

Research Article

Fisher-Information-Matrix Based Analysis of Semiblind MIMO Frequency Selective Channel Estimation

Aditya K. Jagannatham¹ and Bhaskar D. Rao²

¹Department of Electrical Engineering, I.I.T Kanpur, Kanpur 208016, India

²Department of Electrical and Computer Engineering, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0407, USA

Correspondence should be addressed to Aditya K. Jagannatham, adityaj@iitk.ac.in

Received 8 May 2011; Accepted 21 June 2011

Academic Editors: Y. H. Ha, C.-C. Hu, A. Ito, A. Maier, A. Wong, and B. Yuan

Copyright © 2011 A. K. Jagannatham and B. D. Rao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We present a study of semiblind (SB) estimation for a frequency-selective (FS) multiple-input multiple-output (MIMO) wireless channel using a novel Fisher-information matrix (FIM) based approach. The frequency selective MIMO system is modeled as a matrix finite impulse response (FIR) channel, and the transmitted data symbols comprise of a sequence of *pilot* symbols followed by the unknown *blind* symbols. It is demonstrated that the FIM for this system can be expressed as the sum of the blind FIM J^b and pilot FIM J^p . We present a key result relating the rank of the FIM to the number of blindly identifiable parameters. We then present a novel maximum-likelihood (ML) scheme for the semiblind estimation of the MIMO FIR channel. We derive the Cramer-Rao Bound (CRB) for the semiblind scheme. It is observed that the semi-blind MSE of estimation of the MIMO FIR channel is potentially much lower compared to an exclusively pilot-based scheme. Finally, we derive a lower bound for the minimum number of pilot symbols necessary for the estimation of an FIR MIMO channel for any general semi-blind scheme. Simulation results are presented to augment the above analysis.

1. Introduction

Multiple-input multiple-output (MIMO) communication systems [1] have gained widespread popularity as technology solutions for current and future wireless systems such as 3G/4G, WBA (wireless broadband access), and 802.11n for high-speed WLAN (wireless local area network) applications amongst others. The detection performance of the designed MIMO receivers depends on the accuracy of the available channel estimate. Accurate estimates are also valuable for feedback-based schemes such as water-filling power allocation to enhance capacity [2] or design optimum precoders for MIMO transmission [3]. Channel estimation thus plays a key role in the performance gains achievable through deployment of MIMO systems.

In current signal processing research, there are two dominant and widely prevalent approaches for channel estimation. The first one is termed as pilot-based estimation [4, 5] and employs exclusively pilot symbols. These schemes

have a low complexity of implementation and are robust in nature which makes them amenable for implementation in wireless systems. A downside to the employment of pilot-based schemes is that they cause a significant bandwidth overhead since pilot symbols convey no information. This overhead is higher in MIMO systems where the number of channel parameters to be estimated grows as the product of the number of receive and transmit antennas, requiring the transmission of an increasing number of pilot symbols, thus reducing the effective spectral efficiency (defined as the ratio of information bearing bits to total bits per unit bandwidth). The alternative to pilot-based estimation, is blind estimation [6, 7] which employs exclusively the statistical information available about the input symbols and channel. Since, in principle, it employs no training symbols, blind schemes have maximum effective spectral efficiency (ratio of information to total bits equals unity). However, blind schemes are computationally complex and typically identify a channel only up to a scaling, phase or permutation

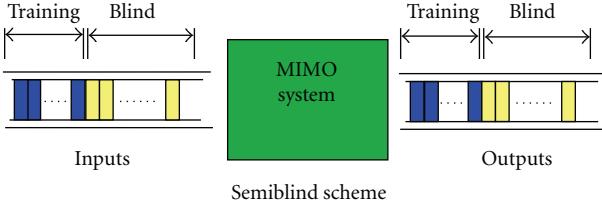


FIGURE 1: Schematic representation of an SB system.

indeterminacy. To avoid the shortcomings of the above schemes, semi-blind (SB) techniques [8–10], which utilize both pilot and blind symbols (Figure 1), have gained popularity. These schemes utilize the statistical information available in the data symbols, thus enhancing the accuracy of the channel estimate, while employing a few pilot symbols to greatly reduce overhead and improve robustness. For instance, in the context of a SIMO wireless channel, it has been shown that with just one pilot symbol placed anywhere in the transmitted data block, the channel can be estimated without any residual indeterminacy [11]. Other interesting semi-blind schemes can be found in works such as [12, 13].

The focus of this work is to present a new semi-blind MIMO estimation scheme and quantify the improvements in bandwidth efficiency and robustness of SB estimation schemes in general. For this purpose, it is necessary to address multiple aspects of SB estimation. The problem of identifiability from blind symbols forms a cornerstone of SB estimation, since SB schemes rely on estimating a significant component of the wireless channel from the blind symbols. Thus, by leveraging the blind statistical information to the greatest extent, they reduce the dependence on pilots or, in other words, achieve a greater accuracy of estimation from the limited pilot data. Therefore, we present a novel approach based on the Fisher information matrix (FIM) to characterize the blind identifiability of the MIMO channel. In a key result presented in this work, we demonstrate that the number of blindly identifiable parameters of the MIMO FIR system is equal to the rank of the FIM. This observation has been employed intuitively in literature without formal justification, such as in [14] where the authors make this observation in the context of a SIMO channel. Identifiability results in the context of SIMO and MIMO channels have been presented in works such as [11, 15], respectively. Such results in the context of systems employing space-time block codes (STBC) can be found in works such as [16, 17]. Our general FIM result can be employed as a unifying framework in such a wide variety of scenarios to characterize the identifiability of the underlying system.

Further, most works on identifiability such as [18] and others consider the transmitted blind symbols to be deterministically unknown in nature, due to the relative ease of analysis of such models. However, in our analysis, we assume that the blind symbols are stochastic in nature for the following reason. In the context of direction of arrival (DOA) estimation, it has been demonstrated in [19] that if C_s , C_d denote the Cramer-Rao bounds (CRB) on the covariance matrices of the stochastic and deterministic estimation schemes, then

$C_d \geq C_s$. In other words, the stochastic model is statistically more efficient and has a lower mean square error (MSE) of estimation than its deterministic counterpart since the stochastic signal model has the advantage that the number of unknown parameters in the system no longer grows with the number of transmitted data symbols. Hence, our estimation schemes which are based on stochastic rather than deterministic blind symbols achieve a lower MSE of estimation by exploiting the statistical information available in the blind symbols.

For the context of MIMO frequency selective channel estimation, consider an $r \times t$ MIMO system (r = no. of receive antennas, t = no. of transmit antennas) with L_h channel taps. This system has $2rtL_h$ real parameters (i.e., rtL_h complex parameters) to be estimated. The FIM-based approach can be used to demonstrate the result that the above MIMO FIR system has t^2 parameters that cannot be identified from exclusively employing blind data. This is a central result for the estimation of MIMO FIR matrices and is derived from the MIMO identifiability results in [15, 20], wherein it has been shown that these t^2 blindly unidentifiable parameters correspond to a $t \times t$ unitary matrix indeterminacy. Motivated by this observation, we present an SB estimation scheme for a FIR MIMO channel, which achieves the dual objective of robustness and spectral efficiency. This scheme employs a whitening-rotation-like decomposition of the FIR MIMO channel, where the whitening matrix is frequency selective. We then utilize the work in [7] where Tugnait and Huang have elaborated a blind algorithm based on linear prediction for the estimation of the frequency-selective whitening matrix of the MIMO frequency selective channel. Utilizing this scheme, in conjunction with an orthogonal pilot sequence-based maximum-likelihood (ML) scheme for the estimation of the unitary matrix indeterminacy [21], we describe a scheme for the semi-blind estimation of a MIMO FIR channel. Thus, we present a novel procedure for SB estimation of the MIMO FIR channel.

The final concern in this endeavor of SB channel estimation is to quantify the estimation accuracy of such schemes for a given number of pilot symbols. Once again, using the FIM-based approach above, a concise result can be found for the MSE of estimation of the SB scheme. We demonstrate that asymptotically, as the block length of data symbols becomes very large, the MSE of SB estimation is directly proportional to the number of blindly unidentifiable parameters or t^2 . Hence, while for exclusively pilot-based estimation the MSE grows proportional to the total number of parameters $2rtL_h$, the MSE bound for SB estimation is only related to a much smaller set of t^2 blindly unidentifiable parameters. Thus, in summary, by exploiting the available blind symbols to the greatest possible extent, that is, by estimating all but a small fraction t^2 of the total $2rtL_h$ parameters from the blind symbols, one can achieve a much lower MSE performance compared to exclusively pilot-based schemes. Finally, we consider the issue of minimum number of pilot symbols necessary for the estimation of the FIR MIMO channel. Indeed, the answer to this question is inherently related to the question of *identifiability* of the MIMO FIR channel, and, using the FIM approach, we

demonstrate the key result that at least t pilot symbols are necessary for the estimation of the MIMO FIR channel.

