ON PRECONDITIONED CONJUGATE GRADIENT METHOD FOR TIME-VARYING OFDM CHANNEL EQUALIZATION

Zijian Tang, Rob Remis

Faculty EEMCS
Delft University of Technology
Mekelweg 4, 2628 CD Delft, The Netherlands
{z.tang, r.f.remis}@tudelft.nl

Magnus Lundberg Nordenvaad

Department of Underwater Research
Swedish Defence Research Agency
SE- 164 90 Stockholm, Sweden
maglun@foi.se

ABSTRACT

We consider using the conjugate gradient (CG) algorithm to equalize a time-varying channel in an orthogonal frequency division multiplexing (OFDM) system. Preconditioning technique to accelerate the convergence of the CG algorithm is discussed, where we show that when the Doppler spread becomes higher, the commonly used diagonal preconditioner, despite its simpleness, can perform even worse than without preconditioner. In such a case, a preconditioner with a more complex structure is proposed.

***Index Terms*—** OFDM, time-varying channel equalization, conjugate gradient, preconditioning

1. INTRODUCTION

High-speed data communications over dispersive channels demands a low-complexity channel equalization method. One of the most successful examples is orthogonal frequency division multiplexing (OFDM), which gains most of its popularity thanks to its ability to transmit signals on separate subcarriers without mutual interference, a fact that mathematically can be represented in the frequency domain by a diagonal channel matrix. This property holds only when the channel stays (almost) constant for at least one OFDM symbol interval. In practice, a time-invariant channel assumption can become invalid due to, e.g., Doppler effects resulting from the motion between the transmitter and receiver. In such a case, the frequency-domain channel matrix is not diagonal but in principle full with the non-zero off-diagonal elements as known as inter-carrier interference (ICI).

Equalizing such channels requires inverting a full matrix, which is very costly. A commonly adopted approach is to approximate the channel matrix as banded [1–3]. Indeed, for narrowband channels subject to realistic Doppler spreads, the ICI is for the most part contributed by nearby subcarriers, implying that most channel energy is still concentrated on the entries close to the main diagonal. With such a banded assumption, the complexity of equalization can still be linear in the number of subcarriers as in time-invariant channel cases. Recently, [4] has proposed using the generalized minimal residual (GMRES) and LSQR algorithms for channel equalization. Just like the conjugate gradient (CG) algorithm [5] to be discussed in this paper, these algorithms invert the channel matrix iteratively. Note that the approximations generated by the CG algorithm belongs to the Krylov subspace. As a result, for a system subject to additive noise, a truncated CG can yield different results than direct matrix inversion. It is well-known that the performance of direct matrix inversion is limited by the condition number of the channel matrix, which usually becomes larger with the increase of the Doppler spread. In comparison, the performance of the CG algorithm turns out to be less susceptible to this problem as we will show in the simulations.

On the other hand, the convergence rate of CG is dependent on the condition number of the channel matrix, which can be accelerated by the use of preconditioning [5]. A good preconditioner is able to enforce the spectrum of the preconditioned channel matrix to cluster around 1. At the same time, the design and implementation of the preconditioner should be simple enough in order to maintain the overall complexity low. [4] designed a circulant preconditioner in the time domain, which is, however, suboptimal, and becomes less effective as Doppler increases. Noting the equivalence between the circulant and diagonal matrices via the discrete Fourier transform (DFT), we design in this paper a diagonal preconditioner directly in the frequency domain following the approach in [6]. However, such a diagonal preconditioner does not necessarily lead to an improved spectral property as claimed by [6], especially in the case of a faster-varying channel, where the energy of the channel matrix will become increasingly dispersed to entries farther away from the main diagonal. At the point when the energy on the off-diagonals is higher than the main diagonal, it turns out that a diagonal preconditioner will cluster the eigenvalues around 0 instead of 1. In such a case, it will be necessary to impose the preconditioner with a more complex structure.

