Middlesex University Research Repository

An open access repository of

Middlesex University research

Decision Feedback Detection for MIMO-ISI Channels: Design Using Fixed and Adaptive Constraints

Xuan Huan Nguyen and Jinho Choi

Abstract— In this correspondence, we investigate the decision feedback detection for multiple-input multiple-output intersymbol interference (MIMO-ISI) channels. Firstly, a novel constrained symbol-by-symbol decision feedback detector (CS-DFD) is proposed, in which a constraint on the feedback filter (FBF) provides robustness against error propagation and outperforms the conventional decision feedback detector (DFD). However, we find that an error floor is observed at high signal-to-noise ratios (SNRs) if a fixed constraint is used. To resolve this problem, we then propose an iterative symbol-by-symbol decision feedback detector (IS-DFD) in which an adaptive constraint is implicitly used to update the DFD's coefficients. Simulation results show that the error floor problem is overcome and the performance becomes satisfactory at high SNRs through iterations.

Index Terms— Multiple-input multiple-output intersymbol interference (MIMO-ISI) channel, decision feedback detection, error propagation, constrained optimization, iterative processing.

I. INTRODUCTION

The idea of using past decisions to mitigate intersymbol interference (ISI) is successfully used in the decision feedback equalizer (DFE) and the decision feedback detector (DFD) [1, ch. 10]. However, the decision errors from the equalizer/detector may result in erroneous cancellation of the postcursor ISI through the feedback filter (FBF) [2] and consequently degrade the overall performance. There have been some ad hoc techniques proposed to deal with the error propagation problem by using erasures [3], error decision threshold [4] or a soft decision device [5]. In [6], some constraints were used in deriving the feedforward filter (FFF) to monitor decision error and ensure ISI cancellation.

In multiple-input multiple-output intersymbol interference (MIMO-ISI) channels, a severe self-interference problem occurs due to ISI as well as co-antenna interference (CAI). Thus, the error propagation becomes more serious and its mitigation has to be considered when designing the MIMO DFD.

The main contributions of this correspondence are as follows:

1) We propose a constrained symbol-by-symbol DFD (CS-DFD) under the minimum mean square error (MMSE) criterion for MIMO-ISI channels via directly solving a convex optimization program. This proposed CS-DFD outperforms the conventional DFD at moderate signal-to-noise ratios (SNRs). However, we find that a fixed constraint (which was also used in [6]) cannot guarantee an optimal trade-off between minimizing the mean square error (MSE) of the detected symbols and mitigating the error propagation. Particularly at high SNRs, where less decision errors occur, the fixed constraint becomes problematic as it limits the capability of the DFD in minimizing the MSE. As a consequence, an error floor eventually appears at high SNRs. To resolve this problem, we consider an iterative DFD as follows.

2) The proposed iterative symbol-by-symbol DFD (IS-DFD) iteratively updates its coefficients based on the severity of the decision error. By finding the variance of the decision error which can be seen as an adaptive constraint, an optimal MMSE DFD can be derived

taking into account the error propagation. As a result, satisfactory performance (without suffering the error floor problem at high SNRs) is eventually obtained as the predicted signal-to-noise-plus-interference ratio (SINR) converges to the true value through several iterations. In addition, as opposed to the block-iterative approaches [9], [10], the proposed IS-DFD performs symbol-by-symbol detection which can avoid the requirement of a large amount of memory to store previous decisions.

II. BACKGROUND

A. MIMO-ISI system model

Assume that there are M_t and M_r transmit and receive antennas, respectively. Quadrature phase-shift keying (QPSK) is used for signalling with the constellation set $\mathcal{Q} = \{\pm 1/\sqrt{2} \pm j/\sqrt{2}\}$. Denote by $b_{m_t}(n) \in \mathcal{Q}$ the data symbol transmitted from the m_tth transmit antenna with $1 \leq m_t \leq M_t$. Assume that the data symbols $\{b_{m_t}(n)\}$ are equally likely (with mean zero and variance one). Let $\mathbf{b}(n) = [b_1(n), b_2(n), ..., b_{M_t}(n)]^T \in \mathcal{Q}^{M_t \times 1}$ and $\mathbf{y}(n) = [y_1(n), y_2(n), ..., y_{M_r}(n)]^T \in \mathcal{C}^{M_r \times 1}$ denote the transmitted and the received signal vectors, respectively, where n represents time index. Note that $\mathcal{Q}^{I \times J}$ and $\mathcal{C}^{I \times J}$ denote the set of all $I \times J$ matrices whose entries belong to $\mathcal{Q}$ and $\mathcal{C}$, respectively, where $\mathcal{C}$ is the set of complex numbers. Denote by $h_{m_r, m_t}(p)$ the pth multipath complex coefficient of the MIMO-ISI channel from the m_tth transmit antenna to the m_rth receive antenna with $1 \leq m_r \leq M_r$. Assuming all channels have finite impulse responses of the same length denoted by P and letting $[\mathbf{H}(p)]_{m_r, m_t} = h_{m_r, m_t}(p)$, then a MIMO-ISI system model is written as

$$\mathbf{y}(n) = \sum_{p=0}^{P-1} \mathbf{H}(p)\mathbf{b}(n-p) + \mathbf{w}(n) \qquad (1)$$

where $\mathbf{H}(p) \in \mathcal{C}^{M_r \times M_t}$ and $\mathbf{w}(n) = [w_1(n), w_2(n), ..., w_{M_r}(n)]^T$ is assumed to be a mean-zero Gaussian noise vector with covariance matrix $E[\mathbf{w}(n)\mathbf{w}^H(n)] = \sigma_w^2 \mathbf{I}$. Defining

