Correspondence

Adaptive Conjugate Gradient DFEs for Wideband MIMO Systems

Aris S. Lalos, Vassilis Kekatos, and Kostas Berberidis

***Abstract*—New adaptive equalization algorithms for wireless systems operating over wideband multiple-input multiple-output (MIMO) channels are proposed. The problem of MIMO decision feedback equalizer (DFE) design is formulated as a set of linear equations with multiple right-hand sides (RHSs) evolving in time. By applying an adaptive modified conjugate gradient algorithm, we derive an equalizer with identical convergence, improved tracking capabilities, but higher computational load as compared to the Recursive Least Squares (RLS) algorithm. To reduce its complexity, two updating strategies of the equalizer filters based on Galerkin projections are employed.**

***Index Terms*—Adaptive equalizers, conjugate gradient methods, decision feedback equalizers, Galerkin method, MIMO systems.**

I. Introduction

Equalization of wireless multiple-input multiple-output (MIMO) frequency-selective channels is a challenging task mainly due to the fact that the respective MIMO equalizers should cope with intersymbol, as well as interstream interference. When the channel is static and has already been estimated by the receiver, a well established solution would be to apply a multicarrier technique, such as a MIMO orthogonal-frequency-division multiplexing (OFDM) system [2]. Even though MIMO OFDM systems offer simplicity in analysis and receiver design, they still suffer from drawbacks related to implementation (peak to average power ratio), identifiability (spectral nulls), and sensitivity to carrier sychronization. Another drawback is that uncoded OFDM has no significant performance gain as the delay spread of the channel increases [3], i.e., uncoded OFDM does not exploit multipath diversity. On the other hand, single-carrier (SC) modulation is a well-proven technology in many existing wireless and wireline applications and has been extensively used in practice (e.g., SC frequency domain equalization has been proposed in 802.16 [4]). Thus, alternative SC approaches for the design of batch MIMO decision feedback equalizers (DFEs) have been proposed in [5] and [6].

However, in MIMO systems with relatively long bursts and under time varying conditions, the involved channel impulse responses change within a burst and, as expected, batch MIMO DFEs fail to equalize the channel. On the other hand, if a MIMO OFDM is adopted, the frame size should be made short and thus cyclic prefix overhead becomes overwhelming. Therefore, to achieve effective channel equalization in such cases, adaptive methods are required. Both a minimum bit error rate (MBER) design [7], [8] and the standard minimum mean-square error (MMSE) design [9]–[11] have been invoked for implementing adaptive MIMO DFEs. The respective equalizers are updated either by using gradient Newton methods (i.e., RLS, LMS-Newton) or by employing stochastic gradient techniques. The main problems appearing in adaptive MIMO equalization, i.e., the increased filter size and the colored noise caused by interstream interference, slow down significantly the performance of stochastic gradient algorithms. On the other hand, the computational requirements of MIMO RLS algorithms increase significantly. Thus, the aim is to seek for adaptive schemes with convergence properties close to RLS but of lower computational cost.

It is known that single-input single-ouptut (SISO) adaptive algorithms based on the CG (conjugate gradient) method [12]–[14] are numerically stable and exhibit convergence properties comparable to RLS with a computational cost that lies between the RLS and LMS algorithms. CG methods were developed for iterative solutions of finite linear equations during the early 1950s [15], [16]. Later, these methods were extended for solving linear equations with multiple right-hand sides [17]. The aim in conjugate gradient methods is to accelerate the slow convergence rate of the steepest descent method while avoiding the involvement of a Hessian matrix associated with Newton methods. To the best of our knowledge, no work has been done towards developing MIMO adaptive equalization algorithms based on the CG method.

In this work, our main goal is to derive a computationally efficient CG-based algorithm for updating the MIMO DFE filters, based on the MMSE design criterion. To this end, we initially formulate the adaptive MIMO DFE problem as a set of linear equations evolving in time. It should be noted that many MIMO adaptive filtering problems, such as system identification and echo cancellation, may have a similar formulation. Therefore, such a problem formulation, allows for the applicability of the proposed schemes to other adaptive filtering problems with simple and straightforward modifications. Then, we apply the idea of [12] to the MIMO case, resulting in a numerically stable scheme (so-called MIMO-modified CG, MIMO-MCG) with convergence properties similar to MIMO-RLS. The main drawback of this scheme is its increased computational cost. Thus, we focus on reducing the cost of the aforementioned algorithm by making use of some linear algebra tools known as Galerkin projections. The Galerkin projection techniques have been proposed for solving linear systems with multiple right hand sides. By incorporating the above ideas to the adaptive MIMO DFE design, we finally derive two new schemes with complexity lower than that of RLS and almost identical convergence properties.

The rest of the correspondence is organized as follows. In Section II, the problem of the adaptive MIMO DFE design is formulated as a set of time-evolving linear equations with multiple RHS. Furthermore, we present the conjugate gradient optimization method for solving single linear systems. In Section III, the new adaptive MIMO DFEs are derived. Their performance is illustrated in Section IV by means of numerical examples, and the work is concluded in Section V.

Manuscript received July 30, 2008; accepted January 15, 2009. First published February 24, 2009; current version published May 15, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Aleksandar Dogandzic. This work was supported in part by the EU FP6 COOPCOM project (#FP6-033533) and in part by the PENED #03ED838 research project, co-financed by national and E.C. funds. This work was presented in part at the IEEE Vehicular Technology Conference, Baltimore, MD, October 2007.

A. S. Lalos and K. Berberidis are with the Department of Computer Engineering and Informatics and RACTI/RU8, University of Patras, 26500 Rio-Patras, Greece (e-mail: lalos@ceid.upatras.gr; berberid@ceid.upatras.gr).