Notation: We use upper (lower) bold face letters to denote matrices (column vectors). $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ represent conjugate, transpose and complex conjugate transpose (Hermitian), respectively. $[\mathbf{x}]_p$ indicates the pth element of the vector $\mathbf{x}$ and $[\mathbf{X}]_{p,q}$ indicates the (p,q)th entry of the matrix $\mathbf{X}$. diag$\{\mathbf{x}\}$ is used to denote a diagonal matrix with $\mathbf{x}$ on the diagonal; $\odot$ represents the Hadamard product. $\mathbf{I}_N$ stands for the $N \times N$ identity matrix and $\mathbf{e}_k$ is reserved for the kth column of $\mathbf{I}_N$; $\mathbf{1}_{M\times N}$ stands for an $M \times N$ all-one matrix, and $\mathbf{W}_K$ for a K-point normalized DFT matrix. Finally, we use $\|\mathbf{X}\|$ ($\|\mathbf{x}\|$) to denote the Frobenius norm of the matrix $\mathbf{X}$ (vector $\mathbf{x}$).

2. SYSTEM MODEL

Let us consider an OFDM system where the discrete-time channel is assumed to be a time-varying causal finite impulse response (FIR) filter with a maximum order L. Using $h_{p,l}$ to denote the time-domain channel gain of the lth tap at the pth time instant, we assume that $h_{p,l} = 0$ if $l < 0$ or $l > L$. Note that this FIR channel model can take transmit/receiver filter, propagation environment, and possible synchronization errors into account.

We use $\mathbf{s}[m]$ to represent the mth OFDM symbol, which contains K data symbols. They are first modulated onto K subcarriers by means of the inverse DFT (IDFT) matrix $\mathbf{W}_K^H$, then concatenated by a cyclic prefix (CP) of length $L_{\text{cp}} \geq L$, and finally sent over the channel. At the receiver, the received samples corresponding to the CP are discarded, and the remaining samples are demodulated by means of the DFT matrix $\mathbf{W}_K$. Mathematically, we can express the received samples during the mth OFDM symbol $\mathbf{y}[m]$ as

$$\mathbf{y}[m] = \mathbf{H}_{\text{F}}[m]\mathbf{s}[m] + \mathbf{z}[m], \tag{1}$$

where $\mathbf{z}[m]$ represents additive noise, and $\mathbf{H}_{\text{F}}[m]$ denotes the channel matrix corresponding to the mth OFDM symbol in the frequency domain, which is related to its time-domain counterpart $\mathbf{H}_{\text{T}}[m]$ as $\mathbf{H}_{\text{F}}[m] = \mathbf{W}_K\mathbf{H}_{\text{T}}[m]\mathbf{W}_K^H$. For the entries of $\mathbf{H}_{\text{T}}[m]$, we have, under the FIR assumption of the channel and letting $L_{\text{cp}} = L$ without loss of generality, $\left[\mathbf{H}_{\text{T}}[m]\right]_{p,q} = h_{m(K+L)+p+L,\text{mod}(p-q,K)}$ with $\text{mod}(a,b)$ standing for the remainder of a divided by b.

Many works approximate time-varying channel taps with a basis expansion model (BEM) (see e.g., [4, 7]). Although this approach is usually exploited for channel estimation, it will still be useful to introduce the BEM here for the reason that will become evident soon. One way of utilizing the BEM is to construct a $K \times (Q+1)$ BEM matrix $\mathbf{B} := [\mathbf{b}_0, \cdots \mathbf{b}_Q]$ with $\mathbf{b}_q$ being its qth column, which is also known as the qth expansion basis. With a sufficiently large Q, the time-variation present in the lth channel tap $\mathbf{h}_l[m] := [h_{m(K+L)}, \cdots, h_{(m+1)(K+L)-1}]^T$ can be approximated by the BEM as:

$$\mathbf{h}_l \approx \mathbf{B}[c_{l,0}, \cdots, c_{l,Q}]^T, \tag{2}$$

with $c_{l,q}[m]$ denoting the qth BEM coefficient for the lth tap corresponding to the mth OFDM symbol.

Because the channel equalization discussed in this paper will be realized for each OFDM symbol individually, we will in the sequel omit the symbol index $[m]$ for the sake of notational ease.

