Turbo Equalization Based on Bi-directional DFE

Seongwook Jeong
Dept. of Electrical and Computer Engineering
University of Minnesota
Minneapolis, Minnesota 55455
Email: jeong030@umn.edu

Jaekyun Moon
Dept. of Electrical and Computer Engineering
University of Minnesota
Minneapolis, Minnesota 55455
Email: moon@umn.edu

***Abstract*—We utilize a pair of decision feedback equalizers (DFEs) operating in opposite directions in turbo equalization setting to remove the effect of intersymbol interference (ISI) at the receiver. With a specific residual interference processing strategy proposed, the bi-directional DFE (BiDFE) are free from any significant error propagation. A diversity combining scheme is also proposed that effectively combines the extrinsic outputs of two opposite direction DFEs to minimize any correlation that may exist between them. A resulting scheme is a low-complexity equalizer that closely approaches the performance of the much more complex BCJR algorithm at the expense of an increased number of decoder-equalizer iterations. The BiDFE turbo equalizer also provides considerably better performance than the well-known soft linear minimum mean square error (MMSE) equalizer in severe ISI channels. The performance advantages are validated with bit-error-rate (BER) simulations and extrinsic information transfer (EXIT) charts analysis.**

I. Introduction

Inter-symbol interference (ISI) arises as the transmitted signal in a digital communication system overlaps with the past and/or future signals. Various equalization methods are introduced in order to cancel out or suppress the ISI terms in the received data sequence at the receiver. Powerful modern equalization methods are based on the turbo equalization principle established in [1], wherein a soft-in soft-out (SISO) detector and a SISO decoder exchange extrinsic information in an iterative fashion until reliable decisions are generated. It has been shown in [1] that for some example channels the detrimental effect of ISI disappears with this approach.

The detector or the equalizer portion of the turbo equalization system is based on the well-known Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm [2]. This algorithm computes the *a posteriori* probability (APP) of the transmitted signal symbols considering the channel and the *a priori* information of the transmitted symbols and is optimum in the sense of achieving the minimum bit error rate (BER) performance. However, the computational complexity of this maximum APP (MAP) equalization via the BCJR algorithm grows exponentially as a function of the channel length and the symbol alphabet set size.

The high computational complexity of MAP equalization has motivated considerable research on numerous suboptimal but low complexity equalization schemes. One suboptimal channel equalization method is the classical decision feedback equalizer (DFE) [3] modified to take soft inputs and generate soft outputs [4]. However, when ISI is severe with the channel response showing nulls or deep valleys in the Nyquist band, turbo equalization based on the DFE does not perform as well as the one based on the BCJR algorithm and even the linear mean square error (LMMSE) equalization [4]. It appears that the inherent error propagation phenomenon for DFE degrades turbo equalization performance. In order to mitigate the error propagation in DFE, many techniques [5], [6], [7] have been investigated. Recently, it has been shown [8], [9], [10] that the employment of normal and time-reversed equalization of the received data sequence with two DFEs along with a proper combining of DFE outputs is very effective for reducing error propagation and improving BER performance. This bi-directional DFE (called BiDFE) algorithm takes advantage of the different decision error and noise distributions at the outputs of the normal and time-reversed DFEs [8], [9].

In this paper, we also investigate SISO DFE for turbo equalization settings. We in particular employ BiDFE for turbo equalization. The novelty of this paper is in 1) a new soft decision feedback method that estimates and suppresses the effect of residual ISI due to potential incorrect past decisions, and in 2) a specific DFE combining strategy that suppresses statistical correlation between the outputs of two opposite direction DFEs before passing the soft equalizer output to the SISO decoder. We show that the resulting performance approaches the performance of the BCJR-based turbo equalizer, easily outperforming the turbo equalizer based on soft LMMSE of [4]. Remarkably, the performance of a time-invariant version of the BiDFE, a lower-complexity method that does not require tap-weight updating as a function of time, also consistently is better than the soft LMMSE scheme of [4] based on a time-varying linear filter, and again comes very close to the optimal turbo equalizer performance.

The remainder of the paper is organized as follows. In Section II, a brief statement of the problem is given. In Section III, we review and then modify the extrinsic information of DFE derived in [4]. We introduce the iterative BiDFE algorithm with the optimally derived combiner of the extrinsic information using the normal and time-reversed DFE outputs in Section IV. In Section V, numerical results and analysis are given. Finally, we draw conclusions in Section VI.

II. System Model

In this paper, it is assumed that the receiver knows the discrete-time baseband channel response accurately and the received data sequence is corrupted by additive white Gaussian noise (AWGN). Given the transmitted sequence of coded bits $\{x_k\}$, the channel output r at time n is

$$r_n = \sum_{k=0}^{L_h-1} h_k x_{n-k} + w_n$$

where w_n is AWGN with variance σ_w^2, and $\{h_k\}$ is an ISI channel impluse response with length L_h.

In turbo equalization, the *a priori* log-likelihood ratio (LLR) is defined as

$$L_a(x_n) \triangleq \ln \frac{\Pr(x_n = +1)}{\Pr(x_n = -1)}.$$

Then, an equalizer computes the *a posteriori* LLR of x_n,

$$L(x_n) \triangleq \ln \frac{\Pr(x_n = +1 \mid \mathbf{r}_n)}{\Pr(x_n = -1 \mid \mathbf{r}_n)}$$

where $\mathbf{r}_n$ is a received sequence sample block at time n. Finally, the extrinsic LLR of x_n is passed to the error-correction code decoder:

$$L_e(x_n) \triangleq L(x_n) - L_a(x_n)$$

The extrinsic information generated by the decoder then is passed to the SISO equalizer as the *a priori* LLR sequence in computing its own extrinsic information sequence.