$$\begin{aligned} \mathbf{b}_n &= [\mathbf{b}^T(n), \mathbf{b}^T(n-1), ..., \mathbf{b}^T(n-N+1)]^T \in \mathcal{Q}^{NM_t \times 1}, \\ \mathbf{y}_n &= [\mathbf{y}^T(n), \mathbf{y}^T(n-1), ..., \mathbf{y}^T(n-L+1)]^T \in \mathcal{C}^{LM_r \times 1}, \\ \mathbf{w}_n &= [\mathbf{w}^T(n), \mathbf{w}^T(n-1), ..., \mathbf{w}^T(n-L+1)]^T \in \mathcal{C}^{LM_r \times 1} \end{aligned}$$

where $N = L + P - 1$ and L denotes the length of the FFF, the system model in (1) can then be rewritten as

$$\mathbf{y}_n = \mathbf{H}\mathbf{b}_n + \mathbf{w}_n. \qquad (2)$$

Here, $\mathbf{H}$ represents the channel filtering matrix of size $[LM_r \times NM_t]$ given by

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}(0) & \cdots & \mathbf{H}(P-1) & & \\ & \ddots & \cdots & \ddots & \\ & & \mathbf{H}(0) & \cdots & \mathbf{H}(P-1) \end{bmatrix}. \qquad (3)$$

B. MMSE DFD for a MIMO-ISI system

Consider a DFD (see Fig. 1) consisting of an FFF and an FBF to equalize a MIMO-ISI channel. Both the FFF and the FBF have a finite number of taps. To detect the transmitted signal with a delay of $(D - 1)$ symbols, we use the output of the DFD as follows:

$$\begin{aligned} \mathbf{d}(n) &= [d_1(n), d_2(n), ..., d_{M_t}(n)]^T \\ &= \sum_{u=0}^{L-1} \mathbf{G}^H(u)\mathbf{y}(n-u) - \sum_{v=D}^{N-1} \mathbf{F}^H(v)\hat{\mathbf{b}}(n-v) \\ &= \mathbf{G}^H \mathbf{y}_n - \mathbf{F}^H \hat{\mathbf{b}}_{n,2} \qquad (4) \end{aligned}$$
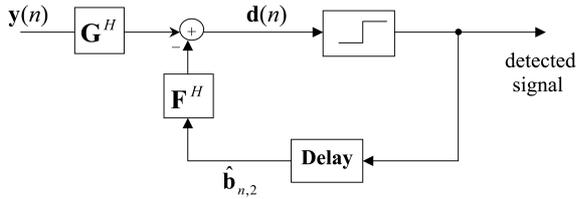
Fig. 1.　Block diagram of decision feedback detector.

where the superscript $(\cdot)^H$ denotes the Hermitian transpose and $\hat{\mathbf{b}}_{n,2}$ denotes the tentative hard-decision vector of $\mathbf{b}_{n,2} = [\mathbf{b}^T(n - D), \mathbf{b}^T(n-D-1), ..., \mathbf{b}^T(n-N+1)]^T \in \mathcal{Q}^{(N-D)M_t \times 1}$. The FFF and FBF weight matrices, $\mathbf{G}$ and $\mathbf{F}$, respectively, are defined as

$$\mathbf{G} = [\mathbf{G}^T(0), \mathbf{G}^T(1), ..., \mathbf{G}^T(L-1)]^T \in \mathcal{C}^{LM_r \times M_t},$$
$$\mathbf{F} = [\mathbf{F}^T(D), \mathbf{F}^T(D+1), ..., \mathbf{F}^T(N-1)]^T \in \mathcal{C}^{(N-D)M_t \times M_t}$$

where $\mathbf{G}(u), u = 0, 1, ..., L-1$, is the $(u+1)$th $(M_r \times M_t)$ filter matrix of the FFF, and $\mathbf{F}(v), v = D, D+1, ..., N-1$, is the $(v-D+1)$th $(M_t \times M_t)$ filter matrix of the FBF. As mentioned above, $\mathbf{d}(n)$ can be considered as an estimate of $\mathbf{b}(n - D + 1)$. Through the FBF, the detected past signal vectors, $\{\hat{\mathbf{b}}(n-D), \hat{\mathbf{b}}(n-D-1), ..., \hat{\mathbf{b}}(n-N+1)\}$, are used to cancel the postcursor ISI components. Under the MMSE criterion, $\mathbf{G}$ and $\mathbf{F}$ can be found so as to minimize the MSE cost function given as

$$\psi(\mathbf{G}, \mathbf{F}) = E\left[\|\mathbf{d}(n) - \mathbf{b}(n - D + 1)\|^2\right]. \tag{5}$$

Assuming the feedback decisions are correct[1] and the channel is perfectly known, the conventional solutions of FFF and FBF [8] are obtained as

$$\mathbf{G}_c = (\mathbf{H}_1 \mathbf{H}_1^H + \sigma_w^2 \mathbf{I})^{-1} \mathbf{H}_D, \tag{6}$$
$$\mathbf{F}_c = \mathbf{H}_2^H \mathbf{G}_c \tag{7}$$

where $\mathbf{H}_1$ is the $LM_r \times DM_t$ submatrix of $\mathbf{H}$ obtained by taking the first DM_t columns, $\mathbf{H}_2$ is the $LM_r \times (N-D)M_t$ submatrix of $\mathbf{H}$ obtained by taking the last $(N-D)M_t$ columns and $\mathbf{H}_D$ is obtained by taking the last M_t columns of $\mathbf{H}_1$.

