

# Initialization Techniques for Improved Convergence of LMS DFEs in Strong Interference Environments

Arun Batra and James R. Zeidler  
Department of Electrical and Computer Engineering  
University of California, San Diego  
La Jolla, California 92093-0409  
Email: abatra@ucsd.edu, zeidler@ece.ucsd.edu

A. A. (Louis) Beex  
DSP Research Laboratory – ECE 0111  
Virginia Tech  
Blacksburg, Virginia 24061-0111  
Email: beex@vt.edu

**Abstract**—The least-mean square (LMS) decision-feedback equalizer (DFE) was previously shown [1], [2] to possess an extended convergence time in an interference limited environment. In [1] it was shown that the convergence time can be significantly reduced by using the received samples and the training data to initialize (data-aided initialization) the LMS weights with an estimate for the Wiener weights. In this paper, two data-aided initialization techniques for equalization in the presence of severe narrowband interference are discussed and compared. The estimate of the Wiener filter is obtained from data-based averages of the autocorrelation matrix and the cross-correlation vector. The first technique is the Multistage Wiener Filter (MSWF) first proposed by Goldstein, *et al.* [3]. This algorithm provides a reduced complexity approach by approximating the Wiener filter in a lower dimensional subspace. The second technique is a parametric approximation to the Direct Matrix Inversion (DMI) solution based on the Gohberg-Semencul formula [4], [5] to obtain the inverse in a computationally efficient fashion. Both techniques were compared in terms of complexity (i.e. the number of multiplications required) and BER performance as compared to the theoretical Wiener filter for the DFE. The MSWF requires fewer training symbols than the approximation to the DMI solution in order to approach the BER performance of the theoretical Wiener filter. The parametric approximation to the DMI solution is computationally efficient but exhibits instability due to assumptions made on the structure of the correlation matrix and when the minimum eigenvalue is close to zero at high SNR.

## I. INTRODUCTION

Severe narrowband interference can cause communications links to become unreliable, requiring receivers to have the ability to quickly mitigate this type of interference. Previous work [6] has shown that adaptive equalizers have the ability to concurrently mitigate narrowband interference and intersymbol interference (ISI). It has also been shown [2] that the convergence time required to successfully implement the concurrent operations of narrowband interference and ISI mitigation becomes excessively long as the interference-to-signal ratio becomes large.

A simple notch filter, with the notch frequency [7] controlled based on the instantaneous frequency of the interference, works well when the interference power is very high;

This work was supported by the Office of Naval Research, Code 313, SPAWAR Systems Center, San Diego, and the UCSD Center for Wireless Communication (UCDG grant # Com 06-10216)

however, it produces excessive signal distortion at moderate and low interference power [8]. The present approach is aimed at extending the good performance of [1] towards lower signal-to-interference (SIR), by addressing the associated convergence issue.

In this paper we will evaluate techniques to initialize the Wiener weights of the adaptive DFE in order to improve convergence. The Wiener weights,  $w_{\text{opt}}$ , minimize the mean squared error (MSE) and provide the means for the equalizer to extract the desired signal from the received signal. When the statistics of the environment are known, these weights can be found by solving the Wiener-Hopf equation,  $\mathbf{R}w_{\text{opt}} = \mathbf{p}$ . In general, this type of *a priori* information is not available at the receiver. One option for obtaining an approximation of the Wiener weights is to use an adaptive algorithm, i.e. least-mean square (LMS), recursive-least squares (RLS), etc [9].

It was previously shown that the LMS DFE requires a large number of training symbols to converge to the Wiener solution in the presence of severe narrowband interference [1], [2], [10]. It was shown in [1] that the convergence time can be significantly reduced by using data-aided initialization. This technique initializes the LMS weights with an estimate of the Wiener weights. These weights are obtained from averages of the autocorrelation matrix and the cross-correlation vector derived from the received samples and the training data. In this paper we will compare the relative BER performance and complexity of two techniques that estimate the Wiener weights in an interference limited environment.

The first approach is based on the concept of Multi-Stage Wiener Filters (MSWFs) as a method that decomposes the conventional Wiener filter into a nested chain of scalar Wiener filters based on orthogonal projections [3]. They can be used to approximate the desired Wiener filter in a lower dimension (i.e a form of rank reduction).

Honig and Xiao [11] observed that the Wiener solution provided by the MSWF is found in the Krylov subspace of the correlation matrix,  $\mathbf{R}$ , and the cross-correlation vector,  $\mathbf{p}$ . This observation led to the use of the Arnoldi algorithm [12] as a means to determine the orthonormal basis vectors. Using the fact that the correlation matrix is Hermitian, Joham, *et al.* [13] noted that the Lanczos algorithm [12] can replace the Arnoldi algorithm when finding the basis vectors. Finally, Dietl, *et al.*

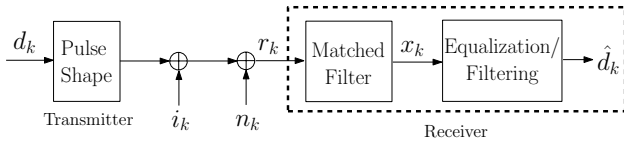


Fig. 1. Discrete-time system model.