V. Kekatos was with the Department of Computer Engineering and Informatics, University of Patras, Greece. He is now with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: kekatos@umn.edu).

Digital Object Identifier 10.1109/TSP.2009.2016245

II. Problem Formulation and Preliminaries

In this section, we initially formulate the adaptive MIMO DFE problem as a set of linear equations evolving in time. It is also shown that the feedforward (FF) and feedback (FB) filters may be treated separately. Also we briefly present the CG optimization technique for solving linear systems.

A. System Model and Problem Formulation

Let us consider a MIMO communication system operating over a frequency selective wireless channel. The system employs M transmit and N receive antennas, with $M \leq N$, while spatial multiplexing is assumed. The signal transmitted through the M antennas at time k can be described by the vector

$$\mathbf{s}(k) = [\, s_1(k) \quad \dots \quad s_M(k) \,]^T \tag{1}$$

where $s_i(k), i = 1, \dots, M$, are i.i.d. symbols of unit variance. By employing a discrete-time complex baseband model, the signal received at the nth antenna can be expressed as

$$x_n(k) = \frac{1}{\sqrt{M}} \sum_{i=1}^{M} \sum_{l=0}^{L} h_{ni}(l) s_i(k-l) + \eta_n(k) \tag{2}$$

where $h_{ni}(l)$ for $l = 0, \dots, L$, is the sampled impulse response between transmitter i and receiver n, $(L+1)$ is the channel length, and $\eta_n(k), n = 1, \dots, N$, are white Gaussian complex noise samples of variance $N_0/2$ per dimension. The samples received at time k can be assembled in the vector

$$\underline{\mathbf{x}}(k) = [\, x_1(k) \quad \dots \quad x_N(k) \,]^T. \tag{3}$$

The intersymbol and interstream interference involved in the system described by (2) can be mitigated through a MIMO DFE [9]. The proposed equalizer architecture is a structure of M MISO DFEs operating in parallel. The ith MISO DFE is designated to extract the ith stream $s_i(k)$, and it consists of a feedforward and a feedback filter of temporal span K_f and K_b taps, respectively. The input of the feedforward filter $\mathbf{f}_i(k)$, for $i = 1, \dots, M$, can be described by the $NK_f \times 1$ vector

$$\mathbf{x}(k) = [\, \underline{\mathbf{x}}^T(k - K_f + 1) \quad \dots \quad \underline{\mathbf{x}}^T(k) \,]^T. \tag{4}$$

Similarly, if $\tilde{d}_i(k)$ denotes the output of the ith DFE, and $d_i(k) = f\{\tilde{d}_i(k)\}$ is the corresponding decision device output, then the input of the feedback filters, $\mathbf{b}_i(k)$ for $i = 1, \dots, M$, can be expressed by the $MK_b \times 1$ vector

$$\mathbf{d}(k) = [\, \underline{\mathbf{d}}^T(k - K_b) \quad \dots \quad \underline{\mathbf{d}}^T(k-1) \,]^T \tag{5}$$

where the $\underline{\mathbf{d}}(k)$ is defined as

$$\underline{\mathbf{d}}(k) = [\, d_1(k) \quad \dots \quad d_M(k) \,]^T.$$

By using the above definitions, the output of the ith DFE can be compactly expressed as

$$\begin{aligned} \tilde{d}_i(k) &= \mathbf{w}_i^H(k)\mathbf{y}(k), \\ \mathbf{w}_i(k) &= [\, \mathbf{f}_i^T(k) \quad \mathbf{b}_i^T(k) \,]^T, \\ \mathbf{y}(k) &= [\, \mathbf{x}^T(k) \quad \mathbf{d}^T(k) \,]^T, \; i = 1, \dots, M. \end{aligned} \tag{6}$$

Notice that all the MISO DFEs have a common input, $\mathbf{y}(k)$, of dimension $K = NK_f + MK_b$.

The MIMO DFE may be found by using a least squares (LS) criterion. Provided that all previous decisions are correct, each equalizer $\mathbf{w}_i(k)$ can be computed as the minimizer of the cost function

$$J(\mathbf{w}, \mathbf{\Phi}(k), \mathbf{z}_i(k)) = \frac{\mathbf{w}^H \mathbf{\Phi}(k) \mathbf{w}}{2} - \mathrm{Re}\left\{ \mathbf{w}^H \mathbf{z}_i(k) \right\} \tag{7}$$

with respect to $\mathbf{w}$.[1] Matrix $\mathbf{\Phi}(k)$ stands for the $K \times K$ exponentially time-averaged input data autocorrelation matrix, and $\mathbf{z}_i(k)$ for the crosscorrelation vector, which can be recursively computed as

$$\mathbf{\Phi}(k) = \lambda \mathbf{\Phi}(k-1) + \mathbf{y}(k)\mathbf{y}^H(k) \tag{8}$$

$$\mathbf{z}_i(k) = \lambda \mathbf{z}_i(k-1) + \mathbf{y}(k) d_i^*(k) \tag{9}$$

and λ is a forgetting factor $(0 < \lambda \leq 1)$.

It can be readily shown starting from (7) that the required minimizers can be derived as

$$\mathbf{\Phi}(k)\mathbf{W}(k) = \mathbf{Z}(k) \tag{10}$$

where $\mathbf{W}(k) = [\, \mathbf{w}_1(k) \quad \dots \quad \mathbf{w}_M(k) \,]$ and $\mathbf{Z}(k) = [\, \mathbf{z}_1(k) \quad \dots \quad \mathbf{z}_M(k) \,]$. Note that (10) is a linear system with multiple RHS.

