Blind OFDM Channel Estimation Using Receiver Diversity

Hao Wang, Ying Lin, and Biao Chen[1]
Department of Electrical Engineering and Computer Science
Syracuse University
121 Link Hall
Syracuse, NY 13244
email: bichen@ecs.syr.edu.

Abstract — **In this paper we exploit receiver diversity for blind channel estimation for OFDM systems with cyclic prefix. The developed channel estimation algorithm has the following advantages compared with existing blind methods: a) It is data efficient — the proposed algorithm can be implemented using a single OFDM block; b) It is computationally efficient — only a single singular value decomposition is required for channel estimation; c) It is symbol independent — unlike many existing algorithms, there is no restriction on the input symbol constellation. Cramer-Rao lower bound for the channel estimation is derived to evaluate the performance. We show through numerical examples that the proposed algorithm compare favorably with an existing subspace algorithm.**

I. Introduction

Because of its resistance to multipath channel fading and spectral efficiency, orthogonal frequency division multiplexing (OFDM) has attracted increasing interest in recent years as a suitable modulation scheme for broadband wireless communication systems, including digital broadcasting and wireless LAN applications.

For coherent OFDM system, reliable estimation of the time dispersive channel is key to achieve the desired performance gain. Training symbol based OFDM method usually requires extra +20% bandwidth therefore consumes too much precious resources. On the other hand, many existing blind OFDM channel estimation is statistical in nature (e.g., second order statistics based as in [1, 2, 3, 4]) which usually requires large number of data blocks. Further, it has limited application in wireless channel involving high mobility (large Doppler spread) as the channel may vary from block to block. Deterministic blind channel estimation, on the other hand, is more data efficient. For example, the finite-alphabet based method in [5] can be implemented using only a single data block. However, the developed algorithm is mostly limited in practice to PSK modulation.

Receiver diversity is another important resource that can be exploited in OFDM channel estimation. In [6, 7] multiple receive antennas are used for channel estimation for OFDM systems without cyclic prefix (CP). In this paper, we also exploit receiver diversity in the form of either multiple antennass or oversampling to allow blind channel estimation for the CP based OFDM systems. The resulting algorithm is simple yet effective. Without restrictions on the input symbol constellation, the proposed method is both data efficient and computationally efficient. We also present some preliminary result on identifiability issue.

[1]This research was supported in part by the CASE center of Syracuse University.

The organization of the paper is as follows. In the next section we introduce the signal model with receiver diversity. In section III, we present a simple blind channel estimation algorithm using channel diversity. Section IV includes results on the identifiability. Cramer-Rao lower bound (CRLB) for the channel estimation is derived in section V while simulation results are given in section VI. We conclude in section VII.

The following notations are frequently used in this paper. The DFT matrix $\mathbf{W}$ can be partitioned as

$$\mathbf{W} = [\mathbf{W}_L | \mathbf{W}_{N-L}] \tag{1}$$

where L is the length of channel impulse response which is assumed known *a priori* in this paper, $\mathbf{W}_L$ is the matrix composed of the first L columns of $\mathbf{W}$, and $\mathbf{W}_{N-L}$ contains the remaining $N-L$ columns. Further we can write

$$\mathbf{W}_L = \begin{bmatrix} \mathbf{u}_1^H \\ \vdots \\ \mathbf{u}_N^H \end{bmatrix} \tag{2}$$

where each $\mathbf{u}_k$ is an L by 1 vector. We use bold face capital letters to denote matrices while bold face small letters to denote vectors.

II. OFDM Signal Model with Receiver Diversity

In OFDM systems with N subcarriers, N information symbols are used to construct one OFDM symbol. Specifically each of the N symbols is used to modulate a subcarrier and the N modulated subcarriers are added together to form an OFDM symbol. Orthogonality among subcarriers are achieved by carefully selecting carrier frequencies such that each OFDM symbol interval contains integer number of periods for all subcarriers. Using discrete-time baseband signal model, one of the most commonly used scheme is the IDFT-DFT (inverse discrete-time Fourier transform — discrete-time Fourier transform) based OFDM system. Guard time, which is cyclically extended to maintain inter-carrier orthogonality in the presence of time-dispersive channel, is inserted that is assumed longer than the maximum delay spread of the channel to totally eliminate inter-block interference [8].

The discrete-time complex baseband OFDM signal is

$$s(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} d_k e^{j2\pi \frac{kn}{N}}$$

where each d_k is used to modulate the subcarrier $e^{j2\pi k/N}$.