[14] derived the relationship between the Conjugate Gradient (CG) algorithm [12], [15], [16] and the Lanczos algorithm for use with MSWFs. This formulation allows the filter weights and the MSE to be updated as each stage is added. The CG implementation of the MSWF also reduces the required complexity by one matrix-vector product as compared to the Lanczos algorithm. The reduced complexity of this approach is illustrated for the implementation of the equalizer for the high speed downlink packet access (HSDPA) receiver in [17], where it was shown that a Krylov equalizer allows a reduction in the computational complexity and storage requirements with almost no loss in performance.

The CG algorithm is a method of the Conjugate Direction family of iterative techniques [16]. It utilizes  $\mathbf{R}$ -orthogonal (or conjugate) search directions, where exactly one step is taken in each direction toward the solution. The solution is guaranteed to be found in  $n$  steps, where  $n$  is the dimension of  $\mathbf{R}$ . This particular algorithm has been shown to be especially useful for solving problems of the type,  $\mathbf{R}\mathbf{w} = \mathbf{p}$ , when the dimensionality of  $\mathbf{R}$  is large and sparse [16], [18].

An alternate technique for finding the inverse of the correlation matrix uses a computationally efficient approximation to the Direct Matrix Inversion (DMI) method. This technique does not find the explicit inverse, instead it uses direct expressions to write down an analytic expression for the inverse. The Gohberg-Semencul formula [4], [5] is a computationally efficient parametric approach for obtaining the inverse [5] of a Toeplitz matrix when the auto-regressive (AR) parameters are known. These AR parameters are obtained using the Levinson algorithm [19]. To find the inverse of the DFE correlation matrix, Schur's complement [20] and the matrix inversion lemma [9] are also needed. From this inverse, we can obtain an approximation for the Wiener solution.

Both techniques provide a significant reduction in computational complexity relative to a direct solution of the Wiener-Hopf equation. In this paper we will compare these two approaches to obtain a solution for the weights of the DFE in terms of computational complexity and performance.

## II. SYSTEM MODEL

A complex baseband representation of a single-carrier communication system is depicted in Fig. 1. The signal of interest is composed of *i.i.d.* symbols, drawn from an arbitrary QAM constellation, with average power equal to  $\sigma_s^2$ . It is passed through a pulse shaping filter that is necessary for bandlimited transmission. This signal is corrupted by narrowband interference and additive white Gaussian noise. A matched filter is employed at the receiver to maximize the signal-to-noise ratio

(SNR) at the output of the filter. Note that the overall frequency response of the pulse shape and the matched filter is assumed to satisfy Nyquist's criterion for no intersymbol interference (ISI) and the filters operate at the symbol rate.

The signal at the input to the equalizer,  $x_k$ , is defined as

$$x_k = d_k + \sqrt{J}e^{j(\Omega kT + \theta)} + n_k, \quad (1)$$

where  $T$  is the symbol duration,  $J$  is the interferer power,  $\Omega$  is the angular frequency of the interferer, and  $\theta$  is a random phase that is uniformly distributed between 0 and  $2\pi$ . The additive noise,  $n_k$  is modeled as a zero-mean Gaussian random process with variance  $\sigma_n^2$ . The signal-to-noise ratio is defined as,  $\text{SNR} = \frac{\sigma_s^2}{\sigma_n^2}$  and the SIR =  $\frac{\sigma_s^2}{J}$ .

It is assumed that the communication signal, interferer, and noise are uncorrelated to each other. The autocorrelation function of the input,  $r_x(m)$ , is defined as

$$\begin{aligned} r_x(m) &= E[x_k x_{k-m}^*] \\ &= (\sigma_s^2 + \sigma_n^2)\delta_m + J e^{j\Omega mT}, \end{aligned} \quad (2)$$

where  $E[\cdot]$  is the expectation operator,  $\delta_m$  is the Kronecker delta function, and  $(\cdot)^*$  indicates conjugation. Note that the elements of the autocorrelation matrix are given by (2).

### A. Theoretical Wiener Solution

The Wiener weights (those that minimize the mean-square error criterion) can be found using the orthogonality principle [9]. For an  $M + 1$ -tap feedforward and an  $M$ -tap feedback equalizer [2],  $2M + 1$  equations are obtained and the weights can be solved using the method described in [6], [21]. The optimal main tap weight ( $w_{\text{DFE},0}$ ), feedforward tap weights ( $w_{\text{DFE},l}$ ), and feedback taps weights ( $f_{\text{DFE},l}$ ) are found to be,

$$w_{\text{DFE},0} = \frac{(\sigma_n^2 + MJ) \sigma_s^2}{(\sigma_s^2 + \sigma_n^2)(\sigma_n^2 + MJ) + \sigma_n^2 J}, \quad (3)$$

$$w_{\text{DFE},l} = \frac{-J \sigma_s^2}{(\sigma_s^2 + \sigma_n^2)(\sigma_n^2 + MJ) + \sigma_n^2 J} e^{-j\Omega lT}, \quad (4)$$

$$f_{\text{DFE},l} = \frac{J \sigma_s^2}{(\sigma_s^2 + \sigma_n^2)(\sigma_n^2 + MJ) + \sigma_n^2 J} e^{-j\Omega lT}, \quad (5)$$

where  $l = 1, \dots, M$ .