1) Computing the FF and FB Filters Separately: Alternatively each one of the systems to be solved in (10) may be written as

$$\begin{bmatrix} \mathbf{\Phi}_{xx}(k) & \mathbf{\Phi}_{xd}(k) \\ \mathbf{\Phi}_{xd}^H(k) & \mathbf{\Phi}_{dd}(k) \end{bmatrix} \begin{bmatrix} \mathbf{f}_i(k) \\ \mathbf{b}_i(k) \end{bmatrix} = \begin{bmatrix} \mathbf{z}_{xd}^i(k) \\ \mathbf{z}_{dd}^i(k) \end{bmatrix} \tag{11}$$

where the submatrices $\mathbf{\Phi}_{xx}(k)$, $\mathbf{\Phi}_{xd}(k)$, $\mathbf{\Phi}_{dd}(k)$ and the vectors $\mathbf{z}_{xd}^i(k)$ and $\mathbf{z}_{dd}^i(k)$ may be computed recursively in a way similar to $\mathbf{\Phi}(k)$ and $\mathbf{z}_i(k)$ in (8) and (9). Thus, the FF and FB filters $\mathbf{f}_i(k)$ and $\mathbf{b}_i(k)$ may be computed separately by solving the following set of equations:

$$\begin{aligned} &\left(\mathbf{\Phi}_{xx}(k) - \mathbf{\Phi}_{xd}(k) \mathbf{\Phi}_{dd}^{-1}(k) \mathbf{\Phi}_{xd}^H(k) \right) \mathbf{f}_i(k) \\ &\quad = \mathbf{z}_{xd}^i(k) - \mathbf{\Phi}_{xd}(k) \mathbf{\Phi}_{dd}^{-1}(k) \mathbf{z}_{dd}^i(k) \end{aligned} \tag{12}$$

$$\mathbf{b}_i(k) = \mathbf{\Phi}_{dd}^{-1}(k) \left(\mathbf{z}_{dd}^i(k) - \mathbf{\Phi}_{xd}^H(k) \mathbf{f}_i(k) \right). \tag{13}$$

Note that matrix $\mathbf{\Phi}_{dd}^{-1}(k)$ could be updated by making use of the matrix inversion lemma [18] which implies that the FB filter could be updated in an RLS fashion. Then the FF filter could be updated by applying any one of the schemes proposed in Section III. This hybrid scheme (*i.e.*,, RLS for the FB and CG for the FF filter) is simply mentioned here as an alternative one and is not further treated in this correspondence.

B. The Conjugate Gradient Method

Before proceeding to the derivation of the new algorithms let us briefly review the conjugate gradient optimization method. This method is an iterative method for solving linear systems of the form $\mathbf{\Phi}\mathbf{w} = \mathbf{z}$, where $\mathbf{\Phi}$ is a $K \times K$ Hermitian positive definite matrix [15]. The CG method minimizes the quadratic function $J(\mathbf{w}, \mathbf{\Phi}, \mathbf{z})$ defined in (7) by iteratively updating the parameters' vector as

$$\mathbf{w}(m) = \mathbf{w}(m-1) + \alpha(m)\mathbf{p}(m). \tag{14}$$

The search direction vectors $\mathbf{p}(m)$ for $m \geq 1$ (where m denotes the iteration step) are designed to be $\mathbf{\Phi}$-orthogonal to each other, i.e.,

[1] Note the cost function in (7) is quadratic with respect to $\mathbf{w}$ for every k even under error propagation conditions.

$\mathbf{p}^H(m)\mathbf{\Phi}\mathbf{p}(l) = 0$ for $m \neq l$. Moreover, the step sizes $\alpha(m)$ are selected as the minimizing arguments of $J(\mathbf{w}(m), \mathbf{\Phi}, \mathbf{z})$ with respect to $\alpha(m)$.

To obtain the $\mathbf{\Phi}$-orthogonal direction vectors, the Gram–Schmidt conjugation process [15] is applied to a set of orthogonal vectors. The CG method selects the successive negative gradients (or residuals) of the cost function, i.e.,

$$\mathbf{g}(m) = -\frac{\partial J(\mathbf{w}(m), \mathbf{\Phi}, \mathbf{z})}{\partial \mathbf{w}^H(m)} = \mathbf{z} - \mathbf{\Phi}\mathbf{w}(m) \tag{15}$$

to be the basis vectors. By using the properties of the gradients, the direction vectors can be updated as

$$\mathbf{p}(m+1) = \mathbf{g}(m) + \beta(m)\mathbf{p}(m) \tag{16}$$

and $\beta(m)$ are chosen to ensure the $\mathbf{\Phi}$-orthogonality among the direction vectors. Finally, it can be shown that the CG method converges in at most K iterations [15].

III. Algorithm Derivation

In this section we derive three new MIMO MCG algorithms. These algorithms are based on CG optimization techniques for solving multiple linear systems of the form $\mathbf{\Phi}\mathbf{w}_i = \mathbf{z}_i$ for $i = 1, \ldots, M$. A straightforward but costly approach, for solving the aforementioned systems of equations, is to treat each system independently and apply to it the CG method presented in Section II-B. However, more sophisticated methods have been proposed in literature [19], [17] for linear systems of the aforementioned form. Those using projections [17] are the most computationally efficient and the most suitable for the problem at hand.

According to the projection methods, one of the linear systems is selected as the "seed" system. This system is solved by using the conventional CG method of Section II-B until it converges to its solution. During these iterations, the rest of the systems do not perform any CG iteration, but rather they use the search direction of the seed system to update their solution. After the seed system is solved, a new seed system is selected from the unsolved ones and the whole procedure is repeated.

In the subsection that follows, we initially apply to the MIMO case the idea of executing one CG iteration [12] per update of the correlation quantities. We treat each one of the M systems that are to be solved independently and we derive an algorithm with convergence properties identical to RLS but of higher computational cost. In order to reduce its complexity we modify the algorithm by properly applying the idea of projections.

