Turbo Equalization with Parametric Uncertainties: Comparison of SNR Estimation Algorithms

Markus A. Dangl and Jürgen Lindner
University of Ulm, Dept. of Information Technology
Albert-Einstein-Allee 43, 89081 Ulm, Germany
Email: {markus.dangl,juergen.lindner}@uni-ulm.de

Abstract

We consider turbo equalization of ISI channels for the practically relevant case when neither channel state information nor SNR knowledge is perfect at the receive side. In particular, we investigate the influence of imperfect (uncertain) training information on the performance of several known SNR estimation algorithms. Numerical results indicate that uncertain training data causes an additional bias that may lead to a performance degradation. We propose a modified, adaptive SNR estimation algorithm with reduced bias in the low SNR regime.

I. Introduction

Iterative processing of extrinsic information is known to be a powerful receiver technique for various estimation and detection problems [1]. Among those, turbo equalization (iterative equalization and decoding) has been successfully employed to combat intersymbol interference (ISI). In the original turbo equalization scheme, proposed by Douillard et al. , an a-posteriori probability (APP) equalizer has been suggested [2]. Later on, approaches with reduced complexity, mainly based on soft interference cancellation, have been presented, for example see [3]–[6]. In this work, we consider turbo equalization schemes where the equalizer follows the unbiased minimum variance (UMV) principle and applies soft interference cancellation ("UMVE-SIC") [7]. This type of equalizer is identical to the minimum mean-square error (MMSE) equalizer in [5] for BPSK modulation, but shows superior performance for multi-amplitude signal constellations.

Our aim in this paper is to investigate the performance of turbo equalization schemes for the interesting and practically relevant case when neither the channel impulse response (CIR) nor the signal-to-noise ratio (SNR) are perfectly known at the receive side. In particular, we focus on the identification of powerful SNR estimators to be included in the iterative (turbo) processing loop. Most turbo equalization schemes require an SNR estimate to compute the output log-likelihood ratios (LLRs) of the equalizer. In our work, we consider transmission schemes where training symbols are regularly multiplexed into the data stream in order to provide initial channel and SNR estimates. Thereby, the soft feedback from the channel decoder can be incorporated into the SNR estimation algorithm to refine the initial training-based SNR estimate (adaptive turbo equalization). In the literature, several approaches of combined turbo equalization and channel estimation (some of them including SNR estimation) have been studied, see [8]–[10] and references therein. In [11], the effect of an SNR mismatch on the performance of Log-MAP and Max-Log-Map equalizers has been investigated, and numerical results for iterative data-aided SNR estimation are shown. A comparison of various SNR estimation algorithms has been carried out in [12] for the AWGN channel. However, to the authors' best knowledge, a study of the influence of imperfect training information on the performance of different SNR estimators has not been addressed yet.

As a key result of our work we observe a performance degradation of the SNR estimators at increasing SNR when imperfect training information such as soft data estimates are utilized. Therefore, we propose a modified SNR estimator with reduced bias when having uncertain training information, leading to superior performance compared to purely training-based SNR estimation. The paper is organized as follows: Section II introduces the system model, and the channel estimation and equalization scheme is briefly described in Section III. Then, the SNR estimation algorithms under consideration are introduced in Section IV. Numerical results and conclusions are discussed in Section V.

II. System Model

We use the following notation: upper case bold letters indicate matrices and lower case bold letters stand for column vectors. $\mathbf{I}_N$ denotes the $N \times N$ identity matrix. The element in ith row and jth column of a matrix $\mathbf{A}$ is denoted by $[\mathbf{A}]_{i,j}$. $(\cdot)^\dagger$ stands for conjugate complex transposition.