III. Derivation of Modified Iterative DFE Algorithm

The classical DFE consists of a linear feedforward filter suppressing the ISI due to the future symbols, a linear feedback filter removing the ISI due to the past symbols through the previously determined symbols, and a decision device. The results of two linear filters are added and then applied to a decision device to determine the current symbol for the equalization of the future symbols.

A. Review of the Conventional Extrinsic LLR Mapping

The MMSE-DFE filter taps in turbo equalization are calculated according to the well-known MMSE criterion, except that the *a priori* information from the decoder is utilized in constructing the average information $\mathrm{E}(x_n)$. In [4], it was shown that MMSE feedforward filter $\mathbf{c}_n$ with length L_c+1 and feedback filter $\mathbf{d}_n$ with length $L_d = L_h - 1$ are time-varying and given by

$$\begin{aligned} \mathbf{c}_n &\triangleq \left[c_{\{n,0\}}, c_{\{n,+1\}}, \ldots, c_{\{n,L_c\}}\right]^T \\ &= \left\{\mathbf{H}\boldsymbol{\Sigma}_n\mathbf{H}^T + (1-z_n)\mathbf{s}\mathbf{s}^T + \sigma_w^2\mathbf{I}\right\}^{-1}\mathbf{s} \quad (1) \\ \mathbf{d}_n &\triangleq \left[d_{\{n,-L_d\}}, d_{\{n,-L_d+1\}}, \ldots, d_{\{n,-1\}}\right]^T \\ &= \mathbf{M}\mathbf{H}^T\mathbf{c}_n \quad (2) \\ y_n &= \mathbf{c}_n^T \cdot (\mathbf{r}_n - \mathbf{H}\bar{\mathbf{x}}_n + \mathrm{E}(x_n)\mathbf{s}) \quad (3) \end{aligned}$$

where $\mathbf{H}$ is a convolutional matrix defined as

$$\mathbf{H} \triangleq \begin{bmatrix} h_{L_h-1} & h_{L_h-2} & \cdots & h_0 & 0 & \cdots & 0 \\ 0 & h_{L_h-1} & h_{L_h-2} \cdots & h_0 & 0 & \cdots & 0 \\ & \ddots & & \ddots & & \ddots & \\ 0 & 0 & \cdots & 0 & h_{L_h-1} & h_{L_h-2} \cdots & h_0 \end{bmatrix},$$

$\boldsymbol{\Sigma}_n \triangleq \mathrm{Diag}(\mathbf{0}_{1\times L_d}, z_n, z_{n+1}, \ldots, z_{n+L_c})$, $z_n \triangleq 1 - \mathrm{E}(x_n)^2$, $\mathbf{s} \triangleq \mathbf{H}[\mathbf{0}_{1\times L_d}, 1, \mathbf{0}_{1\times L_c}]^T$, and $\mathbf{M} \triangleq [\mathbf{I}_{L_d\times L_d}, \mathbf{0}_{L_d\times(L_c+1)}]$. The received sequence is also defined as $\mathbf{r}_n \triangleq [r_n, r_{n+1}, \ldots, r_{n+L_c}]^T$ and $\bar{\mathbf{x}}_n \triangleq [\hat{x}_{n-L_d}, \ldots, \hat{x}_{n-1}, \mathrm{E}(x_n), \ldots, \mathrm{E}(x_{n+L_c})]^T$ where $\hat{x}_n$ is the available decision for x_n based on the *a posteriori* LLR of x_n, i.e., if $L(x_n) \geq 0$, then, $\hat{x}_n = 1$, otherwise, $\hat{x}_n = -1$.

Let us define the anticausal symbol sequence $\mathbf{x}_n \triangleq [x_n, x_{n+1}, \ldots, x_{n+L_c}]^T$, the causal symbol sequence $\mathbf{x}_n^c \triangleq [x_{n-L_d}, x_{n-L_d+1}, \ldots, x_{n-1}]^T$, and the available decision sequence $\hat{\mathbf{x}}_n^c \triangleq [\hat{x}_{n-L_d}, \hat{x}_{n-L_d+1}, \ldots, \hat{x}_{n-1}]^T$. The noise sequence is also defined as $\mathbf{w}_n \triangleq [w_n, w_{n+1}, \ldots, w_{n+L_c}]^T$. Then, the combined filter output y_n can be rewritten as

$$\begin{aligned} y_n &= (\mathbf{c}_n^T\tilde{\mathbf{H}}) \cdot \big(\mathbf{x}_n - \mathrm{E}\{\dot{\mathbf{x}}_n\}\big) + \mathbf{d}_n^T(\mathbf{x}_n^c - \hat{\mathbf{x}}_n^c) + \mathbf{c}_n^T\mathbf{w}_n \\ &= p_{\{n,0\}}x_n + i_n + \sum_{k=1}^{L_c} p_{\{n,k\}}\big(x_{n+k} - \mathrm{E}(x_{n+k})\big) + w_n' \\ &= p_{\{n,0\}}x_n + i_n + v_n \end{aligned} \quad (4)$$