3. LOW-COMPLEXITY EQUALIZATION

For time-varying channels, $\mathbf{H}_{\text{F}}$ becomes full in general, and a direct inversion enlists a cost of $\mathcal{O}(K^3)$. We note that conform Jakes' model [8], the time-variation of the channel causes a channel energy dispersion from the diagonal onwards into the off-diagonals of $\mathbf{H}_{\text{F}}$, whose amount is determined by the Doppler spread. With a realistic Doppler, it is thus reasonable to assume that most of the channel energy is concentrated on the entries of $\mathbf{H}_{\text{F}}$ that are close to and on the main diagonal, and decays fast on the entries farther away. For this reason, many attempts focus on a banded approximation of $\mathbf{H}_{\text{F}}$ to reduce the equalizer complexity [1–3]. To be specific, we write (1) as $\mathbf{y} = \tilde{\mathbf{H}}_{\text{F}}\mathbf{s} + \boldsymbol{\epsilon} + \mathbf{z}$, with

$$\tilde{\mathbf{H}}_{\text{F}} := \mathbf{T}_{Q'} \odot \mathbf{H}_{\text{F}} \tag{3}$$

where $\mathbf{T}_{Q'}$ denotes a $K \times K$ matrix with 1's on its main diagonal, $Q'/2$ sub-diagonals and $Q'/2$ super-diagonals, and 0's on its remaining entries. The resulting banding error is captured in $\boldsymbol{\epsilon}$. With a proper Q' which is usually chosen to be no smaller than Q, it is reasonable to neglect $\boldsymbol{\epsilon}$ in the subsequent equalization.

Different than the direct matrix inversion approach adopted in [2,3], we will use the conjugate gradient (CG) algorithm to invert $\mathbf{H}_{\text{F}}$ iteratively [5]. The solution obtained after the ith iteration, denoted as $\hat{\mathbf{s}}^{\{i\}}$, is constrained in the Krylov subspace which is defined as $\text{span}\{\mathbf{y}, \mathbf{H}\mathbf{y}, \cdots, \mathbf{H}^{i-1}\mathbf{y}\}$. This implies that the solution obtained through iterative algorithms could be different from the solution obtained through direct matrix inversion if $\mathbf{y}$ is subject to an additive noise (as well as a banding error). The basic implementation of CG is described in Table 1.

Step 1. Set $i = 0$, and initialize

$$\mathbf{g}^{\{0\}} = \mathbf{y},$$
$$\mathbf{a}^{\{0\}} = \mathbf{y},$$
$$u^{\{0\}} = \frac{\|\mathbf{g}^{\{0\}}\|^2}{\mathbf{a}^{\{0\}T}\mathbf{H}_{\text{F}}\mathbf{a}^{\{0\}}}, \text{ and}$$
$$\hat{\mathbf{s}}^{\{0\}} = u^{\{0\}}\mathbf{a}^{\{0\}}.$$

Step 2. Increment i, and update the following

$$\mathbf{g}^{\{i\}} = \mathbf{y} - \mathbf{H}_{\text{F}}\hat{\mathbf{b}}^{\{i-1\}}, \tag{4}$$

$$\mathbf{a}^{\{i\}} = \frac{\|\mathbf{g}^{\{i\}}\|^2}{\|\mathbf{g}^{\{i-1\}}\|^2}\mathbf{a}^{\{i-1\|} + \mathbf{g}^{\{i\}},$$

$$u^{\{i\}} = \frac{\|\mathbf{g}^{\{i\}}\|^2}{\mathbf{a}^{\{i\}H}\mathbf{H}_{\text{F}}\mathbf{a}^{\{i\}}}, \tag{5}$$

and

$$\hat{\mathbf{b}}^{\{i\}} = \hat{\mathbf{b}}^{\{i-1\}} + u^{\{i\}}\mathbf{a}^{\{i\}}.$$

Step 3 If the stop criterion is met, then halt the iteration. Otherwise, return to **Step 2**.

Table 1. The CG algorithm.