Consider the communication system depicted in Fig. 1. A vector $\mathbf{q}$ of N_q source bits is encoded by a terminated convolutional code with code rate $R_c = N_q/N_c$ and permuted by an S-random interleaver $\mathbf{\Pi}$ [13] with $S = 18$. The N_c interleaved code bits $\mathbf{c}$ are mapped onto symbols taken from a PSK signal constellation $\mathcal{A}$ with constant unit energy, $E_s = 1$, to form a vector of data symbols $\mathbf{x} = [x_1, x_2, \ldots, x_N]^T$ of length N. Transmission of data is organized in frames $\mathbf{s} = [s_1, s_2, \ldots, s_{P+N}]^T$ that are built from both the N data symbols and P training symbols $\mathbf{p}$ chosen randomly from the set $\{-1; +1\}$. A frame has length $P+N$ and obeys the following pattern: alternately P/K training symbols followed by N/K data symbols are filled in, where $K > 0$ is an integer. Afterwards, these frames are transmitted over a discrete-time, symbol-spaced ISI channel with L

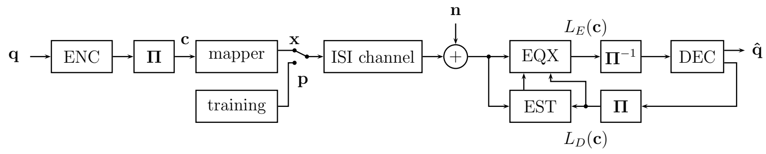

Fig. 1. Communication system employing adaptive turbo equalization at the receive side.

taps and CIR $\mathbf{h} = [h_0, h_1, \ldots, h_{L-1}]^T$, having constant unit energy $E_h = 1$. For conceptual ease, we assume a time-invariant channel and perfect synchronization at the receiver. Thus, the received symbols are obtained by $y_k = \sum_{l=0}^{L-1} h_l s_{k-l} + n_k$. k denotes the discrete (symbol) time index. The n_k are samples from a complex AWGN process with variance $\sigma_n^2 = N_0/2$ per component. Since the channel is not time-varying, all training symbols may be put in the beginning of a frame, i. e., we consider the case of $K = 1$. At the receive side, the training symbols are used to obtain an initial estimate of the CIR and the SNR ("EST"). We define the SNR ρ at the output of a hypothetical channel matched filter[1] as $\rho \triangleq P_s/P_n$, where P_s and P_n denote the average discrete signal and noise energy, respectively. Thus, we obtain $\rho_c = E_s/N_0$ for complex signal constellations and $\rho_r = 2E_s/N_0$ for real-valued signal constellations (independent from the channel, since $E_h = 1$) [12]. After initial channel and SNR estimation, iterative equalization and decoding is performed. Thereby, the extrinsic LLRs $L_D(\mathbf{c})$ of the decoder are possibly used to refine the SNR estimate $\widetilde{\rho}$ and/or channel estimate $\breve{\mathbf{h}} = [\breve{h}_0, \breve{h}_1, \ldots, \breve{h}_{L-1}]^T$.

III. Channel Estimation and Equalization

We employ recursive least squares (RLS) channel estimation with a forgetting factor $\lambda = 0.999$. Details of RLS estimation can be found in [14]. Note that $\lambda < 1$ allows the RLS algorithm to track a possibly time-varying CIR. RLS estimation offers the advantage that it can be easily extended to soft iterative data-aided channel estimation [9], while still providing a moderate complexity in contrast to maximum-likelihood data-aided channel estimation. Thereby, soft data symbol estimates are used to refine the initial training-based channel estimate. These soft symbol estimates are computed as the expected values of the data symbols using the extrinsic LLRs of the decoder given by $L_D(\mathbf{c}) = [L_D(\mathbf{c}_1)^T, \ldots, L_D(\mathbf{c}_N)^T]^T$, where $L_D(\mathbf{c}_i) = [L_D(c_{i,1}), \ldots, L_D(c_{i,m})]^T$ denotes the $m = \log_2 M$ extrinsic LLRs corresponding to the data symbol x_i. The LLR of a code bit c is defined by $L(c) = \ln(P(c = 0)/P(c = 1))$. Thus, for BPSK modulation the soft data symbol estimates are given by $\breve{s}_j = \tanh(L_D(c_i)/2)$, $i = 1, \ldots, N$, where $j = P + i$ denotes the position of the ith data symbol in the transmit frame. For later reference we introduce $v_j = 1 - \breve{s}_j^2$ as the corresponding variance estimate. In the case of general signal constellations, the expressions for soft symbol estimates and corresponding variance estimates can be found in [4]. Note that at positions j where pilot symbols p_j are transmitted, we define accordingly $\breve{s}_j = p_j$ and $v_j = 0$ ($j = 1, \ldots, P$).

