FIR Channel-Shortening Equalizers for MIMO ISI Channels *

Naofal Al-Dhahir †

Keywords : FIR digital filters, Equalizers, MIMO Systems, Mean square error methods

Abstract

Finite–length delay–optimized multi–input multi–output (MIMO) equalizers that optimally shorten the impulse response memory of frequency–selective MIMO channels are derived. The MIMO equalizers are designed to minimize the average energy of the error sequence between the equalized MIMO channel impulse response (CIR) and a MIMO target impulse response (TIR) with shorter memory. Two criteria for optimizing the MIMO TIR are analyzed and compared. The presented analytical framework encompasses a multitude of previously–studied finite–length equalization techniques.

1 Introduction

The combination of maximum likelihood sequence estimation (MLSE) with receiver diversity is an effective technique for achieving high performance over noisy frequency–selective fading channels impaired by co–channel interference [8, 14, 16, 15, 7]. With the addition of transmitter diversity, the resulting multi–input multi–output (MIMO) frequency–selective channel was shown [9, 10] to have a significantly higher capacity than its single–input multi–output (SIMO) or single–input single–output (SISO) counterparts. The use of maximum likelihood multi–user detections techniques on these frequency–selective MIMO channels significantly outperforms single–user detection techniques that treat signals from other users as colored noise. However, MLSE complexity increases exponentially with the number of inputs (or transmit antennas) and with the memory of the MIMO channel [1] making its implementation over severe inter–symbol interference (ISI) channels very costly [6].

*This paper will be presented in part at the 34^{th} Asilomar Conference on Signals, Systems, and Computers, October 2000

†Contact author. Phone : $(973)360-8407$, Fax : $(973)360-8178$, Email : naofal@research.att.com

[1] Defined to be the maximum of the memories of impulse responses between all input and output pairs.

The Discrete Matrix Multitone (DMMT) was shown in [9] to be a practical transceiver structure that asymptotically achieves the MIMO channel capacity when combined with powerful codes. It uses the Discrete Fourier Transform (DFT) in its computationally–efficient Fast Fourier Transform (FFT) implementation to partition the frequency responses of the underlying frequency–selective channels of the MIMO systems into a large number of parallel, independent, and (approximately) memoryless frequency subchannels. To eliminate inter–block and intra–block interference, a cyclic prefix whose length is equal to the MIMO channel memory is inserted in every block. On severe–ISI MIMO channels, the cyclic prefix overhead reduces the achievable DMMT throughput significantly unless a large FFT size is used which in turn increases the computational complexity, processing delay, and memory requirements. An elegant solution is to implement a time–domain equalizer to shorten the channel memory and hence reduce the cyclic prefix overhead. This channel–shortening equalizer (CSE) has been well–studied for single–input single–output (SISO) systems (e.g. see [2] and the references therein). In this paper, we present design procedures for the finite–length minimum–mean–square–error (MMSE) CSE for MIMO channels. As we shall show, this CSE performs spatial and temporal noise whitening, multi–channel matched filtering, and MIMO channel memory shortening simultaneously subject to the finite–number–of–taps constraint.

The rest of this paper is organized as follows. In Section II, we present our input–output model and the mathematical formulation of the MIMO CSE problem. Two MIMO CSE design procedures are derived in Section III and their computational complexities are evaluated. Furthermore, several equalization structures are shown to follow as special cases of the MIMO CSE. Numerical results are presented in Section IV and the paper is concluded in Section V.

Notation

The notation to be adopted throughout this paper conforms to the following convention

- Scalars are denoted in lower case : a.
- Unless otherwise stated, vectors are column vectors and are denoted lower case bold : $\mathbf{x}$.
- $\mathbf{e}_i$ denotes the i^{th} unit vector (has a one in the i^{th} position and zeros everywhere else).
- In situations where the components of the vectors are to be emphasized, the first and last components, separated by a colon, are given as a subscript to the vector : $\mathbf{x}_{k+N_f-1:k-\nu}$.
- Matrices are upper case bold : $\mathbf{A}$.
- $\mathbf{I}_N$ denotes the identity matrix of size N.

- $\mathbf{0}_{N\times M}$ denotes an all–zeros matrix with N rows and M columns.

- $|\mathbf{A}|$ denotes the determinant of matrix $\mathbf{A}$.

- $Trace(\mathbf{A})$ denotes the trace of a matrix $\mathbf{A}$.

- E[.] denotes the expected value operator.

- A diagonal matrix with elements $\{d_0, d_1, \cdots, d_{N_f-1}\}$ on the main diagonal will be denoted by $diag(d_0, d_1, \cdots, d_{N_f-1})$.

- The symbol $(.)^*$ will be used to denote the complex–conjugate transpose of a matrix or a vector and the complex conjugate of a scalar.

- The symbol $(.)^t$ will be used to denote the transpose of a matrix or a vector.

Table 1 summarizes the key matrices used in their paper and their sizes.

2 Analysis

We start in this section by describing the input–output model assumed throughout this paper.