### B. Data-Based Averages

The autocorrelation matrix estimate, based on data, is

$$\hat{\mathbf{R}} = \frac{1}{N_{tr}} \sum_{k=1}^{N_{tr}} \mathbf{u}_k \mathbf{u}_k^H, \quad (6)$$

where  $N_{tr}$  is the number of training symbols and  $(\cdot)^H$  represents conjugate (Hermitian) transpose. The received vector is defined as,

$$\mathbf{u}_k = [x_k, \dots, x_{k-M}, \hat{d}_{k-1}, \dots, \hat{d}_{k-M}]^T. \quad (7)$$

Note that because there are training symbols available,  $\hat{d}_k = d_k$ , and  $(\cdot)^T$  is the transpose operator. Finally, the cross-correlation estimate is given by

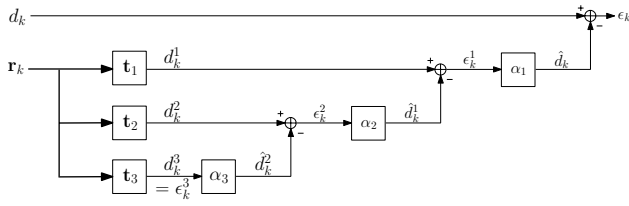


Fig. 2. Multistage Wiener Filter as a Filterbank

$$\hat{\mathbf{p}} = \frac{1}{N_{tr}} \sum_{k=1}^{N_{tr}} \mathbf{u}_k d_k^* \quad (8)$$

### III. MULTISTAGE WIENER FILTER

#### A. Wiener Filter

MSWFs were designed to obtain the Wiener solution without directly finding the inverse of the correlation matrix [3]. An added benefit of this derivation is that an approximate Wiener solution may be found at a lower dimension, thus reducing the complexity of the algorithm. An example of a MSWF as a filter bank can be seen in Fig. 2. Note that  $\mathbf{t}_i$  are the orthonormal basis vectors for the observation space,  $\alpha_i$  are the scalar Wiener filters, and  $d_k$  is the desired value. The error values are defined as  $\epsilon_k^i = d_k^i - \hat{d}_k^i$ .

The basis vectors can be found [13] according to

$$\mathbf{t}_i = \frac{\mathbf{P}_{i-1} \mathbf{P}_{i-2} \mathbf{R} \mathbf{t}_{i-1}}{\|\mathbf{P}_{i-1} \mathbf{P}_{i-2} \mathbf{R} \mathbf{t}_{i-1}\|_2}, \quad (9)$$

where  $\|\cdot\|_2$  is the 2-norm and  $\mathbf{P}_k$  is the projection operation onto the space orthogonal to  $\mathbf{t}_k$ , i.e.,

$$\mathbf{P}_k = \mathbf{I}_{2M+1} - \mathbf{t}_k \mathbf{t}_k^H, \quad (10)$$

where  $\mathbf{I}_{2M+1}$  is the  $(2M+1) \times (2M+1)$  identity matrix. The algorithm is initialized with  $\mathbf{P}_0 = \mathbf{I}_{2M+1}$ ,  $\mathbf{t}_0 = \mathbf{0}$ , and  $\mathbf{t}_1 = \mathbf{p}/\|\mathbf{p}\|_2$ .

Let  $\mathbf{T}^{(D)} = [\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_D]$  be a set of  $D$  basis vectors. The number of basis vectors ( $D$ ) is chosen dependent upon the desired complexity and approximation to the Wiener solution. For  $D = 3$ , (9) and (10) returns 3 basis vectors:

$$\mathbf{T}^{(3)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{M}} e^{-j\Omega T} & 0 \\ \vdots & \vdots & \vdots \\ 0 & \frac{1}{\sqrt{M}} e^{-jM\Omega T} & 0 \\ 0 & 0 & \frac{1}{\sqrt{M}} e^{-j\Omega T} \\ \vdots & \vdots & \vdots \\ 0 & 0 & \frac{1}{\sqrt{M}} e^{-jM\Omega T} \end{bmatrix}. \quad (11)$$

The observation vector,  $d_k^{(3)}$ , is found using the basis given in (11),

$$d_k^{(3)} = \begin{bmatrix} d_k^1 \\ d_k^2 \\ d_k^3 \end{bmatrix} \mathbf{T}^{(3),H} \mathbf{u}_k = \begin{bmatrix} \frac{1}{\sqrt{M}} \sum_{k=1}^M x_{n-k} e^{j\Omega k T} \\ \frac{1}{\sqrt{M}} \sum_{k=1}^M \hat{d}_{n-k} e^{j\Omega k T} \end{bmatrix}. \quad (12)$$

Note that this algorithm is run using training data, so we can replace  $\hat{d}_k$  with the actual transmitted symbols,  $d_k$ .

