REGULARIZED LINEAR EQUALIZATION FOR MULTIPATH CHANNELS WITH IMPERFECT CHANNEL ESTIMATION

Jun Tong and Peter J. Schreier

Signal and System Theory Group, Universität Paderborn, Germany
{jun.tong, peter.schreier}@sst.upb.de

ABSTRACT

This paper deals with different techniques for linear equalization of multipath channels with imperfect channel estimation (CE). We develop a unified framework based on Krylov subspace expansion, which allows us to compare the performance of the conjugate gradient (CG) method, diagonal loading (DL), and a hybrid scheme. Our analysis shows that the DL method generally outperforms its alternatives, but at the cost of higher complexity. However, we also demonstrate that a proper implementation of the low-complexity CG method can also approach the performance of DL. Finally, we show that preconditioning degrades performance when the CE is poor.

Index Terms— Conjugate gradient (CG), diagonal loading (DL), imperfect channel estimation, Krylov subspace, equalization.

1. INTRODUCTION

Linear equalization is widely used for combating the inter-symbol interference (ISI) in multipath channels [1]. Two concerns often arise simultaneously: The complexity can be high when the system is large, and the performance may degrade drastically when the channel estimation (CE) is imperfect. This is even more critical in time-varying environments.

Robust linear equalization aims to alleviate the negative impact of imperfect CE. When *a priori* information about the CE error is available, closed-form solutions to the optimal linear equalization may be derived under certain performance criteria [2]. However, there are still applications where such *a priori* information is absent or inaccurate. This paper considers such cases where regularized linear equalization designed by trying multiple candidates is applied to improve performance. Regularization [3] can be achieved by different approaches such as diagonal loading (DL) [4, 5] and reduced-rank processing [6]. The latter can be realized based on eigendecomposition or Krylov subspace expansion methods such as the conjugate gradient (CG) method [7]. Both the DL and CG methods can noticeably improve performance for imperfect CE, but a detailed comparison of their performance and complexity for regularized linear equalization is still missing.

This paper introduces a framework based on Krylov subspace expansion, which provides a unifying and efficient approach to regularization. Based on this framework, we are able to compare the DL and CG methods and a hybrid scheme for regularized linear equalization of multipath channels. The analysis and numerical examples demonstrate that the DL method performs better but also requires higher complexity. However, we show that the performance gap can be greatly reduced if the CG method is used properly. We also study the effect of preconditioning, an important supplement to Krylov subspace expansion methods. We show that preconditioning is beneficial only when the CE is very accurate.

This work was supported by the Australian Research Council under Discovery Project DP0986916. The Signal and System Theory Group is supported by the Alfried Krupp von Bohlen und Halbach foundation.

2. SYSTEM MODEL

A multiple-antenna communication system operating over a time-invariant or quasi-static multipath channel can be modeled as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{y} \in \mathbb{C}^{N\times 1}$, $\mathbf{H} \in \mathbb{C}^{N\times M}$, $\mathbf{x} \in \mathbb{C}^{M\times 1}$ and $\mathbf{n} \in \mathbb{C}^{N\times 1}$ are the received signal, channel matrix, transmit signal, and noise, respectively. Assume that $\mathbb{E}[\mathbf{x}] = 0$, $\mathbb{E}[\mathbf{x}\mathbf{x}^\dagger] = \mathbf{I}_M$, and the noise, uncorrelated with the signal, has zero mean and covariance matrix $\sigma_n^2\mathbf{I}_N$, where $\mathbb{E}[\cdot]$ denotes expectation, $(\cdot)^\dagger$ conjugate transpose and $\mathbf{I}_N$ an $N \times N$ identity matrix. Let n_t and n_r be the numbers of transmit and receive antennas, respectively.