The stop criterion can be whether or not the discrepancy between $\hat{\mathbf{s}}^{\{i\}}$ and $\hat{\mathbf{s}}^{\{i-1\}}$ drops below a given level. An alternative is to halt after a predetermined number of iterations, a strategy that will be used in the simulations.

It is clear from Table 1 that the computational complexity is determined by (4) and (5). For instance, the operation $\mathbf{a}^{\{i\}H}\mathbf{H}_{\text{F}}\mathbf{a}^{\{i\}}$ costs $\mathcal{O}(K^2)$. This can be alleviated in two ways: 1) replace $\mathbf{H}_{\text{F}}$ by $\tilde{\mathbf{H}}_{\text{F}}$, which is introduced in (3) as a banded approximation to $\mathbf{H}_{\text{F}}$ having a bandwidth of Q'. As a result, the complexity is reduced to $\mathcal{O}(KQ')$ at the cost of inducing an approximation error. 2) It is also possible to resort to the time domain and explore the equality $\mathbf{a}^{\{i\}T}\mathbf{H}_{\text{F}}\mathbf{a}^{\{i\}} = \mathbf{a}^{\{i\}H}\mathbf{W}_K\mathbf{H}_{\text{T}}\mathbf{W}_K^H\mathbf{a}^{\{i\}}$. Seeing that $\mathbf{W}_K^H\mathbf{a}^{\{i\}}$ can be achieved by means of the fast Fourier transform (FFT) at a cost of $\mathcal{O}(K\log K)$, and $\mathbf{H}_{\text{T}}$ is a strictly banded matrix with a bandwidth of L, this approach calls for a complexity of $\mathcal{O}(K\max(L, \log K))$. Obviously, this approach does not suffer from an banded approximation error, but can be more costly if $\max(L, \log K) > Q'$.

Simulation results show that with a proper preconditioning, the CG algorithm converges already after a few iterations even for a very large Doppler spread. As a result, we are still able to implement the iterative channel equalization with a complexity linear in the number of subcarrier K.

4. PRECONDITIONING

For the CG algorithm to converge faster, preconditioning is an indispensable procedure. With a preconditioning matrix $\mathbf{P}$, which is non-singular by assumption, the I/O relationship given in (1) can be rewritten as

$$\mathbf{y} = \mathbf{H}_{\text{F}}\mathbf{P}\mathbf{P}^{-1}\mathbf{s} + \mathbf{z}. \tag{6}$$

The estimate of $\mathbf{s}$ is attained in two steps: 1) $\mathbf{x} = (\mathbf{H}_{\text{F}}\mathbf{P})^{-1}\mathbf{y}$ is computed by the CG algorithm, and afterwards 2) $\hat{\mathbf{s}} = \mathbf{P}\mathbf{x}$. The preconditioner $\mathbf{P}$ is designed to optimize the spectral condition of the matrix product $\mathbf{H}_{\text{F}}\mathbf{P}$. At the same time, its design and implementation should not inflict too much extra effort (otherwise, we can just suffice with a straightforward preconditioner $\mathbf{P} = \mathbf{H}_{\text{F}}^{-1}$). For this

reason, the preconditioner is often designed with some special structure.

One approach is to approximate the time-domain channel matrix $\mathbf{H}_{\mathrm{T}}$ with a circulant matrix $\mathbf{\Lambda}$, which was first proposed in [9]

$$\mathbf{\Lambda} = \arg\min_{\hat{\mathbf{\Lambda}}} \|\mathbf{H}_{\mathrm{T}} - \hat{\mathbf{\Lambda}}\|^2, \tag{7}$$