The first basis vector is chosen to maximize the correlation between the desired signal,  $d_k$ , and the first observed signal,  $d_k^1$ . The second basis vector is orthogonal to the first and is associated with the feedforward side taps. This can be seen by noting that the second observed value,  $d_k^2$ , is a linear combination of past received samples. Notice also that the components of this basis vector contain the phase shifts found in (4) that are needed to cancel the interference. The third basis vector which is orthogonal to the first two, is associated with the feedback taps because the third observed value,  $d_k^3$ , is a linear combination of the fed back symbols. The components have again the phase shifts needed, such that the post-cursor ISI caused by the feedforward side taps can be canceled.

Finally, the scalar Wiener filters can be found as

$$\alpha_i = \frac{r_{\epsilon^i, d^{i-1}}}{\sigma_{\epsilon^i}^2}, \quad (13)$$

where  $r_{\epsilon^i, d^{i-1}} = E[\epsilon_k^i d_k^{*i-1}]$  and  $\epsilon_k^D = d_k^D$ . Using the autocorrelation function given in (2) and the above equations, the scalar Wiener filters are found to be

$$\alpha_3 = 1, \quad (14)$$

$$\alpha_2 = \sigma_s^2 + \sigma_n^2 + MJ, \quad (15)$$

$$\alpha_1 = \frac{(\sigma_n^2 + MJ) \sigma_s^2}{(\sigma_s^2 + \sigma_n^2)(\sigma_n^2 + MJ) + \sigma_n^2 J}. \quad (16)$$

The intermediate estimated observations ( $\hat{d}_k^i$ ) are

$$\hat{d}_k^2 = d_k^3, \quad (17)$$

$$\hat{d}_k^1 = \frac{J\sqrt{M}}{\sigma_n^2 + MJ} [d_k^2 - d_k^3]. \quad (18)$$

The estimate of the desired signal ( $\hat{d}_k$ ) is then found to be

$$\begin{aligned} \hat{d}_k &= \alpha_1 d_k^1 + \alpha_1 \alpha_2 d_k^2 - \alpha_1 \alpha_2 \alpha_3 d_k^3 \\ &= \beta_{\text{DFE}} \begin{bmatrix} \sigma_n^2 + MJ \\ -J e^{-j\Omega T} \\ \vdots \\ -J e^{-jM\Omega T} \\ J e^{-j\Omega T} \\ \vdots \\ J e^{-jM\Omega T} \end{bmatrix}^H \mathbf{u}_k = \mathbf{w}_{\text{DFE}}^H \mathbf{u}_k, \quad (19) \end{aligned}$$

where  $\beta_{\text{DFE}} = \frac{\sigma_s^2}{(\sigma_s^2 + \sigma_n^2)(\sigma_n^2 + MJ) + \sigma_n^2 J}$  and  $\mathbf{w}_{\text{DFE}}$  is the desired Wiener filter. This implies that when the theoretical correlation matrices and the frequency of the interferer are available, the iterative algorithm requires 3 stages to obtain the Wiener filter.

The three subspaces arise from the structure of the DFE, as seen earlier in the basis determination. For the special structure seen in (3)-(5), the feedforward side taps have the same magnitude and the feedback side taps also have the same magnitude. This implies that three values need to be determined: the main

tap, the feedback side taps, and the feedback taps. From the first line of (19),  $\alpha_1$  is the weight needed to scale the current received sample,  $d_k^1 = x_n$ , while the product of  $\alpha_1$  and  $\alpha_2$  provides the weight for the feedforward side taps. Recall that the phase shifts are contained in the basis vector. Finally, noting that the feedforward side taps and the feedback taps are negatives of each other, gives  $\alpha_3 = 1$ , which allows the cancellation of the fed back data symbols.

### B. Complexity

The MSWF described in (9)-(19) uses the knowledge of the theoretical correlation matrices and the frequency of the interference. When this information is unavailable, the CG algorithm described in Algorithm 3 of [14] is used with the received data to provide an estimate of the Wiener filter. The complexity of this algorithm can be determined by counting the number of complex multiplications and divisions needed. This algorithm requires  $4DM^2 + 2(7D + 1)M + 7D + 1$  multiplications and  $2D$  scalar divisions to invert a  $(2M + 1) \times (2M + 1)$  correlation matrix.