The rest of the paper is organized as follows. The MIMO FIR estimation problem is formulated in Section 2. In Section 3, we present results on the SB-FIM followed by Sections 4 and 5 which describe the semi-blind scheme for channel estimation and MSE analysis, respectively. Simulation results are presented in Section 7 followed by conclusions in Section 8. In what follows, $i \in \bar{m}, \bar{n}$ represents $m \leq i \leq n$; $i, m, n \in \mathbb{N}$, where \mathbb{N} denotes the set of natural numbers, $\text{rank}(\cdot)$ is the rank of a matrix, and $\mathcal{N}(\cdot)$ represents the null space of a matrix.

2. Problem Formulation

Consider an L_h tap frequency-selective MIMO channel. The system input-output relation can be expressed as

$$\mathbf{y}(k) = \sum_{i=0}^{L_h-1} H(i)\mathbf{x}(k-i) + \eta(k), \quad (1)$$

where $\mathbf{y}(k), \mathbf{x}(k)$ are the k th received and transmitted symbol vectors, respectively. $\eta(k)$ is spatiotemporally white additive Gaussian noise of variance σ_n^2 , that is, $E\{\eta(k)\eta(l)^H\} = \sigma_n^2 \delta(k-l)\mathbf{I}_r$, where $\delta(k) = 1$ if $k = 0$ and 0 otherwise. Let t, r be the number of transmitters, receivers, and, therefore, $\mathbf{y}(k) \in \mathbb{C}^{r \times 1}$ and $\mathbf{x}(k) \in \mathbb{C}^{t \times 1}$. Each $H(i) \in \mathbb{C}^{r \times t}$, $i \in \overline{0, L_h-1}$ is the MIMO channel matrix corresponding to the i th lag. Also, we assume $r > t$, that is, the number of receivers is greater than the number of transmitters. Let $\{\mathbf{x}_p(1), \mathbf{x}_p(2), \dots, \mathbf{x}_p(L_p)\}$ be a burst of L_p transmitted pilot symbols. The subscript p in the above notation represents pilots. Let $\mathbf{H} \in \mathbb{C}^{r \times L_h t}$ be defined as

$$\begin{aligned} \mathbf{H} &\triangleq [H(0), H(1), \dots, H(L_h-1)], \\ H(i) &\triangleq [\mathbf{h}_1(i), \mathbf{h}_2(i), \dots, \mathbf{h}_t(i)]. \end{aligned} \quad (2)$$

Let the unknown blind information symbols (which yield only statistical information at the receiver) be stacked as $N > L_h$ transmitted symbol vectors in $\mathcal{X}(k)$ described by the system model given below as

$$\begin{aligned} \underbrace{\begin{bmatrix} \mathbf{y}(kN) \\ \mathbf{y}(kN-1) \\ \vdots \\ \mathbf{y}((k-1)N+L_h) \end{bmatrix}}_{\mathcal{Y}(k)} &= \mathcal{H} \begin{bmatrix} \mathbf{x}(kN) \\ \mathbf{x}(kN-1) \\ \vdots \\ \mathbf{x}((k-1)N+1) \end{bmatrix} \\ &+ \begin{bmatrix} \eta(kN) \\ \eta(kN-1) \\ \vdots \\ \eta((k-1)N+L_h) \end{bmatrix}, \end{aligned} \quad (3)$$

where the matrix $\mathcal{H} \in \mathbb{C}^{(N-L_h+1)r \times Nt}$ is the standard block Sylvester channel matrix often employed for the analysis of MIMO FIR channels [14] and is given as

$$\mathcal{H} \triangleq \begin{bmatrix} H(0) & H(1) & H(2) & \cdots & H(L_h-1) & 0 & \cdots \\ 0 & H(0) & H(1) & \cdots & H(L_h-2) & H(L_h-1) & \cdots \\ 0 & 0 & H(0) & \cdots & H(L_h-3) & H(L_h-2) & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (4)$$

The input vector $\mathcal{X}(k) \in \mathbb{C}^{Nt \times 1}$ is the data symbol block, where the length of each input block of data is N symbols long. Such a stacking of the input/output symbols into blocks results in loss of a small number of output information symbols ($L_h - 1$ symbols per block) due to interblock interference (IBI). This model is frequently employed in FIR system studies such as in [22] and is adapted because eliminating the IBI makes the analysis tractable by yielding simplistic-likelihood expressions. Let the transmitted data symbols $\mathbf{x}(k)$ be spatiotemporally white, that is, $E\{\mathbf{x}(k)\mathbf{x}(l)^H\} = \sigma_s^2 \delta(k-l)\mathbf{I}_t$ and the normalized source power $\sigma_s^2 \triangleq 1$. Hence, the covariance of the block input vector $\mathcal{X}(k)$ is given as $\mathcal{R}_{\mathcal{X}} \triangleq E\{\mathcal{X}(k)\mathcal{X}(k)^H\} = \mathbf{I}_{Nt}$. The MIMO transfer function of the FIR channel being defined as $H(z) = \sum_{i=0}^{L_h-1} H(i)z^{-i}$ is assumed to satisfy the following conditions.

- (A.1) $H(z)$ is *irreducible*, that is, $H(z)$ has full column rank for all $z \neq 0$ (including $z = \infty$). It follows that if $H(z)$ is irreducible, the leading coefficient matrix $[\mathbf{h}_1(0), \mathbf{h}_2(0), \dots, \mathbf{h}_t(0)]$ has full column rank (substitute $z = \infty$ in $H(z)$).
- (A.2) $H(z)$ is *column reduced*, that is, the trailing coefficient matrix $[\mathbf{h}_1(L_h-1), \mathbf{h}_2(L_h-1), \dots, \mathbf{h}_t(L_h-1)]$ has full column rank.

In practical applications, since the fading coefficients of the wireless channel matrices arise from the random scattering effects of the ambient propagation environment, the above assumptions are satisfied with very high probability. For a discussion about special scenarios where the above conditions are not satisfied, the reader is referred to works [7, 23]. Next, we present insights into the nature of the above estimation problem.

3. Semiblind Fisher Information Matrix (FIM)

In this section, we formally setup the complex FIM for the estimation of the channel matrix \mathbf{H} and provide insights into the nature of semi-blind estimation. The parameter vector to be estimated $\theta_{\mathcal{H}} \in \mathbb{C}^{2L_h rt \times 1}$ is defined by stacking the complex parameter vector and its conjugate as suggested in [21, 24] as

$$\theta_{\mathcal{H}} \triangleq \begin{bmatrix} \theta_{H(0)} \\ \theta_{H(1)} \\ \vdots \\ \theta_{H(L_h-1)} \end{bmatrix}, \quad \text{where } \theta_{H(i)} \triangleq \begin{bmatrix} \text{vec}(H(i)) \\ \text{vec}(H(i))^* \end{bmatrix} \in \mathbb{C}^{2rt \times 1}, \quad (5)$$

and $\text{vec}(\cdot)$ denotes the standard matrix vector operator which represents a column-wise stacking of the entries of

a matrix into a single column vector. In what follows, $k \in \overline{0, L_h - 1}$, $i \in \overline{1, rt}$. Observe also that $\theta_{\mathcal{H}}^*(2krt+i) = \theta_{\mathcal{H}}((2k+1)rt + i)$. Let L_b blocks of data symbols $\mathcal{X}(p)$, $p \in \overline{1, L_b}$ be transmitted. In addition, let the data symbol vectors $\mathbf{x}(l)$, $l \in \overline{L_p + 1, NL_b + L_p}$ be Gaussian. Then, $\mathcal{R}_{\mathcal{Y}}$, the correlation matrix of the output \mathcal{Y} defined in Section 2 is given as

$$\mathcal{R}_{\mathcal{Y}} = E\{\mathcal{Y}(l)\mathcal{Y}(l)^H\} = \mathcal{H}\mathcal{H}^H + \sigma_n^2\mathbf{I}, \quad (6)$$