Receiver diversity for OFDM systems can be achieved either by employing multiple receiver antennass or via oversampling [9]. In both cases, the discrete-time baseband received signals can be written as

$$\begin{aligned} x_1(n) &= \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} H_1(k)d_k e^{j\frac{2\pi kn}{N}} + v_1(n) \\ x_2(n) &= \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} H_2(k)d_k e^{j\frac{2\pi kn}{N}} + v_2(n) \end{aligned}$$

where $H_i(k)$ is the channel frequency response corresponding to i^{th} channel at subcarrier k and $v_1(n)$ and $v_2(n)$ are both additive white complex Gaussian noise and are uncorrelated with each other. The above signal model can be written in a compact matrix form as

$$\begin{aligned} \mathbf{x}_1 &= \mathbf{W}^H\mathbf{H}_1\mathbf{d} + \mathbf{v}_1 \\ \mathbf{x}_2 &= \mathbf{W}^H\mathbf{H}_2\mathbf{d} + \mathbf{v}_2 \end{aligned}$$

where $\mathbf{W}$ is the DFT matrix as in (1), $\mathbf{H}_i = diag(\mathbf{h}_i)$ with

$$\mathbf{h}_i = [H_i(0), \ldots, H_i(N-1)]^T \tag{3}$$

That is, $\mathbf{H}_i$ is a diagonal matrix with diagonal element $H_i(k)$; and $\mathbf{d} = [d_0, \ldots, d_{N-1}]^T$ is the symbol vector. Taking DFT at the receiver, we have the equivalent frequency domain observation

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{W}\mathbf{x}_1 = \mathbf{H}_1\mathbf{d} + \mathbf{z}_1 \\ \mathbf{y}_2 &= \mathbf{W}\mathbf{x}_2 = \mathbf{H}_2\mathbf{d} + \mathbf{z}_2 \end{aligned}$$

where $\mathbf{z}_1$ and $\mathbf{z}_2$ is statistically identical to $\mathbf{v}_1$ and $\mathbf{v}_2$ because of the unitary property of $\mathbf{W}$, i.e., $\mathbf{z}_1$ and $\mathbf{z}_2$ are both white complex Gaussian and are uncorrelated with each other. Blind channel estimation aims to retrieve both $\mathbf{H}_1$ and $\mathbf{H}_2$ without any knowledge about $\mathbf{d}$. Clearly, a direct approach to estimate the frequency response matrix $\mathbf{H}_i$ is not feasible — the number of unknowns ($6N$ from the three $N \times 1$ complex vectors $\mathbf{h}_1$, $\mathbf{h}_2$, and $\mathbf{d}$) exceeds the number of observations ($4N$ from the two observation vectors $\mathbf{y}_1$ and $\mathbf{y}_2$). However, we note that the actual degrees of freedom associated with $\mathbf{H}_1$ and $\mathbf{H}_2$ are far smaller than N, the OFDM symbol length. This is because the frequency response is simply the DFT of the channel impulse response which is usually assumed to be shorter than the length of the cyclic prefix hence is far smaller than N. With this observation, we can rewrite the signal model as

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{D}\mathbf{h}_1 + \mathbf{z}_1 = \mathbf{D}\mathbf{W}_L\mathbf{g}_1 + \mathbf{z}_1 \\ \mathbf{y}_2 &= \mathbf{D}\mathbf{h}_2 + \mathbf{z}_2 = \mathbf{D}\mathbf{W}_L\mathbf{g}_2 + \mathbf{z}_2 \end{aligned} \tag{4}$$

where $\mathbf{D} = diag(\mathbf{d})$, i.e., it is the diagonal matrix with diagonal element d_k, $\mathbf{h}_i$ is as in (3), and $\mathbf{g}_i$ is the impulse response for the i^{th} channel and is of length L. Here the total number of unkown is $2 \times (N + 2L)$ which is smaller than $4N$, the number of observations. Again, we note here that in this paper we deal exclusively with the case of known channel length L.

III. Blind Channel Estimation

A Noiseless Case

Given the signal model in (4), we first consider the channel estimation in the noiseless case. We emphasize again that by converting the channel estimation from frequency domain to time domain, we have reduced the degree of freedom [10]. Using the above model, we now devise a simple algorithm that can perfectly retrieve the time domain channel in the absence of noise.

Without channel noise, (4) can be written as, in an element by element fashion,

$$\begin{aligned} y_1(k) = d_k \cdot \mathbf{u}_k^H \mathbf{g}_1 \\ y_2(k) = d_k \cdot \mathbf{u}_k^H \mathbf{g}_2 \end{aligned}$$

where $\mathbf{u}_k^H$ is given in (2). Therefore, for $d_k \neq 0$,

$$y_1(k)\mathbf{u}_k^H\mathbf{g}_2 = y_2(k)\mathbf{u}_k^H\mathbf{g}_1$$

The matrix form of the above equation is

$$\mathbf{Y}_1\mathbf{W}_L\mathbf{g}_2 = \mathbf{Y}_2\mathbf{W}_L\mathbf{g}_1$$

where $\mathbf{Y}_1 = diag(\mathbf{y}_1)$ and $\mathbf{Y}_2 = diag(\mathbf{y}_2)$. Equivalently, we have

$$[\mathbf{Y}_2\mathbf{W}_L | -\mathbf{Y}_1\mathbf{W}_L] \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix} = 0 \tag{5}$$

Therefore in the noiseless case, the channel can be retrieved up to a scalar ambiguity by simply finding a solution for the above homogeneous equation. We will address the uniqueness of solution to the above equation (a.k.a., identifiability issue) in the next section. We first discuss below the implementation of the algorithm in the noisy case.