leading to a preconditioner $\mathbf{P} = \mathbf{\Lambda}^{-1}$. Note that [9] aims to find a circulant matrix approximation to a Toeplitz matrix. In our case, although $\mathbf{H}_{\mathrm{T}}$ is not Toeplitz, this method could still be valid for low to moderate Doppler. For the design of $\mathbf{\Lambda}$, we resort to the BEM channel approximation introduced in (2), with which we have $\mathbf{H}_{\mathrm{T}} \approx \sum_{q=0}^{Q} \mathrm{diag}\{\mathbf{b}_q\}\mathbf{C}_q$, where $\mathbf{C}_q$ stands for a $K \times K$ circulant matrix with $[c_{0,q}, \cdots, c_{L,q}, \mathbf{0}_{1\times(K-L-1)}]^T$ as its first column. Further, we introduce two vectors $\mathbf{c}_q$ and $\boldsymbol{\delta}$ such that $\mathrm{diag}\{\mathbf{c}_q\} := \mathbf{W}_K\mathbf{C}_q\mathbf{W}_K^H$ and $\mathrm{diag}\{\boldsymbol{\delta}\} := \mathbf{W}_K\mathbf{\Lambda}\mathbf{W}_K^H$. Armed with these notations, we realize that

$$\begin{aligned}\|\mathbf{H}_{\mathrm{T}} - \hat{\mathbf{\Lambda}}\|^2 &= \|\mathbf{W}_K \sum_{q=0}^{Q} \mathrm{diag}\{\mathbf{b}_q\}\mathbf{C}_q\mathbf{W}_K^H - \mathbf{W}_K\hat{\mathbf{\Lambda}}\mathbf{W}_K^H\|^2, \\ &= \|\sum_{q=0}^{Q} \mathbf{W}_K \mathrm{diag}\{\mathbf{b}_q\}\mathbf{W}_K^H \mathrm{diag}\{\mathbf{c}_q\} - \mathrm{diag}\{\boldsymbol{\delta}\}\|^2.\end{aligned}$$

Obviously, the optimal $\boldsymbol{\delta}$ that minimizes the above is $\sum_{q=0}^{Q} \boldsymbol{\beta}_q \odot \mathbf{c}_q$, where $\boldsymbol{\beta}_q$ equals the diagonal of $\mathbf{W}_K\mathrm{diag}\{\mathbf{b}_q\}\mathbf{W}_K^H$. It is easy to derive that $\boldsymbol{\beta}_q = a_q\mathbf{1}_{K\times 1}$ with $a_q := \frac{\sum_{k=0}^{K-1}[\mathbf{b}_q]_k}{K}$. As a result, we obtain

$$\mathbf{\Lambda} = \sum_{q=0}^{Q} a_q\mathbf{C}_q. \tag{8}$$

Note that [4] uses a preconditioner $\mathbf{P} = \mathbf{C}_0^{-1}$ without giving a clear motivation. Actually, only with a special critically-sampled complex exponential BEM [(C)CE-BEM] assumption [7] for which

$$\mathbf{b}_q := \begin{cases} 1/K[1, \cdots, e^{-\jmath\frac{2\pi}{K}q(K-1)}]^T & \text{if } 0 \le q \le Q/2, \\ 1/K[1, \cdots, e^{\jmath\frac{2\pi}{K}(q-\frac{Q}{2})(K-1)}]^T & \text{otherwise,} \end{cases}$$

then it follows from (8) that $\mathbf{\Lambda} = \mathbf{C}_0$ is a solution to (7). For other BEM assumption, this solution is obviously not optimal.

The circulant preconditioner designed from (7) will become less effective with increasing Doppler spread. Before we address this problem, let us first look at another preconditioner design approach proposed in [6] that seeks the preconditioner directly in the frequency domain, and whose corresponding cost function is formulated as

$$\mathbf{P} = \arg\min_{\hat{\mathbf{P}}} \|\mathbf{H}_{\mathrm{F}}\hat{\mathbf{P}} - \mathbf{I}_K\|^2. \tag{9}$$

In the above, $\mathbf{P}$ is designed to make the eigenvalues of $\mathbf{H}_{\mathrm{F}}\mathbf{P}$ clustered around 1 such that its condition number is expected to be reduced. Obviously, the preconditioner derived from the above will have a similar implementation complexity as (7) if we impose a diagonal structure on $\mathbf{P}$. Equivalently, (9) can be reformulated as

$$\mathbf{p} = \arg\min_{\hat{\mathbf{p}}} \|\mathbf{H}_{\mathrm{F}}\mathrm{diag}\{\hat{\mathbf{p}}\} - \mathbf{I}_K\|^2, \tag{10}$$

from which we can readily have

$$[\mathbf{p}]_k = \frac{[\mathbf{H}_{\mathrm{F}}]_{k,k}^*}{\|\mathbf{H}_{\mathrm{F}}\mathbf{e}_k\|^2}. \tag{11}$$