where $\mathrm{E}\{\dot{\mathbf{x}}_n\} \triangleq [0, \mathrm{E}(x_{n+1}), \mathrm{E}(x_{n+2}), \ldots, \mathrm{E}(x_{n+L_c})]^T$ and $\tilde{\mathbf{H}}$ is a $(L_c+1)\times(L_c+1)$ submatrix of $\mathbf{H}$ formed by the entire rows of the columns from (L_d+1)th to the end. Moreover, $w_n' \triangleq \mathbf{c}_n^T\mathbf{w}_n$, $\mathbf{p}_n \triangleq \left[p_{\{n,0\}}, p_{\{n,1\}}, \ldots, p_{\{n,L_c\}}\right]^T = \tilde{\mathbf{H}}^T\mathbf{c}_n$ and $p_{\{n,0\}} = \mathbf{s}^T\mathbf{c}_n$. The error propagation caused by the mismatched hard decision feedback is denoted as i_n, i.e., $i_n \triangleq \mathbf{d}_n^T(\mathbf{x}_n^c - \hat{\mathbf{x}}_n^c)$ and v_n denotes the sum of the noise and the remaining ISI terms caused by the neighboring symbols: $v_n \triangleq \sum_{k=1}^{L_c} p_{\{n,k\}}\big(x_{n+k} - \mathrm{E}(x_{n+k})\big) + w_n'$. The variance of v_n is also derived in [4],

$$\mathrm{Var}(v_n) = \mathbf{c}_n^T\mathbf{s}(1 - \mathbf{s}^T\mathbf{c}_n) \quad (5)$$

Assuming that the feedback decisions are all correct, i.e., $i_n = 0$, and v_n is AWGN, the LLR is naturally as

$$\acute{L}_e(x_n) = \frac{2p_{\{n,0\}}y_n}{\mathrm{Var}(v_n)}. \quad (6)$$

B. Generation of the Modified Extrinsic Information

While the MAP estimation of i_n is equal to zero, we observe that the variance of i_n, $\mathrm{Var}(i_n)$, cannot be negelected, especially for severe ISI. Since i_n is determined on the basis of observations $\mathbf{y}_n^c \triangleq [y_{n-L_d}, y_{n-L_d+1}, \ldots, y_{n-1}]^T$, the $\mathrm{Var}(i_n)$ is defined as the conditional variance $\mathrm{Var}(i_n \mid \mathbf{y}_n^c)$.

$$\mathrm{Var}(i_n) \triangleq \mathrm{Var}(i_n \mid \mathbf{y}_n^c) = \mathbf{d}_n^T\acute{\boldsymbol{\Sigma}}_n^c\mathbf{d}_n \quad (7)$$

where $\acute{\boldsymbol{\Sigma}}_n^c \triangleq \mathrm{Diag}\left(\acute{z}_{n-L_d}, \acute{z}_{n-L_d+1}, \ldots, \acute{z}_{n-1}\right)$ and $\acute{z}_n \triangleq \mathrm{Var}(x_n \mid y_n) = 1 - \mathrm{E}(x_n \mid y_n)^2$.

Letting $u_n \triangleq v_n + i_n$ and assuming that u_n is AWGN and i_n is independent of x_n and v_n, we modify the LLR of (6) to

$$L_e(x_n) = \frac{2p_{\{n,0\}}y_n}{\mathrm{Var}(u_n)} = \frac{2p_{\{n,0\}}y_n}{\mathrm{Var}(v_n) + \mathrm{Var}(i_n)}. \quad (8)$$

When we compare (8) with (6), the same polarity of extrinsic LLR for x_n is maintained, but its magnitude is decreased in order to account for the i_n term in (4). However, estimating the variance of i_n in (7) via (the probability conversion of) $L(\mathbf{x}_n^c) \triangleq L_e(\mathbf{x}_n^c) + L_a(\mathbf{x}_n^c)$ tends to overestimate the overall noise-interference variance term because of the correlation in APP LLRs; the estimation of $\mathrm{Var}(i_n)$ based on $\acute{L}(\mathbf{x}_n^c) \triangleq \acute{L}_e(\mathbf{x}_n^c) + L_a(\mathbf{x}_n^c)$ works better in BER simulation. Overall, the variance of i_n is obtained through (7) with $\acute{z}_n = 1 - \tanh(\acute{L}(x_n)/2)$.

In essence, we are keeping two different pieces of extrinsic information for x_n, the conventional form of the extrinsic LLR, $\acute{L}_e(x_n)$, for the computation of the residual ISI variance for future processing stages, and the residual-ISI-adjusted extrinsic LLR, $L_e(x_n)$ of (8), to feed the outer decoder as well as for generating hard decisions that will drive the feedback filter. In practice, we make one more exception to the above rule. When the polarities of $L(x_n)$ and $\acute{L}(x_n)$ do not agree (e.g., when $L_a(x_n) = +2$ and the MMSE-DFE algorithm computes $L_e(x_n) = -1$ and $\acute{L}_e(x_n) = -3$), we set $\acute{L}_e(x_n) = L_e(x_n) = -L_a(x_n)$ so that $\acute{L}(x_n) = L(x_n) = 0$. This special provision handles symbols for which the equalizer has a low level of confidence for its decision.