B Noisy Case

In the presence of channel noise, it is clear that equation (5) will not hold. Instead of finding the exact solution we may instead find the right singular vector corresponding to the smallest singular value of the matrix

$$\mathbf{V} = [\mathbf{Y}_2\mathbf{W}_L | -\mathbf{Y}_1\mathbf{W}_L]\,. \tag{6}$$

Equivalently, we may seek to minimize the quadratic form:

$$\min_{\mathbf{g}} \mathbf{g}^H\mathbf{U}\mathbf{g} = \mathbf{g}^H\mathbf{V}^H\mathbf{V}\mathbf{g}$$

where $\mathbf{g} = [\mathbf{g}_1, \mathbf{g}_2]^T$ and $\mathbf{U} = \mathbf{V}^H\mathbf{V}$ is an $2L \times 2L$ matrix and is Hermitian and positive definite. This minimization can be achieved by simply finding the eigenvector corresponding to the smallest eigenvalue of $\mathbf{U}$.

C Multiple OFDM Data Blocks

Most subspace methods require multiple data blocks for them to work. The proposed method works with a single data block yet it can be easily extended to multiple data blocks for enhanced performance.

Two heuristic approaches can be adopted. The first one is to do a channel estimation using each block and then average over the data blocks to smooth out the error. Another approach is to first average over the $\mathbf{V}$ matrix for each block (or equivalently, average over the observations $\mathbf{y}_1$ and $\mathbf{y}_2$), i.e., calculate the $\mathbf{V}$ matrix for each block and then use the average in (6). The second approach is advantageous from the computational point of view — it involves only a single singular value decomposition (SVD) no matter how many data blocks are used. Performance wise, we also find that the latter approach yields much smaller error. This can be explained as following. Averaging over $\mathbf{V}$ allows the smoothing before the SVD, which tends to better smooth out the random channel noise than averaging after SVD.

IV. Identifiability

In this section, we discuss the identifiability issues based on system model (4). The channels are said to be identifiable if in the absence of noise, there is a unique solution (up to a scalar ambiguity) that satisfies the signal model (4). In particular, we propose a sufficient **and** a necessary condition (though *not sufficient and necessary* condition) for channel identifiability using receiver diversity.

Theorem 1 *(sufficient condition) The channel impulse responses $\mathbf{g}_1$ and $\mathbf{g}_2$ can be identified up to a scalar factor if the following conditions hold:*

1. *$H_1(z)$ and $H_2(z)$ do not share common zeros.*
2. *$N \geq 2L-1$*

Proof: In the noiseless case, model (4) becomes

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{D}\mathbf{W}_L\mathbf{g}_1 \\ \mathbf{y}_2 &= \mathbf{D}\mathbf{W}_L\mathbf{g}_2 \end{aligned}$$

Using the notation $\mathbf{u}_k^H$ as in (2), we have

$$\begin{aligned} y_1(k) &= d_k \cdot \mathbf{u}_k^H\mathbf{g}_1 \\ y_2(k) &= d_k \cdot \mathbf{u}_k^H\mathbf{g}_2 \end{aligned}$$

Assume we have another set of channel responses $\tilde{\mathbf{g}}_1$ and $\tilde{\mathbf{g}}_2$ that also satisfy the system model, then

$$\begin{aligned} y_1(k) &= \tilde{d}_k \cdot \mathbf{u}_k^H\tilde{\mathbf{g}}_1 \\ y_2(k) &= \tilde{d}_k \cdot \mathbf{u}_k^H\tilde{\mathbf{g}}_2 \end{aligned}$$

Clearly

$$\begin{aligned} d_k \cdot \mathbf{u}_k^H\mathbf{g}_1 &= \tilde{d}_k \cdot \mathbf{u}_k^H\tilde{\mathbf{g}}_1 \\ d_k \cdot \mathbf{u}_k^H\mathbf{g}_2 &= \tilde{d}_k \cdot \mathbf{u}_k^H\tilde{\mathbf{g}}_2 \end{aligned}$$