2.1 Input-Output Model

We consider the general case of a linear, dispersive, and noisy digital communication system with n_i inputs and n_o outputs. We use the standard complex–valued baseband equivalent signal model. Assuming a temporal oversampling factor of l, the samples at the j^{th} channel output ($1 \leq j \leq n_o$) have the standard form

$$\mathbf{y}_k^{(j)} = \sum_{i=1}^{n_i} \sum_{m=0}^{\nu^{(i,j)}} \mathbf{h}_m^{(i,j)} x_{k-m}^{(i)} + \mathbf{n}_k^{(j)} \ , \tag{1}$$

where $\mathbf{y}_k^{(j)}$ is the j^{th} channel output vector, $\mathbf{h}_m^{(i,j)}$ is the channel impulse response between the i^{th} input and the j^{th} output, whose memory is denoted by $\nu^{(i,j)}$, and $\mathbf{n}_k^{(j)}$ is the noise vector at the j^{th} output. All of these three quantities are $l \times 1$ column vectors corresponding to the l time samples per symbol in the assumed temporally–oversampled channel model. By grouping the received samples from all n_o channel outputs at symbol time k into an $ln_o \times 1$ column vector $\mathbf{y}_k$, we can relate $\mathbf{y}_k$ to the corresponding $n_i \times 1$ column vector of input samples as follows

$$\mathbf{y}_k = \sum_{m=0}^{\nu} \mathbf{H}_m \mathbf{x}_{k-m} + \mathbf{n}_k \ , \tag{2}$$

where $\mathbf{H}_m$ is the MIMO channel matrix coefficient of size $ln_o \times n_i$, and $\mathbf{x}_{k-m}$ is a size $n_i \times 1$ input vector at time $k-m$. The parameter ν is the maximum length of all of the $n_o n_i$ channel impulse responses, i.e., $\nu = max_{i,j}\nu^{(i,j)}$.

Over a block of N_f symbol periods, Equation (2) can be expressed in matrix notation as follows

$$\begin{bmatrix} \mathbf{y}_{k+N_f-1} \\ \mathbf{y}_{k+N_f-2} \\ \vdots \\ \mathbf{y}_k \end{bmatrix} = \begin{bmatrix} \mathbf{H}_0 & \mathbf{H}_1 & \cdots & \mathbf{H}_\nu & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_0 & \mathbf{H}_1 & \cdots & \mathbf{H}_\nu & \mathbf{0} & \cdots \\ \vdots & & & & & & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{H}_0 & \mathbf{H}_1 & \cdots & \mathbf{H}_\nu \end{bmatrix} \begin{bmatrix} \mathbf{x}_{k+N_f-1} \\ \mathbf{x}_{k+N_f-2} \\ \vdots \\ \mathbf{x}_{k-\nu} \end{bmatrix} + \begin{bmatrix} \mathbf{n}_{k+N_f-1} \\ \mathbf{n}_{k+N_f-2} \\ \vdots \\ \mathbf{n}_k \end{bmatrix}, \tag{3}$$

or more compactly

$$\mathbf{y}_{k+N_f-1:k} = \mathbf{H}\mathbf{x}_{k+N_f-1:k-\nu} + \mathbf{n}_{k+N_f-1:k}\ . \tag{4}$$

We define the $n_i(N_f+\nu)\times n_i(N_f+\nu)$ input auto–correlation matrix $\mathbf{R}_{xx} \stackrel{def}{=} \mathrm{E}[\mathbf{x}_{k+N_f-1:k-\nu}\mathbf{x}^*_{k+N_f-1:k-\nu}]$ and the $(n_o l N_f)\times(n_o l N_f)$ noise auto–correlation matrix $\mathbf{R}_{nn} \stackrel{def}{=} \mathrm{E}[\mathbf{n}_{k+N_f-1:k}\mathbf{n}^*_{k+N_f-1:k}]$; *both assumed non-singular.* [2] Then, the input–output cross–correlation and the output auto–correlation matrices are given by

$$\mathbf{R}_{xy} \stackrel{def}{=} \mathrm{E}[\mathbf{x}_{k+N_f-1:k-\nu}\mathbf{y}^*_{k+N_f-1:k}] = \mathbf{R}_{xx}\mathbf{H}^* \tag{5}$$

$$\mathbf{R}_{yy} \stackrel{def}{=} \mathrm{E}[\mathbf{y}_{k+N_f-1:k}\mathbf{y}^*_{k+N_f-1:k}] = \mathbf{H}\mathbf{R}_{xx}\mathbf{H}^* + \mathbf{R}_{nn}\ . \tag{6}$$

2.2 Problem Formulation

Given the MIMO channel matrix $\mathbf{H}$ which has $(\nu+1)$ matrix taps, our objective is to design a MIMO filter with N_f matrix taps, which we denote by the $(n_o l N_f \times n_i)$ matrix $\mathbf{W} \stackrel{def}{=} \begin{bmatrix} \mathbf{W}_0 & \mathbf{W}_1 & \cdots & \mathbf{W}_{N_f-1} \end{bmatrix}^t$, to equalize $\mathbf{H}$ to a MIMO TIR matrix $\mathbf{B} \stackrel{def}{=} \begin{bmatrix} \mathbf{B}_0 & \mathbf{B}_1 & \cdots & \mathbf{B}_{N_b} \end{bmatrix}^t$ with (N_b+1) matrix taps $\mathbf{B}_i$ (each of size $n_i \times n_i$) where $N_b \ll \nu$ (see Figure 1). The MIMO CSE $\mathbf{W}$ is optimized to minimize the equalization mean square error (MSE) defined by $MSE \stackrel{def}{=} trace(\mathbf{R}_{ee})$, where $\mathbf{R}_{ee}$ is the auto–correlation matrix of the equalization error vector $\mathbf{E}_k$ given by [3]

$$\mathbf{E}_k = \tilde{\mathbf{B}}^*\mathbf{x}_{k+N_f-1:k-\nu} - \mathbf{W}^*\mathbf{y}_{k+N_f-1:k}\ . \tag{7}$$

For notational convenience, we define the augmented MIMO TIR matrix $\tilde{\mathbf{B}}^*$ as follows $\tilde{\mathbf{B}}^* \stackrel{def}{=} \begin{bmatrix} \mathbf{0}_{n_i\times n_i\Delta} & \mathbf{B}^*_0 & \mathbf{B}^*_1 & \cdots & \mathbf{B}^*_{N_b} & \mathbf{0}_{n_i\times n_i s} \end{bmatrix} \stackrel{def}{=} \begin{bmatrix} \mathbf{0}_{n_i\times n_i\Delta} & \mathbf{B}^* & \mathbf{0}_{n_i\times n_i s} \end{bmatrix}$ where Δ is the decision delay [3] that lies in the range $0 \le \Delta \le (N_f+\nu-N_b-1)$ and $s \stackrel{def}{=} N_f+\nu-N_b-\Delta-1$.