A drawback of (10) is that if more energy of $\mathbf{H}_{\mathrm{F}}$ is dispersed to the off-diagonals, the resulting eigenvalues of $\mathbf{H}_{\mathrm{F}}\mathrm{diag}\{\mathbf{p}\}$ will cluster around 0 instead of 1 with the consequence that the condition number of the preconditioned channel matrix will increase rather than decrease. To understand this effect, let us use ϵ_1^2 to denote the lowest upper-bound of the residual from (10) such that $\|\mathbf{H}_{\mathrm{F}}\mathbf{e}_k[\mathbf{p}]_k - \mathbf{e}_k\|^2 < \epsilon_1^2$ holds for $k = 0, \cdots, K-1$. [6] shows that the eigenvalues of $\mathbf{H}_{\mathrm{F}}\mathrm{diag}\{\mathbf{p}\}$, denoted as σ_k, should satisfy

$$\sum_{k=0}^{K-1}(1-\sigma_k)^2 \le K\epsilon_1^2,$$

which means that all σ_k should lie inside a disk of $\sqrt{K}\epsilon_1$ centered at 1. On the other hand, following the similar steps as in [6], we use ϵ_0 to denote the smallest positive number such that $\|\mathbf{H}_{\mathrm{F}}\mathbf{e}_k[\mathbf{p}]_k\|^2 < \epsilon_0^2$ holds for $k = 0, \cdots, K-1$. Suppose $\mathbf{U}\mathbf{\Sigma}\mathbf{U}^H$ is a Schur decomposition of $\mathbf{H}_{\mathrm{F}}\mathrm{diag}\{\mathbf{p}\}$ where $\mathbf{U}$ is a unitary matrix, and the diagonal of $\mathbf{\Sigma}$ equals $[\sigma_0, \cdots, \sigma_{K-1}]^T$. Then

$$\sum_{k=0}^{K-1}\sigma_k^2 = \sum_{k=0}^{K-1}|[\mathbf{\Sigma}]_{k,k}|^2 \le \|\mathbf{\Sigma}\|^2 = \|\mathbf{H}_{\mathrm{F}}\mathrm{diag}\{\mathbf{p}\}\|^2 \le K\epsilon_0^2, \tag{12}$$

which implies that all σ_k's at the same time lie inside a disk of radius $\sqrt{K}\epsilon_0$ centered at 0. Obviously, if $\epsilon_0 < \epsilon_1$, then minimizing $\|\mathbf{H}_{\mathrm{F}}\mathrm{diag}\{\mathbf{p}\} - \mathbf{I}_K\|^2$ will at the same time minimize the Frobenius norm of $\mathbf{H}_{\mathrm{F}}\mathrm{diag}\{\mathbf{p}\}$ itself, making the eigenvalues more clustered around 0 rather than around 1. Note that with $\mathbf{p}$ obtained by (10), we can show that

$$\epsilon_1^2 = \max_k \frac{|\sum_{m=k}^{K-1}[\mathbf{H}_{\mathrm{F}}]_{m,k}|^2 - |[\mathbf{H}_{\mathrm{F}}]_{k,k}|^2}{\sum_{m=0}^{K-1}|[\mathbf{H}_{\mathrm{F}}]_{m,k}|^2}, \tag{13}$$

while

$$\epsilon_0^2 = \max_k \frac{|[\mathbf{H}_{\mathrm{F}}]_{k,k}|^2}{\sum_{m=0}^{K-1}|[\mathbf{H}_{\mathrm{F}}]_{m,k}|^2}. \tag{14}$$