Model (1) can be used for both serial and block equalization. In the serial mode, $\mathbf{y}$ denotes a sliding-windowed received signal vector. Let T be the window size, L the number of channel taps, and $\mathbf{H}_l \in \mathbb{C}^{n_r\times n_t}$ the lth tap coefficient matrix. Then the channel matrix can be written as

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_L & \mathbf{H}_{L-1} & \cdots & \mathbf{H}_1 & & & \\ & \mathbf{H}_L & \dots & \mathbf{H}_2 & \mathbf{H}_1 & & \\ & & \ddots & \vdots & \vdots & \ddots & \\ & & & \mathbf{H}_L & \mathbf{H}_{L-1} & \cdots & \mathbf{H}_1 \end{pmatrix} \tag{2}$$

with $N = Tn_r$ and $M = (T + L - 1)n_t$. For each $\mathbf{y}$, n_t symbols transmitted at time $T - T_d$ are jointly estimated if the decision delay is T_d. Block linear equalization can be applied when a guard interval is inserted for each block. In this case,

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_1 & & & & \\ \mathbf{H}_2 & \mathbf{H}_1 & & & \\ \vdots & \vdots & \ddots & & \\ \mathbf{H}_L & \mathbf{H}_{L-1} & \cdots & \mathbf{H}_1 & \\ & \ddots & \ddots & \ddots & \\ & & \mathbf{H}_L & \cdots & \mathbf{H}_1 \\ & & & \ddots & \vdots \\ & & & & \mathbf{H}_L \end{pmatrix}. \tag{3}$$

If the transmit block length at each antenna is $T - L + 1$, then $N = Tn_r$ and $M = (T - L + 1)n_t$. Given $\mathbf{y}$, all the M symbols in $\mathbf{x}$ are jointly estimated. In (2) and (3) $\mathbf{H}$ is block-banded. In many cases, $\mathbf{H}$ is also sparse, with a large L but only a few nonzero $\mathbf{H}_l$.

Assume that $\mathbf{H}$ and σ_n^2 are estimated as $\hat{\mathbf{H}}$ and $\hat{\sigma}_n^2$, respectively,

978-1-4673-0046-9/12/$26.00 ©2012 IEEE

at the receiver. A naive linear equalization scheme treats $\hat{\mathbf{H}}$ and $\hat{\sigma}_n^2$ as perfect. For serial equalization, the mth entry of $\mathbf{x}$ is estimated from $\mathbf{y}$ as $\hat{x}_m = \mathbf{w}_m^\dagger \mathbf{y}$, where $m \in \{(T-T_d-1)n_t+1, \ldots, (T-T_d)n_t\}$, $\mathbf{w}_m$ solves

$$(\hat{\mathbf{H}}\hat{\mathbf{H}}^\dagger + \hat{\sigma}_n^2 \mathbf{I}_N)\mathbf{w}_m = \hat{\mathbf{h}}_m, \qquad (4)$$

and $\hat{\mathbf{h}}_m$ is the mth column of $\hat{\mathbf{H}}$. For block equalization, $\mathbf{x}$ can be estimated as $\hat{\mathbf{x}} = \hat{\mathbf{H}}^\dagger \mathbf{z}$ where $\mathbf{z}$ solves

$$(\hat{\mathbf{H}}\hat{\mathbf{H}}^\dagger + \hat{\sigma}_n^2 \mathbf{I}_N)\mathbf{z} = \mathbf{y}. \qquad (5)$$

It can be verified using the Woodbury matrix identity that an equivalent estimate is the solution to

$$(\hat{\mathbf{H}}^\dagger\hat{\mathbf{H}} + \hat{\sigma}_n^2 \mathbf{I}_M)\hat{\mathbf{x}} = \hat{\mathbf{H}}^\dagger \mathbf{y}. \qquad (6)$$

The naive linear equalization scheme solves (4), (5) or (6) exactly, e.g., using the Cholesky factorization, forward and backward substitution [7]. Its drawback is that the complexity is cubic in the size of the system and the performance can degrade dramatically with imperfect CE.