[2] The analog front-end receive filter must be carefully designed to ensure that $\mathbf{R}_{nn}$ is invertible for the assumed oversampling factor l.

[3] We assume the same decision delay setting for all n_i inputs. Lower MSE is achieved by allowing a variable decision delay across inputs at the expense of increased complexity.

Using the **Orthogonality Principle** which states that $\mathrm{E}[\mathbf{E}_k \mathbf{y}^*_{k+N_f-1:k}] = \mathbf{0}$, it can be shown [3] that the optimum matrix CSE and TIR filters are related by

$$\begin{aligned} \mathbf{W}^*_{opt} &= \tilde{\mathbf{B}}^*_{opt} \mathbf{R}_{xy} \mathbf{R}^{-1}_{yy} &(8)\\ &= \tilde{\mathbf{B}}^*_{opt} \mathbf{R}_{xx} \mathbf{H}^* (\mathbf{H}\mathbf{R}_{xx}\mathbf{H}^* + \mathbf{R}_{nn})^{-1} \\ &= \tilde{\mathbf{B}}^*_{opt} (\mathbf{R}^{-1}_{xx} + \mathbf{H}^* \mathbf{R}^{-1}_{nn} \mathbf{H})^{-1} \mathbf{H}^* \mathbf{R}^{-1}_{nn} \,. &(9) \end{aligned}$$

The last line above shows explicitly that the MIMO CSE consists of a noise whitener $\mathbf{R}^{-1}_{nn}$, a MIMO matched filter $\mathbf{H}^*$, and a bank of FIR channel–shortening filters. It remains to optimize $\tilde{\mathbf{B}}$ such that MSE is minimized. This optimization is carried out in the next section under two criteria.

3 MIMO Channel Shortening Algorithms

The $n_i \times n_i$ error auto–correlation matrix $\mathbf{R}_{ee}$ can be expressed as follows

$$\begin{aligned} \mathbf{R}_{ee} &\stackrel{def}{=} \mathrm{E}[\mathbf{E}_k \mathbf{E}^*_k] \\ &= \tilde{\mathbf{B}}^* (\mathbf{R}_{xx} - \mathbf{R}_{xy} \mathbf{R}^{-1}_{yy} \mathbf{R}_{yx}) \tilde{\mathbf{B}} \\ &= \tilde{\mathbf{B}}^* \mathbf{R}^{\perp} \tilde{\mathbf{B}} \\ &= \mathbf{B}^* \bar{\mathbf{R}} \mathbf{B} \,, &(10) \end{aligned}$$

where $\bar{\mathbf{R}}$ is a submatrix of $\mathbf{R}^{\perp}$ determined by Δ. The optimum MIMO TIR is determined by computing $\mathbf{B}$ that minimizes the trace (or determinant) [4] of $\mathbf{R}_{ee}$ subject to some constraint on its matrix coefficients. We consider two constraints

3.1 Identity Tap Constraint (ITC)

Under the ITC, we restrict the m^{th} matrix coefficient of $\mathbf{B}$ to be equal to the identity matrix. Therefore, we solve the following optimization problem

$$\mathbf{B}^{ITC}_{opt} = argmin_{\mathbf{B}} trace(\mathbf{R}_{ee}) \ \text{ subject to } \ \mathbf{B}^* \mathbf{\Phi} = \mathbf{I}_{n_i} \,, \qquad (11)$$

where $\mathbf{\Phi}^* \stackrel{def}{=} \begin{bmatrix} \mathbf{0}_{n_i \times n_i m} & \mathbf{I}_{n_i} & \mathbf{0}_{n_i \times n_i(N_b - m)} \end{bmatrix}$ and $0 \le m \le N_b$. It can be shown [11] that the optimum MIMO TIR and the corresponding error auto–correlation matrix are given by

$$\begin{aligned} \mathbf{B}^{ITC}_{opt} &= \bar{\mathbf{R}}^{-1} \mathbf{\Phi} (\mathbf{\Phi}^* \bar{\mathbf{R}}^{-1} \mathbf{\Phi})^{-1} &(12)\\ \mathbf{R}^{ITC}_{ee,min} &= (\mathbf{\Phi}^* \bar{\mathbf{R}}^{-1} \mathbf{\Phi})^{-1} \,. &(13) \end{aligned}$$

[4] It can be shown [3] that the same $\mathbf{B}_{opt}$ minimizes the trace and the determinant of $\mathbf{R}_{ee}$.

The delay parameter Δ $(0 \leq \Delta \leq N_f + \nu - N_b - 1)$ that affects $\bar{\mathbf{R}}$ and the index parameter m $(0 \leq m \leq N_b)$ that affects $\mathbf{\Phi}$ are optimized to minimize the trace (or determinant) of $\mathbf{R}_{ee,min}^{ITC}$.