From this we get, through cross multiplication,

$$d_k\tilde{d}_k\left(\mathbf{u}_k^H\mathbf{g}_1\right)\left(\mathbf{u}_k^H\tilde{\mathbf{g}}_2\right) = d_k\tilde{d}_k\left(\mathbf{u}_k^H\mathbf{g}_2\right)\left(\mathbf{u}_k^H\tilde{\mathbf{g}}_1\right)$$

which is equivalent to

$$H_1(k)\tilde{H}_2(k) = \tilde{H}_1(k)H_2(k)$$

for $k = 0, \ldots, N-1$. Notice that $H_i(k)$ and $\tilde{H}_i(k)$ are respectively N point DFT at frequency $2\pi k/N$ for impulse response $\mathbf{g}_i$ and $\tilde{\mathbf{g}}_i$. Correspondingly, we have in the time domain the following identity for N point circular convolution[2]:

$$\mathbf{g}_1 \otimes \tilde{\mathbf{g}}_2 = \tilde{\mathbf{g}}_1 \otimes \mathbf{g}_2$$

Given that $\mathbf{g}_1, \mathbf{g}_2, \tilde{\mathbf{g}}_1, \tilde{\mathbf{g}}_2$ are all vectors of length L, if $N \geq 2L-1$, then N point circular convolution is equivalent to linear convolution. Therefore

$$\mathbf{g}_1 * \tilde{\mathbf{g}}_2 = \tilde{\mathbf{g}}_1 * \mathbf{g}_2$$

or

$$\mathbf{G}_1(z)\tilde{\mathbf{G}}_2(z) = \tilde{\mathbf{G}}_1(z)\mathbf{G}_2(z) \tag{7}$$

Given (7), it is shown in [11] that the channel can be identified up to a scalar factor if $\mathbf{G}_1(z)$ and $\mathbf{G}_2(z)$ do not share any common nulls. Therefore we must have

$$\begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix} = \alpha \begin{bmatrix} \tilde{\mathbf{g}}_1 \\ \tilde{\mathbf{g}}_2 \end{bmatrix}$$

Q.E.D.

[2] We denote circular convolution by $\otimes$ and linear convolution by $*$

Theorem 2 *(necessary condition) If the channel impulse response $\mathbf{g}_1$ and $\mathbf{g}_2$ are identifiable up to a scalar factor, then $N \geq 2L-1$.*

Proof: If the system is identifiable, there will be a unique (up to a scalar ambiguity) solution $\mathbf{g}_1$ and $\mathbf{g}_2$ for (5). Therefore the rank of $\mathbf{V}$ must be $2L-1$, i.e., its null space must have dimension equal to 1. Since $\mathbf{V}$ is an N by $2L$ matrix, we must have $N \geq 2L-1$.

Q.E.D.

V. Cramer-Rao Lower Bound for Channel Estimation

In this section we evaluate the performance of the blind estimation algorithm by deriving the CRLB. The unknown parameter vector is

$$\theta = [Re(\mathbf{g}_1), Re(\mathbf{g}_2), Re(\mathbf{d}), Im(\mathbf{g}_1), Im(\mathbf{g}_2), Im(\mathbf{d})]^T$$

Based on (4), and given that $\mathbf{z}_1$ and $\mathbf{z}_1$ are uncorrelated with each other, the negative log likelihood function can be obtained as, after discarding any irrelevant constant,

$$\begin{aligned} -\ln\Lambda = {} & (\mathbf{y}_1 - \mathbf{D}\mathbf{W}_L\mathbf{g}_1)^H(\mathbf{y}_1 - \mathbf{D}\mathbf{W}_L\mathbf{g}_1) + \\ & (\mathbf{y}_2 - \mathbf{D}\mathbf{W}_L\mathbf{g}_2)^H(\mathbf{y}_2 - \mathbf{D}\mathbf{W}_L\mathbf{g}_2) \end{aligned}$$

From this, the Fisher Information Matrix (FIM) can be derived as

$$\mathbf{F} = 2\begin{bmatrix} Re(\mathbf{F_c}) & -Im(\mathbf{F_c}) \\ Im(\mathbf{F_c}) & Re(\mathbf{F_c}) \end{bmatrix} \tag{8}$$

where

$$\mathbf{F}_c = \frac{1}{\sigma^2}\begin{bmatrix} \mathbf{Q}^H\mathbf{Q} & 0 & \mathbf{Q}^H\mathbf{H}_1 \\ 0 & \mathbf{Q}^H\mathbf{Q} & \mathbf{Q}^H\mathbf{H}_2 \\ \mathbf{H}_1^H\mathbf{Q} & \mathbf{H}_2^H\mathbf{Q} & \mathbf{H}_1^H\mathbf{H}_1 + \mathbf{H}_2^H\mathbf{H}_2 \end{bmatrix},$$

and σ^2 is the noise power and $\mathbf{Q} = \mathbf{D}\mathbf{W}_L$. A detailed derivation of the FIM is provided in the appendix. Note that matrix $\mathbf{F_c}$ is at least rank 1 deficient due to the scalar ambiguity of the channel. To evaluate the CRLB channel estimator, we often consider one element of the channel (e.g., the first element of $\mathbf{g_1}$) as known. After deleting the column and row associated with the known parameter, the remaining matrix will be full rank and CRLB can be evaluated by taking the inverse of that matrix.