3. REGULARIZED LINEAR EQUALIZATION

This section compares various regularized linear equalization schemes in a framework that combines *preconditioning*, *projection*, and *diagonal loading (DL)*. The key step for linear equalization is to solve (4), (5) or (6) which has the form

$$\mathbf{A}\mathbf{t} = \mathbf{b}. \qquad (7)$$

Instead of the exact solution $\mathbf{t} = \mathbf{A}^{-1}\mathbf{b}$, a regularized solution $\mathbf{t}^{\mathrm{reg}}$ to (7) is pursued.

$$\begin{aligned} \mathbf{t}^{\mathrm{reg}} &= \mathbf{P}\mathbf{V}\mathbf{c}^{\mathrm{DL}} & (8)\\ \mathbf{c}^{\mathrm{DL}} &= (\mathbf{T} + \delta \mathbf{I}_K)^{-1}\mathbf{d} & (9)\\ \mathbf{T} &\triangleq \mathbf{V}^\dagger \mathbf{P}^\dagger \mathbf{A}\mathbf{P}\mathbf{V} & (10)\\ \mathbf{d} &\triangleq \mathbf{V}^\dagger \mathbf{P}^\dagger \mathbf{b}. & (11) \end{aligned}$$

The ideas behind these schemes are as follows:

- **Preconditioning:** A preconditioner $\mathbf{P}$ is applied to (7) such that $\mathbf{P}^\dagger\mathbf{A}\mathbf{P}$ has clustered eigenvalues.
- **Projection:** The system $\mathbf{P}^\dagger\mathbf{A}\mathbf{P}(\mathbf{P}^{-1}\mathbf{t}) = \mathbf{P}^\dagger\mathbf{b}$ is projected onto a rank-K subspace with K orthonormal basis vectors specified by the columns of $\mathbf{V}$, leading to a linear system $\mathbf{T}\mathbf{c} = \mathbf{d}$ with $\mathbf{T}$ and $\mathbf{d}$ given by (10) and (11), respectively.
- **DL:** A DL solution (9) to $\mathbf{T}\mathbf{c} = \mathbf{d}$ is computed and a reduced-rank solution to (7) is then obtained using (8), giving

  $$\mathbf{t}^{\mathrm{reg}} = \mathbf{P}\mathbf{V}(\mathbf{V}^\dagger\mathbf{P}^\dagger\mathbf{A}\mathbf{P}\mathbf{V} + \delta\mathbf{I})^{-1}\mathbf{V}^\dagger\mathbf{P}^\dagger\mathbf{b}, \qquad (12)$$

  where δ is the DL factor.

The above formulation is very general. For example, for the principal component analysis (PCA) method, $\mathbf{P} = \mathbf{I}$, $\delta = 0$ and $\mathbf{V}$ consists of the K principal eigenvectors of $\mathbf{A}$.

3.1. Krylov Subspace Projection

For a matrix $\mathbf{A}$ and a vector $\mathbf{b}$, the order-K Krylov subspace is defined as $\mathcal{K}_K(\mathbf{A}, \mathbf{b}) \triangleq \mathrm{span}(\mathbf{b}, \mathbf{A}\mathbf{b}, \ldots, \mathbf{A}^{K-1}\mathbf{b})$. Given a preconditioner $\mathbf{P}$, the following Lanczos procedure can be used to generate an orthonormal basis for $\mathcal{K}_K(\mathbf{P}^\dagger\mathbf{A}\mathbf{P}, \mathbf{P}^\dagger\mathbf{b})$ and a tridiagonal $\mathbf{T}$. First, set $\mathbf{v}_0 = \mathbf{0}$, $\beta_1\mathbf{v}_1 = \mathbf{P}^\dagger\mathbf{b}$ where $\beta_1 = ||\mathbf{P}^\dagger\mathbf{b}||$. Then, for $i = 1, 2, \ldots, K$, compute

$$\begin{aligned} \alpha_i &= \mathbf{v}_i^\dagger \mathbf{P}^\dagger \mathbf{A}\mathbf{P}\mathbf{v}_i \\ \beta_{i+1}\mathbf{v}_{i+1} &= \mathbf{P}^\dagger\mathbf{A}\mathbf{P}\mathbf{v}_i - \alpha_i\mathbf{v}_i - \beta_i\mathbf{v}_{i-1} \qquad (13) \end{aligned}$$

recursively, where $\{\beta_i\}$ are chosen to make $\{\mathbf{v}_i\}$ unit length. This procedure yields $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_K]$ and a tridiagonal matrix

$$\mathbf{T} = \mathbf{V}^\dagger\mathbf{P}^\dagger\mathbf{A}\mathbf{P}\mathbf{V} = \begin{bmatrix} \alpha_1 & \beta_2 & & \\ \beta_2 & \ddots & & \\ & \ddots & & \beta_K \\ & & \beta_K & \alpha_K \end{bmatrix}. \qquad (14)$$

With tridiagonal $\mathbf{T}$, (9) can be solved with $9K$ operations [7, p. 157].