3.2 Ortho-Normality Constraint (ONC)

Under the ONC, we constrain $\mathbf{B}$ to have orthonormal rows, i.e., $\mathbf{B}^*\mathbf{B} = \mathbf{I}_{n_i}$.[5] Therefore, we solve the following optimization problem

$$\mathbf{B}_{opt}^{ONC} = argmin_{\mathbf{B}} trace(\mathbf{R}_{ee}) \text{ subject to } \quad \mathbf{B}^*\mathbf{B} = \mathbf{I}_{n_i} \ . \tag{14}$$

If we define the eigen–decomposition

$$\bar{\mathbf{R}} \stackrel{def}{=} \mathbf{U}\mathbf{\Sigma}\mathbf{U}^* = \mathbf{U} diag(\sigma_0, \cdots, \sigma_{n_i(N_b+1)-1})\mathbf{U}^* \ , \tag{15}$$

where $\sigma_0 \geq \sigma_1 \geq \cdots \geq \sigma_{n_i(N_b+1)-1}$, then the optimum MIMO TIR and the resulting error auto–correlation matrix are given by (see appendix for a proof)

$$\begin{aligned} \mathbf{B}_{opt}^{ONC} &= \mathbf{U}\left[\begin{array}{ccc} \mathbf{e}_{n_i N_b} & \cdots & \mathbf{e}_{n_i(N_b+1)-1} \end{array}\right] && (16) \\ \mathbf{R}_{ee,min}^{ONC} &= diag(\sigma_{n_i N_b}, \cdots, \sigma_{n_i(N_b+1)-1}) \ . && (17) \end{aligned}$$

In words, the optimum MIMO TIR matrix is given by the n_i eigenvectors of $\bar{\mathbf{R}}$ that correspond to its n_i smallest eigenvalues. The delay parameter Δ $(0 \leq \Delta \leq N_f + \nu - N_b - 1)$ that affects $\bar{\mathbf{R}}$ is optimized to minimize the trace (or determinant) of $\mathbf{R}_{ee,min}^{ONC}$. Under both criteria, the performance measure adopted in this paper is equalization SNR defined as follows

$$SNR_{eq} \stackrel{def}{=} \frac{\frac{1}{n_i(N_f+\nu)} trace(\mathbf{R}_{xx})}{\frac{1}{n_i} trace(\mathbf{R}_{ee,min})} \ . \tag{18}$$

We conclude this section by noting that several equalization structures follow as special cases of the finite–length MIMO CSE presented here including

- The finite–length MIMO MMSE-DFE studied in [3] follows as a special case of ITC by setting $\mathbf{B}_0 = \mathbf{I}_{n_i}$ and $\mathbf{\Phi} = \left[\begin{array}{cc} \mathbf{I}_{n_i} & \mathbf{0}_{n_i \times n_i N_b} \end{array}\right]^t$.

- The finite–length MIMO MMSE-LE follows as a special case of ITC by setting $\mathbf{B}_0 = \mathbf{I}_{n_i}$ and $\mathbf{B}_i = \mathbf{0}_{n_i \times n_i}$ $(1 \leq i \leq N_b)$.

[5]This constraint implies that the energy of the aggregate channel impulse response into any particular output j formed by concatenating the channel impulse responses from all n_i inputs to output j is equal to unity for all $1 \leq j \leq n_o$. Moreover, these aggregate channel impulse responses are orthogonal for different outputs.

- The finite–length MIMO partial response LE follows as a special case by setting $\mathbf{B}$ equal to a desired *fixed* TIR.
- The finite–length single–user receive–diversity–combining MMSE–DFE studied in [13, 12] follows as a special case of ITC by setting $n_i = 1$ and $\mathbf{B}_0 = 1$.
- The finite–length MMSE CSE for SISO systems (see [2] and the references therein) follows as a special case by setting $n_i = n_o = 1$. In this case, the ITC and the ONC criteria become identical to the unit–tap and the unit–energy constraints of [2], respectively.

3.3 Computational Complexity

In this subsection, we enumerate the computational tasks required to compute $\mathbf{B}_{opt}$ and $\mathbf{W}_{opt}$ under both ONC and ITC and their complexities as measured by the number of complex multiplies. We assume uncorrelated input and noise processes.

1. Computation of $\bar{\mathbf{R}}$ requires inversion of the *block-Toeplitz* matrix $\mathbf{R}_{yy}$ with a complexity of $O(l^3 n_o^3 N_f^2)$ operations using the efficient algorithm in [1]. Computation of the matrix product $\mathbf{R}_{xy}\mathbf{R}_{yy}^{-1}\mathbf{R}_{yx}$ requires $n_i(N_f+\nu)(l n_o N_f)^2$ complex multiplies.
2. Under ITC, computing $\mathbf{B}_{opt}^{ITC}$ using (12) requires $O((n_i(N_b+1))^3)$ operations to compute $(\bar{\mathbf{R}})^{-1}$, $O(n_i^3)$ operations to compute $(\mathbf{\Phi}^*\bar{\mathbf{R}}^{-1}\mathbf{\Phi})^{-1}$, and $n_i^3(N_b+1)$ complex multiplies to compute $\mathbf{B}_{opt}^{ITC}$.
3. Under ONC, the complexity of computing $\mathbf{B}_{opt}^{ONC}$ using Equation (16) is dominated by the complexity of the eigen–decomposition in Equation (15) which requires $O(n_i^3(N_b+1)^3)$ operations.
4. Finally, computing $\mathbf{W}_{opt}$ using (9) requires $n_i(N_f+\nu)(n_o l N_f)^2 + n_i^2(N_b+1)^2$ complex multiplies.

It is worth mentioning that this complexity estimate assumes a given value of Δ, i.e., the complexity of searching for the optimum Δ is not included. We found that the complexities in the MIMO CSE and TIR computation under both ONC and ITC were comparable.