where $\mathcal{R}_{\mathcal{Y}} \in \mathbb{C}^{(N-L_h+1)r \times (N-L_h+1)r}$. Hence, employing the zero IBI assumption as defined in the section above, the log-likelihood expression for the semi-blind scenario is given by a sum of the blind and pilot log-likelihoods as

$$\mathcal{L}(\mathcal{Y}; \theta_{\mathcal{H}}) = \mathcal{L}_b + \mathcal{L}_p, \quad (7)$$

where \mathcal{L}_b , the Gaussian log-likelihood of the blind symbols, is given as

$$\mathcal{L}_b = -\sum_{k=1}^{L_b} \text{tr}(\mathcal{Y}(k)^H \mathcal{R}_{\mathcal{Y}}^{-1} \mathcal{Y}(k)) - L_b \ln \det \mathcal{R}_{\mathcal{Y}}, \quad (8)$$

and \mathcal{L}_p , the least-squares log-likelihood of the pilot part, is given as

$$\mathcal{L}_p = -\frac{1}{\sigma_n^2} \sum_{i=1}^{L_p} \left\| \mathbf{y}_p(i) - \sum_{j=0}^{L_h-1} H(j) \mathbf{x}_p(i-j) \right\|^2. \quad (9)$$

Hence, the FIM for the sum likelihood is given as

$$J_{\theta_{\mathcal{H}}} = J^b + J^p, \quad (10)$$

where $J^b, J^p \in \mathbb{C}^{2rtL_h \times 2rtL_h}$ are the FIMs for the blind and pilot symbol bursts, respectively, which are defined by the likelihoods $\mathcal{L}_b, \mathcal{L}_p$ [21, 25]. This splitting of the FIM into pilot and blind components is similar to the approaches employed in [26, 27] and can be considered as representing the blind and pilot information components available for the estimation of the wireless channel. In the semi-blind scheme, we wish to make complete use of the blind and pilot information for channel estimation, as against an exclusively pilot-based scheme which employs only pilot information. This forms the central basis for the development of semi-blind estimation schemes. Next, we present a general result on the properties of the FIM before we apply it to the problem at hand in the succeeding sections.

3.1. FIM: A General Result. In this section, we present an interesting property of an FIM-based analysis by demonstrating a relation between the rank of the FIM and the number of unidentifiable parameters. Let $\bar{\alpha} \in \mathbb{C}^{k \times 1}$ be the complex parameter vector of interest. As described in [21, 24], for estimation of complex parameters, we employ a stacking of $\bar{\alpha}$ as $\bar{\theta} = [\bar{\alpha}^T, \bar{\alpha}^H]^T \in \mathbb{C}^{n \times 1}$ where $n \triangleq 2k$. Let $p(\bar{w}; g(\bar{\theta}))$ be the pdf of the observation vector \bar{w} parameterized by $g(\bar{\theta})$, where $g(\bar{\theta}) : \mathbb{C}^{n \times 1} \rightarrow \mathbb{C}^{l \times 1}$ is a function of the parameter vector $\bar{\theta}$. Similar to stacking $\bar{\alpha}, \bar{\alpha}^*$, let the function $f(\bar{\theta}) : \mathbb{C}^{n \times 1} \rightarrow \mathbb{C}^{m \times 1}, m \triangleq 2l$ be defined as

$$f(\bar{\theta}) = \begin{bmatrix} g(\bar{\theta}) \\ g^*(\bar{\theta}) \end{bmatrix}. \quad (11)$$

Given the log-likelihood $\mathcal{L}(\bar{w}; \bar{\theta}) \triangleq \ln p(\bar{w}; f(\bar{\theta}))$, the FIM $J_{\bar{\theta}} \in \mathbb{C}^{n \times n}$ is given [25] as

$$J_{\bar{\theta}} \triangleq -E\left\{ \frac{\partial^2 \mathcal{L}(\bar{w}; \bar{\theta})}{\partial \bar{\theta} \partial \bar{\theta}^H} \right\}. \quad (12)$$

Let $f(\bar{\theta})$ be an identifiable function of the parameter $\bar{\theta}$, that is, the FIM with respect to $f(\bar{\theta})$ has full rank. We then have the following result.

Lemma 1. Let $p(\bar{w}; f(\bar{\theta}))$, be the pdf of the observation vector \bar{w} , and let $f(\bar{\theta}) : \mathbb{C}^{n \times 1} \rightarrow \mathbb{C}^{m \times 1}$ be a function of the parameter vector $\bar{\theta}$ satisfying the following conditions.

- (C.1) Let $f(\bar{\theta})$ itself be an identifiable function of the parameter $\bar{\theta}$, that is, the FIM with respect to $f(\bar{\theta})$ has full rank.
- (C.2) Let $\text{rank}(\mathcal{N}(\partial f(\bar{\theta})/\partial \bar{\theta})) = d$, or, in other words, the dimension of the null space of $\partial f(\bar{\theta})/\partial \bar{\theta}$ is d .

Under the above conditions, the FIM $J(\bar{\theta}) \in \mathbb{C}^{n \times n}$ is rank deficient and, moreover,

$$\text{rank}(J(\bar{\theta})) = n - d. \quad (13)$$

Proof. Let $p(\bar{w}; f(\bar{\theta}))$ be the pdf of the observations \bar{w} . The derivative of the log-likelihood with respect to the parameter vector $\bar{\theta}$ is given as

$$\frac{\partial}{\partial \bar{\theta}} \ln p(\bar{w}; f(\bar{\theta})) = \frac{\partial}{\partial f(\bar{\theta})} \ln p(\bar{w}; f(\bar{\theta})) \frac{\partial f(\bar{\theta})}{\partial \bar{\theta}}. \quad (14)$$

The unconstrained FIM for the estimation of the parameter vector $\bar{\theta}$ is given as

$$\begin{aligned} J(\bar{\theta}) &= E\left\{ \left(\frac{\partial}{\partial \bar{\theta}} \ln p(\bar{w}; f(\bar{\theta})) \right)^T \frac{\partial}{\partial \bar{\theta}} \ln p(\bar{w}; f(\bar{\theta})) \right\} \\ &= \left(\frac{\partial f(\bar{\theta})}{\partial \bar{\theta}} \right)^T \underbrace{E\left\{ \left(\frac{\partial}{\partial f(\bar{\theta})} \ln p(\bar{w}; f(\bar{\theta})) \right)^T \frac{\partial}{\partial f(\bar{\theta})} \ln p(\bar{w}; f(\bar{\theta})) \right\}}_{J(f(\bar{\theta}))} \frac{\partial f(\bar{\theta})}{\partial \bar{\theta}}. \end{aligned} \quad (15)$$

Hence, from the condition C.2 above, it follows that $\text{rank}(J(\bar{\theta})) = n - d$. \square

Thus, the rank of the FIM is deficient by d , which is the number of unidentifiable parameters. We now provide a deeper insight into the above result that connects the nature of the FIM to the number of unidentifiable parameters. Explicitly, let $\bar{\theta}$ be reparameterized by the real parameter vector $\bar{\xi} \triangleq [\bar{\xi}_1^T, \bar{\xi}_2^T]^T$, $\bar{\xi}_1 \in \mathbb{R}^{d' \times 1}$, $\bar{\xi}_2 \in \mathbb{R}^{d \times 1}$ as $\bar{\theta}(\bar{\xi})$. Let $f(\bar{\theta}(\bar{\xi})) \in \mathbb{C}^{m \times 1}$ satisfy the property

$$\frac{\partial f(\bar{\theta}(\bar{\xi}))}{\partial \bar{\xi}_2} = \frac{\partial f(\bar{\theta})}{\partial \bar{\theta}} \frac{\partial \bar{\theta}}{\partial \bar{\xi}_2} = \mathbf{0}_{m \times d}, \quad (16)$$

or, in other words, the function $f(\bar{\theta})$ remains unchanged as the parameter vector $\bar{\theta}$ varies over the d dimensional constrained manifold $\bar{\theta}(\bar{\xi}_2)$, and thus $\partial f(\bar{\theta})/\partial \bar{\theta}$ has at least a d dimensional null space. The parameter vector $\bar{\xi}_2$ is the unconstrained parameterization of the constraint manifold and represents the unidentifiable parameters. This implies that each parameter θ_i in $\bar{\theta}$ is identifiable only up to d degrees of freedom owing to the unidentifiability of the $\bar{\xi}_2$ component which is of dimension d . This result has interesting applications, especially in the investigation of identifiability issues in the context of blind and semi-blind wireless channel estimation. The implications of this result in the context of semi-blind MIMO channel estimation are explored in the next section where we examine the rank of the semi-blind FIM and derive further insights into the nature of the estimation problem.