## IV. DIRECT MATRIX INVERSION

### A. Wiener Filter

The Wiener solution filter weights can be estimated from the estimated correlation matrices by implicitly solving

$$\hat{\mathbf{R}}\mathbf{w}_{\text{DMI}} = \hat{\mathbf{p}}. \quad (20)$$

If the inverse is found using Gaussian elimination,  $\mathcal{O}((2M + 1)^3)$  multiplications for an  $(2M + 1) \times (2M + 1)$  matrix are required. Instead a method that requires less complexity is evaluated here. An expression for the inverse can be obtained more efficiently using direct formulas, such as the Levinson algorithm [19], the Gohberg-Semencul formula [4], [5], Schur's complement [20], the matrix inversion lemma [9], and the Toeplitz structure of the correlator matrix. We will assume that  $\hat{\mathbf{p}}$  is given by its theoretical value,

$$\mathbf{p} = [\sigma_s^2, 0, \dots, 0]^T, \quad (21)$$

based on the fact that the components of (2) are uncorrelated. Note that the signal power is assumed normalized to unity.

### B. Complexity for Toeplitz Matrices

A simple procedure has been proposed for efficiently finding the inverse of an  $(M + 1) \times (M + 1)$  Toeplitz matrix [12, and references therein] due to its special structure. The first step uses the Levinson algorithm [19] to obtain an estimate for the autoregressive (AR) parameters. This technique requires  $M^2 + 2M + 1$  complex multiplications [12]. These AR parameters are then used with the Gohberg-Semencul [4], [5] formula to obtain the inverse of the matrix. This final step uses  $M^2/2 + 3M/2$  complex multiplications. In all,  $3M^2/2 + 7M/2 + 1$  complex multiplies are necessary to write down the inverse of an  $(M + 1) \times (M + 1)$  Toeplitz matrix. The complexity for finding the inverse of the DFE correlation matrix is found in the next section.

### C. Complexity for the DFE

The theoretical autocorrelation matrix can be partitioned as follows,

$$\mathbf{R} = E[\mathbf{u}_k \mathbf{u}_k^H] = \begin{bmatrix} \mathbf{R}_{xx} & \mathbf{Q}_{dx}^H \\ \mathbf{Q}_{dx} & \sigma_s^2 \mathbf{I}_M \end{bmatrix}. \quad (22)$$

where  $\mathbf{I}_M$  is the  $M \times M$  identity matrix,  $\mathbf{R}_{xx} = E[\mathbf{x}_k \mathbf{x}_k^H]$  is the autocorrelation of the feedforward section, and  $\mathbf{Q}_{dx} = E[\mathbf{d}_k \mathbf{x}_k^H]$  is the cross-correlation of the feedforward section (received signal,  $\mathbf{x}_k$ ) and the feedback section (fed back training symbols,  $\mathbf{d}_k$ ) and is given by

$$\mathbf{Q}_{dx} = [\mathbf{0}_M \quad \sigma_s^2 \mathbf{I}_M], \quad (23)$$

where  $\mathbf{0}_M$  is the  $M \times 1$  zero vector. A technique using Schur's complement [20, pp. 264-265] can be used to invert  $\mathbf{R}$ ,

$$\begin{aligned} \mathbf{R}^{-1} &= \begin{bmatrix} \mathbf{R}_{xx} & \mathbf{Q}_{dx}^H \\ \mathbf{Q}_{dx} & \mathbf{I}_M \end{bmatrix}^{-1}, \\ &= \begin{bmatrix} \mathbf{S}_C^{-1} & -\mathbf{S}_C^{-1} \mathbf{Q}_{dx}^H \\ -\mathbf{Q}_{dx} \mathbf{S}_C^{-1} & \mathbf{I}_M + \mathbf{Q}_{dx} \mathbf{S}_C^{-1} \mathbf{Q}_{dx}^H \end{bmatrix}, \end{aligned} \quad (24)$$

where  $\mathbf{S}_C$  is Schur's complement and is given by

$$\mathbf{S}_C = \mathbf{R}_{xx} - \mathbf{Q}_{dx}^H \mathbf{Q}_{dx}. \quad (25)$$

Using the matrix inversion lemma [9],

$$\begin{aligned} [\mathbf{R}_{xx} - \mathbf{Q}_{dx}^H \mathbf{Q}_{dx}]^{-1} &= \mathbf{R}_{xx}^{-1} - \mathbf{R}_{xx}^{-1} \mathbf{Q}_{dx}^H \\ &\quad \times (\mathbf{Q}_{dx} \mathbf{R}_{xx}^{-1} \mathbf{Q}_{dx}^H - \mathbf{I}_M)^{-1} \\ &\quad \times \mathbf{Q}_{dx} \mathbf{R}_{xx}^{-1}. \end{aligned} \quad (26)$$

Now we are interested in only finding the inverse of Schur's complement which is an  $(M + 1) \times (M + 1)$  matrix. The reduction in complexity for this method arises from the assumption of knowledge of both  $\mathbf{Q}_{dx}$  and  $\mathbf{I}_M$  of (22). Note that this assumption will result in a degradation in the approximated Wiener filter when there is not an adequate number of training symbols. The assumption holds for a large number of training symbols because the estimated correlation matrix will be closer to the theoretical matrix.