Equalization is carried out with the UMVE-SIC. A detailed derivation of turbo equalization with the UMVE-SIC for block transmission schemes can be found in [15]. Here, we consider an implementation of the UMVE-SIC as a linear filter with $N_f = N_1 + N_2 + 1$ time-varying filter coefficients $\mathbf{w}_k = [w_{k,-N_1}, w_{k,-N_1+1}, \ldots w_{k,N_2}]^T$ that can be easily derived from [15], following the approach in [5]. Let us consider equalization of the kth symbol in the lth ($l = 1, 2, \ldots$) iteration of the turbo equalization scheme, where an iteration covers a single equalization and decoding step. The equalizer operates on a part of the received data vector, $\mathbf{y}_k \triangleq [y_{k-N_2}, y_{k-N_2+1}, \ldots, y_{k+N_1}]^T$, the channel estimate $\breve{\mathbf{h}}$, noise power estimate $\widetilde{P}_n$, and uses a vector of soft data symbol estimates $\breve{\mathbf{s}}_{\backslash k} = [\breve{s}_{k-L-N_2+1}, \ldots, \breve{s}_{k-1}, \breve{s}_{k+1}, \ldots, \breve{s}_{k+N_1}]^T$ with corresponding variance estimates $\mathbf{v}_{\backslash k} = [v_{k-L-N_2+1}, \ldots, v_{k-1}, v_{k+1}, \ldots, v_{k+N_1}]^T$, where the kth symbol is excluded according to the turbo principle. Let $\mathbf{H}$ denote the $N_f \times (N_f + L - 1)$ channel convolution matrix [5] based on the CIR $\mathbf{h}$ so that $\mathbf{y}_k = \mathbf{H}\mathbf{s}_k + \mathbf{n}_k$ with $\mathbf{s}_k \triangleq [s_{k-L-N_2+1}, s_{k-L-N_2+2}, \ldots, s_{k+N_1}]^T$ and $\mathbf{n}_k \triangleq [n_{k-N_2}, n_{k-N_2+1}, \ldots, n_{k+N_1}]^T$. Furthermore, let $\mathbf{H}_{\backslash k}$ be $\mathbf{H}$ without the $(N_2 + L)$th column $\mathbf{g} \triangleq [\mathbf{H}]_{:,N_2+L}$.

Then, after soft interference cancellation based on $\breve{\mathbf{s}}_{\backslash k}$, the remaining ISI is further reduced by the linear filter with coefficients $\mathbf{w}_k$ yielding finally the symbol estimate $\widetilde{x}_k$ of the data symbol x_k[2]

$$\widetilde{x}_k = \mathbf{w}_k^\dagger(\mathbf{y}_k - \mathbf{H}_{\backslash k}\breve{\mathbf{s}}_{\backslash k}) \quad \text{with} \quad \mathbf{w}_k^\dagger = \frac{\mathbf{g}^\dagger \mathbf{\Phi}_k^{-1}}{\mathbf{g}^\dagger \mathbf{\Phi}_k^{-1}\mathbf{g}} \quad \text{and} \quad \mathbf{\Phi}_k = \mathbf{H}_{\backslash k}\mathrm{Diag}(\mathbf{v}_{\backslash k})\mathbf{H}_{\backslash k}^\dagger + P_n \mathbf{I}_{N_f}, \tag{1}$$

where P_n is given by $N_0/2$ for real-valued and N_0 for complex-valued signal constellations, respectively. The extrinsic output LLRs $L_E(\mathbf{c}) = [L_E(\mathbf{c}_1)^T, \ldots, L_E(\mathbf{c}_N)^T]^T$ with $L_E(\mathbf{c}_k) = [L_E(c_{k,1}), \ldots, L_E(c_{k,m})]^T$ of the equalizer are finally obtained by assuming a Gaussian distribution of $\widetilde{x}_k$ with the output variance estimate $2\sigma_{E,k}^2 = 1/(\mathbf{g}^\dagger \mathbf{\Phi}_k^{-1}\mathbf{g})$ [15]. Since in our case the CIR $\mathbf{h}$ and the noise power P_n are unknown, we replace in (1) $\mathbf{H}_{\backslash k}$ with $\breve{\mathbf{H}}_{\backslash k}$, $\mathbf{g}$ with $\breve{\mathbf{g}}$ (similar definitions, but now based on the CIR estimate $\breve{\mathbf{h}}$), and P_n with a noise power estimate $\widetilde{P}_n$.

