} I_{(\gamma,\Pi)}(X; \overline{R}) \tag{18}$$

Problem Eq. (18) can be solved by first finding optimized $\Pi^*(\gamma)$ for reasonable values of γ:

$$\Pi^*(\gamma) = \arg\max_{\Pi} I_{(\gamma,\Pi)}(X; \overline{R}), \tag{19}$$

Then γ^* can be selected as

$$\gamma^* = \arg\max_{\gamma} I_{(\gamma,\Pi^*(\gamma))}(X; \overline{R}) \text{ and } \Pi^* = \Pi^*(\gamma^*), \tag{20}$$

which yields the optimal combination of (Π^*, γ^*).

Solving Eq. (19) for an optimized index assignment $\Pi^*(\gamma)$ is a combinatorial optimization problem. Given the fact, that the number of possible index assignments for an N-Bit quantization is given by $(2^N)!$, from which $(2^N - 1)!/N!$ yield distinct results (Azami et al., 1996), one has therefor to resort to heuristic optimization algorithms for $N > 3$. A meta-heuristic which had already turned out efficient for other combinatorial problems is simulated annealing (e.g. Reeves, 1993; Farvardin, 1990); it is employed in this work to solve Eq. (19).

The result of this optimization is illustrated in Fig. 4 (left), where the MI $I(X; \overline{R})$ is shown as a function of the scaler γ for optimized index assignment Π^* and NBC Π_{NBC}. In general it can be noticed, that the memory channel causes a reduction of achievable MI $I(X; \overline{R})$ compared to the reference $I(X; R)$. Also, it can be observed that the MI depends on γ and that the optimal γ is different in all considered cases. This implies that even if the index assignment is not modified, the scaling factor should be adapted to the memory channel, because otherwise, a performance reduction may take place: For example, taking $N = 4$, the optimal scaler for $p_e = 0$ ($I(X; R)$, "Ref") is about 2, which leads to $I(X; \overline{R}) = 0.370$ (Π^* and Π_{NBC}), while actually greater MI would be possible by selecting $\gamma = 3$. The figure also shows that a larger quantization width than actually required to accurately represent $\tilde{R}$ allows for a larger gain due to the optimized index assignment: While for $N = 4$ there is only little improvement of Π^* over Π_{NBC}, a larger quantization width of $N = 6$ enables a significantly increased $I(X; \overline{R})$. The reason for this can be found in the larger redundancy available in R, which provides more degrees of freedom to the index assignment optimization, and thus yields a more effective protection against errors induced by the memory channel.
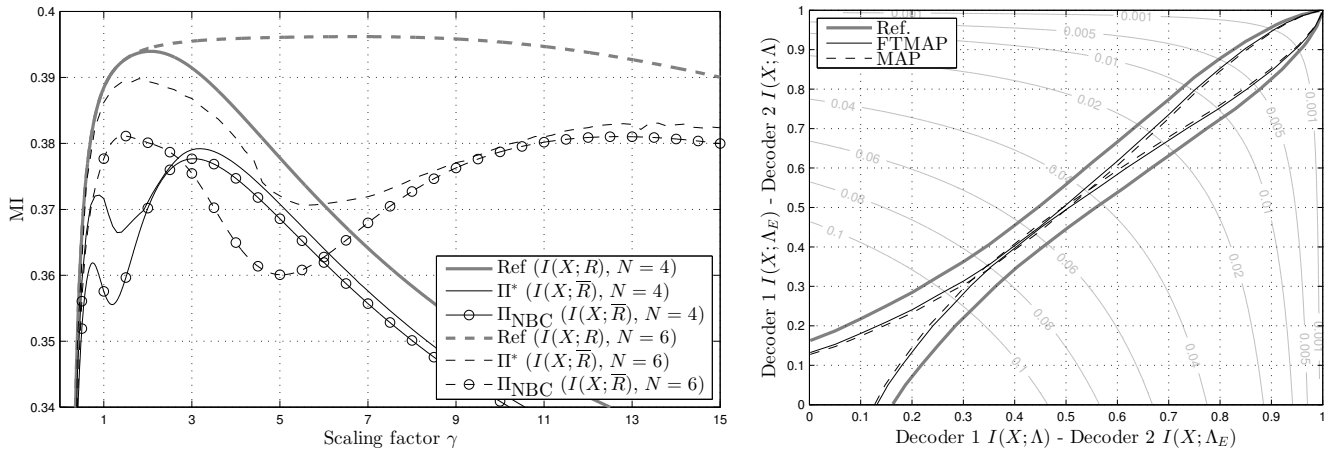
Figure 4. Left: Mutual information after receive buffer for NBC and optimized index assignment for $N = 4$ and $N = 6$ ($E_b/N_0 = 0.5\,\text{dB}$, $p_e = 0.01$, $R_c = 1/3$). Right: EXIT chart of conventional and FT Turbo decoder with unreliable receive buffer for $N = 4$ ($p_e = 0.01$, $E_b/N_0 = 0.25\,\text{dB}$, NBC, BER estimate as contour lines).
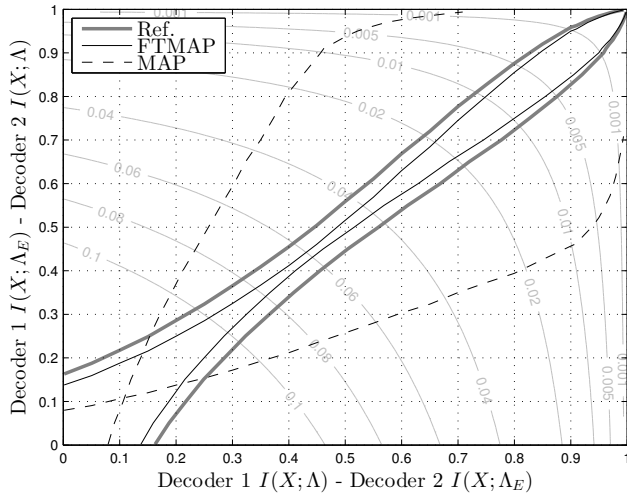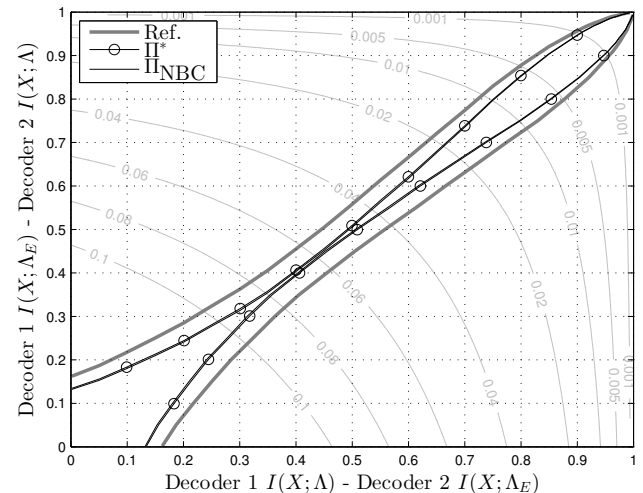