Obviously, if

$$|[\mathbf{H}_{\mathrm{F}}]_{k,k}|^2 < \sum_{m=0,m\neq k}^{K-1}|[\mathbf{H}_{\mathrm{F}}]_{m,k}|^2, \text{ for } k = 0, \cdots, K-1\ , \tag{15}$$

then a diagonal preconditioner will cluster the eigenvalues in a "wrong" area. (15) implies that the sum of the off-diagonal power in each column is higher than the power on the diagonal, which tends to be more likely with the increase of the Doppler spread. In such cases, a diagonal $\mathbf{P}$ is not sufficient, and a more complex structure will be necessary. We mentioned before that the energy of time-varying channels complying with Jakes' model is usually concentrated on the entries close to the main diagonal. Hence, t is reasonable to impose a banded structure on $\mathbf{P}$ as well. Accordingly, we modify the cost function (9) as follows

$$\mathbf{P} = \arg\min_{\hat{\mathbf{P}}} \|\mathbf{H}_{\mathrm{F}}\hat{\mathbf{P}} - \mathbf{I}_K\|^2, \text{ s.t. } \hat{\mathbf{P}} = \mathbf{T}_D \odot \hat{\mathbf{P}}\ . \tag{16}$$

where $\mathbf{T}_D$ is introduced in (3) with the integer D being a design parameter.

To illustrate the performance of the preconditioners, we plot in Fig. 1 the eigenvalues of $\mathbf{H}_{\mathrm{F}}$, and the eigenvalues resulting from a diagonal preconditioner and from a tridiagonal preconditioner, respectively. The channel is assumed to have a maximum normalized Doppler spread of $\nu_{\mathrm{D}} = \frac{c}{f}vT = 0.5$, where v denotes the vehicle velocity, f the carrier frequency, T the OFDM symbol duration, and c the speed of the propagation medium. Refer to Section 5 for more details of the channel parameters.

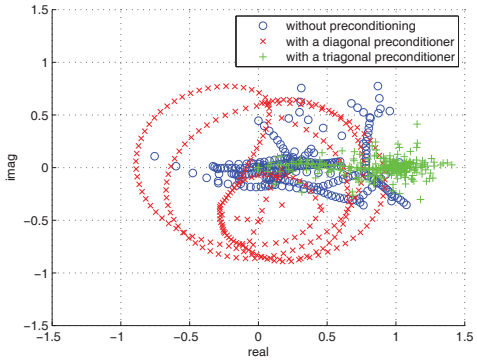

Fig. 1. Eigenvalues with and without a diagonal preconditioner.

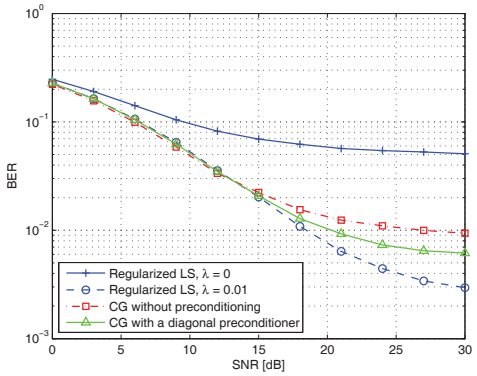

Fig. 2. BER performance for channels with $\nu_{\rm D} = 0.2$.

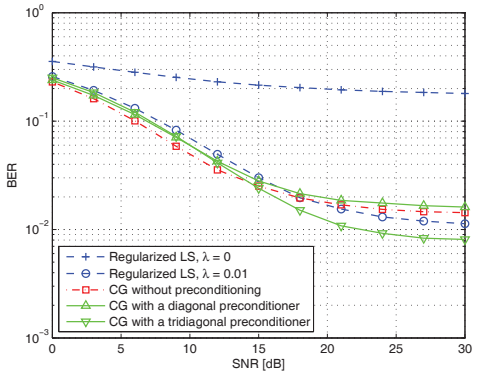

Fig. 3. BER performance for channels with $\nu_{\rm D} = 0.5$.

5. SIMULATIONS AND DISCUSSIONS

In the following simulations, the time-varying channels are generated that conforms to Jakes' Doppler profile [8] using the channel generator given in [10]. We assume $L+1 = 6$ channel taps, which are mutually uncorrelated with variance $\sigma_l^2 = 1/\sqrt{L+1}$, and consider an OFDM system with 256 subcarriers. A quadrature phase-shift keying (QPSK) symbol is modulated onto each subcarrier.