A. Adaptive MIMO Modified Conjugate Gradient

By generalizing the adaptive MCG algorithm of [12] to the MIMO case, the solution of (10) can be time updated as

$$\mathbf{W}(k) = \mathbf{W}(k-1) + \mathbf{P}(k)\mathbf{A}(k) \tag{17}$$

where the columns of $\mathbf{P}(k)$ are the search directions for each of the M systems, and $\mathbf{A}(k)$ is a $M \times M$ diagonal matrix having as the ith diagonal element, $\alpha_i(k)$, the step size of the corresponding system. Each one of the systems in (10) is treated independently, and its solution is updated by executing only one CG iteration per update of the correlation quantities. Thus, a time instant corresponds to an iteration of the algorithm. The gradients of the systems at the kth time can be derived by (17), (8), and (9) as

$$\mathbf{G}(k) = \mathbf{Z}(k) - \mathbf{\Phi}(k)\mathbf{W}(k) = \mathbf{T}(k) - \mathbf{\Phi}(k)\mathbf{P}(k)\mathbf{A}(k) \tag{18}$$

where $\mathbf{T}(k)$ is defined as

$$\mathbf{T}(k) = \lambda\mathbf{G}(k-1) + \mathbf{y}(k)\mathbf{e}^H(k) \tag{19}$$

and

$$\mathbf{e}(k) = \mathbf{s}(k) - \mathbf{W}^H(k-1)\mathbf{y}(k) \tag{20}$$

is the *a priori* estimation error. The step sizes $\alpha_i(k)$ are selected as the minimizing arguments of $J(\mathbf{w}_i(k), \mathbf{\Phi}(k), \mathbf{z}_i(k))$, i.e.,

$$\alpha_i(k) = \frac{\mathbf{p}_i^H(k)\mathbf{t}_i(k)}{\mathbf{p}_i^H(k)\mathbf{\Phi}(k)\mathbf{p}_i(k)}, \quad i = 1, \ldots, M \tag{21}$$

and $\mathbf{t}_i(k)$ is the ith column of $\mathbf{T}(k)$. Then, the search directions for the next update are computed as

$$\mathbf{P}(k+1) = \mathbf{G}(k) + \mathbf{P}(k)\mathbf{B}(k) \tag{22}$$

where $\mathbf{B}(k)$ is again a $M \times M$ diagonal matrix. By employing the Polak–Ribiere method [12], the diagonal elements of $\mathbf{B}(k)$ can be computed as

$$\beta_i(k) = \frac{(\mathbf{g}_i(k) - \mathbf{g}_i(k-1))^H \mathbf{g}_i(k)}{\mathbf{g}_i^H(k-1)\mathbf{g}_i(k-1)}, \quad i = 1, \ldots, M, \quad k \geq 2 \tag{23}$$

where $\mathbf{g}_i(k)$ is the ith column of the gradient matrix $\mathbf{G}(k)$. The proposed algorithm, called hereafter MIMO-MCG, is summarized in Table I. Notice that it can be viewed as the application of the SISO-MCG algorithm of [12] to each of the linear systems of (10) independently. Following standard practice in DFE design, a decision delay should be inserted between equalizer decisions and transmitted symbols. As in [5], we consider a decision delay parameter Δ common for all streams, and set it to $\Delta = K_f - 1$. Hence, the decision $d_i(k)$ corresponds to symbol $s_i(k - \Delta)$.

By comparing MIMO-MCG to the MIMO-RLS algorithm, three remarks should be made: a) Following the rationale of [12], it can be shown that the MIMO-MCG converges to the solution of the systems in (10). Moreover, as it will be shown by simulations, its convergence performance in terms of mean square error (MSE) is identical to that of the RLS algorithm, which is optimum in the least squares sense. b) In contrast to the RLS algorithm, the matrix $\mathbf{\Phi}^{-1}$ is not needed. Note that in rapid time-varying conditions where smaller values of λ are used, and/or in high-dimensional MIMO DFE systems, the size of the autocorrelation matrix may become comparable to the effective memory of the system, $1/(1-\lambda)$. As verified by simulations in Section IV, the involvement of an ill-conditioned $\mathbf{\Phi}^{-1}$ leads MIMO-RLS to numerical instability and reduces its tracking capabilities, whereas MIMO-MCG remains robust under such conditions. c) As shown in Table III, MIMO-MCG is more computationally demanding than MIMO-RLS. To reduce its complexity, we incorporate Galerkin projections into the MCG algorithm.

TABLE I
SUMMARY OF MIMO-MCG

Initialization: $\mathbf{G}(0) = \mathbf{W}(0) = \mathbf{0}_{K\times M}$, $\mathbf{B}(0) = \mathbf{0}_M$, $\mathbf{P}(1) = \mathbf{y}(1)\mathbf{s}^H(1)$, and $\mathbf{\Phi}(0) = \delta\mathbf{I}_K$, where δ is a small positive constant (δ=0.001).

1) Update matrix $\mathbf{\Phi}(k)$ (8).

2) Update matrix $\mathbf{T}(k)$ (19).

3) Compute the step sizes by using (21).

4) Update the equalizer filters from (17).

5) Update the gradients (18).

6) Compute the $\beta_i(k)$ for $i = 1, \ldots, M$, from (23).

7) Update the search directions as in (22).