Figure 5. EXIT chart of conventional and FT Turbo decoder with unreliable receive buffer for $N = 6$ ($p_e = 0.01$, $E_b/N_0 = 0.25\,\text{dB}$, NBC, BER estimate as contour lines).

Figure 6. EXIT chart for NBC and optimized index assignment for $N = 4$ ($p_e = 0.01$, $E_b/N_0 = 0.25\,\text{dB}$, FT decoder).

3.3 EXIT chart

The influence of optimized index assignment and quantization width N on the decoding behaviour is first studied using EXIT charts. Consider Fig. 4 (right), where the performance of conventional Turbo decoder ("MAP") and FT decoder ("FTMAP") using NBC index assignment (Π_{NBC}) are compared for two different quantization widths $N = 4$ and $N = 6$ and $p_e = 0.01$. As a reference the chart for a decoder with $p_e = 0$ and identical quantization settings is shown ("Ref.") – given the width of the tunnel, it can be expected that the decoder will converge at this E_b/N_0 level. If, however, the received values experience additional distortion due to the memory channel, then the tunnel becomes much smaller and

decoding performance will be deteriorated. In Fig. 4 (right), where a quantization width of $N = 4$ is used, this impact appears like a small SNR degradation on the communication channel and manifests as a reduction of the EXIT tunnel. It can also be observed, that in this case, there is only very little difference between the conventional decoder and the fault tolerant decoder, and only very close observation reveals the slight improvement due to latter. Convergence to BERs smaller than about 0.08 cannot be expected for both decoders.