Let us summarize our LLR generation method: 1) The extrinsic information $\acute{L}_e(x_n)$ is set according to either (6) or by $\acute{L}_e(x_n) = -L_a(x_n)$, if the polarities of $L(x_n)$ and $\acute{L}(x_n)$ do not agree. 2) $\mathrm{Var}(i_n)$ is computed using past values of $\acute{L}(x_n) = \acute{L}_e(x_n) + L_a(x_n)$. 3) The extrinsic information $L_e(x_n)$ of (8) based on $\mathrm{Var}(i_n)$ as obtained in Step 2) is used to form APP LLR $L(x_n) = L_e(x_n) + L_a(x_n)$ for making hard decisions that drive the feedback filter. 4) Modify the extrinsic information $L_e(x_n)$ so that it is either (8) or simply given by $L_e(x_n) = -L_a(x_n)$, when the polarities of $L(x_n)$ and $\acute{L}(x_n)$ do not agree. This modified extrinsic information feeds the outer decoder.

C. Reducing Complexity

A low-complexity turbo equalization scheme is achieved with the time-varying filters $\mathbf{c}_n$ and $\mathbf{d}_n$ replaced by the time-invariant filters $\mathbf{c}$ and $\mathbf{d}$ as we set $\mathrm{E}(x_n) = 0$ for all n [4]:

$$\begin{aligned} \mathbf{c} &\triangleq [c_0, c_{+1}, \ldots, c_{L_c}]^T \\ &= \left(\mathbf{H}\boldsymbol{\Sigma}\mathbf{H}^T + \sigma_w^2 \mathbf{I}\right)^{-1} \mathbf{s} \qquad (9) \\ \mathbf{d} &\triangleq [d_{-L_d}, d_{-L_d+1}, \ldots, d_{-1}]^T \\ &= \mathbf{M}\mathbf{H}^T \mathbf{c} \qquad (10) \end{aligned}$$

where $\boldsymbol{\Sigma} \triangleq \mathrm{Diag}(\mathbf{0}_{1\times L_d}, \mathbf{1}_{1\times(L_c+1)})$.

Then, the DFE with the time-invariant filters $\mathbf{c}$ and $\mathbf{d}$ yields an equalized symbol,

$$\begin{aligned} y_n &= \mathbf{c}^T \cdot \left(\mathbf{r}_n - \mathbf{H}\bar{\mathbf{x}}_n + \mathrm{E}(x_n)\mathbf{s}\right) \\ &= p_0 x_n + i_n + v_n \qquad (11) \end{aligned}$$

where $p_0 = \mathbf{s}^T\mathbf{c}$ and $i_n = \mathbf{d}^T(\mathbf{x}_n^c - \hat{\mathbf{x}}_n^c)$. The noise variance of v_n and the variance of i_n with the time-invariant filters are also given,

$$\begin{aligned} \mathrm{Var}(v_n) &= \mathbf{c}^T \left(\mathbf{H}\boldsymbol{\Sigma}_n\mathbf{H}^T - z_n \mathbf{s}\mathbf{s}^T + \sigma_w^2 \mathbf{I}\right)\mathbf{c} \qquad (12) \\ \mathrm{Var}(i_n) &= \mathbf{d}^T \acute{\boldsymbol{\Sigma}}_n^c \mathbf{d}. \qquad (13) \end{aligned}$$

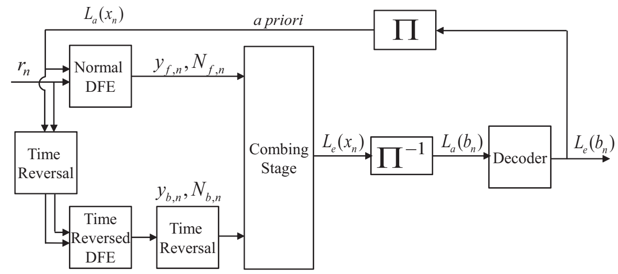


Fig. 1: Iterative Equalization Scheme based on BiDFE

IV. Derivation of Iterative BiDFE Algorithm

We now discuss an iterative BiDFE algorithm. The idea of BiDFE is already motivated in [8], [9] by the fact that DFE can be performed on the reversed received sequence using the time-reversed channel response. Iterative equalization schemes based on BiDFE are shown in Fig. 1. Basically, the channel equalizer is a SISO equalizer which employs the normal DFE, the time-reversed DFE, and an LLR combining block. The received data sequence is equalized in both directions by the two DFEs, and the extrinsic information from two DFEs are combined and passed to the error correction code decoder. We show that a proper combining of the two sets of extrinsic information can suppress error propagation and noise further and generate more reliable extrinsic information for the outer decoder.