The algorithm now requires finding the inverse of Schur's complement and it is found through the following steps, ignoring multiplications by 1:

- Let  $\mathbf{Z} = \mathbf{R}_{xx}^{-1}$ . The resulting matrix ( $\mathbf{Z}$ ) happens to be Toeplitz and is found using the procedure discussed in the previous section, thus requiring  $3M^2/2 + 7M/2 + 1$  complex multiplications. Note that it is not always the case that the inverse of a Toeplitz matrix is Toeplitz [22].
- Let  $\mathbf{Y} = \mathbf{Q}_{dx} \mathbf{R}_{xx}^{-1} \mathbf{Q}_{dx}^H$ . This step does not involve any multiplications; instead,  $\mathbf{Y}$  is equal to the lower-right  $M \times M$  submatrix of  $\mathbf{Z}$ . Note that  $\mathbf{Y}$  is still Toeplitz.
- Let  $\mathbf{X} = (\mathbf{Q}_{dx} \mathbf{R}_{xx}^{-1} \mathbf{Q}_{dx}^H - \mathbf{I}_M)^{-1}$ . This inverse is found in the same manner as the first step, since the term in parentheses is Toeplitz. This step requires  $3M^2/2 + M/2$  complex multiplications. The matrix  $\mathbf{X}$  is again Toeplitz.

- Let  $\mathbf{W} = \mathbf{Q}_{dx} \mathbf{X} \mathbf{Q}_{dx}^H$ . This step does not require any multiplications.  $\mathbf{W}$  is equal to  $\mathbf{X}$  with an additional row of zeros as the top of the matrix and an additional column of zeros at the front of the matrix.
- Finally,  $\mathbf{V} = \mathbf{R}_{xx}^{-1} \mathbf{W} \mathbf{R}_{xx}^{-1} = \mathbf{Z} \mathbf{W} \mathbf{Z}$ . The matrix  $\mathbf{V}$  has the desirable properties that the lower-right  $M \times M$  submatrix is Hermitian Toeplitz and the top row and first column are related via Hermitian symmetry.  $\mathbf{V}$  can be found by performing  $M(4M + 1)$  complex multiplications.

Finally, the components of  $\mathbf{R}^{-1}$  are found with the following steps:

- $\mathbf{S}_C^{-1}$  is found by performing  $7M^2 + 5M + 1$  complex multiplications as discussed in the previous steps.
- $\mathbf{S}_C^{-1} \mathbf{Q}_{dx}^H$  is the inverse of the Schur complement (found in the previous step) with the first column eliminated.
- $\mathbf{Q}_{dx} \mathbf{S}_C^{-1}$  is the inverse of the Schur complement with the first row eliminated.
- $\mathbf{Q}_{dx} \mathbf{S}_C^{-1} \mathbf{Q}_{dx}^H$  is the inverse of the Schur complement with both the first row and first column eliminated.

A total of  $7M^2 + 5M + 1$  complex multiplications are needed to determine the inverse. The approximation for the Wiener solution is then found by multiplying this inverse by the cross-correlation vector. Recalling that the signal power is normalized to unity,  $\mathbf{p}$  simply selects the first column of the inverse. No additional multiplications are necessary to find the approximation of the Wiener filter.

#### D. Eigenvalues of Schur's Complement

The minimum eigenvalue of (25) for the case of  $M > 1$  is found to be  $\lambda_{\min} = \sigma_n^2$  with multiplicity  $M - 1$ . When this eigenvalue approaches zero (i.e SNR approaches infinity), the inverse may diverge as shown in Sec. V.

### V. RESULTS

In the simulation results to follow, a BPSK constellation is considered, the SIR = -20 dB, and the interferer frequency is located at DC ( $\Omega = 0$ ). The equalizer order is set to,  $M = 3$ , where the DFE is formed by a 4-tap feedforward filter and a 3-tap feedback filter. Each packet is made up of 10,000 symbols. Note that the received samples and the training symbols are used to form the estimates of the correlation matrices given in (6) and (8). The estimated weights are then used to calculate the BER in a decision-directed mode over the remaining symbols.

Fig. 3 demonstrates the BER for the theoretical Wiener filter and the full-rank MSWF as a function of the number of training symbols. The MSWF DFE suffers when the estimated correlation matrices are poor, due to the lack of training symbols. When this number is increased, the performance of the CG MSWF approaches that of the theoretical Wiener filter. For this scenario, both 250 and 500 training symbols provide a good approximation, which entails a 2.5-5% of overhead for this scenario.

Fig. 4 is a plot of the BER for the theoretical Wiener filter and the reduced-rank MSWF as a function of the number of

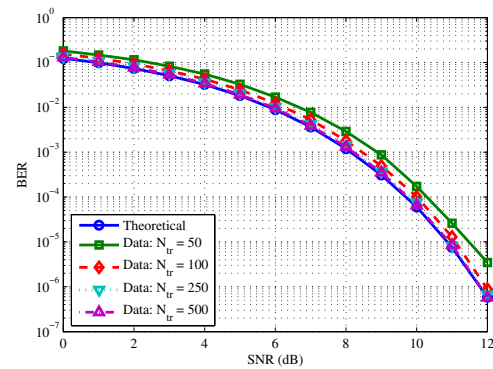


Fig. 3. BER performance for the theoretical Wiener filter and the full-rank CG MSWF for a varying number of training symbols, SIR = -20 dB.