It is interesting to consider the situation when $\mathbf{g}_1$ and $\mathbf{g}_2$ share a common zero at a subcarrier frequency, say k_0. In this case matrix $\mathbf{F_c}$ will have an all zero row and column at the corresponding input symbol location, i.e., the row and column corresponding to d_{k_0}. Therefore even if we assume $g_1(1)$ is known, hence its corresponding row and column is deleted from the FIM, the remaining FIM is still not full rank. One explanation could be that because of the common zero at a subcarrier frequency, the corresponding symbol d_{k_0} is clearly not identifiable. However, it is found numerically that after getting rid of the row and column corresponding to d_{k_0}, the remaining FIM is still rank deficient — which implies that the channel itself may not be identifiable. This observation suggests that a possible necessary condition for channel identifiability is that $\mathbf{g}_1$ and $\mathbf{g}_2$ do not share common zeros at subcarrier frequencies. Notice this condition is weaker than that stated in the sufficient condition where $\mathbf{g}_1$ and $\mathbf{g}_2$ do not share any common zeros without regard to their possible locations.

VI. Simulations

We provide some numerical examples in this section to evaluate the performance of the proposed method and we compare this method to the subspace method in [2]. We use normalized root mean square error (NRMSE) as the performance criterion:

$$NRMSE = \frac{1}{\|\mathbf{g}\|}\sqrt{\frac{1}{M_c L}\sum_{m=1}^{M_c}\|\hat{\mathbf{g}}-\mathbf{g}\|^2}$$

where M_c is the the number of Monte Carlo runs, L is the channel length, $\hat{\mathbf{g}}$ is the channel estimate, and $\mathbf{g}$ is the true channel. We used $N = 16$, $L = 5$, $M_c = 1000$ and 16 QAM modulation scheme which are the same as the ones in [2]. The channel impulse responses are

$$\begin{aligned}\mathbf{g}_1 &= [-.40-.17i, .11+.06i, -.10+.12i, .66-.50i, -.24+.16i]^T\\ \mathbf{g}_2 &= [-.16-.10i, .52-.10i, .14+.01i, .50+.57i, -.25+.14i]^T\end{aligned} \quad (9)$$

Since the subspace method in [2] does not require channel diversity, we evaluate the performance of channel estimation for $\mathbf{g}_1$, i.e., $\mathbf{g} = \mathbf{g}_1$ in the NRMSE expression. Because of the scalar ambiguity, we set $\hat{g}(1) = g(1)$ in calculating the NRMSE. The results are plotted in Figure 1.

In the simulation, the subspace method uses 60 blocks while the diversity method uses only 30 blocks. Clearly, the diversity based method performs better than the subspace method. We also note here that the gain also depends on the second channel impulse response $\mathbf{g}_2$, though we do find through extensive simulation that in almost all cases, the diversity based channel estimation provides substantial performance gain over the subspace method provided no common zeros exist for the two channels.

In Figure 2, we plot the NRMSE for a situation where both $\mathbf{g}_1$ and $\mathbf{g}_2$ have channel nulls at some subcarriers, though they do not share common nulls. The channel impulse responses are chosen as

$$\begin{aligned}\mathbf{g}_1 &= [.47+.21i, -.28+.18i, .03+.10i, .77+.05i, -.02-.08i]^T\\ \mathbf{g}_2 &= [.38+.15i, .30-.36i, .03+.22i, .67-.05i, .13-.30i]^T\end{aligned} \quad (10)$$

It is easy to verify that $\mathbf{g}_1$ has nulls at the 3rd and 9th subcarrier, while $\mathbf{g}_2$ has nulls at the 4th and 7th subcarrier. Simulation results show that channels nulls do not affect the performance of the diversity based estimator, as long as no identical nulls exist for both channels. Indeed, if the two channels have common nulls at a subcarrier, the performance of the diversity method is very poor. This supports the conjecture that no common zero at subcarrier frequency is also a necessary condition for identifiability.

Finally, we plot the mean square error (MSE) of channel estimation using diversity method along with the CRLB. Parameter settings are the same as in (9). We assume the first element of $\mathbf{g}_1$ is known therefore the first row and column of the matrix in $\mathbf{F}_c$ is deleted and the corresponding CRLB is numerically evaluated by taking the inverse of the remaining FIM. We compare the MSE and CRLB for $\mathbf{g}_2$ and the MSE is obtained as usual:

$$MSE = \frac{1}{M_c}\sum_{m=1}^{M_c}\|\hat{\mathbf{g}_2}-\mathbf{g}_2\|^2$$

We use only a single OFDM block in this scenario. The results are given in Figure 3. It can be seen that the MSE of the proposed method is fairly close to the CRLB.