$L_E(\mathbf{c})$ is fed to the decoder after deinterleaving ($\mathbf{\Pi}^{-1}$) which computes extrinsic LLRs $L_D(\mathbf{c})$ and initiates the next iteration.

[1]Since the receiver does not know the CIR, the exact channel matched filter cannot be implemented.

[2]$\mathrm{Diag}(\cdot)$ creates a diagonal matrix from a vector argument.

IV. SNR ESTIMATION

The SNR estimation algorithms under consideration operate on the channel output symbols after (approximately) collecting the energy of the useful signal part, and (imperfect) cancellation of the ISI. Note that the noise sequence after this procedure is assumed to be still white, thus, this approach could be implemented by using a matched filter followed by a whitening filter with subsequent ISI cancellation. Since we have a system model on symbol basis, we may alternatively express the resulting symbols by first (imperfect) cancellation of the complete useful signal part and finally adding the estimated transmit signal weighted with the estimated energy yielding

$$\widetilde{y}_k = y_k - \sum_{l=0}^{L-1} \check{h}_l \breve{s}_{k-l} + \sqrt{E_{\check{h}}}\breve{s}_k \quad \text{with} \quad E_{\check{h}} = \sum_{l=0}^{L-1} |\check{h}_l|^2. \tag{2}$$

Thus, for perfect channel estimation $\check{\mathbf{h}} = \mathbf{h}$ and perfect symbol estimates $\breve{s}_k = s_k$, $\widetilde{y}_k$ in (2) would correspond to received symbols when transmitting over an AWGN channel yielding $\widetilde{y}_k = s_k + n_k$. In the first iteration, SNR estimation is based on the received symbols corresponding to known (training) symbols, only. In further iterations, soft data symbol estimates from the decoder may be exploited to extend estimation over the received data symbols. We compare three different standard SNR estimators; two of them are discussed in detail in [12]: A maximum-likelihood estimator (ML-E) with reduced bias and a second-and-fourth-order-moments estimator (M2M4-E). For real-valued symbols we have

$$\widetilde{\rho}_{\mathrm{ML}} = \frac{\left(\frac{1}{T}\sum\limits_{k=L}^{L+T-1} \widetilde{y}_k \breve{s}_k\right)^2}{\frac{1}{T-3}\sum\limits_{k=L}^{L+T-1} \widetilde{y}_k^2 - \frac{1}{T(T-3)}\left(\sum\limits_{k=L}^{L+T-1} \widetilde{y}_k \breve{s}_k\right)^2} \quad \text{and} \quad \widetilde{\rho}_{\mathrm{M2M4}} = \frac{\frac{1}{2}\sqrt{6M_2^2 - 2M_4}}{M_2 - \frac{1}{2}\sqrt{6M_2^2 - 2M_4}}, \tag{3}$$

where $M_2 = 1/T \sum_{k=L}^{L+T-1} \widetilde{y}_k^2$ and $M_4 = 1/T \sum_{k=L}^{L+T-1} \widetilde{y}_k^4$. The corresponding expressions for complex-valued symbols can be found in [12]. Estimation is based on in total T symbols; the first $L-1$ received symbols are not used for estimation due to the channel memory. In contrast to the M2M4-E, the ML-E is a data-aided estimator and may utilize both training symbols and extrinsic soft estimates of the data symbols. In addition, we include an approach that directly estimates the noise power by simply averaging over the squared magnitude of the estimated noise samples that corresponds to the minimum variance unbiased estimator of a nonrandom parameter [16, Ch. IV.C]. We refer to this estimator as the direct estimator (D-E)

$$\widetilde{\rho}_{\mathrm{D}} = \frac{E_{\check{h}}}{\widetilde{P}_{n,\mathrm{MVUE}}} \quad \text{with} \quad \widetilde{P}_{n,\mathrm{MVUE}} = \frac{1}{T-1}\sum_{k=L}^{L+T-1} |y_k - \sum_{l=0}^{L-1} \check{h}_l \breve{s}_{k-l}|^2. \tag{4}$$

Note that $\widetilde{P}_{n,\mathrm{MVUE}}$ in (4) has been applied to noise power estimation in [11]. The estimate $\widetilde{P}_n$ of the noise power P_n required for equalization is obtained from (3) or (4) by $E_{\check{h}}/\widetilde{\rho}$.