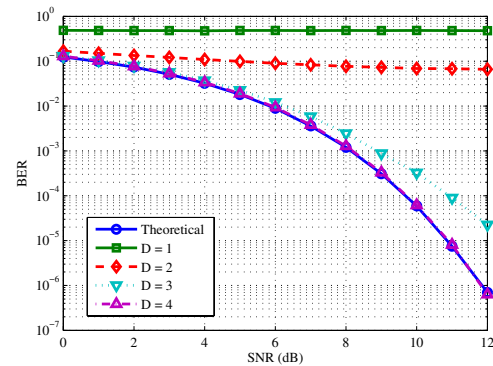


Fig. 4. BER performance for the theoretical Wiener filter and the reduced-rank CG MSWF for a varying number of stages, SIR = -20 dB.

implemented stages, with  $N_{tr} = 500$ . When employing the theoretical correlation matrices with the MSWF algorithm, it was previously shown that 3 stages are needed to obtain the Wiener solution. In this plot for  $D = 1, 2$ , the performance is very poor, indicating that the MSWF solution is far from the desired. An improvement is obtained when using  $D = 3$  stages, however toward the high SNR region, the performance deviates from the ideal. An additional stage ( $D = 4$ ) improves the performance and it is comparable to that of the theoretical Wiener filter.

Fig. 5 shows the BER for the theoretical Wiener filter and the parametric approximation to the DMI solution as a function of the number of training symbols. The parametric approximation to the DMI solution provides a good approximation to the desired correlation matrices at low SNR for all sizes of training symbols. However, as the SNR increases, two forms of degradation for the parametric approximation to DMI solution are noticed. The first deviation from the theoretical curve occurs because of the assumptions made on the structure of the correlation matrix. Recall that we are attempting to invert,  $\hat{\mathbf{S}}_C = \hat{\mathbf{R}}_{xx} - \hat{\mathbf{Q}}_{dx}^H \hat{\mathbf{Q}}_{dx}$ , however, because we assume that  $\hat{\mathbf{Q}}_{dx} = \mathbf{Q}_{dx}$  and  $\hat{\mathbf{I}}_M = \mathbf{I}_M$  in (22), we are actually solving  $\hat{\mathbf{S}}_C = \hat{\mathbf{R}}_{xx} - \mathbf{Q}_{dx}^H \mathbf{Q}_{dx}$ . These assumptions, the forced Toeplitz structure of  $\hat{\mathbf{R}}_{xx}$ , and the use of the ideal cross-correlation vector ( $\mathbf{p}$ ) lead to a loss of information

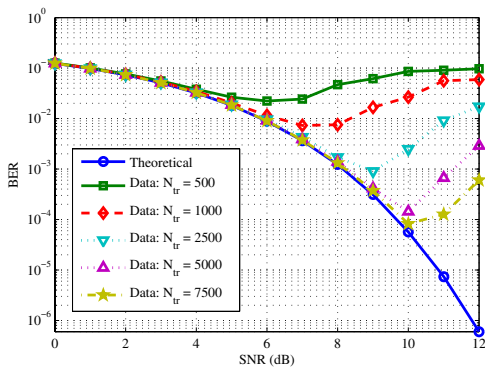


Fig. 5. BER performance for the theoretical Wiener filter and the parametric approximation to the DMI solution for a varying number of training symbols, SIR = -20 dB.

when inverting the matrix and finding the approximate Wiener solution. Note that the initial deviation from the theoretical curve occurs earlier for smaller sets of training symbols. As more training is employed, the estimate of the correlation matrix approaches the theoretical matrix, which allows the assumptions made to hold and provide better performance. The second degradation occurs at higher SNR values and is due the minimum eigenvalue of the data-based estimate of Schur's complement. This eigenvalue given in Sec. IV-D is close to zero in this SNR region causing the inversion process to suffer. The MSWF algorithm seen in Figs. 3 and 4 is not affected in a similar manner because the CG implementation is less sensitive to eigenvalue spread [16], [18] of  $\mathbf{R}$ .

For the parameters used, the parametric approximation to the DMI solution is less complex, needing 79 multiplications, while the CG MSWF requires 347 multiplications and 8 scalar divisions. Note, however, that the impact of these additional computations for the CG MSWF on performance is shown to be significant since the requirement of additional training symbols and reduced reliability is a major system limitation.