However, increasing the quantization width to $N = 6$ Bit, as shown Fig. 5, leads to a significant difference between both decoders: The conventional decoder exhibits a substantial performance degradation, which is indicated by a very early crossing of the EXIT curves. In fact even for very good

a priori information ($I(X;\Lambda) \to 1$), the decoder is not capable of producing improved extrinsic information, such that decoding performance is expected to be strongly degraded. The FT decoder on the other hand shows an improved EXIT chart compared to the $N = 4$ case: While the tunnel is still considerably smaller than for the reference, the curves do not intersect and given a sufficient number of iterations the decoder may be able to achieve reference BER performance. Thus it can be concluded that while an increased quantization width amplifies the impact of the memory channel and degrades performance of a conventional decoder, it also enables improved error-resilience due to increased redundancy in quantizer output. Exploitation of this redundancy requires the FT decoder though.

As suggested by Fig. 4 (left) a higher quantization width and an optimized index assignment w.r.t. Eq. (18) can yield further improvement. Figures 6 and 7 therefore compare EXIT charts of the FT decoder for NBC (Π_{NBC}) and optimized index assignments (Π^*). In Fig. 6 a quantization width of $N = 4$ is employed. The limited redundancy in R and smaller degrees of freedom during the optimization do not allow for any gain of the optimized index assignment over NBC in this case. Both charts are identical. But with N increased by two Bit, the optimization becomes more effective as shown in Fig. 7. The tunnel nearly approaches the reference, indicating that the decoder will converge with less iterations than in the NBC case.

A general observation from the presented EXIT charts is that an unreliable receive buffer leads to degradation in the lower to medium part of the chart. This is due to the fact that during the ongoing decoding process the extrinsic LLRs become more important than the received values. Regarding the BER performance, we thus expect a degradation of decoding threshold but only marginal impact on the error floor, because low BERs are represented by the very upper part of the chart, where the received values only have small impact.

3.4 BER

Figure 8 shows performance results for a Turbo decoder with unreliable receive buffer in terms of the BER. The scaling factor γ is computed as a linear function of E_b/N_0. The coefficients of this function have been found by linear fit of the solution of Eq. (20) for a suitable range of E_b/N_0 and for each of the considered index assignments (cf. Geldmacher and Götze, 2012).

In Fig. 8 (left) the performance of FT and conventional Turbo decoder are compared for $p_e = 0.01$, where both decoders employ NBC. The number of quantization bits is varied from 4 to 6, that is $N = 4, 5, 6$. First of all it can be observed, that the additional distortion due to the memory channel increases the decoding threshold, but does not affect the error floor – thus it acts like a decreased SNR on the communications channel as noted before. When comparing both decoders, it can be noticed that the FT decoder outper-
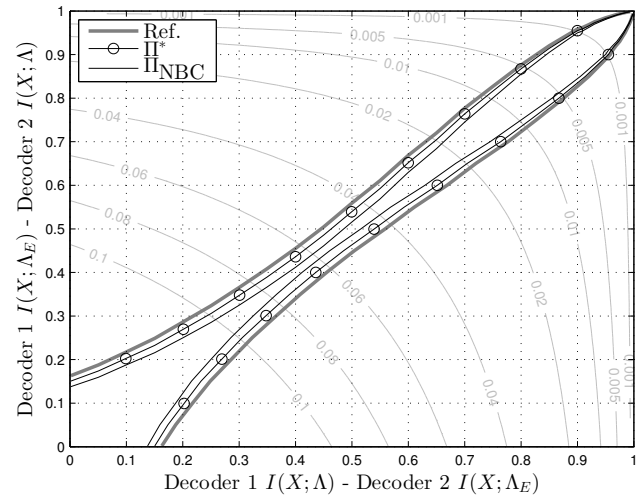

Figure 7. EXIT chart for NBC and optimized index assignment for $N = 6$ ($p_e = 0.01$, $E_b/N_0 = 0.25\,\text{dB}$, FT decoder).

forms the conventional decoder. Furthermore it can be seen that an increased quantization width slightly improves error-resilience: Baseline performance is approached earlier if 5 or 6 Bit are employed instead of only 4 Bit.

To take into account the impact of the index assignment we now compare performance of the FT decoder with optimized index assignment and the conventional decoder with NBC. For $p_e = 0.01$, the plot in Fig. 8 (middle) shows the influence of an increased quantization width for this scenario. A significantly improved performance can be observed for the FT decoder with higher quantization width: Actually, for $N = 6$ the FT decoder completely compensates the disturbance due to the memory channel for BERs smaller than 10^{-5}, since the higher redundancy involved with the larger quantization width can be exploited more efficiently by the index assignment optimization. This is in accord with the conclusion from the EXIT chart analysis in the previous section.