The CG algorithm assumes full channel knowledge, and is designed to halt after 5 iterations. Next to the CG algorithm, we will also compare the performance of a regularized least-squares (LS) equalizer, which is similar to that considered in [2, 3] whose data estimate $\hat{\mathbf{s}}$ is given by

$$\hat{\mathbf{s}} = (\mathbf{H}_{\rm F}^H \mathbf{H}_{\rm F} + \lambda \mathbf{I}_K)^{-1} \mathbf{H}_{\rm F}^H \mathbf{y},$$

where λ is known as the regularization parameter. For an ill-conditioned $\mathbf{H}_{\rm F}$, which is often the case with faster-varying channels, the presence of λ reduces the error floor especially for high SNR. In the simulations, we examine both $\lambda = 0.01$ and $\lambda = 0$ (no regularization). In order to further reduce the complexity, all the equalizers are based on a banded approximation of the true channel matrix $\tilde{\mathbf{H}}_{\rm F} = \mathbf{T}_{Q'} \odot \mathbf{H}_{\rm F}$ with $Q' = 9$.

Test Case 1. We first look at a relatively slow time-varying channel case with a normalized Doppler spread $\nu_{\rm D} = 0.2$. The BER vs. SNR performance is plotted in Fig. 2. It is clear to see that the LS equalizer without regularization suffers from a very high noise floor. Further, the diagonally preconditioned CG algorithm outperforms that without preconditioning, although it is still inferior to the LS equalizer with proper regularization.

Test Case 2. For a highly time-varying channel case with a normalized Doppler spread $\nu_{\rm D} = 0.5$, the BER vs. SNR performance is plotted in Fig. 3 where we can see that the significance of the regularization parameter λ to the LS equalizer is even more pronounced. The CG algorithm with a diagonal preconditioner performs obviously even worse than without preconditioning. When equipped with a tridiagonal preconditioner, the CG algorithm outperforms the LS equalizer.

6. REFERENCES

[1] P. Schniter, "Low-complexity equalization of OFDM in doubly-selective channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 4, pp. 1002–1011, Apr 2004.

[2] L. Rugini, P. Banelli, and G. Leus, "Simple equalization of time-varying channels for OFDM," *IEEE Communications Letters*, vol. 9, no. 7, pp. 619–621, July 2005.

[3] L. Rugini, P. Banelli, and G. Leus, "Low-complexity banded equalizers for OFDM systems in Doppler spread channels," *EURASIP Journal on Applied Signal Processing*, pp. Article ID 67404, 13 pages, 2006.

[4] T. Hrycak, S. Das, G. Matz, and H. G. Feichtinger, "Low complexity equalization for doubly selective channels modeled by a basis expansion," *IEEE Transactions on Signal Processing*, vol. 58, no. 11, pp. 5706–5719, Nov. 2010.

[5] G. H. Golub and C. F. van Loan, *Matrix Computations*, The Johns Hopkins University Press, 1989.

[6] M. J. Grote and T. Huckle, "Parallel preconditioning with sparse approximate inverses," *SIAM Journal on Scientific Computing*, vol. 18, no. 3, pp. 838–853, May 2007.

[7] X. Ma, G. Giannakis, and S. Ohno, "Optimal training for block transmissions over doubly-selective fading channels," *IEEE Transactions on Signal Processing*, vol. 51, no. 5, pp. 1351–1366, May 2003.

[8] W. C. Jakes, *Microwave Mobile Channels*, New York: Wiley, 1974.

[9] T. F. Chan, "An optimal circulant preconditioner for Toeplitz systemss," *SIAM Journal on Scientific and Statistical Computing*, vol. 9, no. 4, pp. 766–771, 1988.

[10] Y. R. Zheng and C. Xiao, "Simulation models with correct statistical properties for Rayleigh fading channels," *IEEE Transactions on Communications*, vol. 51, no. 6, pp. 920–928, June 2003.