A. Combining Extrinsic Information

Without loss of generality, the unbiased equalizer output [11] corresponding to the transmitted coded symbol X_n, where $X_n = x_n$, from the forward DFE and backward DFE can be represented respectively as

$$\begin{aligned} Y_{f,n} &= X_n + I_{f,n} + V_{f,n} = X_n + U_{f,n} \\ Y_{b,n} &= X_n + I_{b,n} + V_{b,n} = X_n + U_{b,n} \end{aligned}$$

For notational simplicity, we further drop the time index n with an understanding that processing remains identical as n progresses: $Y_f = X + U_f$ and $Y_b = X + U_b$. The noise U_f and U_b are assumed to be zero mean Gaussian random variables which are independent with the coded data X but correlated to each other with the correlation coefficient ρ. Our strategy is to whiten the noise U_f and U_b before combining. The noise correlation matrix $\mathbf{R}$ is defined as,

$$\mathbf{R} \triangleq \begin{bmatrix} \mathrm{Var}(U_f) & \mathrm{E}(U_f U_b) \\ \mathrm{E}(U_f U_b) & \mathrm{Var}(U_b) \end{bmatrix} = \begin{bmatrix} N_f & \rho\sqrt{N_f N_b} \\ \rho\sqrt{N_f N_b} & N_b \end{bmatrix}$$

where $N_f \triangleq \mathrm{Var}(U_f)$ and $N_b \triangleq \mathrm{Var}(U_b)$. Then, the eigenvalues of the noise correlation matrix, $\lambda_{1,2}$, with their corresponding normalized eigenvectors $\mathbf{g}_{1,2}$ are given as

$$\lambda_{1,2} = \frac{(N_f + N_b) \pm \sqrt{(N_f - N_b)^2 + 4\rho^2 N_f N_b}}{2}$$

$$\mathbf{g}_1 \triangleq \frac{1}{\sqrt{g_{11}^2 + g_{21}^2}} \begin{bmatrix} g_{11} \\ g_{21} \end{bmatrix}, \ \mathbf{g}_2 \triangleq \frac{1}{\sqrt{g_{12}^2 + g_{22}^2}} \begin{bmatrix} g_{12} \\ g_{22} \end{bmatrix}$$

where $g_{11,12} = \frac{1}{2}\left[(N_f - N_b) \pm \sqrt{(N_f - N_b)^2 + 4\rho^2 N_f N_b}\right]$ and $g_{21} = g_{22} = \rho\sqrt{N_f N_b}$. Therefore, the noise correlation matrix $\mathbf{R}$ is non-singular unless $\rho = \pm 1$. If $\mathbf{R}$ is non-singular, $\mathbf{R}$ can be expanded as $\mathbf{R} = \mathbf{G}\mathbf{\Lambda}\mathbf{G}^{-1}$ where $\mathbf{G} = \left[\mathbf{g}_1\, \mathbf{g}_2\right]$ and $\mathbf{\Lambda} = \mathrm{Diag}(\lambda_1, \lambda_2)$. Since $\mathbf{G}$ is a unitary matrix, the noise whitening matrix is $\mathbf{A} \triangleq \mathbf{G}^{-1} = \mathbf{G}^T$. Then, let us define the equalized output vector $\mathbf{Y} \triangleq [Y_f, Y_b]^T$ and its uncorrelated output vector $\mathbf{Y}' \triangleq [Y_f', Y_b']^T = \mathbf{A}\mathbf{Y}$ generating the noise correlation matrix $\mathbf{R}' \triangleq \mathbf{A}\mathbf{R}\mathbf{A}^T = \mathbf{\Lambda}$. The extrinsic information of X can now be expressed as

$$
\begin{aligned}
L_e(X) &\triangleq \ln\frac{\Pr(Y_f', Y_b' \mid X = +1)}{\Pr(Y_f', Y_b' \mid X = -1)} \\
&= \ln\frac{\Pr(Y_f' \mid X = +1)}{\Pr(Y_f' \mid X = -1)} + \ln\frac{\Pr(Y_b' \mid X = +1)}{\Pr(Y_b' \mid X = -1)} \\
&= \frac{2\left(N_b - \rho\sqrt{N_f N_b}\right)}{(1-\rho^2)\, N_f N_b} Y_f + \frac{2\left(N_f - \rho\sqrt{N_f N_b}\right)}{(1-\rho^2)\, N_f N_b} Y_b
\end{aligned}
\tag{14}
$$

For the singular noise correlation matrix $\mathbf{R}$ (i.e., $\rho = +1$), $N_f = N_b = N$ and $Y_f = Y_b = Y$. Consequently, the extrinsic information of X becomes $L_e(X) = 2Y/N$. If $\rho = -1$, $U_f = -U_b$ and we can cancel out the noise perfectly by averaging the outputs: $(Y_f + Y_b)/2$. The extrinsic information of X in this case is $L_e(X) = +\inf$ when $(Y_f + Y_b)/2 \geq 0$ while $L_e(X) = -\inf$ when $(Y_f + Y_b)/2 < 0$.

B. Estimation of Correlation Coefficient

The noise correlation coefficient between the normal (forward) and the time-reversed (backward) DFE is defined as

$$
\begin{aligned}
\rho_n &\triangleq \frac{\mathrm{E}(U_{f,n} U_{b,n})}{\sqrt{\mathrm{Var}(U_{f,n})}\sqrt{\mathrm{Var}(U_{b,n})}} \\
&= \frac{\mathrm{E}((I_{f,n} + V_{f,n})(I_{b,n} + V_{b,n}))}{\sqrt{N_{f,n}}\sqrt{N_{b,n}}}.
\end{aligned}
$$

Unfortunately, it is difficult to compute the correlation coefficient analytically. However, assuming that the noise is stationary, we have $\rho_n = \rho$, and the correlation coefficient can be estimated through time-averaging:

$$
\hat{\rho} = \frac{\sum\left\{(Y_{f,n} - \hat{X}_{f,n})(Y_{b,n} - \hat{X}_{b,n})\right\}}{\sqrt{\sum(Y_{f,n} - \hat{X}_{f,n})^2}\sqrt{\sum(Y_{b,n} - \hat{X}_{b,n})^2}} \tag{15}
$$

where the summations are over some reasonably large finite window. Note that the hard decisions for the transmitted symbols in normal and time-reversed DFEs might be different; in estimating the correlation coefficient, we only consider those noise samples $U_{f,n}$ and $U_{b,n}$ for which $\hat{X}_{f,n}$ and $\hat{X}_{b,n}$ are identical.