III. Constrained Symbol-by-Symbol Decision Feedback Detection

Practically, we can have erroneous decisions in the DFD, i.e., $\hat{\mathbf{b}}_{n,2} \neq \mathbf{b}_{n,2}$. This can increase the chance of introducing an error in estimating the next symbol vector $\mathbf{b}(n-D+1)$. This is called error propagation and it causes a sequence of errors in detecting future symbols when the present decision is erroneous.

Assuming the FBF is conventionally given as $\mathbf{F} = \mathbf{H}_2^H \mathbf{G}$, we will optimize the FFF to mitigate the effect of error propagation. With erroneous decisions, the MSE cost function in (5) can be rewritten as

$$\psi(\mathbf{G}) = E\left[\left\|\mathbf{G}^H\mathbf{H}_1\mathbf{b}_{n,1} + \mathbf{G}^H\mathbf{H}_2\mathbf{e}_{n,2} + \mathbf{G}^H\mathbf{w}_n \right.\right.$$
$$\left.\left. - \mathbf{b}(n - D + 1)\right\|^2\right] \tag{8}$$

where

$$\mathbf{b}_{n,1} = [\mathbf{b}^T(n), \mathbf{b}^T(n-1), ..., \mathbf{b}^T(n-D+1)]^T,$$
$$\mathbf{e}_{n,2} = \mathbf{b}_{n,2} - \hat{\mathbf{b}}_{n,2}.$$

We can see that the FBF coefficient matrix $\mathbf{F} = \mathbf{H}_2^H \mathbf{G}$ becomes the weight on the error term such as $\mathbf{G}^H\mathbf{H}_2\mathbf{e}_{n,2}$ in (8). However,

[1] This assumption leads to the fact that the MMSE DFD has no immunity against the error propagation.

if perfect feedback is obtained (i.e., $\mathbf{e}_{n,2} = \mathbf{0}$), this error term is annulled, and (8) becomes

$$\tilde{\psi}(\mathbf{G}) = E\left[\left\|\mathbf{G}^H\mathbf{H}_1\mathbf{b}_{n,1} + \mathbf{G}^H\mathbf{w}_n - \mathbf{b}(n - D + 1)\right\|^2\right].$$

Similar to [6] where a constraint is imposed on the FFF of the DFE to avoid excessive noise enhancement, we impose a quadratic inequality constraint on the FBF to directly restrict the impact of the error term $\mathbf{G}^H\mathbf{H}_2\mathbf{e}_{n,2}$. Assuming that the decision errors are uncorrelated at different time instants and uncorrelated from different transmitters and their mean-square values towards different transmitters are equal, the constraint on the FBF can then be expressed as

$$\mathrm{tr}\left\{\mathbf{G}^H\mathbf{H}_2\mathbf{H}_2^H\mathbf{G}\right\} \leq \gamma M_t \tag{9}$$

where $\mathrm{tr}\{.\}$ denotes the trace of a matrix and γ is the constraint level ($\gamma > 0$). If decision errors occur frequently, the constraint needs to be tight and thus γ should be close to zero. On the other hand, a larger value of γ can be taken if less decision errors are expected. It is difficult to choose the optimal γ which closely depends on the operating SNR. However, from our simulation observation, a reasonable choice of γ falls approximately in the range $0.5 < \gamma < 1$ for moderate SNRs. Letting $\alpha = \gamma M_t$, the constrained MIMO MMSE-DFD problem can be written as

$$\mathbf{G}_{op} = \arg\min_{\mathbf{G}} \tilde{\psi}(\mathbf{G}) \tag{10}$$
$$\text{subject to} \quad \mathrm{tr}\left\{\mathbf{G}^H\mathbf{H}_2\mathbf{H}_2^H\mathbf{G}\right\} \leq \alpha.$$

In [6], the author did not directly solve the optimization problem. It is observed that (10) is a convex quadratic optimization program. Using the Lagrangian multiplier method, we obtain

$$\mathbf{G}_{op} = \arg\min_{\mathbf{G}}\max_{\lambda \geq 0}\left\{\tilde{\psi}(\mathbf{G}) + \lambda\left(\mathrm{tr}\left\{\mathbf{G}^H\mathbf{H}_2\mathbf{H}_2^H\mathbf{G}\right\} - \alpha\right)\right\}$$
$$= \arg\min_{\mathbf{G}}\max_{\lambda \geq 0} \phi(\mathbf{G}, \lambda) \tag{11}$$

where

$$\phi(\mathbf{G}, \lambda) = \mathrm{tr}\left\{\mathbf{G}^H\mathbf{A}(\lambda)\mathbf{G}\right\} - \mathrm{tr}\left\{\mathbf{G}^H\mathbf{H}_D\right\}$$
$$- \mathrm{tr}\left\{\mathbf{H}_D^H\mathbf{G}\right\} + M_t - \lambda\alpha,$$
$$\mathbf{A}(\lambda) = \mathbf{H}_1\mathbf{H}_1^H + \lambda\mathbf{H}_2\mathbf{H}_2^H + \sigma_w^2\mathbf{I}$$

and λ is a Lagrange multiplier. Since $\mathbf{G}^H\mathbf{A}(\lambda)\mathbf{G}$ is a positive semidefinite matrix, the duality can be applied to (11) [7, ch. 5], thus leading to the solution of λ as

$$\lambda_{op} = \arg\max_{\lambda \geq 0}\min_{\mathbf{G}} \phi(\mathbf{G}, \lambda)$$
$$= \arg\max_{\lambda \geq 0} \phi(\mathbf{G}_m(\lambda), \lambda) \tag{12}$$

where

$$\mathbf{G}_m(\lambda) = \arg\min_{\mathbf{G}} \phi(\mathbf{G}, \lambda) = \mathbf{A}^{-1}(\lambda)\mathbf{H}_D. \tag{13}$$