The impact of the error probability p_e is shown in Fig. 8 (right) for $N = 6$. Taking a BER working point of 10^{-5} it can be seen that error rates of $p_e = 0.005$ and $p_e = 0.01$ are completely compensated by the FT decoder with optimized index assignment, while for $p_e = 0.05$ a loss of about 0.4 dB can be observed. The performance of the conventional decoder on the other hand is significantly degraded.

In order to quantify the loss of coding gain $\Delta E_b/N_0$ associated with a certain p_e at a given working point, Fig. 9 compares the performance of a conventional decoder and a FT decoder with each $N = 4$ Bit and NBC, and a FT decoder with $N = 6$ and optimized index assignment. Taking for example a relatively strong memory channel distortion of about $p_e = 0.05$ it can be seen that the conventional decoder with NBC and $N = 4$ requires an additional 1.7 dB to achieve a BER of 10^{-5} compared to a Turbo decoder with error-free buffer memory. Employing the FT decoder with $N = 4$ and
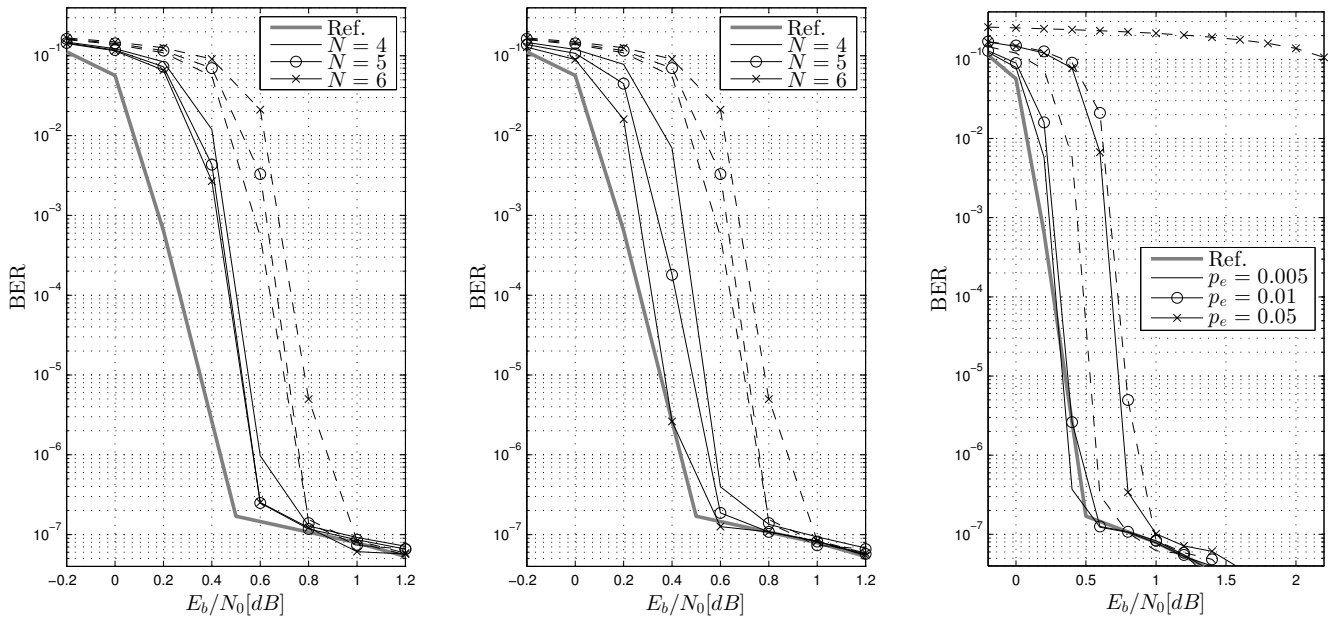
Figure 8. Left: Conventional ("$--$") vs. FT ("$-$") decoder ($p_e = 0.01$, NBC). Middle, right: NBC with conventional decoder ("$--$") vs. optimized index assignment with FT decoder ("$-$") ($p_e = 0.01$ (middle), $N = 6$ (right)).
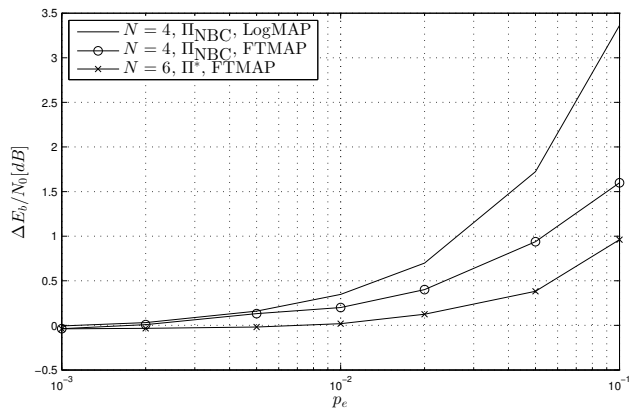