C. Reducing Sensitivity of Combiner to Estimation Error

Let us consider the effect of errors in estimating ρ on extrinsic information. Write $\hat{\rho} = \rho + \Delta_\rho$ where Δ_ρ is the estimation error. Then, the sensitivity of the combiner in (14) to the estimation error is defined as

$$
\begin{aligned}
\mathrm{S}(\rho) &\triangleq \left|\frac{\partial L_e(X)}{\partial \rho}\right| \\
&= \left|\frac{2\left(2\rho N_b - (1+\rho^2)\sqrt{N_f N_b}\right)}{\left(1-\rho^2\right)^2 N_f N_b} Y_f \right. \\
&\quad \left. + \frac{2\left(2\rho N_f - (1+\rho^2)\sqrt{N_f N_b}\right)}{\left(1-\rho^2\right)^2 N_f N_b} Y_b \right|
\end{aligned}
\tag{16}
$$

which approaches infinity as $\rho \to \pm 1$. This means that the combiner of (14) is unfortunately very sensitive to the correlation estimator error, as the magnitude of the correlation becomes large.

The sensitivity of the combiner can be reduced if we assume that the normal DFE and the time-reversed DFE produce the same noise variance, i.e., $N = N_f = N_b = (N_f + N_b)/2$. This assumption is reasonable when the same feedforward and feedback filter length is used in both DFEs. Then, from (14), the combined extrinsic information for X for non-singular $\mathbf{R}$ is simply given as

$$
L_e(X) = \frac{2(Y_f + Y_b)}{(1+\rho)N} \tag{17}
$$

with the sensitivity to the correlation estimation error

$$
\mathrm{S}(\rho) = \left|\frac{2(Y_f + Y_b)}{(1+\rho)^2 N}\right|. \tag{18}
$$

Although the sensitivity of this combiner to the estimation error also goes to infinity as $\rho \to -1$, it shows more robustness than the previous combiner as $\rho \to +1$ since $\lim_{\rho \to +1} \mathrm{S}(\rho) = |(Y_f + Y_b)/2N|$. Moreover, the sensitivity of this combiner as $\rho \to 0$ is also smaller than that of the combiner in (14) since $N = (N_f + N_b)/2 \geq \sqrt{N_f N_b}$.

V. Simulation Results

In this section, simulation results of several iterative equalization schemes are presented. The transmitted symbols are encoded with a recursive rate-$1/2$ convolutional code encoder with parity generator $(1 + D^2)/(1 + D + D^2)$ with 4096-bit interleaver and are modulated by BPSK so that $x_n \in \{-1, 1\}$. We also assume that the noise is AWGN and the noise variance and the channel information is perfectly known to the receiver. The impulse response of the ISI channel $\mathbf{h} = [0.229 \quad 0.459 \quad 0.688 \quad 0.459 \quad 0.229]^T$ investigated in [4], [7] is used for evaluating the performance of iterative equalizers. The channel $\mathbf{h}$ is a severe ISI channel because the spectral characteristic of the channel $\mathbf{h}$ possesses nulls over the Nyquist band. Finally, the decoder is implemented using the BCJR algorithm. Only the SISO equalizer changes from one scheme to another. MMSE-DFE with 17 feedforward taps and 4 feedback taps is used for both the normal and the time-reversed DFEs while the LMMSE equalizer uses 21 taps.

The 6 different equalizer types are simulated in the paper. The notation "TV-" denotes equalizers with time-varying filters while "TIV-" indicates those with time-invariant filters.

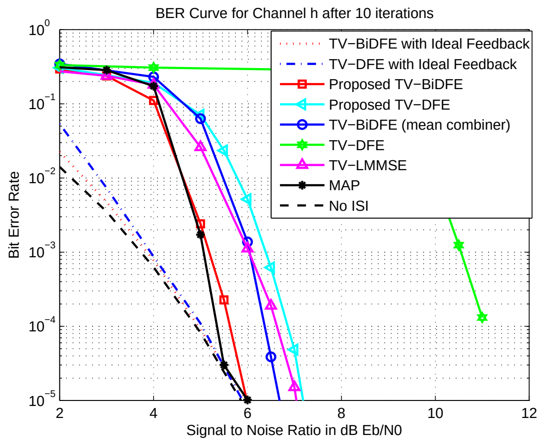


Fig. 2: BER Curve on the Channel $\mathbf{h}$ after 10 Iterations with Time-varying Filters