3.2. Preconditioning

One approach to reducing the rank for Krylov subspace expansion is to use a preconditioner $\mathbf{P}$ to cluster the eigenvalues of $\mathbf{P}^\dagger\mathbf{A}\mathbf{P}$. Consider (4) and (5) where $\mathbf{A} = \hat{\mathbf{H}}\hat{\mathbf{H}}^\dagger + \hat{\sigma}_n^2\mathbf{I}_N$, which can be approximated by the covariance matrix of the observed vector for a cyclic-prefixed system operating over the same channel:

$$\widetilde{\mathbf{A}} = \mathbf{F}_{\mathrm{Rx}}(\mathbf{G}\mathbf{G}^\dagger + \hat{\sigma}_n^2\mathbf{I}_N)\mathbf{F}_{\mathrm{Rx}}^\dagger, \qquad (15)$$

where $\mathbf{F}_{\mathrm{Rx}}$=$\mathbf{F}\otimes\mathbf{I}_{n_r}$, $\otimes$ denotes Kronecker product, $\mathbf{F}$ is the unitary $T\times T$ discrete Fourier transform matrix, and $\mathbf{G} \triangleq \mathrm{diag}(\mathbf{G}_1, \mathbf{G}_2, \ldots, \mathbf{G}_T)$ is block-diagonal with blocks $\mathbf{G}_k = \sum_{l=1}^{L} \hat{\mathbf{H}}_l e^{-\frac{j2\pi(k-1)(l-1)}{T}}$ of size $n_r \times n_t$, which can be computed using $n_r n_t$ FFTs of T points. A preconditioner is constructed as

$$\mathbf{P} = \mathbf{F}_{\mathrm{Rx}}(\mathbf{G}\mathbf{G}^\dagger + \hat{\sigma}_n^2\mathbf{I}_N)^{-1/2}, \qquad (16)$$

which can be implemented based on the Cholesky factors of $\mathbf{G}_k\mathbf{G}_k^\dagger + \hat{\sigma}_n^2\mathbf{I}_{n_r}$, $k = 1, 2, \ldots, T$. Preconditioner (16) is, in essence, a circulant preconditioner [11] constructed from the CE.

3.3. Analysis of Regularization Effect and Complexity

Different regularized linear equalization schemes can now be discussed using the above framework. A globally optimal solution to $(\mathbf{P}, K, \delta)$ leads to high complexity. This paper analyzes and compares three suboptimal solutions with relatively low complexity.

3.3.1. Conjugate Gradient (CG)

Without preconditioning and diagonal loading ($\mathbf{P} = \mathbf{I}, \delta = 0$),

$$\mathbf{t}^{\mathrm{reg}} = \mathbf{V}(\mathbf{V}^\dagger\mathbf{A}\mathbf{V})^{-1}\mathbf{V}^\dagger\mathbf{b} \qquad (17)$$

is the same as a solution obtained by applying the CG (or other equivalent) algorithm [8, 9, 10] to $\mathbf{A}\mathbf{t} = \mathbf{b}$. As K increases, approximate solutions in different Krylov subspaces $\mathcal{K}_K(\mathbf{A}, \mathbf{b})$ are obtained. For perfect CE, the performance improves with K until $\mathcal{K}_K(\mathbf{A}, \mathbf{b})$ converges. For imperfect CE, a semi-convergence behavior may be observed: At the beginning, the performance improves with K by including more basis vectors specified by $\{\mathbf{v}_i\}$; after crossing the optimal rank $K^\star$, the performance degrades with K due to the increasing contribution from less reliable directions contaminated by the CE error. The $K^\star$ can be determined by estimating the mean squre error (MSE) using training or the decisions on $\mathbf{x}$. It is often the case that $K^\star$ corresponds to the first local minimum of the MSE versus K curve. In this case, the maximum K to be tested is only slightly larger than $K^\star$ and the overall complexity is linear in $K^\star$. The CG method usually requires lower complexity than the Cholesky factorization method for linear equalization in large systems.