B. Galerkin Projection-Based MIMO-MCG

Recall that according to MIMO-MCG, all linear systems, $\mathbf{\Phi}(k)\mathbf{w}_i(k) = \mathbf{z}_i(k)$, $i = 1, \ldots, M$, are constantly updated by the MCG algorithm. By using the idea of Galerkin projections, an approximate solution can be obtained by updating via MCG just a single system j (seed system) at each time instant k, while the others are updated through Galerkin projections, i.e., they use the search direction of the seed system, $\mathbf{p}_j(k)$, to update their solution as

$$\mathbf{w}_i(k) = \mathbf{w}_i(k-1) + \alpha_i(k)\mathbf{p}_j(k) \tag{24}$$

where the step size $\alpha_i(k)$ is again selected as the minimizing argument of $J(\mathbf{w}_i(k), \mathbf{\Phi}(k), \mathbf{z}_i(k))$

$$\alpha_i(k) = \frac{\mathbf{p}_j^H(k)\mathbf{t}_i(k)}{\mathbf{p}_j^H(k)\mathbf{\Phi}(k)\mathbf{p}_j(k)} \tag{25}$$

and $\mathbf{p}_j(k)$ is the search direction of the seed system. Thus, all systems take different steps, $a_i(k)$ for $i = 1, \ldots, M$, in the same direction $\mathbf{p}_j(k)$. Note that, during the MCG update of the seed system j at the kth time instant, a new direction that is $\mathbf{\Phi}$-orthogonal to $\mathbf{p}_j(k)$ is followed. This one will be used for updating the solution of the jth system the next time that this system is selected again as the seed system. Note that MCG updating leads to more accurate approximations than Galerkin projections. Extensive simulations have confirmed that substituting Galerkin projections for MCG updates results in a slight degradation in performance.

Considering the selection of the seed system, we have studied the following two alternatives.

1) Circular Scheme (MIMO-MCG-PI): A straightforward choice would be to select the seed system at each time instant k, in a round robin fashion. Thus, the seed system that is MCG updated at the kth time instant is selected according to $j = \mathrm{mod}(k, M) + 1$. Furthermore, the selection can be considered random given that each system is selected in a circular way without taking into account issues related to the current solution.

2) Maximum Update Scheme (MIMO-MCG-PII): A more sophisticated scheduling technique takes into account the "need" for an MCG update. One way of measuring this "need" is to measure for each system the absolute difference between the filters at two successive previous time instances and select the seed system as the one that diverged more. The seed system can therefore be selected as

$$\begin{aligned} j &= \arg\max_i \|\mathbf{w}_i(k-1) - \mathbf{w}_i(k-2)\|^2 \\ &= \arg\max_i |\alpha_i(k-1)|^2 . \end{aligned} \tag{26}$$

It has been verified by simulations that this policy leads to an improved performance as compared to the round robin one, without paying any increase in complexity. Finally, it should be mentioned that its performance is very close to the optimum performance obtained in the least square sense (RLS algorithm).

The two schemes described above are presented in Table II, where the initialization can be performed as in Table I. Since only a single system is MCG-updated at each time instant, while the non-seed systems are updated according to efficient Galerkin projections, both schemes have a computational complexity of $O(K^2 + MK)$. The complexities for all the proposed algorithms, as well as the MIMO-RLS of [9] implemented in a square-root fashion to avoid numerical instability [10], are presented in Table III.

At this point, it should be mentioned that projections can be also applied in time dimension as well. To be more specific, the direction vector $\mathbf{p}_j(k)$, that is used for updating each one of the weight coefficient vectors, may be also used for updating the systems solutions $\mathbf{w}_i(k)$ for some next time steps as well, as described in the time projection scheme presented in [1]. The schemes presented above, along with the time projection scheme presented in [1], provide a flexible framework in MIMO adaptive equalization design to trade efficiency for performance.

IV. SIMULATION RESULTS

The performance of the proposed equalizers was evaluated through computer simulations. We considered a system transmitting uncoded QPSK symbols of duration T_s=0.25 μs over a wireless channel modeled according to the UMTS Vehicular Channel Model A [20]. Following standard DFE design practice [5], the feedforward and feedback filters had a temporal span of K_f= 20 and K_b= 10 taps.

Initially, to study the convergence of the equalizers, the channel was kept static for an interval of $2000T_s$. An M=N= 3 antenna configuration operating at $\mathrm{SNR} = 16$ dB was simulated, while the system was constantly in training mode. Six different MIMO DFE algorithms were tested: a) a MIMO-RLS of [9] implemented in a square root RLS

TABLE II
SUMMARY OF PROJECTION-BASED SCHEMES (MIMO-MCG-PI & II)

A. Select the Seed System

PI: $j = \mod(k, M) + 1$

PII: $j = \arg\max_i \|a_i(k-1)\|^2$

B. MCG Update the Seed System

1) Update matrix $\mathbf{\Phi}(k)$ by using (8).
2) $\mathbf{t}_j(k) = \lambda \mathbf{g}_j(k-1) + \mathbf{y}(k)e_j^*(k)$.
3) Compute the step size $\alpha_j(k)$ from (21).
4) $\mathbf{w}_j(k) = \mathbf{w}_j(k-1) + \alpha_j(k)\mathbf{p}_j(k)$.
5) $\mathbf{g}_j(k) = \mathbf{t}_j(k) - \alpha_j(k)\mathbf{\Phi}(k)\mathbf{p}_j(k)$.
6) Compute the $\beta_j(k)$ from (23).
7) Update the direction vector as

PI: $\mathbf{p}_j(k+M) = \mathbf{g}_j(k) + \beta_j(k)\mathbf{p}_j(k)$.

PII: $\mathbf{p}_j(l) = \mathbf{g}_j(k) + \beta_j(k)\mathbf{p}_j(k)$,

where $l > k$ is the next time instant this system will be selected again as the seed system.