As a figure of merit for the performance of the SNR estimators we consider the normalized MSE defined by $\mu = \mathrm{E}\{(\widetilde{\rho} - \rho)^2/\rho^2\}$, as well as the normalized bias given by $b = \mathrm{E}\{(\widetilde{\rho} - \rho)/\rho\}$.

V. NUMERICAL RESULTS AND DISCUSSION

Figs. 2 and 3 show numerical results for the normalized MSE performance and the normalized bias of the SNR estimators under consideration. As an example for an ISI channel we choose channel (b) with the CIR $\mathbf{h} = [0.407\, 0.815\, 0.407]^T$ taken from [17]. One-shot training-based channel estimation with the RLS estimation algorithm is employed. Here, the decoder's output LLRs are modeled as i. i. d. Gaussian random variables for an arbitrarily chosen average mutual information of $\mathrm{I}\big(\mathbf{c}; L_D(\mathbf{c})\big) = 0.9$, see [18]. Therefore, note that the equalization and decoding step has been omitted, since both SNR and channel estimator require only received symbols and possibly soft feedback from the decoder. The number of training symbols is $P = 30$, and BPSK modulation with a data block length of $N_c = N = 1024$ is applied. When only training symbols are used for SNR estimation, the ML-E and D-E perform similarly as expected, see [16]. Interestingly, we observe that a-priori information improves the normalized MSE performance only in a certain SNR range. We observe an asymptotic (SNR$\rightarrow \infty$) MSE of 1 and an asymptotic bias of -1 if uncertain training information (soft data estimates) is used (all dashed lines except for *). This behavior can be intuitively justified by the following proposition concerning the much simpler case of an AWN channel.

Proposition 1: The asymptotic normalized MSE μ_∞ and asymptotic normalized bias b_∞ of the D-E for an AWGN channel with BPSK modulation and imperfect training information is given by $\mu_\infty = 1$ and $b_\infty = -1$, respectively.

Proof: Let $y_k = s_k + n_k$ be the received symbols and let $\breve{s}_k$ denote the (imperfect) symbol estimates for the transmit symbols s_k with average mutual information $\mathrm{I}\big(s_k; \breve{s}_k\big) < 1$. The noise power of the AWGN process with samples n_k is $P_n = N_0/2$. With increasing number of symbol estimates the time average in (4) arbitrarily closely approximates the corresponding expected value, i. e., we have $\widetilde{P}_n \rightarrow \mathrm{E}\{(y_k - \breve{s}_k)^2\}$. Exploiting that s_k and $\breve{s}_k$ are both statistically independent from the noise process we obtain after some manipulations $\widetilde{P}_n = 1 + P_n + E_{\breve{s}} - 2a$, where we defined a as $a \triangleq \mathrm{E}\{s_k \breve{s}_k\}$ and $E_{\breve{s}} \triangleq \mathrm{E}\{\breve{s}_k^2\}$.

Note that a and $E_{\breve{s}}$ are constants that depend only on the average mutual information $\mathrm{I}(s_k; \breve{s}_k)$. The SNR ρ is given by $\rho = 1/P_n$. Recalling the definition of the normalized bias we obtain $b = \widetilde{\rho}/\rho - 1 = P_n/(1 + P_n + E_{\breve{s}} - 2a) - 1$ and thus $b_\infty = \lim_{\rho\to\infty} b = -1$. The normalized MSE can be written as $\mu = (\widetilde{\rho}/\rho)^2 - 2\widetilde{\rho}/\rho + 1$ that finally yields $\mu_\infty = 1$. ■