3.2. Blind FIM. We now apply the above result to our problem of MIMO FIR channel estimation. We start by investigating the properties of the blind FIM J^b . Let the block Toeplitz parameter derivative matrix $\mathcal{E}(k) \in \mathbb{C}^{(N-L_h+1)r \times Nt}$ be defined employing complex derivatives as

$$\mathcal{E}(krt+i) \triangleq \frac{\partial \mathcal{H}}{\partial \theta_{\mathcal{H}}^{2krt+i}}. \quad (17)$$

From the results for the Fisher information matrix of a complex Gaussian stochastic process [28], J^b defined in (10) above is given as

$$\begin{aligned} J_{2krt+i, 2lrt+j}^b &= J_{(2l+1)rt+j, (2k+1)rt+i}^b \\ &= L_b \text{tr} \left(\mathcal{E}(krt+i) \mathcal{H}^H \mathcal{R}_y^{-1} \mathcal{H} \mathcal{E}(lrt+j)^H \mathcal{R}_y^{-1} \right), \\ J_{2krt+i, (2l+1)rt+j}^b &= \left(J_{(2l+1)rt+j, 2krt+i}^b \right)^* \\ &= L_b \text{tr} \left(\mathcal{E}(krt+i) \mathcal{H}^H \mathcal{R}_y^{-1} \mathcal{E}(lrt+j) \mathcal{H}^H \mathcal{R}_y^{-1} \right), \end{aligned} \quad (18)$$

where $J_{k,l}^b$ denotes its (k, l) th element. We can now apply the result in Lemma 1 above to this FIM matrix J^b , and we have the following result on the rank of the blind FIM for the MIMO FIR channel. Under the assumptions above,

it is known [15, 20] that $H(z)$ can be identified up to a constant $t \times t$ unitary matrix from second-order statistical information. Such a unitary matrix has t^2 real parameters, [21]. Hence, from Lemma 1 above, we have the following result.

Theorem 2. *Let the MIMO FIR channel transfer function $H(z)$ satisfy (A.1) and (A.2) above. Then, the rank upper-bound on the blind FIM $J^b \in \mathbb{C}^{2rtL_h \times 2rtL_h}$ defined in (18) above is given as*

$$\text{rank}(J^b) = 2rtL_h - t^2. \quad (19)$$

In fact, a basis for the $t^2 \times 1$ null space $\mathcal{N}(J^b)$ is given by $U(\mathbf{H}) \in \mathbb{C}^{2rtL_h \times t^2}$ as

$$U(\mathbf{H}) \triangleq \begin{bmatrix} U(H(0)) \\ U(H(1)) \\ \vdots \\ U(H(L_h)) \end{bmatrix}, \quad (20)$$

where the matrix function $U(H) : \mathbb{C}^{r \times t} \rightarrow \mathbb{C}^{2rt \times t^2}$ for the matrix $H = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_t]$ is defined as

$$U(H) = \begin{bmatrix} -\mathbf{h}_1^* & -\mathbf{h}_2^* & -\mathbf{h}_3^* & \cdots & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{h}_1^* & -\mathbf{h}_2^* & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ \mathbf{h}_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{h}_2 & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{h}_1 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{h}_2 & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{h}_1 & \cdots & \mathbf{0} & \mathbf{0} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (21)$$

Proof. See Appendix A. \square

Thus, it is clear that MIMO FIR impulse response of the channel can be estimated up to an indeterminacy of t^2 real parameters from the statistical or blind information. This result has significant implications for estimation of the MIMO channel. As r, L_h increase, the number of real parameters in the system, which is equal to $2rtL_h$ (the dimension of matrix $J_{\theta_{\mathcal{H}}}$), increases manyfold. However, the number of parameters that cannot be identified from blind symbols may be as small as t^2 implying that a wealth of data can be identified from the blind symbols without any need for pilots. This reduction in the number of parameters to be estimated from pilots results in a significant decrease in the MSE of estimation, a result which will be rigorously justified in Section 5. Next, we present an algorithm for the SB estimation of the MIMO FIR channel matrix.

4. Semiblind Estimation: Algorithm

As shown above, since only t^2 (of the total $2rtL_h$) parameters cannot be identified from blind data, they can be identified

from the pilot symbols. These t^2 parameters correspond to a unitary matrix. More precisely, Let $H(z) \in \mathbb{C}^{r \times t}(z)$ be the $r \times t$ irreducible channel transfer matrix, let the input-output system model be as shown in Section 2. Then, $H(z)$ can be identified up to a unitary matrix from the blind statistical information, that is, without the aid of any pilots. In the following discussion, let $\hat{\theta}$ denote the estimate of the quantity θ . The matrix $\mathbf{H} \in \mathbb{C}^{r \times L_h t}$ can be expressed as

$$\mathbf{H} = \mathbf{W}(\mathbf{I}_{L_h} \otimes Q^H), \quad (22)$$

where $\mathbf{W} \triangleq [W(1), W(2), \dots, W(L_h - 1)]$.

From the above result, the matrices $\widehat{W}(i)$, $i \in \overline{0, L_h - 1}$ can be estimated from the blind second-order statistical information, that is, from the correlation lags $R_y(j) \triangleq \mathbb{E}\{\mathbf{y}(j)\mathbf{y}(0)^H\}$, $j \in \overline{0, L_h - 1}$, without the aid of pilot symbols. In the flat fading channel case ($L_h = 1$), this can be done by a simple Cholesky decomposition of the instantaneous output correlation matrix $R_y(0)$ (as $R_y(0) = \widehat{W}(0)\widehat{W}(0)^H$). However, for the case of frequency selective channels, estimating the matrices $W(i)$ is not straight forward and a scheme based on designing multiple delay linear predictors is given in [7] (Set $n_a = 0$, $d = n_b = L_h - 1$ and it follows that $\widehat{W}(i) = \widetilde{F}_i$). It thus remains to compute the unitary matrix $\widehat{Q} \in \mathbb{C}^{t \times t}$, that is, \widehat{Q} is such that $\widehat{Q}\widehat{Q}^H = \widehat{Q}^H\widehat{Q} = \mathbf{I}$ and Q has very few parameters (t^2 real parameters, [21]). In the next section, we present algorithms for the estimation of this unitary Q indeterminacy from the transmitted pilot symbols.

4.1. Orthogonal Pilot ML (OPML) for Q Estimation. We now describe a procedure to estimate the unitary matrix Q from an orthogonal pilot symbol sequence X_p . Let $X_p(i)$, $i \in \overline{0, L_h - 1}$ be defined as $X_p(i) \triangleq [\mathbf{x}_p(L_h - i), \mathbf{x}_p(L_h - i + 1), \dots, \mathbf{x}_p(L_p - i)]$. Let X_p^o the pilot matrix be defined as

$$X_p^o \triangleq \begin{bmatrix} X_p(0) \\ X_p(1) \\ \vdots \\ X_p(L_h - 1) \end{bmatrix}. \quad (23)$$

The least squares cost function for the constrained estimation of the unitary matrix Q can then be written as

$$\left\| Y_p - \sum_{i=0}^{L_h-1} W(i)\widehat{Q}^H X_p(i) \right\|^2, \quad \text{subject to } \widehat{Q}\widehat{Q}^H = \mathbf{I}_t. \quad (24)$$

Let the pilot matrix X_p^o be orthogonal, that is, $X_p^o(X_p^o)^H = L_p \mathbf{I}_{L_h t}$. The cost minimizing \widehat{Q} is then given as

$$\widehat{Q} = UV^H, \quad \text{where } U\Sigma V^H = \text{SVD} \left(\sum_{i=0}^{L_h-1} X(i)Y^H W(i) \right). \quad (25)$$

Proof. It follows from an extension of the result in [21]. \square

FIGURE 2: Paley Hadamard matrix.