## VI. CONCLUSION

We have investigated two different computationally efficient techniques for obtaining an estimate of the Wiener weights for the initialization of the decision-feedback equalizer (DFE) from estimates of the correlation matrices in the presence of severe narrowband interference. The methods are compared in terms of complexity and performance relative to that of the theoretical Wiener filter for the DFE. The parametric approximation to the DMI solution is the lower complexity method ( $7M^2 + 5M + 1$  multiplications), however, the performance is degraded due to assumptions on the correlation matrix ( $\mathbf{R}$ ) that are only met when a large number of training symbols is used. Further degradation occurs when the minimum eigenvalue gets small (at high SNR) causing the inverse to possibly diverge. The CG MSWF algorithm proves to be a better option than the parametric approximation to the DMI solution even though its complexity is larger ( $4DM^2 + 2(7D + 1)M + 7D + 1$  multiplications and  $2D$  scalar divisions) because it needs fewer

training symbols to obtain performance comparable to that of the theoretical DFE Wiener filter for all SNR values.

## REFERENCES

- [1] A. Batra, T. Ikuma, J. R. Zeidler, A. A. Beex, and J. G. Proakis, "Mitigation of unknown narrowband interference using instantaneous error updates," in *Conference Record of the 38th Asilomar Conference on Circuits Systems and Computers*, vol. 1, Pacific Grove, CA, Nov. 2004, pp. 115–119.
- [2] A. Batra, J. R. Zeidler, and A. A. Beex, "Mitigation of narrowband interference using adaptive equalizers," in *Proceedings of the European Signal Processing Conference (EUSIPCO)*, Florence, Italy, Sept. 2006.
- [3] J. S. Goldstein, I. S. Reed, and L. L. Scharf, "A multistage representation of the Wiener filter based on orthogonal projections," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2943–2959, Nov. 1990.
- [4] I. C. Gohberg and A. A. Semencul, "On the inversion of finite Toeplitz matrices and their continuous analogues (in Russian)," *Mat. Issled.*, vol. 2, pp. 201–203, 1972.
- [5] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*, 1st ed. Philadelphia, PA: Prentice Hall, 1997.
- [6] L.-M. Li and L. B. Milstein, "Rejection of CW interference in QPSK systems using decision-feedback filters," *IEEE Trans. Commun.*, vol. COM-31, pp. 473–483, Apr. 1983.
- [7] B.-S. Chen, T.-Y. Yang, and B.-H. Lin, "Adaptive notch filter by direct frequency estimation," *Signal Processing*, vol. 27, pp. 161–176, May 1992.
- [8] P. Shan and A. A. Beex, "Time-varying filtering using full spectral information for soft-cancellation of FM interference in spread spectrum communications," in *2nd IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Annapolis MD, May 1999, pp. 321–324.
- [9] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, NJ: Prentice Hall, 2002.
- [10] A. Batra, J. R. Zeidler, and A. A. (Beex), "A two-stage approach for improving the convergence of least-mean-square adaptive equalizers in the presence of severe narrowband interference," *EURASIP Journal on Applied Signal Processing*, In Press.
- [11] M. L. Honig and W. Xiao, "Performance of reduced-rank linear interference suppression," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1928–1946, July 2001.
- [12] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed. Baltimore: The Johns Hopkins Univ. Press, 1984.
- [13] M. Joham and M. D. Zoltowski, "Interpretation of the multi-stage nested Wiener filter in the Krylov subspace framework," Purdue University/Munich University of Technology, Tech. Rep. TR-ECE-00-51/TUM-LNS-TR-00-6, 2000. [Online]. Available: [http://cobweb.ecn.purdue.edu/~mikedz/research/msnwf\\_tutorial.pdf](http://cobweb.ecn.purdue.edu/~mikedz/research/msnwf_tutorial.pdf)
- [14] G. Dietl, M. D. Zoltowski, and M. Joham, "Recursive reduced-rank adaptive equalization for wireless communications," *Proc. of SPIE*, vol. 4395, pp. 16–27, Apr. 2001.
- [15] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of Research of the National Bureau of Standards*, vol. 49, no. 6, pp. 409–432, Dec. 1952.
- [16] J. R. Shewchuk, "An introduction to the conjugate gradient method without the agonizing pain," Aug. 1994, unpublished. [Online]. Available: <http://www.cs.cmu.edu/~quake-papers/painless-conjugate-gradient.pdf>
- [17] C. Dumard, F. Kaltenberger, and K. Freudenthaler, "Low-cost approximate LMMSE equalizer based on krylov subspace methods for HSDPA," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1610–1614, May 2007.
- [18] L. N. Trefethen and D. Bau, *Numerical Linear Algebra*, 1st ed. Philadelphia, PA: Soc for Industrial & Applied Math, 1997.
- [19] N. Levinson, "The Wiener RMS error criterion in filter design and prediction," *Journal of Mathematical Physics*, vol. 25, pp. 261–278, 1947.
- [20] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. Upper Saddle River, NJ: Prentice Hall, 2000.
- [21] J. R. Zeidler, E. H. Satorius, D. M. Chabries, and H. T. Wexler, "Adaptive enhancement of multiple sinusoids in uncorrelated noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-26, pp. 240–254, June 1978.
- [22] R. M. Gray, "Toeplitz and circulant matrices: A review," *Foundations and Trends in Communications and Information Theory*, vol. 2, pp. 155–239, 2006.