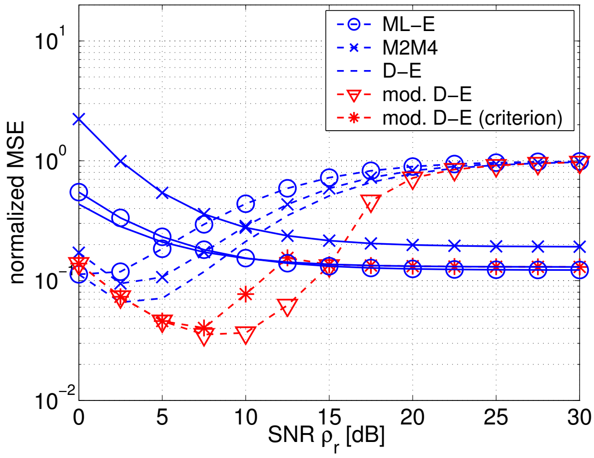

Fig. 2. Normalized MSE vs. SNR. Channel (b). One-shot channel estimation with training symbols. Solid lines: SNR estimation with training symbols. Dashed lines: SNR estimation with training and soft data symbol estimates.

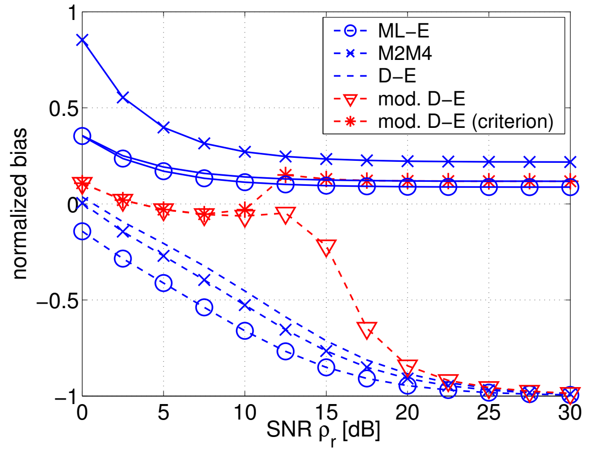

Fig. 3. Normalized bias vs. SNR. Channel (b). One-shot channel estimation with training symbols. Solid lines: SNR estimation with training symbols. Dashed lines: SNR estimation with training and soft data symbol estimates.

Motivated by this result we consider an improved adaptive version of the D-E, where the uncertainty of the training information is taken into account. $E_{\breve{s}}$ can be approximated by the time average of the N variance estimates of the soft symbols, i. e., $E_{\breve{s}} \approx 1 - \bar{v}$, where $\bar{v} = 1/N \sum_{k=P+1}^{P+N} v_k$. Moreover, assuming correct signs of the LLRs we have $a \approx \mathrm{E}\{|\breve{s}_k|\}$ which could be again approximated using $\bar{v}$ yielding $a \approx 1 - \bar{v}$. Thus, one obtains $\widetilde{P}_n \approx P_n + \bar{v}$, i. e., the D-E overestimates the noise power for uncertain training information. Subtracting the average variance $\bar{v}$ from $\widetilde{P}_n$ yields a modified adaptive (depending on the reliability of the decoder's output LLRs) version of the D-E that shows an improved MSE performance and a reduced bias in the low SNR regime according to Figs. 2 and 3. Since several approximations are employed and finite block lengths are considered one has to be careful with this subtraction. Therefore, we apply the modified D-E only, when the condition $\widetilde{P}_n > \alpha\bar{v}$ is satisfied. We found that $\alpha = 1.3$ yields satisfactory results. We remark that the idea of an correction of the initial noise power estimate $\widetilde{P}_n$ can also be found in [11], however, our approach has the advantage that it is adaptive.

The results in Figs. 2 and 3 suggest that also uncertain training information could be used adaptively, i. e., only if an improvement of the initial training-based SNR estimate can be expected. This idea corresponds to the suggestion made in [9] for soft iterative data-aided channel estimation. Figs. 2 and 3 depict the results for the modified D-E that uses soft data estimates only, when the criterion $\beta\widetilde{P}_{n,\mathrm{tr}} > \bar{v}$ is satisfied, where $\widetilde{P}_{n,\mathrm{tr}}$ denotes the noise power estimate for perfect training information. β is similar like α a heuristic factor, required due to the impossibility of an exact analysis. Thus, if the estimated noise power exceeds the variance estimate $\bar{v}$ it becomes more likely that the overall performance can be improved by the unreliable training information. Here, we choose $\beta = 1.5$. With this approach the poor asymptotic behavior of the modified D-E is circumvented.