For block equalization, applying the CG algorithm to (5) and (6) yield different regularization effect. With (5), a regularized solution $\mathbf{z}^{\text{reg}}$ to (5) is first obtained and $\mathbf{x}$ is then estimated as $\hat{\mathbf{x}} = \hat{\mathbf{H}}^\dagger \mathbf{z}^{\text{reg}}$. The operation $\hat{\mathbf{H}}^\dagger \mathbf{z}^{\text{reg}}$ compromises the regularizaton effect since it can re-introduce the error due to the imperfect CE. Consequently, the scheme based on (6), which regularizes the solution at the final stage, outperforms the one using (5). When $\hat{\sigma}_n^2 = 0$ and (6) is used, the CG solution is equivalent to the LSQR algorithm [10].

3.3.2. Diagonal Loading (DL)

The Krylov subspace expansion also yields an efficient approach to implementing the DL method. Without preconditioning ($\mathbf{P} = \mathbf{I}$),

$$\mathbf{t}^{\text{reg}} = \mathbf{V}(\mathbf{V}^\dagger \mathbf{A}\mathbf{V} + \delta\mathbf{I})^{-1}\mathbf{V}^\dagger \mathbf{b} \tag{18}$$

gives a reduced-rank solution to the diagonally loaded system

$$(\mathbf{A} + \delta\mathbf{I})\mathbf{t} = \mathbf{b}, \tag{19}$$

which resides in a rank-K Krylov subspace $\mathcal{K}_K(\mathbf{A} + \delta\mathbf{I}, \mathbf{b}) = \text{span}(\mathbf{b}, (\mathbf{A} + \delta\mathbf{I})\mathbf{b}, \ldots, (\mathbf{A} + \delta\mathbf{I})^{K-1}\mathbf{b}) = \mathcal{K}_K(\mathbf{A}, \mathbf{b})$. As $\mathbf{V}$ and $\mathbf{T} = \mathbf{V}^\dagger \mathbf{A}\mathbf{V}$ are independent of δ, the Lanczos procedure in Section 3.1 can be shared by different δ. Then for different DLF δ, one only needs to solve (9) using Cholesky factorization (or other methods) with a complexity linear in K.

This provides a low-complexity alternative to the eigendecomposition approach to implementing DL. (The latter has been suggested in [4].) Firstly, computing an orthorgonal basis vector for $\mathcal{K}_K(\mathbf{A}, \mathbf{b})$ is significantly cheaper than computing an eigenvector of $\mathbf{A}$, especially for a sparse $\mathbf{A}$ [7]. Secondly, the rank K required for solving (19) is much smaller than the eigendecomposition approach if $\mathbf{A}$ has clustered eigenvalues.

The DL method can outperform the CG method in terms of regularization effect. Let us model the imperfect CE by $\hat{\mathbf{H}} = \mathbf{H} + \tilde{\mathbf{H}}$ and $\hat{\sigma}_n^2 = \sigma_n^2 + \tilde{\sigma}_n^2$, where $\tilde{\mathbf{H}}$ and $\tilde{\sigma}_n^2$ denote the errors in the estimates of channel matrix and noise variance, respectively. Assume that $\tilde{\mathbf{H}}$ and $\tilde{\sigma}_n^2$ are independent of $\mathbf{x}$, $\mathbf{n}$, $\mathbf{H}$, and σ_n^2, and $\tilde{\mathbf{H}}$ has zero-mean. Consider the equalizer $\hat{\mathbf{x}} = \mathbf{W}^\dagger \mathbf{y}$. Then the total MSE averaged over $\tilde{\mathbf{H}}$, $\tilde{\sigma}_n^2$, $\mathbf{x}$ and $\mathbf{n}$, conditioned on $\hat{\mathbf{H}}$ and $\hat{\sigma}_n^2$, is