Finally, $\widehat{\mathbf{H}}$ is given as $\widehat{\mathbf{H}} \triangleq \widehat{\mathbf{W}}(\mathbf{I}_{L_h} \otimes \widehat{Q}^H)$. It now remains to demonstrate a scheme to construct the orthogonal pilot matrix X_p^o which is treated next.

4.2. Orthogonal Pilot Matrix Construction. An orthogonal pilot matrix in the context of MIMO FIR channels can be constructed by employing the Paley Hadamard (PH) orthogonal matrix structure shown in Figure 2, and such a scheme has been described in [29]. Another scheme based on signal constellations derived from the roots of unity is presented in [5]. The PH matrix has blocks of shifted orthogonal rows (illustrated with the aid of rectangular boundaries), thus giving it the block Sylvester structure. Each transmit stream of orthogonal pilots for the FIR system can be constructed by considering the “L”-shaped block shown in the figure and removing the prefix of L_h (channel length) symbols at the receiver. Thus, the pilot matrix for orthogonal pilots X_p^o is given as

$$X_p^o \triangleq \begin{bmatrix} \mathbf{x}_p(L_h) & \mathbf{x}_p(L_h + 1) & \cdots & \mathbf{x}_p(L_p) \\ \mathbf{x}_p(L_h - 1) & \mathbf{x}_p(L_h) & \cdots & \mathbf{x}_p(L_p - 1) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_p(1) & \mathbf{x}_p(2) & \cdots & \mathbf{x}_p(L_p - L_h + 1) \end{bmatrix}. \quad (26)$$

Orthogonal pilots have shown to be optimally suited for MIMO channel estimation in studies such as [30, 31]. However, the Sylvester structure of FIR pilot matrices further constrains the set of orthogonal pilot symbol streams compared to flat-fading channels. As the number of channel taps increases, employing a PH matrix to construct an orthogonal pilot symbol stream implies choosing a PH matrix with a large dimension. This in turn implies an increase in the length of the pilot symbol sequence and hence a larger overhead in communication. This problem can be alleviated by employing a nonorthogonal pilot symbol sequence which results in slightly suboptimal estimation performance but enables the designer to choose any pilot length desired. This iterative general maximum-likelihood (IGML) scheme for channel estimation using nonorthogonal pilots is described in [21] for flat-fading channels and can be extended to FIR channels in a straight forward manner. Experimental results

have shown its performance to be comparable to the above OPML scheme.

5. Semiblind MSE of Estimation

Consider now the asymptotic performance of the semi-blind scheme from an FIM perspective. As the amount of blind information increases (which does not increase the pilot overhead of the system), the variance of estimation of the covariance matrices $R_y(i)$ progressively decreases to zero, implying that the blindly identifiable parameters (such as the whitening matrix) can be estimated accurately. Thus, intuitively, the SB estimation problem reduces to the constrained estimation of the t^2 blindly unidentifiable parameters from the pilot symbols, a proposition which is rigorously justified below. Hence, the limiting MSE is equal to the MSE for the complex constrained estimation of the t^2 blindly unidentifiable parameters. SB techniques can therefore yield a far lesser MSE of estimation than an exclusively pilot-based scheme as illustrated by the following result.

Theorem 3. Let $J^p = (L_p/\sigma_n^2)\mathbf{I}_{2rtL_h}$, which is achieved by the orthogonal pilot matrix X_p^o . Then, as the number of blind symbol transmissions increases ($L_b \rightarrow \infty$), the semi-blind CRB $J_{\theta_{\mathcal{H}}}^{-1}$ approaches the CRB for the exclusive estimation of the t^2 blindly unidentifiable parameters. Further, the semi-blind MSE bound, given by the trace of the CRB matrix, converges to

$$\lim_{L_b \rightarrow \infty} \mathbb{E}\left\{\|\hat{\mathbf{H}} - \mathbf{H}\|_F^2\right\} \geq \frac{1}{2} \text{tr}(J_{\theta_{\mathcal{H}}}^{-1}) = \left(\frac{\sigma_n^2}{2L_p}\right)t^2, \quad (27)$$

which depends only on t^2 , the number of blindly unidentifiable parameters.

Proof. Given the fact that $J^p = (L_p/\sigma_n^2)\mathbf{I}_{2rtL_h}$, the semi-blind FIM can be expressed as

$$J_{\theta_{\mathcal{H}}} = \frac{L_p}{\sigma_n^2} \mathbf{I}_{2rtL_h \times 2rtL_h} + L_b \tilde{J}^b, \quad (28)$$

where \tilde{J}^b is the blind FIM corresponding to a single observed blind symbol block \mathcal{Y} and is given as $\tilde{J}^b \triangleq (J^b/L_b)$, where the blind FIM J^b is defined in (18). From Lemma 1, it can be seen that \tilde{J}^b is rank deficient and $\text{rank}(\tilde{J}^b) = \text{rank}(J^b) = 2rt - t^2$. Let the eigen-decomposition of \tilde{J}^b be given as $\tilde{J}^b = E_b \Lambda_b E_b^H$, where $\Lambda_b \in \mathbb{C}^{(2rt-t^2) \times (2rt-t^2)}$ is a diagonal matrix. Then,

$$\begin{aligned} J(\theta_{\mathcal{H}}) &= \frac{L_p}{\sigma_n^2} [E_b^\perp, E_b] [E_b^\perp, E_b]^H + L_b E_b \Lambda_b E_b^H \\ &= [E_b, E_b^\perp] \begin{bmatrix} \frac{L_p}{\sigma_n^2} \mathbf{I} + L_b \Lambda_b & \mathbf{0} \\ \mathbf{0} & \frac{L_p}{\sigma_n^2} \mathbf{I} \end{bmatrix} [E_b, E_b^\perp]^H. \end{aligned} \quad (29)$$

Hence, the CRB $J^{-1}(\theta_{\mathcal{H}})$ is given as

$$J^{-1}(\theta_{\mathcal{H}}) = [E_b, E_b^\perp] \begin{bmatrix} \left(\frac{L_p}{\sigma_n^2} \mathbf{I} + L_b \Lambda_b\right)^{-1} & \mathbf{0} \\ \mathbf{0} & \frac{\sigma_n^2}{L_p} \mathbf{I} \end{bmatrix} [E_b, E_b^\perp]^H. \quad (30)$$

As the number of blind symbols $L_b \rightarrow \infty$, the diagonal matrix $(L_p/\sigma_n^2) \mathbf{I} + L_b \Lambda_b \rightarrow \mathbf{0}_{(2rt-t^2) \times (2rt-t^2)}$ in the above expression. Thus the semi-blind bound approaches the complex constrained Cramer-Rao bound (CC-CRB) [21] given as

$$\lim_{L_b \rightarrow \infty} J^{-1}(\theta_{\mathcal{H}}) = \frac{\sigma_n^2}{L_p} E_b^\perp E_b^{\perp H}. \quad (31)$$

In fact, the bound on the MSE is clearly seen to be given as

$$\begin{aligned} \mathbb{E}\left\{\|\hat{\theta}_{\mathcal{H}} - \theta_{\mathcal{H}}\|_F^2\right\} &\geq \frac{\sigma_n^2}{L_p} \text{tr}(E_b^\perp E_b^{\perp H}) \\ &\Rightarrow 2\left(\mathbb{E}\left\{\|\hat{\mathbf{H}} - \mathbf{H}\|_F^2\right\}\right) \geq \frac{\sigma_n^2}{L_p} \text{tr}(E_b^\perp E_b^{\perp H}) \\ \mathbb{E}\left\{\|\hat{\mathbf{H}} - \mathbf{H}\|_F^2\right\} &\geq \frac{1}{2} \frac{\sigma_n^2}{L_p} (2rt - (2rt - t^2)) \\ &= \left(\frac{\sigma_n^2}{2L_p}\right) t^2, \end{aligned} \quad (32)$$

which is the constrained bound for the estimation of the MIMO channel matrix \mathbf{H} . \square

Thus, the bound for the MSE of estimation and hence the asymptotic MSE of the maximum-likelihood estimate of the channel matrix \mathbf{H} with the aid of blind information, are directly proportional to t^2 .