Once λ_{op} is found, the FFF solution is obtained by

$$\mathbf{G}_{op} = \mathbf{G}_m(\lambda_{op}).$$

In order to solve for λ_{op}, we look back at the convex quadratic optimization problem in (10) and it can be observed that the FFF solution can only be one of the two following options. The first option is

$$\mathbf{G}_{op} = \arg\min_{\mathbf{G}} \tilde{\psi}(\mathbf{G}) = \mathbf{G}_c, \text{ i.e., } \lambda_{op} = 0$$

if $\mathbf{G}_c$ satisfies the constraint $\mathrm{tr}\left\{\mathbf{G}_c^H\mathbf{H}_2\mathbf{H}_2^H\mathbf{G}_c\right\} \leq \alpha$. In the case of this constraint not being satisfied, the solution becomes

$$\mathbf{G}_{op} = \mathbf{G}_m(\lambda_{op})$$

where λ_{op} is the solution of $\text{tr}\left\{\mathbf{G}_m^H(\lambda_{op})\mathbf{H}_2\mathbf{H}_2^H\mathbf{G}_m(\lambda_{op})\right\} = \alpha$. To find out λ_{op}, we consider

$$
\begin{aligned}
f(\lambda) &= \text{tr}\left\{\mathbf{G}_m^H(\lambda)\mathbf{H}_2\mathbf{H}_2^H\mathbf{G}_m(\lambda)\right\} \\
&= \text{tr}\left\{\mathbf{H}_D^H\mathbf{A}^{-1}(\lambda)\mathbf{H}_2\mathbf{H}_2^H\mathbf{A}^{-1}(\lambda)\mathbf{H}_D\right\}.
\end{aligned}
$$

Since $\mathbf{A}(\lambda)$ is a Hermitian matrix and its diagonal elements are proportional to λ, $f(\lambda)$ is a monotonically decreasing function of λ. When $\lambda = \lambda_0 = 0$, the solution becomes a conventional DFD. Generally we have $f(\lambda_0 = 0) \geq \alpha$. If we have $\lambda_1 \geq \lambda_0$ such that $f(\lambda_1) \leq \alpha \leq f(\lambda_0)$ then we know λ_{op} must satisfy $\lambda_0 \leq \lambda_{op} \leq \lambda_1$. Thus, a well known golden section search (GSS) algorithm is applicable to iteratively search for λ_{op}. From our observation, the optimal λ_{op} falls between $\lambda = 0$ (i.e., the conventional DFD) and $\lambda = 1$ (i.e., the FFF is the linear MMSE filter solution). As shown in Fig. 2 (the simulation parameters are given in Section V-B), λ_{op}'s obtained in the CS-DFD scheme for three values of $\gamma = 0.5, 0.6$ and 0.7 are far smaller than 1. The choice of $\lambda_1 = 1$ is therefore confirmed to be large enough for the GSS algorithm.
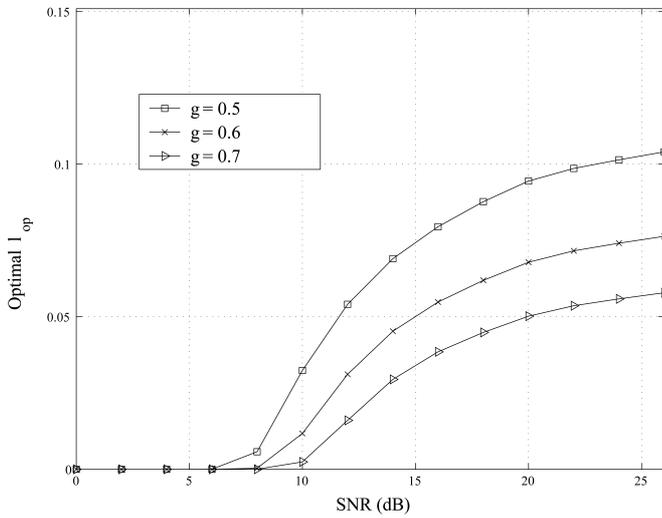

Fig. 2. The optimal λ_{op} obtained by the GSS algorithm in the CS-DFD scheme with $\gamma = 0.5, 0.6$, and 0.7 and channel length $P = 6$.

IV. ITERATIVE SYMBOL-BY-SYMBOL DECISION FEEDBACK DETECTION

In the CS-DFD approach, the value of the constraint parameter γ is fixed. Thus, if γ is found to be good for low SNRs (more decision errors), it would be unsatisfactory for higher SNRs (less decision errors). In other words, the trade-off between the degree of freedom (to minimize the MSE) and the degree of constraint (to mitigate the error propagation) in solving the optimization problem would not be optimally satisfied with a fixed constraint. Our design in this section is inspired by how severe the decision error is. The constraint parameter γ should adapt to the decision error rate and, therefore, should generally be an increasing function of SNR. However, it is difficult to find out the optimal γ which depends on the operating SNR. Fortunately, instead of finding γ, we have an alternative way to cope with the error propagation by incorporating the variance of the decision error in the design of the DFD. To some extent, the variance of the decision error plays a crucial role in the CS-DFD approach via λ and this will be explained later in the section. Note that the variance of the decision error is needed to determine the DFD coefficients but,

at the same time, the estimation of the variance of the decision error is based on the DFD coefficients. We solve this "*circle*" problem by employing an iterative scheme in which convergence is obtained when the predicted variance of the decision error approaches the true value. In this section, we use the notation $\text{diag}(\mathbf{X})$ to denote a diagonal matrix whose diagonal is the diagonal of matrix $\mathbf{X}$, and use $\text{diag}(\mathbf{x})$ to denote a diagonal matrix whose diagonal is vector $\mathbf{x}$.