Next we give examples for the BER performance of our adaptive turbo equalization scheme. Since according to Figs. 2 and 3 the best performance in the case of imperfect training information is obtained with the D-E, we focus on this estimator in the following. We apply a channel code of rate $1/2$, memory 2, with generators $(1 + D^2, 1 + D + D^2)$. The number of source bits is $N_q = 510$, yielding $N_c = N = 1024$ for BPSK modulation. The number of turbo iterations is set to 5. In the first example shown in Fig. 4, we consider data transmission over channel (b). The numbers of filter coefficients in the equalizer are $N_1 = 5$ and $N_2 = 3$, the number of training symbols is $P = 30$. The gain of the modified D-E over the D-E is in the order of 0.5 dB, but could be further enhanced when applying a more complicated criterion to compensate for the uncertainty of the training information (not shown). The usage of the criterion $\beta\widetilde{P}_{n,\mathrm{tr}} > \bar{v}$ did not show an improvement compared to the modified D-E, therefore corresponding results are omitted. We remark that the degradation of the modified D-E observed in Figs. 2 and 3 for high SNR is not visible in Fig. 4, since we operate in the low SNR regime (here ρ_r corresponds directly to E_b/N_0 values) and with varying amount of a-priori information. However, the D-E with soft iterative channel estimation indeed shows a slightly worse performance than the purely training-based estimation approach. Furthermore, note that the gap from the modified D-E to perfect SNR knowledge is only about 0.5 dB when using soft iterative channel estimation.

In the second example we consider a scenario with a larger amount of ISI using channel (c) with $\mathbf{h} = [0.227\ 0.46\ 0.688\ 0.46\ 0.227]^T$ from [17]. The numbers of filter coefficients in the equalizer are now $N_1 = 9$ and $N_2 = 5$,

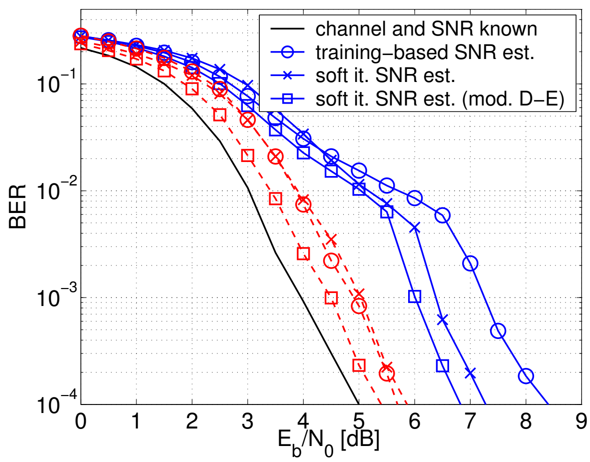

Fig. 4. BER vs. E_b/N_0. Channel (b) and BPSK modulation with $P = 30$ training symbols. Solid lines: one-shot training-based channel estimation. Dashed lines: soft iterative data-aided channel estimation. SNR estimation with the D-E/modified D-E. 5 turbo iterations.

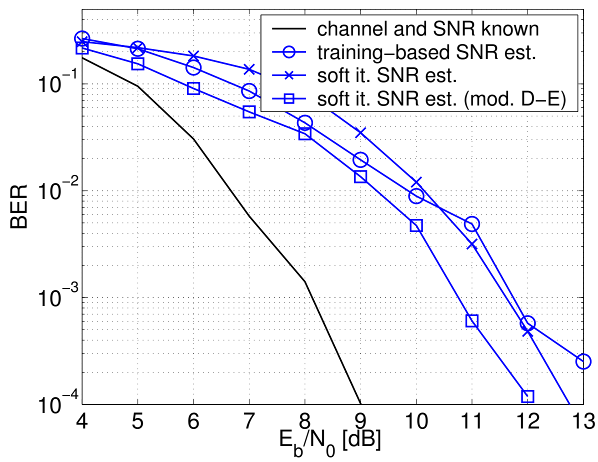

Fig. 5. BER vs. E_b/N_0. Channel (c) and BPSK modulation using $P = 100$ training symbols. Soft iterative data-aided channel estimation. SNR estimation with the D-E/modified D-E. 5 turbo iterations.

the number of training symbols is $P = 100$. All other parameters are chosen as described above. The BER performance is shown in Fig. 5. Here, the improvement of the modified D-E over the D-E is more prominent. However, there is still a large gap of about 3 dB to the curve corresponding to perfect SNR and channel knowledge. Note that the curves with one-shot training-based channel estimation have been omitted, since the performance has been only slightly worse compared to iterative channel estimation.