4 Numerical Results

The CIR's used in our numerical results are unit–energy symbol–spaced (i.e. $l = 1$) FIR filters with 7 taps (i.e. $\nu = 6$) generated as complex zero–mean uncorrelated Gaussian random variables. The input and noise processes are assumed to be uncorrelated. The performance results are calculated by averaging over 100 channel realizations. Figure 2 depicts the variation of the equalization SNR, as defined by Equation (18), versus delay under both ONC and ITC constraints. We assume a 2–input 2–output system with input SNR equal to 20 dB on all 4 channels. The CIR's are equalized to length–3 TIR's (i.e. $N_b = 2$) using a MIMO CSE with $N_f = 12$ and $l = 1$. Three main observations can be made based on Figure 2. First, a suboptimum choice of the delay parameter Δ can result in substantial performance degradation. Second, at any given delay, using the ONC results in better performance than the ITC. Third, as long as Δ is optimized, the effect of optimizing the index parameter under the ITC on performance is marginal. In Figures 3 and 4, we plot the variation of the equalization SNR versus N_f and N_b under ONC with optimized delay. As expected, increasing N_f and N_b results in improved performance which comes at the expense of the increased complexity in implementing the additional equalizer taps and the increased number of states in the MLSE receiver or increased cyclic prefix overhead in DMMT.

5 Conclusions

We derived a general framework for analyzing finite–length MIMO MMSE equalizers that shorten the memory of linear frequency–selective MIMO channels in addition to performing noise whitening and multi–channel matched filtering. Closed–form expressions for the optimum FIR delay–optimized MIMO equalizer and MIMO TIR were derived under an orthonormality and an identity–tap constraint on the MIMO TIR. The orthonormality constraint was shown to result in a lower equalization mean square error. The presented framework accommodates fractionally–spaced equalizers, arbitrary decision delay setting, colored noise, transmit/receive diversity, and variable multipath delay spread across the different diversity paths. Applications to turbo equalization are discussed in [5, 4].

A Proof of Equations (16) and (17)

Starting from Equations (14) and (15), we have

$$trace(\mathbf{R}_{ee}) = \sum_{k=0}^{n_i-1} \mathbf{e}_k^* \mathbf{B}^* \mathbf{U} \mathbf{\Sigma} \mathbf{U}^* \mathbf{B} \mathbf{e}_k \stackrel{def}{=} \sum_{k=0}^{n_i-1} \mathbf{b}_k^* \mathbf{U} \mathbf{\Sigma} \mathbf{U}^* \mathbf{b}_k ,$$

where $\mathbf{b}_k^* \stackrel{def}{=} \mathbf{B}\mathbf{e}_k$. Therefore,

$$\begin{aligned} trace(\mathbf{R}_{ee}) &= \sum_{k=0}^{n_i-1} \sum_{i=0}^{n_i(N_b+1)-1} \sigma_i \mathbf{b}_k^* \mathbf{U} \mathbf{e}_i \mathbf{e}_i^* \mathbf{U}^* \mathbf{b}_k \\ &= \sum_{k=0}^{n_i-1} \sum_{i=0}^{n_i(N_b+1)-1} \sigma_i |\mathbf{e}_i^* \mathbf{U}^* \mathbf{b}_k|^2 \end{aligned}$$

$$\stackrel{def}{=} \sum_{k=0}^{n_i-1} MSE_k \ .$$

At any given Δ, to minimize this sum, we must choose $\mathbf{b}_k$ to minimize each term MSE_k. Since the eigenvalues σ_i are arranged in descending order, it follows that $trace(\mathbf{R}_{ee})$ is minimized by setting $\mathbf{b}_k = \mathbf{U}\mathbf{e}_{(n_i N_b)+k}$ which is equivalent to (16) and results in

$$trace(\mathbf{R}_{ee,min}^{ONC}) = \sum_{k=0}^{n_i-1} \sum_{m=0}^{n_i(N_b+1)-1} \sigma_m |\mathbf{e}_m^* \mathbf{e}_{(n_i N_b)+k}|^2 = \sum_{k=0}^{n_i-1} \sigma_{(n_i N_b)+k} \ , \tag{19}$$

which is identical to (17).

Acknowledgment

The author would like to thank A.F. Naguib, S.N. Diggavi, R. Jana, V. Tarokh, and A.R. Calderbank of AT&T Shannon Laboratory, N. Seshadri of Broadcom, and G. Bauch of Munich University of Technology for many interesting discussions on the subject of this paper.

List of Tables

List of Figures

References

[1] H. Akaike. "Block Toeplitz Matrix Inversion". *SIAM J. Appl. Math.*, 24(2), March 1973.

[2] N. Al-Dhahir and J.M.Cioffi. "Efficiently-Computed Reduced-Parameter Input-Aided MMSE Equalizers for ML Detection : A Unified Approach". *IEEE Trans. Information Theory*, pages 903–915, May 1996.

[3] N. Al-Dhahir and Ali H. Sayed. "A Computationally-Efficient FIR MMSE-DFE for Multi-User Communications". In *Asilomar Conference on Signals, Systems, and Computers*, October 1999.

[4] G. Bauch and N. Al-Dhahir. "Iterative Equalization and Decoding with Channel Shortening Filters for Space-Time Coded Modulation". To be presented at VTC Fall 2000.

[5] G. Bauch and N. Al-Dhahir. "Reduced-Compelxity Turbo Equalization with Multiple Transmit/Receive Antennas over Fading Multipath Channels". In *Conference on Information Sciences and Systems*, pages WP3–13 to WP3–18, March 2000.

[6] G. Bauch and A. Naguib. "MAP Equalization of Space-Time Coded Signals over Frequency-Selective Channels". In *WCNC*, September 1999.