Assume that the FBF is still conventionally obtained (i.e., $\mathbf{F} = \mathbf{H}_2^H\mathbf{G}$). Denoting by $e_i(n) = b_i(n) - \hat{b}_i(n)$ the decision error of symbol $b_i(n)$, $i = 1, 2, ..., M_t$, we make the following assumption

$$
(AS1): E[e_i(n)e_{i'}^*(n')] = \sigma_{e_i}^2\delta(n - n')\delta(i - i')
$$

where $\sigma_{e_i}^2$ is the variance of $e_i(n)$, $\delta(\cdot)$ is the Kronecker delta, and $(\cdot)^*$ denotes complex conjugation. Denoting by $\mathbf{Q} = E[\mathbf{e}_{n,2}\mathbf{e}_{n,2}^H]$ the autocorrelation matrix of decision errors and from (AS1), we have

$$
\mathbf{Q} = \text{diag}([\mathbf{q}_1^T, \mathbf{q}_2^T, ..., \mathbf{q}_{N-D}^T]^T)
$$

where $\mathbf{q}_1 = \mathbf{q}_2 = ... = \mathbf{q}_{N-D} = [\sigma_{e_1}^2, \sigma_{e_2}^2, ..., \sigma_{e_{M_t}}^2]^T$. Thus, the solution of the FFF can be obtained by

$$
\begin{aligned}
\mathbf{G} &= \arg\min_{\mathbf{G}}\psi(\mathbf{G}) \\
&= (\mathbf{H}_1\mathbf{H}_1^H + \mathbf{H}_2\mathbf{Q}\mathbf{H}_2^H + \sigma_w^2\mathbf{I})^{-1}\mathbf{H}_D. \quad (14)
\end{aligned}
$$

Note that if $\sigma_{e_1}^2 = \sigma_{e_2}^2 = ... = \sigma_{e_{M_t}}^2 = \sigma_e^2$, then $\mathbf{Q} = \sigma_e^2\mathbf{I}$. In this case, the solution in (14) is equivalent to the solution in (13) with $\lambda = \sigma_e^2$. If there are more decision errors, σ_e^2 (i.e., λ) should be larger. Therefore, the variance of the decision error can be seen as an adaptive constraint used in the CS-DFD, where we assume that the error probabilities of symbols transmitted from different antennas are the same (i.e., $\sigma_{e_1}^2 = \sigma_{e_2}^2 = ... = \sigma_{e_{M_t}}^2$). However, this may not always be the case as channel characteristics from each transmit antenna can be different. This turns out to be crucial in designing the optimal DFD as our simulations (not included here due to the space limitation) have shown that the variances $\{\sigma_{e_i}^2\}$ are significantly different when the channel impulse responses are randomly generated. Thus, instead of using a single constraint as in the CS-DFD approach, we employ the solution in (14) with the estimate of the autocorrelation matrix of decision errors $\mathbf{Q}$.

Now, we attempt to find $\sigma_{e_i}^2$. Note that

$$
\begin{aligned}
\sigma_{e_i}^2 &= E[e_i(n)e_i^*(n)] \\
&= E[(b_i(n) - \hat{b}_i(n))(b_i(n) - \hat{b}_i(n))^*] \\
&= 2(1 - E[b_i(n)\hat{b}_i^*(n)]) \quad (15)
\end{aligned}
$$

where the correlation $E[b_i(n)\hat{b}_i^*(n)]$ can be either theoretically approximated as $E[b_i(n)\hat{b}_i^*(n)] = 1 - P_{s,i}$ (for QPSK signalling) [9] where $P_{s,i}$ is the error probability of symbol $b_i(n)$ or numerically found using the input-decision correlation method described in [10]. The output of the DFD can now be rewritten as

$$
\begin{aligned}
\mathbf{d}(n) &= \mathbf{G}^H\mathbf{H}_D\mathbf{b}(n - D + 1) + \mathbf{G}^H\mathbf{H}_{1,D}\mathbf{b}_{n,D} \\
&\quad + \mathbf{G}^H\mathbf{H}_2\mathbf{e}_{n,2} + \mathbf{G}^H\mathbf{w}_n \\
&= \mathbf{D}\mathbf{b}(n - D + 1) + \mathbf{v}(n) \quad (16)
\end{aligned}
$$

where $\mathbf{H}_{1,D}$ is a submatrix of $\mathbf{H}_1$ obtained by deleting the last M_t columns, $\mathbf{b}_{n,D}$ is the subvector of $\mathbf{b}_{n,1}$ obtained by deleting the last M_t elements, $\mathbf{D} = \text{diag}(\mathbf{G}^H\mathbf{H}_D)$, and

$$
\begin{aligned}
\mathbf{v}(n) &= (\mathbf{G}^H\mathbf{H}_D - \mathbf{D})\mathbf{b}(n - D + 1) + \mathbf{G}^H\mathbf{H}_{1,D}\mathbf{b}_{n,D} \\
&\quad + \mathbf{G}^H\mathbf{H}_2\mathbf{e}_{n,2} + \mathbf{G}^H\mathbf{w}_n. \quad (17)
\end{aligned}
$$