Figure 9. Loss of coding gain for a target BER of 10^{-5} as a function of p_e for different scenarios.

NBC reduces this loss to about 0.95 dB. Further improvement is achieved by spending 2 additional Bits and employing an optimized index assignment: In this case the loss of coding gain is reduced to 0.4 dB. It general it can be observed from the figure that the latter configuration outperforms the first two and can even completely compensate the impact of the memory channel for p_e up to 1 %. If only $N = 4$ Bit are employed, there still is a benefit from the FT decoder for $p_e \geq 0.005$, since it reduces the loss of coding gain by about 40 % to 50 %.

4   Conclusions

Increasing integration depths of integrated circuits may lead to higher susceptibility against process variations and soft error events. Similarly, aggressive voltage scaling can yield unreliable operation of logic and memory of an integrated circuit. Developing signal processing algorithms that can to some extend deal with unreliable underlying hardware is thus an emerging problem. In this paper a Turbo decoder with an unreliable receive buffer was studied. This buffer is modeled as a binary symmetric channel that acts on the quantized received values. It can thus be seen as an additional discrete memoryless channel that is cascaded to the actual communication channel. The interaction of both channels, represented by quantizer and index assignment, was identified as an important impact factor on the resilience of the Turbo decoder against errors originating from the memory channel. For a decoder with a transition metric adapted to the actual statistical model of the channel cascade, a joint optimization of both components yields improved error-resilience without increasing the computational complexity of the decoder.

References

Abdallah, R. and Shanbhag, N.: Error-Resilient Low-Power Viterbi Decoder Architectures, IEEE T. Signal Proces., 57, 4906–4917, 2009.

Alles, M., Brack, T., and Wehn, N.: A Reliability-Aware LDPC Code Decoding Algorithm, in: IEEE 65th Vehicular Technology Conference (VTC2007-Spring), 1544–1548, 2007.

Azami, S. B. Z., Duhamel, P., and Rioul, O.: Joint Source-Channel Coding: Panorama of Methods, in: Proc. of CNES Workshop on Data Compression, 1996.

Bahl, L., Cocke, J., Jelinek, F., and Raviv, J.: Optimal decoding of linear codes for minimizing symbol error rate, IEEE T. Inform. Theory, 20, 284–287, 1974.

Baumann, R.: Soft errors in advanced computer systems, IEEE Des. Test Comput., 22, 258–266, 2005.

Borkar, S.: Designing reliable systems from unreliable components: the challenges of transistor variability and degradation, IEEE Micro, 25, 10–16, 2005.

Brehm, C., May, M., Gimmler, C., and Wehn, N.: A Case Study on Error Resilient Architectures for Wireless Communication, in: Architecture of Computing Systems (ARCS 2012), 7179, 13–24, 2012.

Cover, T. and Thomas, J.: Elements of Information Theory, Wiley, 1991.

Djahromi, A. K., Eltawil, A. M., Kurdahi, F. J., and Kanj, R.: Cross Layer Error Exploitation for Aggressive Voltage Scaling, in: 8th Int. Symp. on Quality Electronic Design (ISQED '07), 192–197, 2007.

Farvardin, N.: A study of vector quantization for noisy channels, IEEE T. Inform. Theory, 36, 799–809, 1990.

Geldmacher, J. and Götze, J.: On Fault Tolerant Decoding of Turbo Codes, in: International Symposium on Turbo Codes & Iterative Information Processing (ISTC2012), Gothenburg, Sweden, 2012.

Geldmacher, J. and Götze, J.: EXIT-Optimized Index Assignments for Turbo Decoders with Unreliable LLR Transfer, IEEE Commun. Lett., 17, 992–995, 2013.

Ghosh, S. and Roy, K.: Parameter Variation Tolerance and Error Resiliency: New Design Paradigm for the Nanoscale Era, P. IEEE, 98, 1718–1751, 2010.