[7] S. Diggavi and A. Paulraj. "Performance of Multisensor Adaptive MLSE in Fading Channels". In *Vehicular Technology Conference*, pages 2148–2152, 1997.

[8] G. Bottomley and K. Jamal. "Adaptive Arrays and MLSE Equalization". In *Vehicular Technology Conference*, pages 50–54, 1995.

[9] G. Raleigh and J. Cioffi. "Spatio-Temporal Coding for Wireless Communication". *IEEE Transactions on Communications*, pages 357–366, March 1998.

[10] G.J.Foschini and M.J.Gans. "On Limits of Wireless Communication in a Fading Environment when Using Multiple Antennas". *Wireless Personal Communication*, pages 311–335, March 1998.

[11] S. Kay. *Fundamentals of Statistical Signal Processing : Estimation Theory*. Prentice Hall, 1993.

[12] S. Lin and V.K. Prabhu. "Optimum Diversity Combining with Finite-Tap Decision Feedback Equalization in Digital Cellular Mobile Radio". In *International Conference on Communications*, June 1997.

[13] N. Lo, D.D.Falconer, and A. Sheikh. "Adaptive Equalization and Diversity Combining for Mobile Radio using Interpolated Channel Estimates". *IEEE Transactions on Vehicular Technology*, pages 636–645, August 1991.

[14] J. Modestino and V. Eyuboglu. "Integrated Multielement Receiver Structures for Spatially Distributed Interference Channels". *IEEE Transactions on Information Theory*, pages 195–219, March 1986.

[15] A. Paulraj and B. Ng. "Space-Time Modems for Wireless Personal Communications". *IEEE Personal Communications*, pages 36–48, February 1998.

[16] K. Scott, E. Olasz, and A. Sendyk. "Diversity Combining with MLSE Equalization". *IEE Proceedings - Communications*, pages 105–108, April 1998.

Matrix	Name	Size
$\mathbf{H}$	Channel Matrix	$ln_oN_f \times n_i(N_f+\nu)$
$\mathbf{R}_{xx}$	Input Auto–Correlation Matrix	$n_i(N_f+\nu) \times n_i(N_f+\nu)$
$\mathbf{R}_{nn}$	Noise Auto–Correlation Matrix	$ln_oN_f \times ln_oN_f$
$\mathbf{R}_{xy}$	Input–Output Cross–Correlation Matrix	$n_i(N_f+\nu) \times ln_oN_f$
$\mathbf{R}_{yy}$	Output Auto–Correlation Matrix	$ln_oN_f \times ln_oN_f$
$\mathbf{W}$	Channel Shortening Equalizer Matrix	$n_olN_f \times n_i$
$\tilde{\mathbf{B}}$	Augmented Target Impulse Response Matrix	$n_i(N_f+\nu) \times n_i$
$\mathbf{R}_{ee}$	Error Auto–Correlation Matrix	$n_i \times n_i$
$\mathbf{R}^{\perp}$	$\stackrel{def}{=} \mathbf{R}_{xx} - \mathbf{R}_{xy}\mathbf{R}_{yy}^{-1}\mathbf{R}_{xy}^{*}$	$n_i(N_f+\nu) \times n_i(N_f+\nu)$
$\bar{\mathbf{R}}$	Defined in Equation (10)	$n_i(N_b+1) \times n_i(N_b+1)$

Table 1: Summary of key matrices used in the paper and their sizes

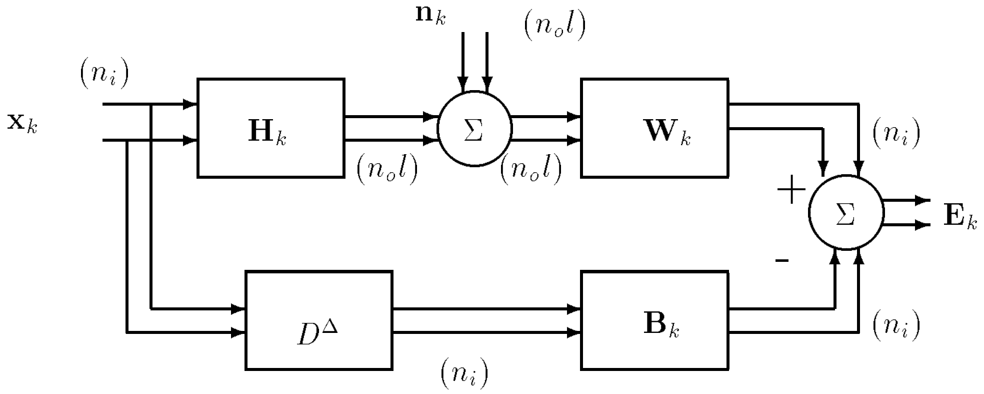

Figure 1: Block diagram of the multi-input multi-output channel shortening equalizer (dimensions of various signals are indicated between the brackets)