Letting

$$\begin{aligned} \mathbf{R}_v &= E[\mathbf{v}(n)\mathbf{v}^H(n)] \\ &= \mathbf{G}^H(\mathbf{H}_{1,D}\mathbf{H}_{1,D}^H + \mathbf{H}_2\mathbf{Q}\mathbf{H}_2^H + \sigma_w^2\mathbf{I})\mathbf{G} \\ &\quad + (\mathbf{G}^H\mathbf{H}_D - \mathbf{D})(\mathbf{G}^H\mathbf{H}_D - \mathbf{D})^H \\ &= \mathbf{G}^H\mathbf{H}_D - \mathbf{G}^H\mathbf{H}_D\mathbf{D}^H - \mathbf{D}\mathbf{H}_D^H\mathbf{G} + \mathbf{D}\mathbf{D}^H, \end{aligned}$$

the SINR of the detector output, $d_i(n)$, becomes

$$SINR_i = \frac{(\mathbf{g}_i^H\mathbf{h}_{D,i})^2}{[\mathbf{R}_v]_{ii}} \qquad (18)$$

where $\mathbf{g}_i$ and $\mathbf{h}_{D,i}$ are the ith columns of $\mathbf{G}$ and $\mathbf{H}_D$, respectively, and $[\mathbf{R}_v]_{ij}$ denotes the (i,j)th element of matrix $\mathbf{R}_v$. Since $[\mathbf{R}_v]_{ii} = \mathbf{g}_i^H\mathbf{h}_{D,i} - (\mathbf{g}_i^H\mathbf{h}_{D,i})^2$, we can further simplify (18) as

$$SINR_i = \frac{\mathbf{g}_i^H\mathbf{h}_{D,i}}{1 - \mathbf{g}_i^H\mathbf{h}_{D,i}}. \qquad (19)$$

As in [9], in order to find out $P_{s,i}$, the noise-plus-interference term $\mathbf{v}(n)$ is assumed to be Gaussian. From this, we have [1, p. 269]

$$P_{s,i} = 1 - \left(1 - Q(\sqrt{SINR_i})\right)^2 \qquad (20)$$

where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-t^2/2}\mathrm{d}t$.

We can now apply an iterative algorithm for estimating a set of error variances $\{\sigma_{e_i}^2\}$ and obtain the coefficients $\mathbf{G}$, as follows (we use the subscript $_{(l)}$ to denote the lth iteration):

1. For the first iteration:

- Set initial $\sigma_{e_i}^2 = 1$ for all $i = 1, 2, ..., M_t$, and therefore, $\mathbf{Q}_{(1)} = \mathbf{I}$. Thus, for the first iteration, the FFF solution $\mathbf{G}_{(1)}$ is a linear MMSE equalizer.
- $SINR_{i,(1)}$ and $P_{s,i,(1)}$ are now available from (19) and (20), respectively.

2. For the lth iteration ($l \geq 2$):

- Update $\sigma_{e_i,(l)}^2 = 2(1 - E[b_i(n)\hat{b}_i^*(n)]_{(l-1)})$.
- $\mathbf{G}_{(l)}$, $SINR_{i,(l)}$ and $P_{s,i,(l)}$ are available from (14), (19), and (20), respectively.

Initially the variance of the decision error $\sigma_{e_i}^2$ is set to 1, thus the resultant FFF solution is the linear MMSE solution. The value of $\sigma_{e_i}^2$ will then decrease and be closer to its true value after each iteration. Since $SINR_i$ in (19) is a monotonically decreasing function of $\sigma_{e_i}^2$, it will increase as $\sigma_{e_i}^2$ decreases. The convergence is obtained when the predicted $SINR_i$ approaches the true value. Note that the iterative algorithm is used only to obtain the optimal coefficients of the DFD, and the DFD still works on online signal processing basis (symbol-by-symbol detection). This avoids the significant amount of memory required to store the decisions from previous iteration as is usually the case for the block-iterative approaches (see [9], [10]).

V. PERFORMANCE EVALUATION

A. Complexity performance: CS-DFD versus IS-DFD

The complexity of the DFD includes the pre-processing complexity of the coefficients $\mathbf{G}$, $\mathbf{F}$ and the on-line processing complexity for filtering/recovering the signal symbols. As the latter is the same for both the CS-DFD and IS-DFD, we focus on the pre-processing complexity only. The CS-DFD requires one update of $\mathbf{G}$, $\mathbf{F}$ and $f(\lambda)$ while the IS-DFD requires the update of $\mathbf{G}$, $\mathbf{F}$, and SINR for each iteration. The computation of $\mathbf{G}$ and $\mathbf{F}$ requires $[\cdot]_{LM_r}^{-1} + M_t L^2 M_r^2$ and $L(N-D)M_r M_t^2$ complex multiplications (CMs), respectively, where $[\cdot]_{LM_r}^{-1}$ denotes the complexity of the inversion of an $LM_r \times LM_r$ matrix. The numbers of CMs required for the computation of $f(\lambda)$ and SINR are $[\cdot]_{LM_r}^{-1} + (M_t+1)L^2 M_r^2$ and $LM_r M_t$, respectively. So, assuming N_1 and N_2 iterations are required for the GSS algorithm