$$\begin{aligned}
&\mathbb{E}_{\tilde{\mathbf{H}},\tilde{\sigma}_n^2,\mathbf{x},\mathbf{n}}\left[||\hat{\mathbf{x}} - \mathbf{x}||^2|\hat{\mathbf{H}}, \hat{\sigma}_n^2\right] \\
=\ &\mathbb{E}_{\tilde{\mathbf{H}},\tilde{\sigma}_n^2,\mathbf{x},\mathbf{n}}\left[\mathbf{x}^\dagger(\mathbf{W}^\dagger\mathbf{H}-\mathbf{I})^\dagger(\mathbf{W}^\dagger\mathbf{H}-\mathbf{I})\mathbf{x}+\mathbf{n}^\dagger\mathbf{W}^\dagger\mathbf{W}\mathbf{n}|\hat{\mathbf{H}}, \hat{\sigma}_n^2\right] \\
=\ &\text{tr}\left(\mathbb{E}_{\tilde{\mathbf{H}}}\left[(\mathbf{H}^\dagger\mathbf{W} - \mathbf{I})(\mathbf{W}^\dagger\mathbf{H} - \mathbf{I})\right]\right) + \mathbb{E}_{\tilde{\sigma}^2}[\sigma_n^2]\text{tr}\left(\mathbf{W}\mathbf{W}^\dagger\right) \\
=\ &\text{tr}\left(\left(\hat{\mathbf{H}}\hat{\mathbf{H}}^\dagger + \mathbb{E}_{\tilde{\mathbf{H}}}[\tilde{\mathbf{H}}\tilde{\mathbf{H}}^\dagger] + (\hat{\sigma}_n^2 + \mathbb{E}_{\tilde{\sigma}^2}[\tilde{\sigma}^2])\mathbf{I}\right)\mathbf{W}\mathbf{W}^\dagger\right) \\
&-\text{tr}\left(\hat{\mathbf{H}}\mathbf{W}^\dagger + \mathbf{W}\hat{\mathbf{H}}^\dagger\right) + M.
\end{aligned} \tag{20}$$

It can be verified that the optimal equalizer that minimizes (20) is

$$\mathbf{W} = \left(\hat{\mathbf{H}}\hat{\mathbf{H}}^\dagger + \mathbb{E}_{\tilde{\mathbf{H}}}[\tilde{\mathbf{H}}\tilde{\mathbf{H}}^\dagger] + (\hat{\sigma}_n^2 + \mathbb{E}_{\tilde{\sigma}^2}[\tilde{\sigma}^2])\mathbf{I}\right)^{-1}\hat{\mathbf{H}}. \tag{21}$$

If the entries of $\tilde{\mathbf{H}}$ are uncorrelated and $\mathbb{E}_{\tilde{\mathbf{H}}}[\tilde{\mathbf{H}}\tilde{\mathbf{H}}^\dagger] = \epsilon\mathbf{I}$ for some ϵ, the optimal equalizer can be found by trying different DL factors. In this sense, the DL method can outperform the CG method.

The parameters K and δ can be determined by a two-stage procedure. First, a K large enough is chosen (e.g., by the discrepancy criterion) such that $\mathbf{A}\mathbf{t} = \mathbf{b}$ can be solved almost exactly. Then δ is found from a number of candidates, similarly to the search of K for the CG method. In this paper, δ is chosen from 21 candidates $\frac{\text{tr}(\mathbf{T})}{K}10^{-0.4p}$, $p = 0, 1, \ldots, 20$. The resulting complexity of DL is linear in K and higher than the CG method since the rank K needed is usually larger than the optimal $K^\star$ for the CG method.

3.3.3. Hybrid Regularization

The CG and DL methods use either K or δ to control the regularization effect. The CG method may suffer from relatively poor performance while the DL method requires higher complexity. A hybrid scheme that yields a compromise between these methods is as follows: First, choose a K as the optimal $K^\star$ as in Section 3.3.1. Then δ is optimized with $K = K^\star$.

3.3.4. Impact of Preconditioning

Preconditioning aims to reduce the rank required for solving (7) exactly. This can reduce complexity for the case with highly accurate CE. With imperfect CE, preconditioning may compromise the regularization effect. Intuitively, this is because the linear transform by $\mathbf{P} \neq \mathbf{I}$ mixes the signal and noise subspaces.