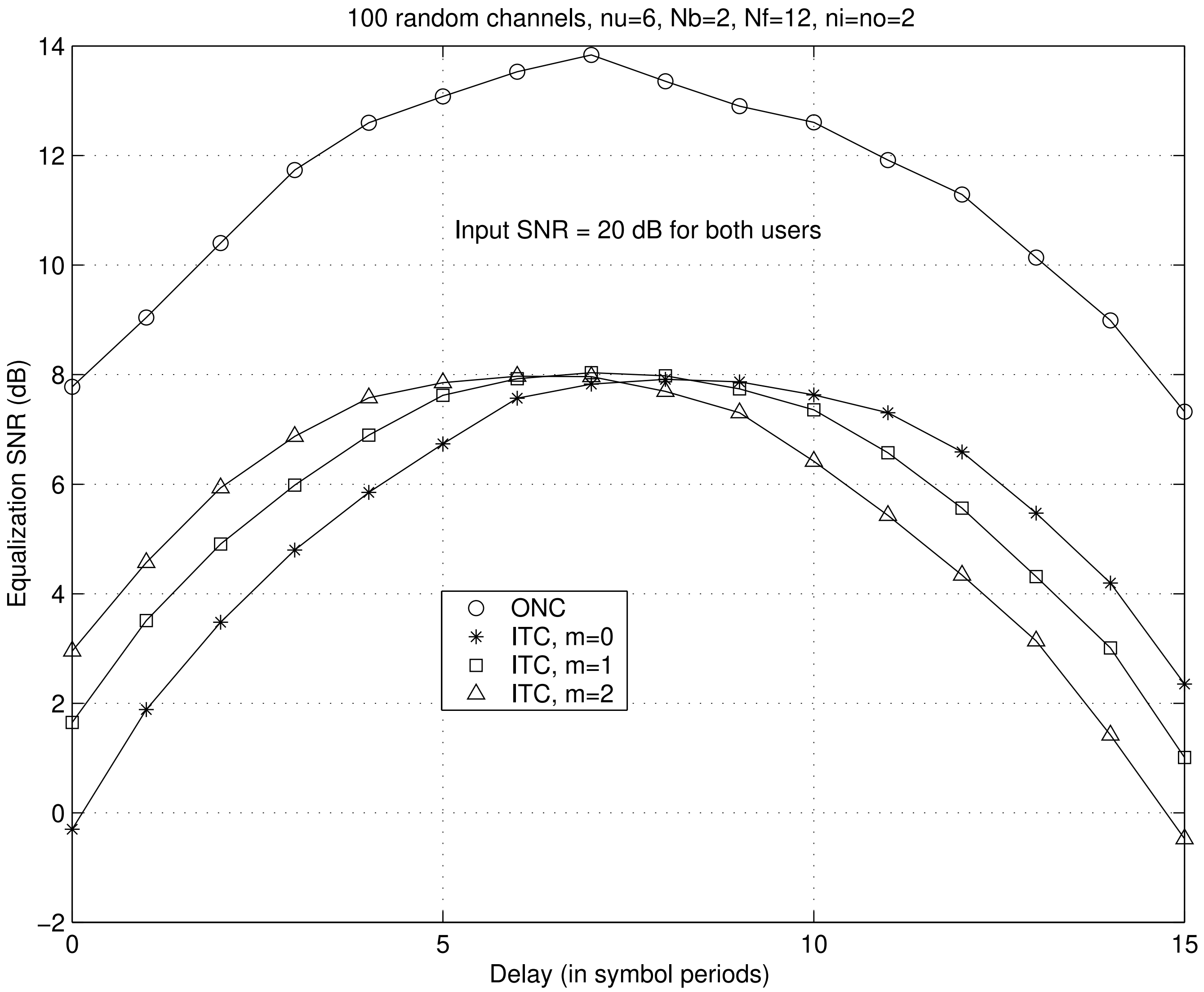

Figure 2: Variation of the equalization SNR of the MIMO CSE under the ONC and ITC constraints versus delay for $\nu = 6$, $N_b = 2$, and $N_f = 12$