C. Project the Non-Seed Systems $(i = 1, \ldots, M,\ i \neq j)$

1) $\mathbf{t}_i(k) = \lambda \mathbf{g}_i(k-1) + \mathbf{y}(k)e_i^*(k)$.
2) Compute $\alpha_i(k)$ from (25).
3) Update the equalizer filter according to (24).
4) $\mathbf{g}_i(k) = \mathbf{t}_i(k) - \alpha_i(k)\mathbf{\Phi}(k)\mathbf{p}_j(k)$.

TABLE III
COMPARISON OF COMPLEXITIES

Algorithm	Complex Multiplication-Add Operations
MIMO-RLS	$\frac{9}{2}K^2+9MK+O(K)$
MIMO-MCG	$\left(2M+\frac{3}{2}\right)K^2+17MK+O(K)$
MIMO-MCG-P-(I or II)	$\frac{7}{2}K^2+11MK+O(K)$

fashion to avoid numerical instability [10], b) a MIMO normalized LMS (MIMO-NLMS) algorithm, c) the MIMO-MCG of Table I, d) the MIMO-MCG-PI of Table II, e) the MIMO-MCG-PII of Table II, and f) a scheme that executes only one MCG update every M iterations and does not perform any projection (MIMO-MCG-NOP). In Fig. 1(a), the MSE, averaged over all streams, is plotted. By setting the parameter λ of the MCG algorithms equal to the forgetting factor of the RLS and adjusting accordingly the step size of the NLMS, all the equalizers were tuned to converge at the same steady state error. As it is shown, the MIMO-MCG curve coincides with that of MIMO-RLS, while MIMO-MCG-PI and MIMO-MCG-PII lie very close to it. The PII scheme is better than the PI as it was expected, while the MIMO-MCG-NOP and MIMO-NLMS exhibit very slow convergence. Note that the projections can offer significant improvement in performance at a limited additional cost.

The convergence of the projection schemes can be intuitively attributed to the following: i) given that the sequence of seed search directions among the systems is uncorrelated and by ensuring that each search direction of a seed system is perpendicular to the next one of this system, the subspace where all filters lie is adequately spanned and ii) the line search performed during the Galerkin projections in (25) is crucial.

Error propagation effects in decision-directed mode and the impact of MCG updates frequency were studied by simulating a system that operates over a 6×6 static channel. A training period of $512T_s$ was employed. As it can be seen in Fig. 1(b), all equalizers are robust to error propagation effects. Moreover, by performing the MCG update less frequently [every $M = 6$ iterations instead of 3 in Fig. 1(a)], a slight degradation in performance can be observed.

The tracking performance of the algorithms was studied by simulating a system that operates over a 3×3 slow fading channel for an interval of 8000 symbols. A normalized Doppler frequency $f_D T_s = 5.5\ 10^{-5}$ was simulated by using the Jakes method. It should be noted that within this interval and for the specific Doppler spread

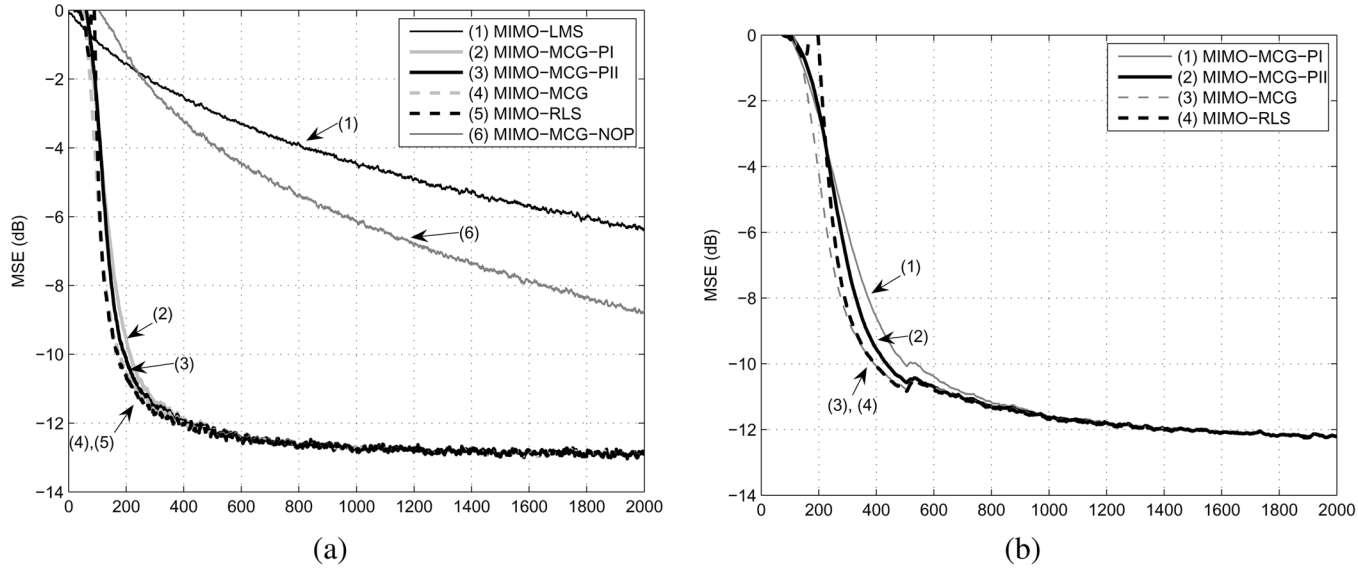

Fig. 1. Convergence of adaptive equalizers at SNR $=$ 16 dB. (a) Always trained system over a 3×3 static MIMO channel ($\lambda = 0.9989$). (b) Training and decision directed mode over a 6×6 static MIMO channel ($\lambda = 0.9994$).