3.4. Examples

The discussions above are now validated by examples. Assume that Q out of the L tap coefficient matrices $\mathbf{H}_l$ are nonzero, the entries of the nonzero $\mathbf{H}_l$ are i.i.d., complex Gaussian with variance $1/Q$ and the imperfect CE introduces uncorrelated errors to the estimates of nonzero $\mathbf{H}_l$, as modeled in Section 3.3.2. The CE error is characterized by $\mathbb{E}[||\tilde{\mathbf{H}}_l||^2]/\mathbb{E}[||\mathbf{H}_l||^2]$. In Fig. 1, a serial equalization scheme with mismatched noise variance information $\hat{\sigma}_n^2 = 0$ is simulated. Only the best achievable performance and the corresponding ranks for each regularization scheme are shown. Practical parameter determination is not considered here due to space limitations. For a fixed preconditioner $\mathbf{P}$, the overall complexity for each regularization scheme is linear in the average rank shown in Fig. 1b. For a fixed rank, the complexity with preconditioning is about twice that without preconditioning. Fig. 2 considers the case with perfect estimate of noise variance, i.e., $\hat{\sigma}_n^2 = \sigma_n^2$. Fig. 3 shows a block equalization example with $\hat{\sigma}_n^2 = 0.2\sigma_n^2$. The simulation results agree well with the discussions in the previous sections.

4. CONCLUSIONS

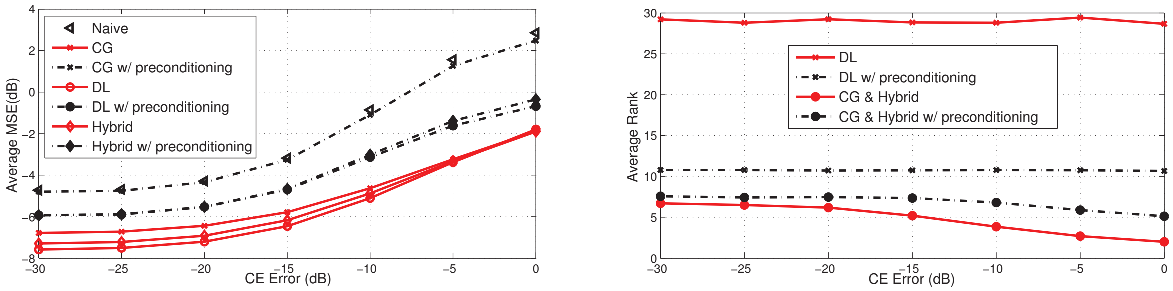
This paper has discussed regularized linear equalization for multipath channels with imperfect channel estimation. It is shown that Krylov subspace expansion offers a unified framework for various regularization techniques. The analysis and examples demonstrate that the CG method achieves a good performance-complexity tradeoff; the DL method provides noticable performance gains if the estimation of channel and noise variance is poor; and preconditioning is useful only when the channel estimation is reasonably accurate.

5. REFERENCES

[1] J. Proakis and M. Salehi, *Digital Communications*, 5th ed., New York: McGraw-Hill, 2008.

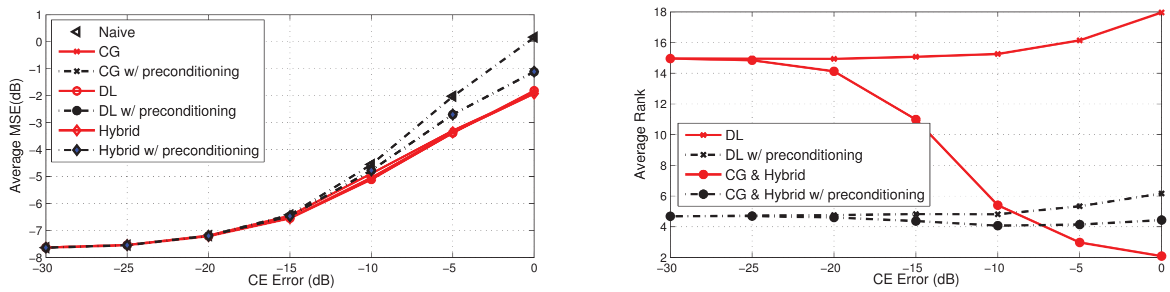
[2] Y. Rong, S. A. Vorobyov, and A. B. Gershman, "Robust linear receivers for multiaccess space-time block-coded MIMO systems: A probabilistically constrained approach," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1560–1569, Aug. 2006.