Contrast this result with the MSE of estimation exclusively using an exclusively pilot-based scheme (i.e., a scheme which does not leverage the blind data like the semi-blind scheme). This MSE is given as $(1/2) \text{tr}(\{J^p\}^{-1}) = (\sigma_n^2/2L_p)2rtL_h$ and is proportional to $2rtL_h$, the total number of real parameters. Hence, the SB estimate which has an asymptotic MSE lower by a factor of $2(r/t)L_h$ can potentially be very efficient compared to exclusive pilot only channel estimation schemes. For instance, in a MIMO system with $r = 4$, $t = 2$, and $L_h = 2$ channel taps, the potential reduction in MSE by employing a semi-blind estimation procedure is $2(r/t)L_h = 9$ dB as demonstrated in Section 7. Thus, the SB estimation scheme can result in significantly lower MSE. Finally, it is worth mentioning that the above results, which are derived employing the Gaussian data symbol distribution in Section 3, are in close agreement with the performance of a system employing a discrete signal constellation such as quadrature phase-shift keying (QPSK), as illustrated in the simulation results of Section 7.

6. Pilots and FIM

In this section, we obtain a lower bound for the minimum number of pilot symbols necessary to achieve regularity or a full rank FIM $J_{\theta_{\mathcal{H}}}$, that is, for the SB identifiability of the MIMO channel $H(z)$. Recall that $\{\mathbf{x}_p(1), \mathbf{x}_p(2), \dots, \mathbf{x}_p(L_p)\}$ are the L_p transmitted pilot symbols. Then, the FIM of the pilot symbols J^p is given as $J^p = \sum_{i=1}^{L_p} J^p(i)$, where $J^p(i)$ is the FIM contribution from the i th pilot symbol transmission.

Given complex vectors in $\mathbb{C}^{t \times 1}$, let the matrix function ${}^iV_j : (\mathbb{C}^{t \times 1}, \mathbb{C}^{t \times 1}) \rightarrow \mathbb{C}^{2rt \times 2rt}$ be defined as

$${}^iV_j \triangleq \begin{bmatrix} \mathbf{x}_p(i)\mathbf{x}_p(j)^H \otimes I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{x}_p(i)^*\mathbf{x}_p(j)^T \otimes \mathbf{I}_r \end{bmatrix}, \quad \text{if } i, j > 0, \quad (33)$$

and ${}^iV_j = \mathbf{0}_{2rt \times 2rt}$ if $i \leq 0$ or $j \leq 0$. After some manipulations, it can be shown that the FIM contribution $J^p(i) \in \mathbb{C}^{2rtL_h \times 2rtL_h}$ is given as

$$J^p(i) = \frac{1}{\sigma_n^2} \begin{bmatrix} {}^iV_i & {}^iV_{i-1} & \cdots & {}^iV_{i-L_h+1} \\ {}^{i-1}V_i & {}^{i-1}V_{i-1} & \cdots & {}^{i-1}V_{i-L_h+1} \\ \vdots & \vdots & \ddots & \vdots \\ {}^{i-L_h+1}V_i & {}^{i-L_h+1}V_{i-1} & \cdots & {}^{i-L_h+1}V_{i-L_h+1} \end{bmatrix}. \quad (34)$$

The following result bounds the rank of the semi-blind (pilot + blind) FIM $J_{\theta_{\mathcal{R}}}$.

Theorem 4. Let $L_p \leq t$ pilot symbols $\mathbf{x}_p(1), \mathbf{x}_p(2), \dots, \mathbf{x}_p(L_p)$ be transmitted and let the matrix $H(0)$ be full column rank as assumed above. A rank upper bound of the sum (pilot + blind) FIM $J_{\theta_{\mathcal{R}}}$ defined in (10) above is given as

$$\text{rank}(J_{\theta_{\mathcal{R}}}) \leq 2rtL_h - (t - L_p)^2, \quad 0 \leq L_p \leq t \quad (35)$$

or, in other words, a lower bound on the rank deficiency is given as $(t - L_p)^2$.

Proof. See Appendix B. \square

The above result gives an expression for the rank upper bound of the MIMO FIR Fisher information matrix for each transmitted pilot symbol. Since identifiability requires a full rank FIM, it thus presents a key insight into the number of pilot symbols needed for identifiability of the MIMO FIR system as shown next.

6.1. Pilots and Identifiability. From the above result, one can obtain a lower bound on the number of pilot symbols necessary for SB identifiability of the MIMO channel. This result is stated below.

Lemma 5. The number of pilot symbol transmissions L_p should at least equal the number of transmit antennas t for the FIM $J_{\theta_{\mathcal{R}}}$ to be full rank and hence the MIMO FIR system in (1) to be identifiable.

Proof. It is easy to see from (35) that, for $L_p < t$,

$$\text{rank}(J_{\theta_{\mathcal{R}}}) = 2rtL_h - (t - L_p)^2 < 2rtL_h, \quad (36)$$

that is, strictly less than full rank. As the number of pilot symbols increases, for $L_p = t$, $\text{rank}(J_{\theta_{\mathcal{R}}}) \leq 2rtL_p$, where $2rtL_p$ is the dimension of $J_{\theta_{\mathcal{R}}}$ and therefore represents full rank. Hence, at least t pilot symbols are necessary for the identifiability of the MIMO FIR wireless channel. \square

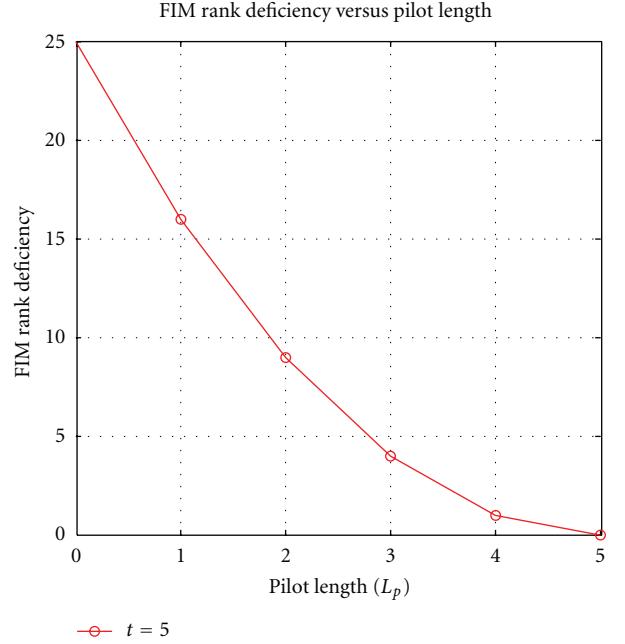


FIGURE 3: Rank deficiency of the complex MIMO FIM versus number of transmitted pilot symbols (L_p) for a 6×5 MIMO FIR system of length $L_h = 5$.

Thus at least t pilot symbols are necessary for the system to become identifiable. One has to observe that in the case of semi-blind estimation potentially fewer number (t^2) of parameters need to be estimated. Hence, even though semi-blind schemes necessitate the transmission of t pilot symbols, the accuracy of estimation of such a scheme can be higher owing to the fact that they estimate fewer parameters from the limited pilot symbols. This improvement in MSE performance has been quantified in Section 5 where we presented results about the asymptotic performance of the SB estimator using the above FIM framework. Next, we present results of simulation studies.

7. Simulation Results

In this section, we present results of simulation experiments to illustrate the salient aspects of the work described above. In a majority of our simulations below, we consider a 4×2 MIMO FIR channel with 2 taps, that is, $L_h = 2$, $r = 4$, and $t = 2$. Each of the elements of \mathbf{H} is generated as a zero-mean circularly symmetric complex Gaussian random variables of unit variance, that is, a Rayleigh fading wireless channel. The orthogonal pilot sequence is constructed from Paley Hadamard matrices by employing the scheme in Section 4.1. The transmitted symbols, both pilot and blind (data), are assumed to be drawn from a quadrature phase shift keying (QPSK) symbol constellation [32].