We conclude that by taking into account the uncertainty of imperfect training information, the SNR estimation can be improved. Future work will consider the case of multi-amplitude signal constellations and time-varying channels.

Acknowledgment

The authors would like to thank Christian Sgraja for many valuable discussions, as well as the anonymous reviewers for useful suggestions. The support by the COST 289 Action is gratefully acknowledged.

References

[1] J. Hagenauer, "The turbo principle: Tutorial introduction and state of the art," in *Proc. Int. Symp. on Turbo-Codes*, Brest, France, Sept. 1997, pp. 1–11.
[2] C. Douillard, M. Jézéquel, C. Berrou, A. Picart, P. Didier, and A. Glavieux, "Iterative correction of intersymbol interference: Turbo-equalization," *European Trans. Telecomm.*, vol. 6, no. 5, pp. 507–511, Sept./Oct. 1995.
[3] A. Glavieux, C. Laot, and J. Labat, "Turbo equalization over a frequency selective channel," in *Proc. Int. Symp. on Turbo-Codes*, Brest, France, Sept. 1997, pp. 96–102.
[4] C. Sgraja, A. Engelhart, W. G. Teich, and J. Lindner, "Combined equalization and decoding for BFDM packet transmission schemes," in *Proc. 1st Int. OFDM-Workshop*, Hamburg, Germany, Sept. 1999, pp. 19–1ff.
[5] M. Tüchler, R. Koetter, and A. Singer, "Turbo equalization: Principles and new results," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 754–767, May 2002.
[6] J. Egle and J. Lindner, "Iterative joint equalization and decoding based on soft Cholesky equalization for general complex valued modulation symbols," in *Proc. 6th Int. Symp. on DSP for Communication Systems*, Sydney-Manly, Australia, Jan. 2002, pp. 163–170.
[7] A. Bury and J. Lindner, "Block transmission equalizers using constrained minimum variance filters with application to MC-CDM," in *Proc. IEEE ISSSTA*, New Jersey, USA, Sept. 2000, pp. 159–163.
[8] L. M. Davis, I. B. Collings, and P. Hoeher, "Joint MAP equalization and channel estimation for frequency-selective and frequency-flat fast-fading channels," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2106–2114, Dec. 2001.
[9] M. Tüchler, R. Otnes, and A. Schmidbauer, "Performance of soft iterative channel estimation in turbo equalization," in *Proc. IEEE ICC*, New York, USA, Apr./May 2002.
[10] J. Garcia-Frias and J. D. Villasenor, "Combined turbo detection and decoding for unknown ISI channels," *IEEE Trans. Commun.*, vol. 50, no. 1, pp. 79–85, Jan. 2003.
[11] S. Talakoub and B. Shahrrava, "Turbo equalization with iterative online SNR estimation," in *Proc. IEEE WCNC*, New Orleans (LA), USA, Mar. 2005.
[12] D. R. Pauluzzi and N. C. Beaulieu, "A comparison of SNR estimation techniques for the AWGN channel," *IEEE Trans. Commun.*, vol. 48, no. 10, pp. 1681–1691, Oct. 2000.
[13] S. Dolinar and D. Divsalar, "Weight distributions for turbo codes using random and nonrandom permutations," *JPL TDA Progress Report*, pp. 56–65, Aug. 1995.
[14] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, New Jersey: Prentice Hall, 2002.
[15] M. A. Dangl, C. Sgraja, and J. Lindner, "Block turbo equalization for imperfect channel state information," in *Proc. IEEE Int. Symp. Inform. Theory*, Adelaide, Australia, Sept. 2005, to appear.
[16] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York: Springer, 1998.
[17] J. G. Proakis, *Digital Communications*, 3rd ed. McGraw-Hill, 1998.
[18] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.