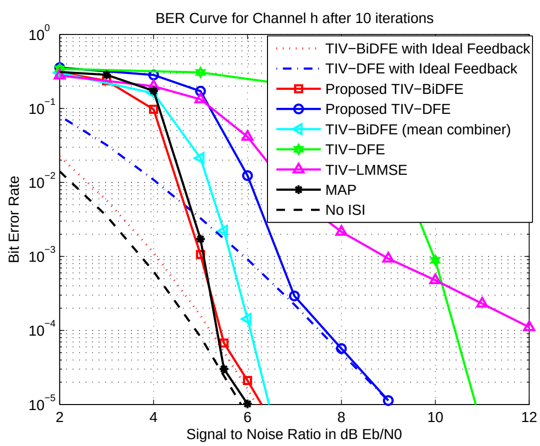
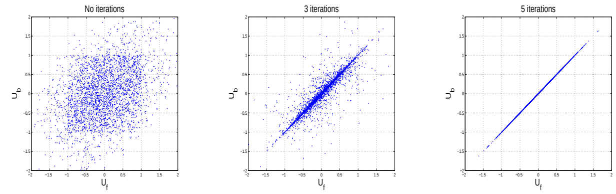


Fig. 3: BER Curve on the Channel $\mathbf{h}$ after 10 iterations with Time-invariant Filters

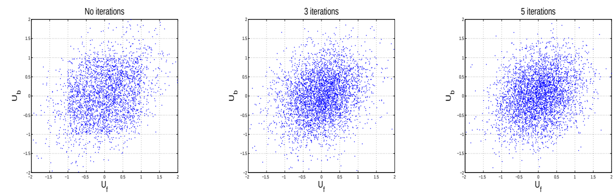
For instance, the "TV-LMMSE" in the legend indicates the LMMSE equalizer with a time-varying filter. The "Proposed DFE" denotes the iterative DFE approach which adopts the proposed dual LLR generation while the "DFE" uses the conventional LLR mapping of (6). The "Proposed BiDFE" is the iterative BiDFE algorithm which is described in Section IV. In other words, in this scheme, dual LLR generation is used for both normal and time-reversed DFEs along with the proposed combiner of (17) in conjunction with the noise correlation coefficient of (15). The "BiDFE (mean combiner)" is the iterative BiDFE algorithm with the conventional LLR mapping and the mean combiner, $L_e(X) = (L_{e,f}(X) + L_{e,b}(X))/2$ (of [10]), simulated for performance comparison purposes. Finally, the "MAP" is the optimal equalizer implemented via the BCJR algorithm.

Fig. 2 shows the performance of several turbo equalizers with time-varying filters after 10 iterations. TV-DFE with the conventional LLR mapping shows poor performance but once the proposed dual LLR generation is used ("Proposed TV-DFE"), the DFE performance becomes nearly as good as the TV-LMMSE method of [4] at low BERs. The "Proposed TV-BiDFE" is considerably better than the TV-BiDFE based on a mean combiner, approaching the performance of the MAP scheme. Note that only MAP and Proposed TV-BiDFE completely close the gap to the zero-ISI performance (at low BERs).



(a) (b) (c)

Fig. 4: Noise Correlation of "Proposed TV-BiDFE" after (a) No Iteration (b) 3 Iterations and (c) 5 Iterations: The correlation is described with $U_f = Y_f - \hat{X}_f$ and $U_b = Y_b - \hat{X}_b$ only when $\hat{X}_f = \hat{X}_b$ on the horizontal and vertical axis respectively.



(a) (b) (c)

Fig. 5: Noise Correlation of "Proposed TIV-BiDFE" after (a) No Iteration (b) 3 Iterations and (c) 5 Iterations: The correlation is described with $U_f = Y_f - \hat{X}_f$ and $U_b = Y_b - \hat{X}_b$ only when $\hat{X}_f = \hat{X}_b$ on the horizontal and vertical axis respectively.

Fig. 3 shows the BER performance of time-invariant filter based turbo equalizers. As the figure shows, the "Proposed TIV-DFE" also shows the superior performance to the "TIV-DFE". Moreover, the performance of "Proposed TIV-BiDFE" is similar to the performance of the MAP equalizer while achieving a very low computational complexity based on the use of time-invariant filters. Also notice that both "Proposed TIV-DFE" and "Proposed TIV-BiDFE" achieve decision-error-free performance at low BERs, indicating the error propagation effect has been nearly eliminated using the proposed dual LLR generation method.

The noise correlation in one block of coded data bits is described in Fig. 4 and 5, at different iteration numbers at a 7 dB SNR on $\mathbf{h}$. The correlation coefficient of "Proposed TV-BiDFE" goes to 1 as the number of interations increases because the *a priori* information from the decoder becomes reliable, and the time-varying filters in the normal and the time-reversed DFEs produce essentially the same equalized output sequences. On the other hand, the correlation coefficient of "Proposed TIV-BiDFE" actually decreases as the number