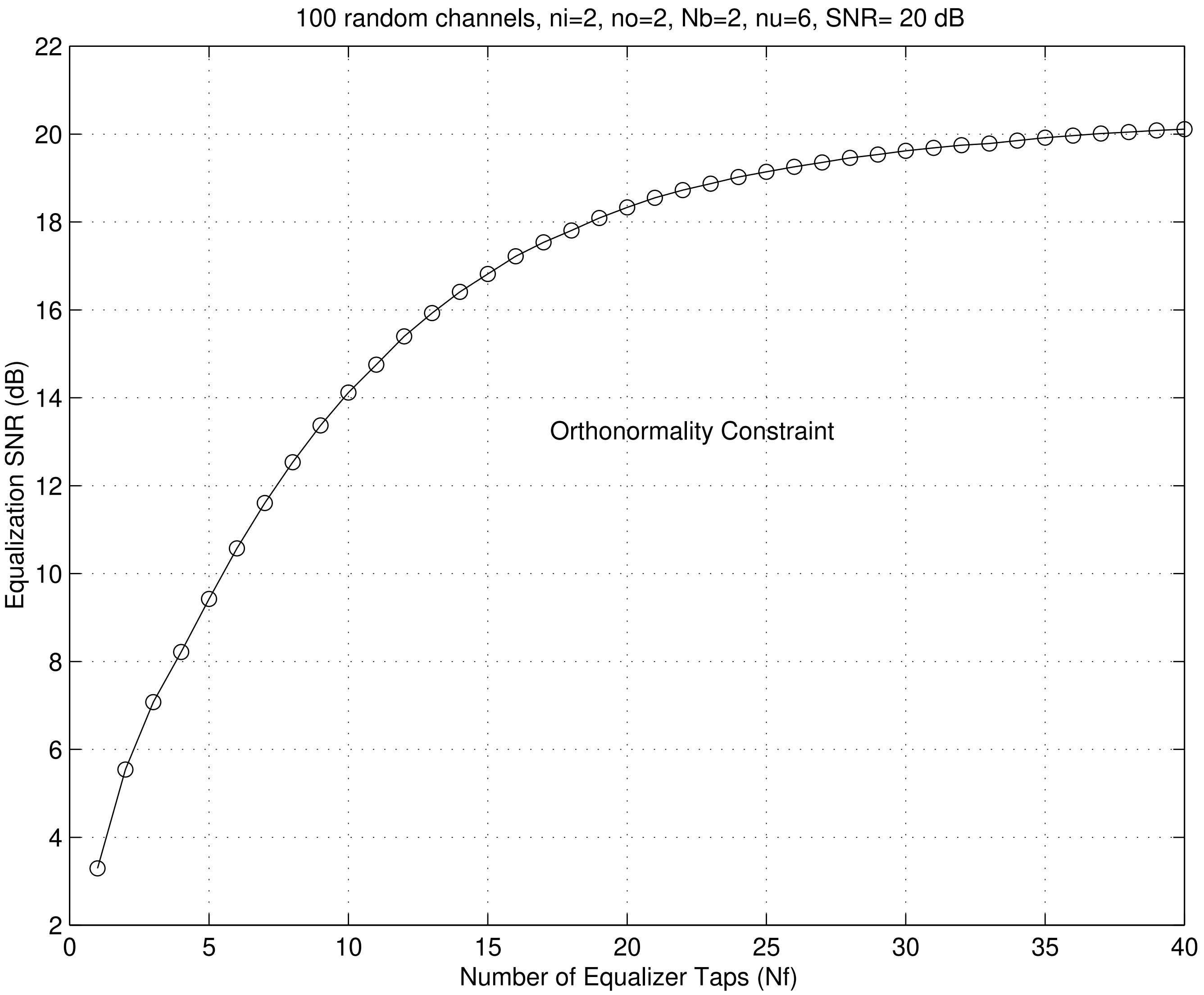

Figure 3: Variation of the equalization SNR of the MIMO CSE under the ONC constraint versus N_f for $\nu = 6$ and $N_b = 2$

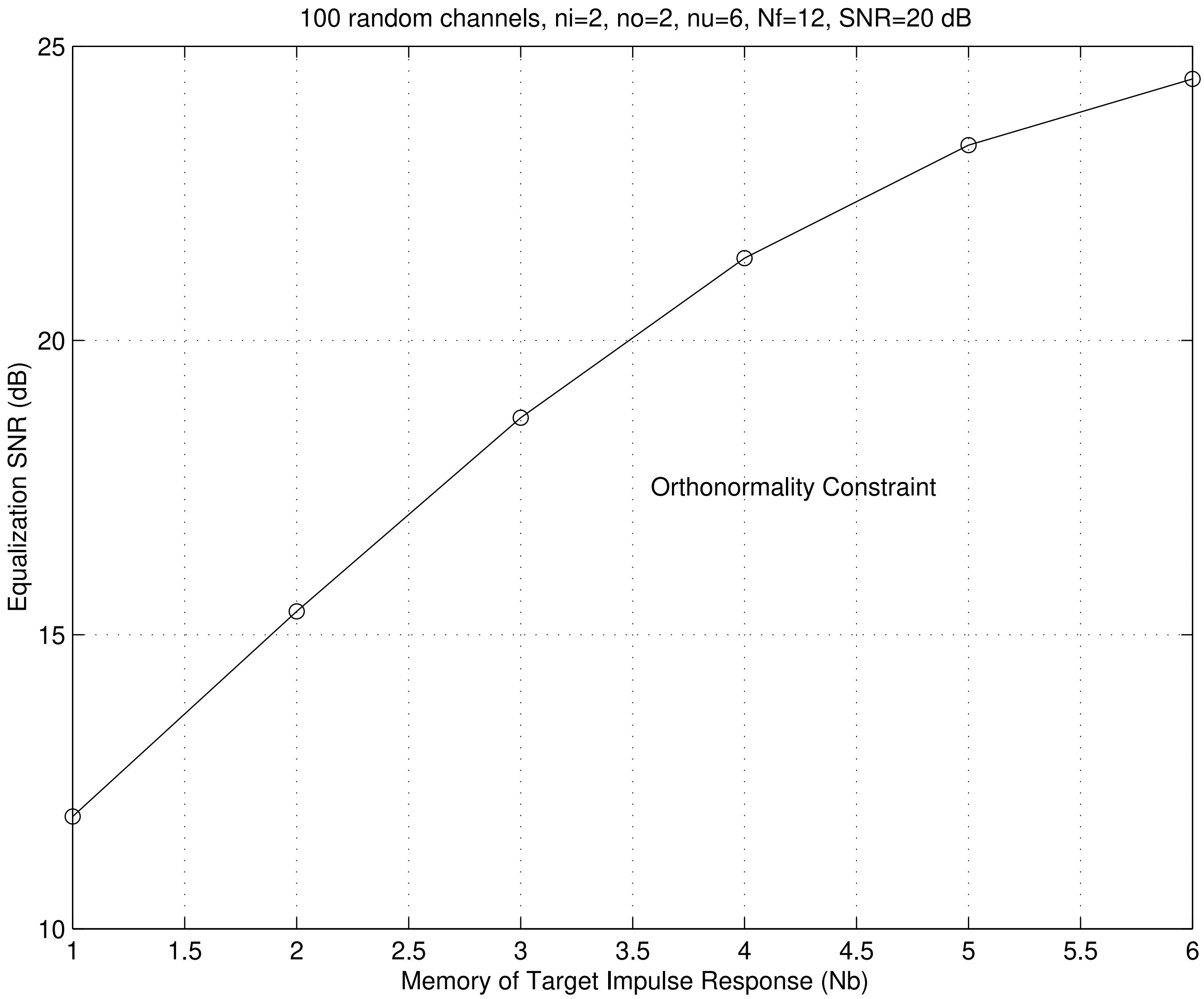

Figure 4: Variation of the equalization SNR of the MIMO CSE under the ONC constraint versus N_b for $\nu = 6$ and $N_f = 12$