Experiment 1. In Figure 3, we plot the rank deficiency of the FIM of a 6×5 MIMO FIR system ($r = 6, t = 5$) with $L_h = 5$ channel taps as a function of the number of transmitted pilot

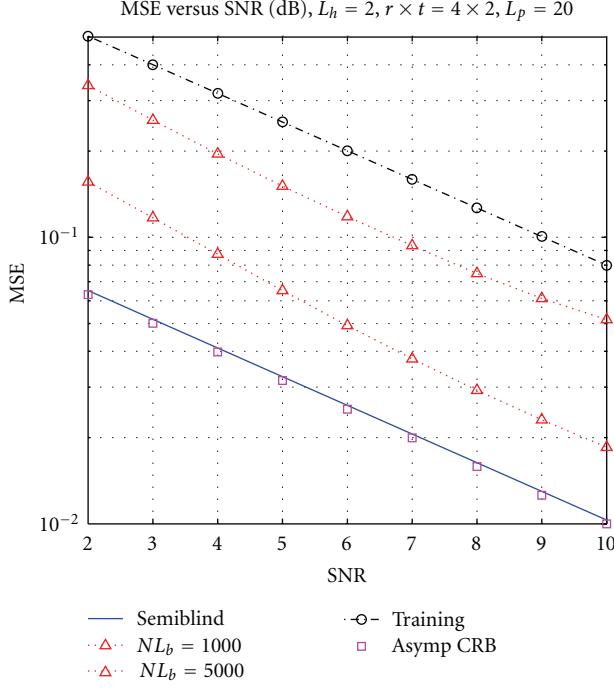
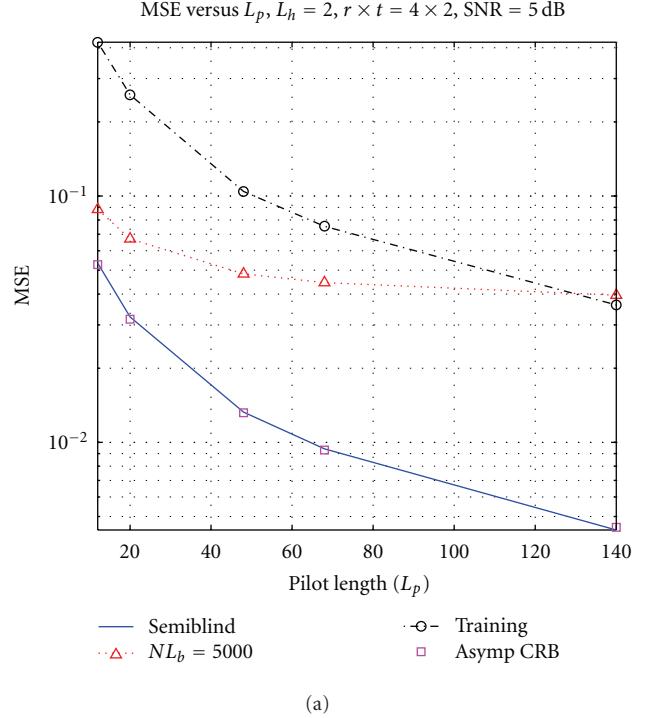


FIGURE 4: MSE versus SNR in a 4×2 MIMO channel with $L_h = 2$ channel taps, $L_p = 20$ pilot symbols.

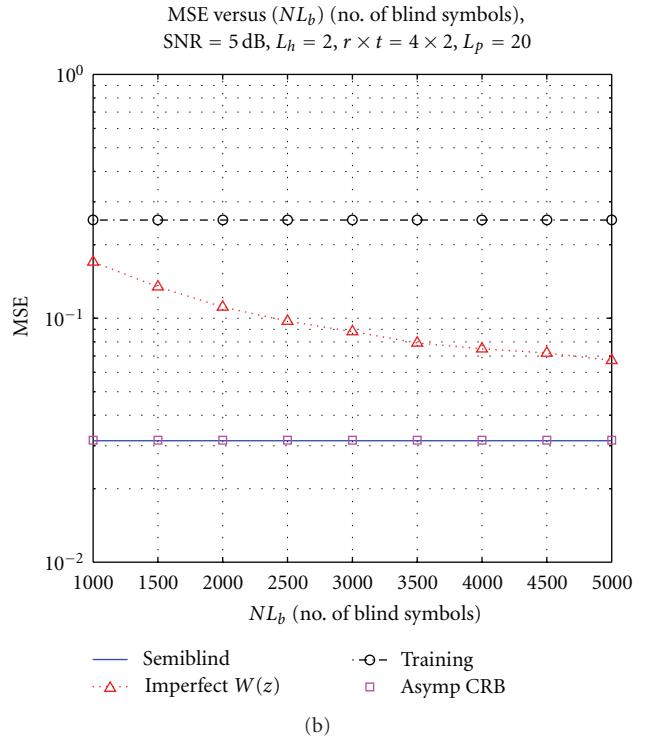
symbols L_p . The rank was computed for 100 realizations of randomly generated Rayleigh fading MIMO channels, and the rank deficiency observed was precisely [25, 16, 9, 4, 1, 0] for $L_p = [0, 1, 2, 3, 4, 5]$ transmitted pilot symbols, respectively. Hence, rank deficiency 25 for $L_p = 0$ verifies that the assumptions (A.1), (A.2) (in turn Lemma 1) hold with overwhelming probability in the case of randomly generated MIMO channels. Further, for $1 \leq L_p \leq 5$, the rank deficiency is given as $(5 - L_p)^2$ which additionally verifies the bound in (35) for FIM rank deficiency as a function of number of pilot symbols.

Experiment 2. In Figure 4, we plot the MSE versus SNR when the whitening matrix $W(z)$ is estimated from $NL_b = 1000, 5000$ blind received symbols employing the linear prediction-based scheme from [7]. The Q matrix is estimated from $L_p = 20$ orthogonal pilot symbols employing the semi-blind scheme in Section 4.1. For comparison, we also plot the MSE when \mathbf{H} is estimated exclusively from training using least-squares [21, 25], the asymptotic complex constrained CRB (CC-CRB) given by (27), and the MSE of estimation with the genie-assisted case of perfect knowledge of $W(z)$. It can be observed that the MSE progressively decreases towards the complex constrained CRB as the number of blind symbols increases. Also observed as illustrated in Theorem 3, the asymptotic SB estimation error is $10\log(32/4) = 9$ dB lower than the pilot-based scheme as illustrated in Section 5.

In Figure 5(a), we plot the MSE performance of the competing estimation schemes above for different transmitted pilot symbol lengths L_p and 5000 transmitted QPSK



(a)



(b)

FIGURE 5: MSE performance for estimation of a 4×2 MIMO frequency-selective channel. (a) MSE versus L_p and (b) MSE versus number of blind symbols.

data symbols (blind received symbols). As illustrated in Section 4.1, we employ Paley Hadamard matrices to construct the orthogonal pilot sequences. Since such matrices exist only for certain lengths L_p , we plot the performance for $L_p = 12, 20, 48, 68, 140$ pilot symbols. The asymptotic semi-

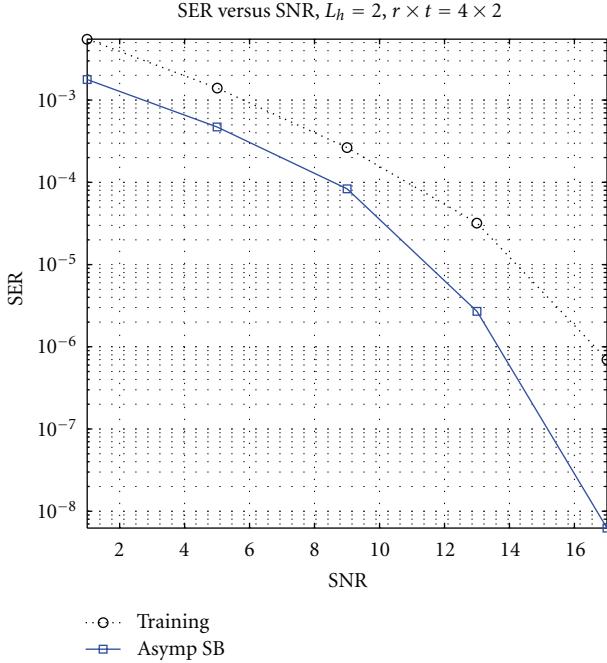


FIGURE 6: Symbol error rate (SER) versus SNR for QPSK symbol transmission of a 4×2 MIMO frequency selective channel with $L_h = 2$ channel taps.

blind performance is 9 dB lower in MSE as seen above. Also, for a given number of blind symbols, the performance gap in MSE of performance of the semi-blind scheme with $W(z)$ estimation and that of the training scheme progressively decreases. This is due to the fact that more blind symbols are required to accurately estimate the whitening matrix $W(z)$ for the MSE performance of the semi-blind scheme to be commensurate with the performance improvement of the pilot-based scheme. Finally, in Figure 5(b), we plot the performance of the competing schemes for different number of blind symbols in the range 1000–5000 QPSK symbols and $L_p = 12$ pilot symbols. The performance of the SB scheme with $W(z)$ estimated can be seen to progressively improve as the number of received blind symbols increases.