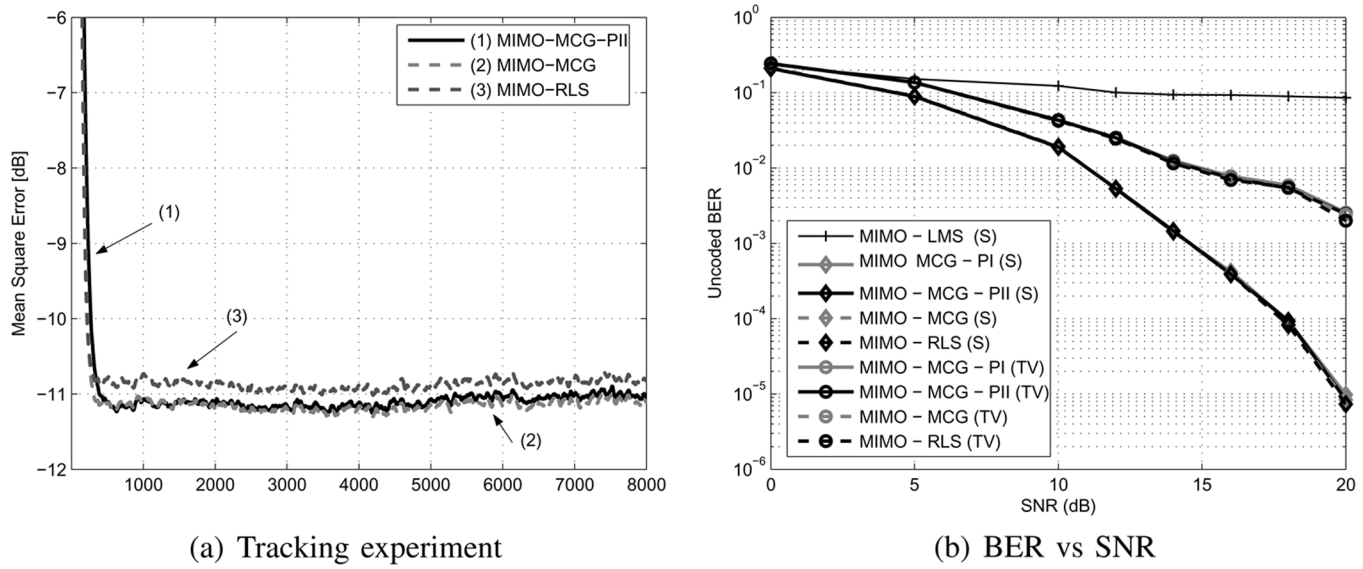

Fig. 2. (a) Convergence of adaptive equalizers for a 3×3 time-varying MIMO system at SNR $=$ 16 dB ($\lambda = 0.98$) and (b) uncoded BER curves for a 2×2 static (S) and time-varying (TV) systems.

the amplitude of the channel taps may change significantly (i.e., 75% amplitude variation and $7\pi/8$ phase rotation) which implies that the use of adaptive algorithms is imperative. As illustrated in Fig. 2(a), MIMO-MCG and MIMO-MCG-PII successfully track channel variations, with the latter exhibiting a slightly inferior performance. On the other hand, MIMO-RLS, even in its square root implementation and even though it was tuned to yield the same tracking performance, converges at a higher MSE due to the involvement of matrix $\mathbf{\Phi}^{-1}$ which for small λ (short effective memory) becomes ill-conditioned.

The BER performance of the proposed equalizers operating over static and time-varying frequency selective channels has also been studied. We consider a $M = N = 2$ system, trained for $512T_s$. For the time-varying case a normalized Doppler frequency $f_D T_s = 5.5\ 10^{-5}$ was simulated. The curves in Fig. 2(b) indicate that the proposed algorithms can operate efficiently under the simulated severe channel selectivity conditions. As expected, the tracking curve is worse than that in the static case.

V. Conclusion

Three adaptive algorithms for updating a MIMO DFE have been developed. By extending the algorithm of [12], we derived an adaptive CG MIMO DFE. To reduce its complexity, we employed the idea of Galerkin projections, and two schemes of convergence close to RLS have been proposed. As shown by simulations, all the new equalizers can be successfully employed in practical MIMO wideband systems. Moreover, the proposed algorithms can be easily modified so as to be applicable to other signal processing problems as well which can be formulated as linear systems of equations with multiple right hand sides evolving in time.

REFERENCES

[1] V. Kekatos, A. Lalos, and K. Berberidis, "Adaptive conjugate gradient DFEs for wideband MIMO systems using Galerkin projections," presented at the IEEE Vehicular Technology Conf. (VTC), Baltimore, MD, Oct. 2007.

[2] G. Stuber, J. Barry, S. Mclaughlin, Y. Li, M. Ingram, and T. Pratt, "Broadband MIMO-OFDM wireless communications," *Proc. IEEE*, vol. 92, pp. 271–294, Feb. 2004.

[3] X. Zhu and R. Murch, "Layered space-frequency equalization in a single-carrier MIMO system for frequency-selective channels," *IEEE Trans. Wireless Commun.*, vol. 3, no. 3, pp. 701–708, May 2004.

[4] *Amendment to the 802.16 Standard Air Interface for Fixed Broadband Wireless Access Systems*, 802.16a/b, Jun. 2001.

[5] N. Al-Dhahir and A. H. Sayed, "The finite length multi-input multi-output MMSE-DFE," *IEEE Trans. Signal Process.*, vol. 48, no. 10, pp. 2921–2936, Oct. 2000.

[6] R. Dinis, R. Kalbasi, D. Falconer, and A. H. Banihashemi, "Iterative layered space-time receivers for single-carrier transmission over severe time-dispersive channels," *IEEE Commun. Lett.*, vol. 8, pp. 579–581, Sep. 2004.