VII. Conclusions

In this paper, bandwidth efficient channel identification algorithm utilizing receiver diversity is proposed. In the noiseless case, the algorithm can perfectly retrieve the channels up to a scalar factor. In the presence of noise, the algorithm has very low complexity — only a single SVD is needed independent of the number of OFDM blocks used. Some identifiability results are obtained and the effect of channel nulls on subcarriers is also discussed. Unlike many existing methods, this new approach is data efficient — it can be implemented using a single data block, although more data will certainly enhance the estimation performance. Furthermore, the proposed algorithm imposes no restriction on the input symbol constellation. Cramer-Rao Lower Bound of the channel estimation based on the diversity model is also derived. Simulation is conducted to verify its performance advantage over an existing blind algorithm.

A. Derivation of the Fisher Information Matrix

Consider the signal model as in (4), the unknown parameter vector is

$$\boldsymbol{\theta} = [Re(\mathbf{g}_1), Re(\mathbf{g}_2), Re(\mathbf{d}), Im(\mathbf{g}_1), Im(\mathbf{g}_2), Im(\mathbf{d})]^T$$

Apparently the FIM, denoted by $\mathbf{F}$, is of dimension $2N + 4L$ by $2N + 4L$. Define

$$\boldsymbol{\mu} = \begin{bmatrix}\mathbf{D}\mathbf{W}_L\mathbf{g}_1\\ \mathbf{D}\mathbf{W}_L\mathbf{g}_2\end{bmatrix}$$

to be the mean value of the observation vector $[\mathbf{y}_1, \mathbf{y}_2]^T$ that is otherwise Gaussian distributed. Each element of FIM can be written as, given that the noise covariance matrix is $\sigma^2\mathbf{I}$,

$$\mathbf{F}(i,j) = \frac{2}{\sigma^2}Re\left[\left(\frac{\partial\boldsymbol{\mu}}{\partial\theta_i}\right)^H\left(\frac{\partial\boldsymbol{\mu}}{\partial\theta_j}\right)\right]$$

Define $\tilde{\boldsymbol{\theta}} = [\mathbf{g}_1, \mathbf{g}_2, \mathbf{d}]^T$. In matrix form, $\mathbf{F}$ can be written as [12, 13]

$$\mathbf{F} = 2\begin{bmatrix} Re(\mathbf{F}_c) & -Im(\mathbf{F}_c)\\ Im(\mathbf{F}_c) & Re(\mathbf{F}_c)\end{bmatrix}$$

where each element of $\mathbf{F_c}$ is

$$\mathbf{F_c}(i,j) = \frac{1}{\sigma^2}\left[\left(\frac{\partial\boldsymbol{\mu}}{\partial\tilde{\theta}_i}\right)^H\left(\frac{\partial\boldsymbol{\mu}}{\partial\tilde{\theta}_j}\right)\right]$$

Write $\mathbf{F_c}$ in partitioned matrix form

$$\mathbf{F_c} = \frac{1}{\sigma^2}\begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13}\\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23}\\ \mathbf{A}_{31} & \mathbf{A}_{32} & \mathbf{A}_{33}\end{bmatrix}$$

Let $\mathbf{Q} = \mathbf{D}\mathbf{W}_L$, we can obtain each block of the $\mathbf{F}_c$ matrix as following

$$\begin{aligned}
\mathbf{A}_{11} &= \frac{\partial\boldsymbol{\mu}^H}{\partial\mathbf{g}_1}\frac{\partial\boldsymbol{\mu}}{\partial\mathbf{g}_1^H} = [\mathbf{W}_L^H\mathbf{D}^H \;\; 0][\mathbf{D}\mathbf{W}_L \;\; 0]^T = \mathbf{Q}^H\mathbf{Q}\\
\mathbf{A}_{12} &= \frac{\partial\boldsymbol{\mu}^H}{\partial\mathbf{g}_1}\frac{\partial\boldsymbol{\mu}}{\partial\mathbf{g}_2^H} = [\mathbf{W}_L^H\mathbf{D}^H \;\; 0][0 \;\; \mathbf{D}\mathbf{W}_L]^T = 0\\
\mathbf{A}_{13} &= \frac{\partial\boldsymbol{\mu}^H}{\partial\mathbf{g}_1}\frac{\partial\boldsymbol{\mu}}{\partial\mathbf{d}^H} = [\mathbf{W}_L^H\mathbf{D}^H \;\; 0][\mathbf{H}_1 \;\; 0]^T = \mathbf{Q}^H\mathbf{H}_1\\
\mathbf{A}_{21} &= \frac{\partial\boldsymbol{\mu}^H}{\partial\mathbf{g}_2}\frac{\partial\boldsymbol{\mu}}{\partial\mathbf{g}_1^H} = [0 \;\; \mathbf{W}_L^H\mathbf{D}^H][\mathbf{D}\mathbf{W}_L \;\; 0]^T = 0
\end{aligned}$$