Fig. 3. BER performance comparison of the proposed decision feedbacks detectors (DFDs) with conventional DFD, DFD in [6] and ideal feedback DFD when channel length $P = 6$.
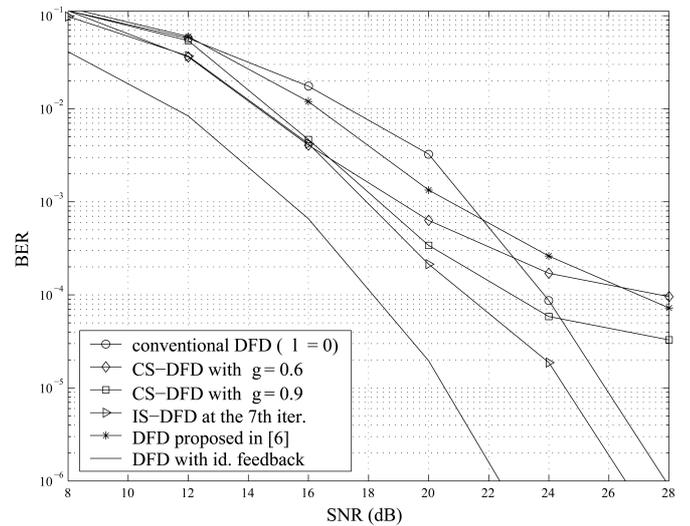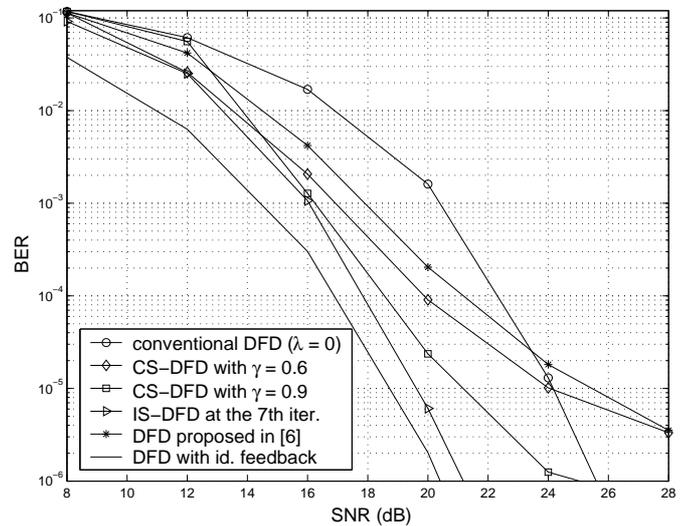

Fig. 4. BER performance comparison of the proposed decision feedbacks detectors (DFDs) with conventional DFD, DFD in [6] and ideal feedback DFD when channel length $P = 12$.

and the IS-DFD, respectively, a complexity comparison for the CS-DFD and IS-DFD is summarized in Table I.

If we approximate $[\cdot]_{LM_r}^{-1} \approx L^3 M_r^3$ and let $L = D = P = 6$, $M_t = M_r = 4$, $N_1 = 5$ (according to our simulation, the GSS algorithm generally converges after 5 iterations), it is seen from Table I that the IS-DFD requires more computation than the CS-DFD does.

B. Numerical simulation

In our simulation, we apply the proposed DFD designs to a MIMO-ISI system employing QPSK modulation. The channel impulse response with a length of $P = 6$ and 12 has been randomly generated. We assume that $h_{m_r,m_t}(p)$ is a complex zero-mean Gaussian random variable with variance one and spatially and temporally uncorrelated. The SNR is defined as

$$SNR = \frac{\sum_{m_t=0}^{M_t-1}\sum_{p=0}^{P-1} E\left[|h_{m_r,m_t}(p)|^2\right]}{\sigma_w^2}.$$

TABLE I
COMPARISON OF PRE-PROCESSING COMPLEXITY: CS-DFD VERSUS IS-DFD

Algorithm	Complex multiplications	$N_2 = 7$	$N_2 = 10$	$N_2 = 15$
CS-DFD	$(N_1 + 1)[\cdot]_{LM_r}^{-1} + (M_t + N1M_t + N1)L^2M_r^2 + L(N - D)M_rM_t^2$	101568	101568	101568
IS-DFD	$(N_2)[\cdot]_{LM_r}^{-1} + N_2M_tL^2M_r^2 + N_2L(N - D)M_rM_t^2 + N_2LM_rM_t$	127008	181440	272160

Let $L = D = 6$ and $M_t = M_r = 4$ be used. For the CS-DFD, two fixed constraint values $\gamma = 0.6$ and 0.9 are used. For the IS-DFD, the set of coefficients $\{\mathbf{G}_{(l)}, \mathbf{F}_{(l)}\}$ obtained after the lth iteration is used for the DFD, thus the resulting bit error rate (BER) behaviour after each iteration can be observed. BERs are obtained using $200 \times (3 \times SNR_{dB} + 1)$ simulation runs where SNR_{dB} is the value of SNR in dB. Note that a new channel impulse response and a new set of 4000 transmitted QPSK symbols are regenerated after each run.

Figures 3 and 4 shows the BER performances of the proposed detection schemes with channel length of $P = 6$ and 12, respectively. The DFD with ideal feedback and the DFD using the method proposed by Tian in [6] are also shown for comparison. We applied Tian's method, where both equality and quadratic inequality constraints are imposed on the FFF to preserve the signal energy and to avoid excessive noise enhancement, to the same MIMO-ISI channel.

Figure 3 shows that the CS-DFD scheme outperforms the conventional DFD at a range of moderate SNRs (i.e., 8-24dB). For example, a difference of 3dB between the conventional DFD and the CS-DFD (with $\gamma = 0.9$) is found at the BER of 10^{-3}. As we impose constraint on the FBF which directly restricts the weight of error term in the feedback, our proposed CS-DFD is slightly better compared to Tian's method. However, when the simulation is carried out at high SNRs ($\geq$ 24dB), the BER error floor is observed for both the CS-DFD and Tian's method as the fixed constraints are no longer efficient. In other words, an effective constraint at a moderate SNR can be an improper constraint at a high SNR. In this case, the IS-DFD scheme must be chosen as it implicitly adapts to the severity of the decision errors as shown by the converged BER performance (at the 7th iteration).

It is worth noting that the proposed schemes perform better with larger channel length. The assumptions made to (9) in the CS-DFD scheme are more well-justified with larger P. The performance of iterative DFD in the IS-DFD scheme is less dependant on the particular channel realization as the predicted SINR can better converge to the true value when the number of taps of the FFF and FBF filter is large (i.e., large P). This is shown in Fig. 4 that the performance of the proposed schemes with $P = 12$ is closer to the ideal-feedback performance when comparing to the case $P = 6$ in Fig. 3.