[7] S. Chen, L. Hanzo, and A. Livingstone, "MBER space-time decision feedback equalization assisted multiuser detection for multiple antenna aided SDMA systems," *IEEE Trans. Signal Process.*, vol. 54, no. 8, pp. 3090–3098, Aug. 2006.

[8] R. d. Lamare and R. Sampaio-Neto, "Adaptive MBER decision feedback multiuser receivers in frequency selective fading channels," *IEEE Commun. Lett.*, vol. 7, no. 2, pp. 73–75, Feb. 2003.

[9] A. Maleki-Tehrani, B. Hassibi, and J. M. Cioffi, "Adaptive equalization of multiple-input multiple-output (MIMO) channels," in *Proc. IEEE Int. Conf. Communications*, New Orleans, LA, Jun. 2000, pp. 1670–1674.

[10] V. Kekatos, A. A. Rontogiannis, and K. Berberidis, "Cholesky factorization—Based adaptive blast DFE for wideband MIMO channels," *EURASIP J. Adv. Signal Process.*, vol. 2007, no. 45789, pp. 1–11, 2007.

[11] V. Kekatos, K. Berberidis, and A. Rontogiannis, "A block adaptive frequency domain MIMO DFE for wideband channels," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Honolulu, HI, Apr. 2007, pp. 197–200.

[12] P. S. Chang and A. N. Willson, "Analysis of conjugate gradient algorithms for adaptive filtering," *IEEE Trans. Signal Process.*, vol. 48, no. 2, pp. 409–418, Feb. 2000.

[13] J. S. Lim and C. K. Un, "Block conjugate gradient algorithms for adaptive filtering," *Signal Process.*, vol. 55, no. 1, pp. 65–77, 1996.

[14] A. S. Lalos and K. Berberidis, "An efficient conjugate gradient method in the frequency domain: Application to channel equalization," *Signal Process.*, vol. 88, no. 1, pp. 99–116, 2008.

[15] R. Fletcher, *Practical Methods of Optimization*, 2nd ed. Chichester, U.K.: Wiley, 2000.

[16] M. Hestenes, *Conjugate Direction Methods in Optimization*. New York: Springer, 1980.

[17] T. F. Chan and M. K. Ng, "Galerkin projection methods for solving multiple linear systems," *SIAM J. Sci. Comput.*, vol. 21, no. 3, pp. 836–850, 2000.

[18] A. H. Sayed, *Fundamentals of Adaptive Filtering*. New York: Wiley, 2003.

[19] I. Groh, G. Dietl, and W. Utschick, "Preconditioned and rank-flexible block conjugate gradient implementations of MIMO Wiener decision feedback equalizers," presented at the IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP), Toulouse, France, May 2006.

[20] Universal Mobile Telecommunications System (UMTS), "Selection procedures for the choice of radio transmission technologies of the UMTS," ETSI, Sophia-Antipolis, France, Tech. Rep. 101.112, Apr. 1998.

On Fourier Interpolation Error for Band-Limited Signals

Zhengguang Xu, Benxiong Huang, and Kewei Li

***Abstract*—Fourier interpolation is a useful tool to reconstruct continuous signals from discrete samples. This correspondence derives Fourier interpolation expression from Fourier series for periodic signals and investigates its error expression and upper bound for general band-limited signals. Two numerical examples are given to interpret the meaning of the formulas.**

***Index Terms*—Error bound, error expression, Fourier interpolation.**

I. INTRODUCTION

For band-limited signals, the Shannon sampling theorem makes it possible to denote continuous signals by discrete samples, therefore the discrete signal has found wide applications in signal processing fields [1]. However, some special applications, for example, Empirical Mode Decomposition [2], [3], requires the exact location of extrema in the signal. With the limited sampling rate, the extrema are difficult to locate, thus we should reconstruct the continuous signal from the discrete samples. The mostly used technology is Shannon interpolation as follows:

$$s(t) = \sum_{n=-\infty}^{\infty} s(n\Delta t)\, \frac{\sin \frac{\pi}{\Delta t}(t - n\Delta t)}{\frac{\pi}{\Delta t}(t - n\Delta t)}$$

where Δt is sampling interval and $s(n\Delta t)$ are samples. Although Shannon interpolation is capable of recovering the original signal for a band-limited signal, it still has two drawbacks: 1) an infinite number of samples is required in Shannon interpolation so that a finite sequence will cause truncation error [4], [5] and 2) the extrema in the reconstructed signal are obtained by searching all the interpolation sequence, which is complex in computation and limited in location precision. A new technology based on Fourier interpolation for finding the extrema is proposed in [6]. Fourier interpolation expression is

$$s(t) = \sum_{k=-\infty}^{+\infty} c_k e^{jk2\pi/Tt}$$

where c_k are Fourier coefficients and T is duration of the signal. Compared with Shannon interpolation, Fourier interpolation has two advantages: 1) Fourier interpolation can recover the band-limited periodic signal perfectly by finite samples and 2) some formulas can be used to locate the extrema exactly [6]. However, Fourier interpolation also has a drawback that Fourier interpolation will cause errors if the signal is nonperiodic. Some literature studied the property of Fourier interpolation [7]–[9], but few of them focused on the error of Fourier interpolation for nonperiodic signals. This correspondence studies the problem, and investigates the error expression and upper bound of Fourier interpolation for general band-limited signals.

Manuscript received August 05, 2008; accepted January 20, 2009. First published February 24, 2009; current version published May 15, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Thierry Blu. This work was supported by a grant from the National Natural Science Foundation of China (No. 60832002).

The authors are with the Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: xray.hust@gmail.com).

Color versions of one or more of the figures in this correspondence are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2009.2016263