[3] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion.* Philadelphia, PA: SIAM, 1998.



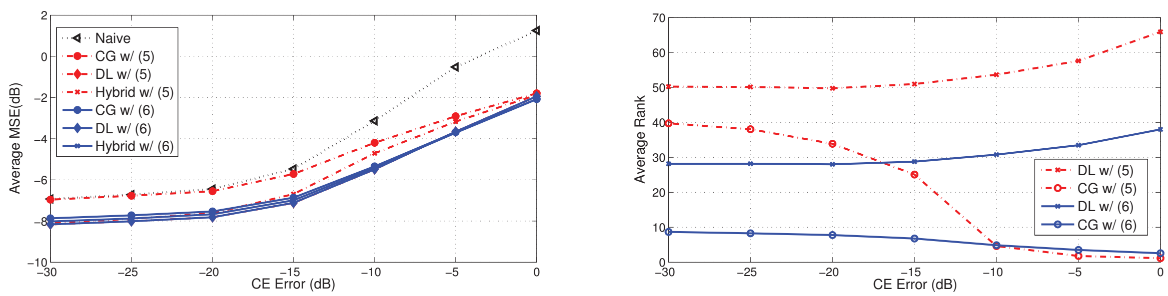
(a) Average MSE per symbol (b) Average of the rank K

Fig. 1: Serial equalization for multipath channels with $n_t = n_r = 2, L = 20, Q = 6, T = 128, M = 294, N = 256$, SNR $\triangleq 1/\sigma_n^2 = 10$ dB, $\hat{\sigma}_n^2 = 0$ and decision delay $T_d = 64$. The DL scheme outperforms CG but requires higher complexity. Preconditioning degrades performance.



(a) Average MSE per symbol (b) Average of the rank K

Fig. 2: Serial equalization for multipath channels with the same parameters as those in Fig. 1 except $\hat{\sigma}_n^2 = \sigma_n^2$. Preconditioning reduces the complexity with minor performance degradation when the CE is accurate. Without preconditioning, the CG, DL and hybrid schemes achieve similar performance, with DL performing slightly better.



(a) Average MSE per symbol (b) Average of the rank K

Fig. 3: Block equalization for multipath channels with $n_t = n_r = 2, L = 20, Q = 6, T = 128, M = 216, N = 256$, SNR $\triangleq 1/\sigma_n^2 = 10$ dB and $\hat{\sigma}_n^2 = 0.2\sigma_n^2$. The CG and hybrid methods based on (6) outperform those based on (5). The DL scheme based on (6) has the same performance as that based on (5) but requires lower complexity. With (6), the CG, DL and hybrid schemes achieve similar performance, significantly outperforming the naive solution.

[4] J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1702–1715, Jul. 2003.

[5] X. Mestre and M. A. Lagunas, "Finite sample size effect on minimum variance beamformers: Optimum diagonal loading factor for large arrays," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 69–82, Jan. 2006.

[6] L. L. Scharf, "The SVD and reduced rank signal processing," *Sig. Process.*, vol. 25, pp. 113–133, 1991.

[7] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., The Johns Hopkins University Press, 1996.

[8] H. Ge, M. Lundberg, and L. L. Scharf, "Warp convergence in conjugate gradient Wiener filters," in *Proc. 3rd IEEE Int.Workshop Sensor Array and Multichannel Signal Process. (SAM)*, Barcelona, Spain, Jul. 2004.

[9] R. Guan and T. Strohmer, "Krylov subspace algorithms and circulant-embedding method for efficient wideband single-carrier equalization," *IEEE Trans. Sig. Process.*, vol. 56, no. 6, pp. 2483–2495, June 2008.

[10] T. Hrycak, S. Das, G. Matz, and H. G. Feichtinger, "Low-complexity equalization for doubly selective channels modelled by a basis expansion," *IEEE Trans. Sig. Process.*, vol. 58, no. 11, pp. 5706–5719, Nov. 2010.

[11] R. H. Chan and M. K. Ng, "Conjugate gradient methods for Toeplitz systems," *SIAM Review*, vol. 38, no. 3, pp. 427–482, Sep. 1996.