$$
\begin{aligned}
\mathbf{A}_{22} &= \frac{\partial \boldsymbol{\mu}^H}{\partial \mathbf{g}_2}\frac{\partial \boldsymbol{\mu}}{\partial \mathbf{g}_2^H} = [0 \;\; \mathbf{W}_L^H \mathbf{D}^H][0 \;\; \mathbf{D}\mathbf{W}_L]^T = \mathbf{Q}^H\mathbf{Q} \\
\mathbf{A}_{23} &= \frac{\partial \boldsymbol{\mu}^H}{\partial \mathbf{g}_2}\frac{\partial \boldsymbol{\mu}}{\partial \mathbf{d}^H} = [0 \;\; \mathbf{W}_L^H \mathbf{D}^H][0 \;\; \mathbf{H}_2]^T = \mathbf{Q}^H\mathbf{H}_2 \\
\mathbf{A}_{31} &= \frac{\partial \boldsymbol{\mu}^H}{\partial \mathbf{d}}\frac{\partial \boldsymbol{\mu}}{\partial \mathbf{g}_1^H} = [\mathbf{H}_1^H \;\; \mathbf{H}_2^H][\mathbf{D}\mathbf{W}_L \;\; 0]^T = \mathbf{H}_1^H\mathbf{Q} \\
\mathbf{A}_{32} &= \frac{\partial \boldsymbol{\mu}^H}{\partial \mathbf{d}}\frac{\partial \boldsymbol{\mu}}{\partial \mathbf{g}_2^H} = [\mathbf{H}_1^H \;\; \mathbf{H}_2^H][0 \;\; \mathbf{D}\mathbf{W}_L]^T = \mathbf{H}_2^H\mathbf{Q} \\
\mathbf{A}_{33} &= \frac{\partial \boldsymbol{\mu}^H}{\partial \mathbf{d}}\frac{\partial \boldsymbol{\mu}}{\partial \mathbf{d}^H} = [\mathbf{H}_1^H \;\; \mathbf{H}_2^H][\mathbf{H}_1 \;\; \mathbf{H}_2]^T = \mathbf{H}_1^H\mathbf{H}_1 + \mathbf{H}_2^H\mathbf{H}_2
\end{aligned}
$$

Finally

$$
\mathbf{F}_c = \frac{1}{\sigma^2}\begin{bmatrix} \mathbf{Q}^H\mathbf{Q} & 0 & \mathbf{Q}^H\mathbf{H}_1 \\ 0 & \mathbf{Q}^H\mathbf{Q} & \mathbf{Q}^H\mathbf{H}_2 \\ \mathbf{H}_1^H\mathbf{Q} & \mathbf{H}_2^H\mathbf{Q} & \mathbf{H}_1^H\mathbf{H}_1 + \mathbf{H}_2^H\mathbf{H}_2 \end{bmatrix}
$$

References

[1] B. Muquet and M. de Courville, "Blind and semi-blind channel identification methods using second order statistics for OFDM systems," in *Proc. International Conference on Acoustic Speech and Signal Processing*, Phoenix, AZ, March 1999, vol. 5, pp. 2745–2748.

[2] X. Cai and A.N. Akansu, "A subspace method for blind channel identification in OFDM systems," in *Proc. ICC'2000*, New Brunswick, New Jersey, March 2000, vol. 2, pp. 929–933.

[3] X. Zhuang, Z. Ding, and A.L. Swindlehurst, "A statistical subspace method for blind channel identification in OFDM communications," in *Proc. 2000 International Conference on Acoustic Speeach and Signal Processing*, Istanbul, Turkey, June 2000, vol. 5, pp. 2493–2496.

[4] C. Li and S. Roy, "Subspace based blind channel estimation for OFDM by exploiting virtual carrier," in *Proc. CLOBECOM'01*, San Antonio, TX, Nov 2001, vol. 1, pp. 295–299.

[5] S. Zhou and G.B. Giannakis, "Finite-alphabet based channel estimation for OFDM and related multicarrier systems," *IEEE Trans. Communications*, pp. 1402–1414, August 2001.

[6] H. Ali, J.H. Manton, and Y. Hua, "A SOS subspace method for blind channel identification and equalization in bandwidth efficient OFDM systems based on receive antenna diversity," in *Proc. 11th IEEE Signal Processing Workshop on Statistical Signal Processing*, Singapore, August 2001, pp. 401–404.

[7] C. Li and S. Roy, "A subspace blind channel estimation method for OFDM systems without cyclic prefix," in *Proc. VTC'01 Fall*, Atlantic City, NJ, Oct 2001, vol. 4, pp. 2148–2152.

[8] R. van Nee and R. Prasad, *OFDM For Multimedia Wireless Communications*, Artech House, Boston, MA, 2000.