Gimmler-Dumont, C., Brehm, C., and Wehn, N.: Reliability study on system memories of an iterative MIMO-BICM system, in: IEEE/IFIP 20th International Conference on VLSI and System-on-Chip (VLSI-SoC), 255–258, 2012.

Hegde, R. and Shanbhag, N.: Soft digital signal processing, IEEE T. VLSI Syst., 9, 813–823, 2001.

Hussien, A., Khairy, M., Khajeh, A., Eltawil, A., and Kurdahi, F.: Combined Channel and Hardware Noise Resilient Viterbi Decoder, in: Asilomar Conf. on SS&C, Pacific Grove, CA, 2010.

Hussien, A., Khairy, M., Khajeh, A., Eltawil, A., and Kurdahi, F.: A Class of Low Power Error Compensation Iterative Decoders, in: IEEE Global Telecommunications Conference (GLOBECOM 2011), Houston, TX, USA, 2011.

Karakonstantis, G., Roth, C., Benkeser, C., and Burg, A.: On the exploitation of the inherent error resilience of wireless systems under unreliable silicon, in: Proceedings of the 49th Annual Design Automation Conference, DAC '12, 510–515, New York, NY, USA, 2012.

Khajeh, A., Amiri, K., Khairy, M., Eltawil, A., and Kurdahi, F.: A Unified Hardware and Channel Noise Model for Communication Systems, in: IEEE Global Communications Conference (GLOBECOM 2010), 2010.

Khairy, M. S., Shen, C.-A., Eltawil, A. M., and Kurdahi, F.: Error resilient MIMO detector for memory-dominated wireless communication systems, in: IEEE Global Communications Conference (GLOBECOM), 3566–3571, 2012.

Kleeberger, V. B., Gimmler-Dumont, C., Weis, C., Herkersdorf, A., Mueller-Gritschneder, D., Nassif, S. R., Schlichtmann, U., and Wehn, N.: A cross-layer technology-based study of how memory errors impact system resilience, IEEE Micro, 33(4), 46–55, 2013.

Liu, Y., Zhang, T., and Hu, J.: Design of Voltage Overscaled Low-Power Trellis Decoders in Presence of Process Variations, IEEE T. VLSI Syst., 17, 439–443, 2009.

Makhzan, M., Khajeh, A., Eltawil, A., and Kurdahi, F.: Limits on voltage scaling for caches utilizing fault tolerant techniques, in: 25th International Conference on Computer Design (ICCD 2007), 488–495, 2007.

May, M., Alles, M., and Wehn, N.: A case study in reliability-aware design: a resilient LDPC code decoder, in: Proceedings of the Conference on Design, Automation and Test in Europe (DATE '08), 456–461, New York, NY, USA, 2008.

Mitra, S., Seifert, N., Zhang, M., Shi, Q., and Kim, K.: Robust system design with built-in soft-error resilience, Computer, 38, 43–52, 2005.

Mitra, S., Brelsford, K., and Sanda, P.: Cross-layer resilience challenges: Metrics and optimization, in: Design, Automation Test in Europe Conference Exhibition (DATE), 1029–1034, 2010.

Mitra, S., Brelsford, K., Kim, Y., Lee, H., and Li, Y.: Robust System Design to Overcome CMOS Reliability Challenges, IEEE Journal of Emerging and Selected Topics in Circuits and Systems, 1, 30–41, 2011.

Novak, C., Studer, C., Burg, A., and Matz, G.: The effect of unreliable LLR storage on the performance of MIMO-BICM, in: Asilomar Conf. on SS&C, 736–740, 2010.

Reeves, C. (Ed.): Modern heuristic techniques for combinatorial problems, John Wiley & Sons, Inc., New York, NY, USA, 1993.

Roth, C., Benkeser, C., Studer, C., Karakonstantis, G., and Burg, A.: Data mapping for unreliable memories, in: 50th Annual Allerton Conference on Communication, Control, and Computing, 2012.

Shanbhag, N., Abdallah, R., Kumar, R., and Jones, D.: Stochastic computation, in: 47th ACM/IEEE Design Automation Conference (DAC), 859–864, 2010.

Yoshimoto, S., Amashita, T., Okumura, S., Nii, K., Yoshimoto, M., and Kawaguchi, H.: Bit-Error and Soft-Error Resilient 7T/14T SRAM with 150-nm FD-SOI Process, IEICE T. Fund. Electr., E95-A, 1359–1365, 2012.