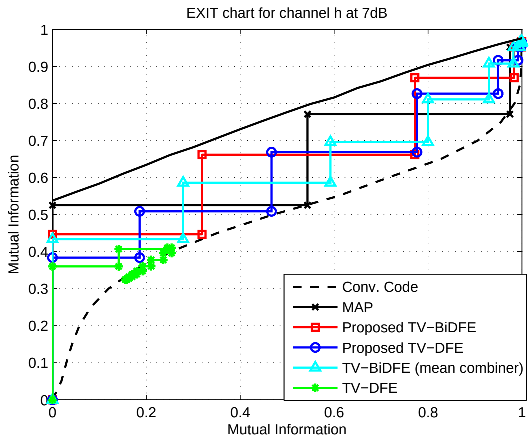


Fig. 6: EXIT Chart on the Channel $\mathbf{h}$ at 7dB with Time-varying Filters

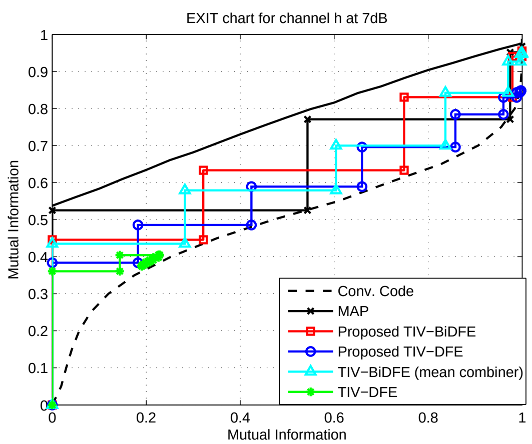


Fig. 7: EXIT Chart on the Channel $\mathbf{h}$ at 7dB with Time-invariant Filters

of iterations increases, and the noise correlation coefficient converges to that of "TIV-BiDFE with Ideal Feedback". This is because the decision feedback errors disappear. Note that the filter coefficients in both DFEs do not change with the *a priori* information.

Fig. 6 shows the extrinsic information transfer (EXIT) chart [12] corresponding to time-varying-filter based equalizers for $\mathbf{h}$ at a 7 dB SNR while Fig. 7 shows the similar EXIT charts for time-invariant-filter-based schemes. The trajectories of "TV-DFE" and "TIV-DFE" move up for the first couple of iterations, but later, they move down due to error propagation and inadequate extrinsic LLR generations that cannot handle error propagation. However, the trajectories of "Proposed TV-DFE" and "Proposed TIV-DFE" keep moving up as the number of iterations increases, clearing indicating the advantage and effectiveness of the proposed dual LLR generation method. Moreover, the trajectory of the "Proposed TV-BiDFE" and "Proposed TIV-BiDFE" also shows a clear trajectory from 0 bits of mutual information to 1 bit of mutual information with a less number of iteration runs than "Proposed TV-DFE", "Proposed TIV-DFE", "TV-BiDFE (mean combiner)", or "TIV-BiDFE (mean combiner)".

We do notice, however, that the proposed BiDFE scheme requires more iterations in achieving the full performance, relative to the MAP equalizer. Nevertheless, the proposed BiDFE method offers a reasonable tradeoff between complexity and performance.

VI. Conclusion

In this paper, we proposed low-complexity turbo equalization methods based on DFE and BiDFE structures. The proposed dual LLR generation designed to reduce error propagation indeed provide decision-error-free performance in DFEs in turbo equalizer settings. When further employing an LLR combining method that estimates the correlation between the forward and backward DFE outputs and whitens them, the BiDFE performance approaches the performance of the "optimal" turbo equalizer based on the much more complex BCJR channel equalizer. The proposed LLR generation and combining methods remain effective even when a time-invariant constraint is imposed on the feedforward and feedback filters of the DFEs. Overall, the proposed BiDFE method based on time-invariant filter taps provides the excellent performance-complexity tradeoff for severe ISI channels where the linear SISO equalizer is not effective.

References

[1] C. Douillard, M. Jezequel, C. Berrou, A. Picart, P. Didier, and A. Glavieux "Iterative correction of intersymbol interference: Turbo equalization," *European Trans. Telecommun.*, vol. 6, no. 5, pp.507-511, Sep.-Oct. 1995.

[2] L. Bahl, J. Cocke, F. Jelinek, AND J. Raviv. "Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate," *IEEE Trans. Inform. Theory*, vol. IT-20, pp.284-287, Mar. 1974.

[3] M. Austin "Decision-feedback equalization for digital communication over dispersive channels," *M.I.T. RES. Lab. Electron., Tech. Rep.461*, Aug. 1967.

[4] M. Tüchler, R. Kötter, and A. Singer "Turbo equalization: principles and new results," *IEEE Trans. on Signal Processing*, vol.50, pp.754-767, May 2002

[5] A. Fertner "Improvement of bit-error-rate in decision feedback equalizer by preventing decision-error propagation," *IEEE Trans. Signal Processing*, vol 46, pp.1872-1877, Jul. 1998.

[6] S. Ariyavisitakul and Y. Li "Joint coding and decision feedback equalization for broadband wireless channels," *IEEE Trans. Selected Areas in Communications*, vol 16, pp.1670 - 1678, Dec. 1998

[7] R. R. Lopes and J. R. Barry "The Soft-Feedback Equalization for Turbo Equalization of Highly Dispersive Channels," *IEEE Trans. Communications*, vol 54, no. 5, May 2006.

[8] J. Balakrishnan and C.R. Johnson, Jr. "Bidirectional Decision Feedback Equalizer: Infinite Length Results," *In Proc. Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2001.

[9] J. Nelson, A. Singer, U. Madhow and C. McGahey "BAD: bidirectional arbitrated decision-feedback equalization," *IEEE Trans. Commun.*, vol. 53, no. 2, Feb. 2005.

[10] J. Jiang, C. He, E. M. Kurtas and K. R. Narayanan "Performance of soft feedback equalization over magnetic recording channels," *In Proc. Intermag*, San Diego, CA, May 2006.

[11] J. Cioffi, G. Dudevior, M. Eyuboglu, and G. Forney "MMSE decision-feedback equalizers and coding - Part I and II," *IEEE Trans. Commun.*, vol. 43, pp.2582-2604, Oct. 1995.

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