[9] L. Tong and S. Perreau, "Multichannel blind channel estimation: From subspace to maximum likelihood methods," IEEE Proceedings, 1998, pp. 1951–1968, Oct 1998.

[10] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system," *IEEE Trans. Consumer Electronics*, vol. 44, no. 3, pp. 1122–1128, August 1998.

[11] G. Xu, H. Liu, L. Tong, and T. Kailath, "A Least-Squares Approach to Blind Channel Identification," *IEEE Trans. Signal Processing*, vol. SP-43, no. 12, pp. 2982–2993, December 1995.

[12] P. Stoica and R. Moses, *Introduction to Spectral Analysis*, Prentice Hall, Upper Saddle River, NJ, 1997.

[13] Y. Hua, "Fast maximum likelihood for blind identification of multiple FIR channels," *IEEE Trans. Signal Processing*, vol. SP-44, no. 3, pp. 661–672, March 1996.

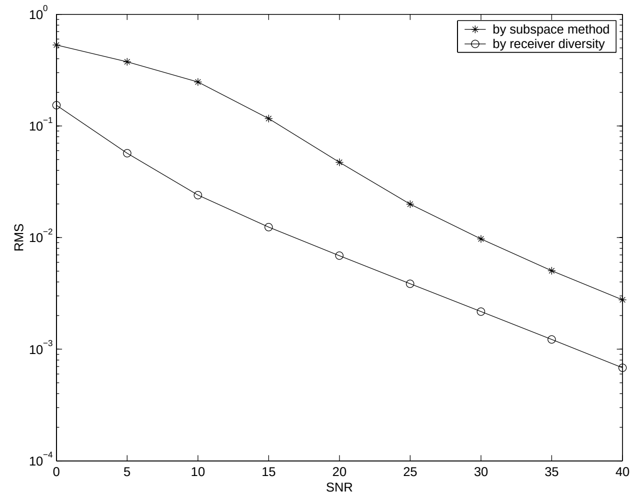

Figure 1: Normalized RMSE for the blind OFDM channel estimation using the diversity scheme and the subspace scheme in [2]. The channel impulse responses are specificed in (9) where the NRMSE for channel $\mathbf{g}_1$ is used for comparison.

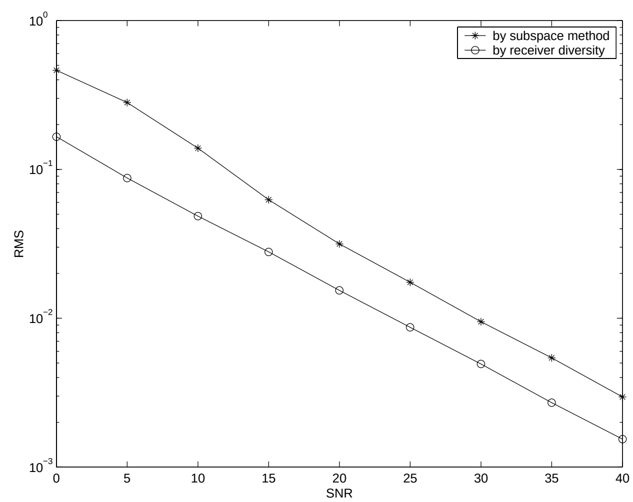

Figure 2: Normalized RMSE for the blind OFDM channel estimation using the diversity scheme and the subspace scheme in [2]. The channel impulse responses are specificed in (10) where the NRMSE for channel $\mathbf{g}_1$ is used for comparison. In this example there are certain channel nulls for both $\mathbf{g}_1$ and $\mathbf{g}_2$ at some subcarrier frequencies.

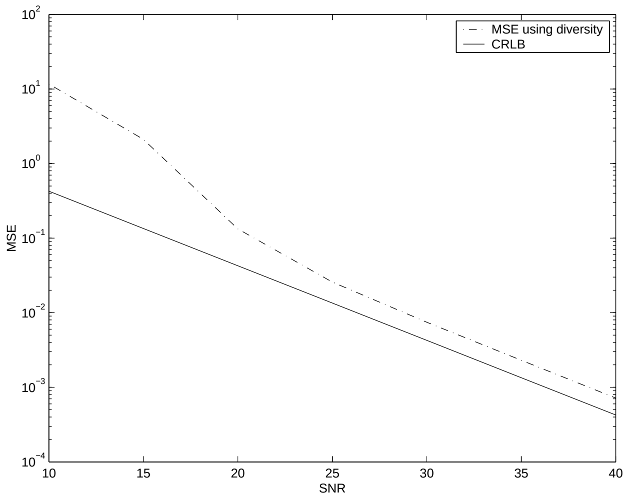

Figure 3: MSE for the blind OFDM channel estimation using diversity scheme and the corresponding CRLB. The channel impulse responses are specified in (9) and the MSE and CRLB are computed for $\mathbf{g}_2$ under that the assumption that $g_1(1)$ is known.