Figure 5 shows the BER performance of the IS-DFD scheme after each iteration. At high SNRs, the performance of the IS-DFD scheme is satisfactory (no error floor). Fig. 6 further illustrates the convergence behaviour through the average SINR from the output of the IS-DFD scheme with different SNRs. This average SINR is obtained as $(1/M_t)E\left[\sum_{i=1}^{M_t} SINR_i\right]$. The expectation here is replaced by the average value when running a number of realizations. From Fig. 6, we also observe that convergence is obtained after the 7th iteration.

VI. CONCLUSION

We have proposed novel symbol-by-symbol MMSE DFD designs for MIMO-ISI channels. By solving a convex optimization program, the CS-DFD was introduced with robustness against error propagation at moderate SNRs, which was verified by the superior BER performance compared to the conventional DFD. However, to resolve the error floor problem at high SNRs, from which the CS-DFD suffers,
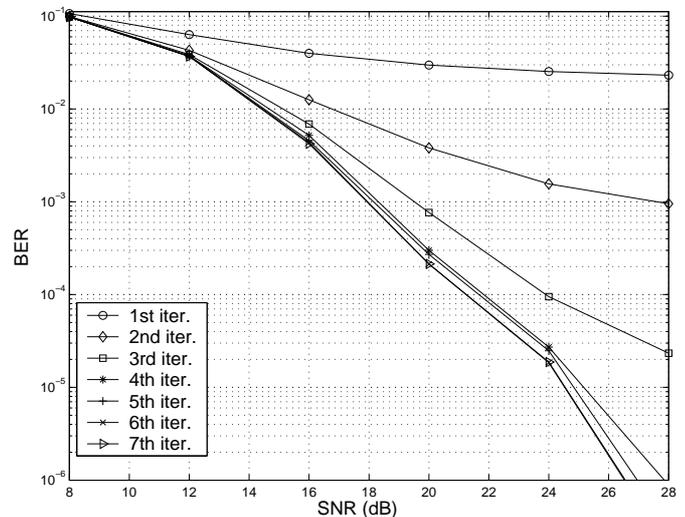

Fig. 5. Convergence behaviour of the iterative symbol-by-symbol DFD (IS-DFD) with channel length $P = 6$.
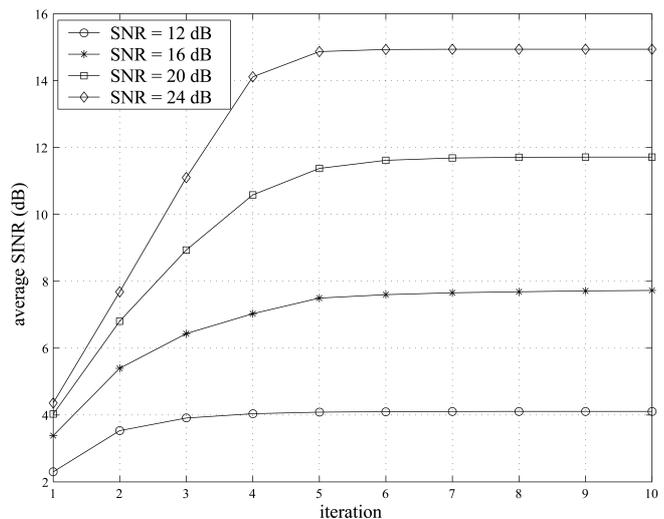

Fig. 6. Average output SINR (with different SNRs) in the IS-DFD approach with channel length $P = 6$.

we proposed the IS-DFD where its coefficients were iteratively updated by considering the severity of the decision error. Simulation showed that the IS-DFD provides a better BER performance at an expense of higher computational load required for iterations.

REFERENCES

[1] J.G. Proakis, *Digital communications*, 4th ed. New York, NY: McGraw-Hill, 2001.

[2] P. Monsen, "Theoretical and measured performance of a DFE modem on a fading multipath channel," *IEEE Trans. Commun.*, vol. COM-25, pp. 1144-1153, Oct. 1977.

[3] M. Chiani, "Introducing erasures in decision-feedback equalization to reduce error propagation," *IEEE Trans. Commun.*, vol. 45, pp. 757-769, Jul. 1997.

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

[5] M. Reuter, J. C. Allen, J. R. Zeidler, and R. C. North "Mitigating error propagation effects in a decision feedback equalizer," *IEEE Trans. Commun.*, vol. 49, pp. 2028-2041, Nov. 2001.

[6] Z. Tian, "Mitigating error propagation in decision-feedback equalization for multiuser CDMA," *IEEE Trans. Commun.*, vol. 52, pp. 525-529, Apr. 2004.

[7] S. Boyd and L. Vandenberghe, *Convex optimization*, Cambridge, England: Cambridge University Press, 2004.

[8] J. Choi, "Equalization and semi-blind channel estimation for space-time block coded signals over a frequency-selective fading channel," *IEEE Trans. Signal Process.*, vol. 52, pp. 774-785, Mar. 2004.

[9] A. M. Chan and G. W. Wornell, "A class of block-iterative equalizers for intersymbol interference channels: fixed channel results," *IEEE Trans. Commun.*, vol. 49, pp. 1966-1976, Nov. 2001.

[10] Y. C. Liang, S. Sun, and C. K. Ho, "Block-iterative generalized decision feedback equalizers (BI-GDFE) for large MIMO systems: Algorithm design and asymtotic performance analysis," *IEEE Trans. Signal Process.*, vol. 54, pp. 2035-2048, Jun. 2006.