Experiment 3. We compare the symbol error rate (SER) performance of the training and semi-blind channel estimation schemes. At the receiver, we employ a stacking as in (3) of 7 received symbol vectors $\mathbf{y}(k)$ followed by a MIMO minimum mean-square error (MMSE) detector [1] constructed from the MIMO channel matrix \mathcal{H} . In Figure 6, we plot the SER of detection of the transmitted QPSK symbols versus SNR in the range 2–16 dB. It can be seen that the asymptotic semi-blind estimator has a 1–2 dB improvement in detection performance over the exclusive training-based scheme. The SER drops from around 1×10^{-1} at 2 dB to 1×10^{-8} at 16 dB. Thus, an SB-based estimation scheme can potentially yield significant throughput gains when employed for the estimation of the wireless MIMO frequency-selective channel.

8. Conclusion

In this work, we have investigated the rank properties of the semi-blind FIM of a L_h tap $r \times t$ ($r > t$) MIMO FIR channel. The MIMO channel transfer function $H(z)$ can be decomposed as $H(z) = W(z)Q^H$, where the whitening transfer function $W(z)$ can be estimated from the blind symbols alone. A constrained semi-blind estimation scheme has been presented to estimate the unitary matrix Q from pilot symbols along with an algorithm to achieve an orthogonal pilot matrix structure for MIMO frequency selective channels using Paley Hadamard matrices. From an asymptotic MSE analysis, it has been demonstrated that the semi-blind scheme achieves a significantly lower MSE than an exclusively pilot-based scheme. It has also been demonstrated that at least t pilot symbol transmissions are necessary to achieve a full-rank FIM (and hence identifiability). Simulation results demonstrate the performance of the proposed semi-blind scheme.

Appendices

A. Proof of Theorem 2

Proof. Consider the result for the simpler case of the flat fading channel, that is, $L_h = 1$. Then, $\mathbf{H} = \mathcal{H} = H(0) = H \in \mathbb{C}^{r \times t}$. Let the system be parameterized as $\bar{\theta} = \text{vec}(H)$ and, hence, from the discussion above,

$$\bar{\theta} \triangleq \begin{bmatrix} \text{vec}(H) \\ \text{vec}(H)^* \end{bmatrix}. \quad (\text{A.1})$$

From the system model described in (1), the pdf of the observation vector $\mathbf{y} \in \mathbb{C}^{r \times 1}$ is given by $\mathbf{y} \sim \mathcal{N}(\mathbf{0}, g(\bar{\theta}))$, where $g(\bar{\theta}) \in \mathbb{C}^{r \times 1}$ is the output correlation defined as $g(\bar{\theta}) \triangleq \text{vec}(HH^H + \sigma_n^2 \mathbf{I})$. Consider a reparameterization of the channel matrix H by the real parameter vector $\bar{\xi}$ as $H(\bar{\xi}) = W(\bar{\xi}_1)Q(\bar{\xi}_2)$, where $W(\bar{\xi}_1) \in \mathbb{C}^{r \times t}$ is also known as a whitening matrix and $Q(\bar{\xi}_2) \in \mathbb{C}^{t \times t}$ is a unitary matrix. $g(\bar{\theta})$ is now a many to one mapping since $g(\bar{\theta}(\bar{\xi})) = \text{vec}(W(\bar{\xi}_1)W(\bar{\xi}_1)^H + \sigma_n^2 \mathbf{I})$ and

$$f(\bar{\theta}(\bar{\xi})) = \begin{bmatrix} \text{vec}\left(W(\bar{\xi}_1)W(\bar{\xi}_1)^H + \sigma_n^2 \mathbf{I}\right) \\ \left(\text{vec}\left(W(\bar{\xi}_1)W(\bar{\xi}_1)^H + \sigma_n^2 \mathbf{I}\right)\right)^* \end{bmatrix}, \quad (\text{A.2})$$

which is independent of the parameter vector $\bar{\xi}_2$. Hence, it is a many to one mapping since for all unitary matrices $Q(\bar{\xi}_2)$ we have

$$\frac{\partial f(\bar{\theta}(\bar{\xi}))}{\partial \bar{\xi}_2} = \mathbf{0}_{r^2 \times t^2}, \quad (\text{A.3})$$

since $\bar{\xi}_2 \in \mathbb{R}^{t^2 \times 1}$ (t^2 is the number of real parameters to characterize a $t \times t$ unitary matrix). Hence, $d = t^2$, and, from Lemma 1, the first result follows. The proof for the general frequency selective case follows from the result in [15].

For the second result, let $\Phi \triangleq (HH^H + \sigma_n^2 I)^{-1}$. Let the blind FIM J^b be block partitioned as

$$J^b \triangleq \begin{bmatrix} J_{11}^b & J_{12}^b \\ J_{21}^b & J_{22}^b \end{bmatrix}. \quad (\text{A.4})$$

It can be verified from (18) that $J_{21}^b = {J_{12}^b}^H$ and $J_{22}^b = {J_{11}^b}^T$. The block components of the FIM are given as

$$J_{11}^b = \begin{bmatrix} \mathbf{h}_1^H \Phi \mathbf{h}_1 \Phi_{11} & \mathbf{h}_1^H \Phi \mathbf{h}_1 \Phi_{21} & \cdots & \mathbf{h}_1^H \Phi \mathbf{h}_t \Phi_{r1} \\ \mathbf{h}_1^H \Phi \mathbf{h}_1 \Phi_{12} & \mathbf{h}_1^H \Phi \mathbf{h}_1 \Phi_{22} & \cdots & \mathbf{h}_1^H \Phi \mathbf{h}_t \Phi_{r2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{h}_1^H \Phi \mathbf{h}_1 \Phi_{1r} & \mathbf{h}_1^H \Phi \mathbf{h}_1 \Phi_{2r} & \cdots & \mathbf{h}_1^H \Phi \mathbf{h}_t \Phi_{rr} \\ \mathbf{h}_2^H \Phi \mathbf{h}_1 \Phi_{11} & \mathbf{h}_2^H \Phi \mathbf{h}_1 \Phi_{21} & \cdots & \mathbf{h}_2^H \Phi \mathbf{h}_t \Phi_{r1} \\ \mathbf{h}_2^H \Phi \mathbf{h}_1 \Phi_{12} & \mathbf{h}_2^H \Phi \mathbf{h}_1 \Phi_{22} & \cdots & \mathbf{h}_2^H \Phi \mathbf{h}_t \Phi_{r2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{h}_t^H \Phi \mathbf{h}_1 \Phi_{1r} & \mathbf{h}_t^H \Phi \mathbf{h}_1 \Phi_{2r} & \cdots & \mathbf{h}_t^H \Phi \mathbf{h}_t \Phi_{rr} \end{bmatrix}, \quad (\text{A.5})$$

which can be written succinctly as $(H^H \Phi H) \otimes \Phi^T$. Similarly, J_{12}^b is given as

$$J_{12}^b = \begin{bmatrix} \chi_{11}\chi_{11} & \chi_{12}\chi_{11} & \cdots & \chi_{11}\chi_{21} & \cdots & \chi_{1r}\chi_{t1} \\ \chi_{11}\chi_{12} & \chi_{12}\chi_{12} & \cdots & \chi_{11}\chi_{22} & \cdots & \chi_{1r}\chi_{t2} \\ \vdots & \vdots & \ddots & \cdots & \ddots & \vdots \\ \chi_{11}\chi_{1r} & \chi_{12}\chi_{1r} & \cdots & \chi_{11}\chi_{2r} & \cdots & \chi_{1r}\chi_{tr} \\ \chi_{21}\chi_{11} & \chi_{22}\chi_{11} & \cdots & \chi_{21}\chi_{21} & \cdots & \chi_{2r}\chi_{t1} \\ \chi_{21}\chi_{12} & \chi_{22}\chi_{12} & \cdots & \chi_{21}\chi_{22} & \cdots & \chi_{2r}\chi_{t2} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \chi_{21}\chi_{1r} & \chi_{22}\chi_{1r} & \cdots & \chi_{21}\chi_{2r} & \cdots & \chi_{2r}\chi_{tr} \\ \chi_{t1}\chi_{11} & \chi_{t2}\chi_{11} & \cdots & \chi_{t1}\chi_{21} & \cdots & \chi_{tr}\chi_{t1} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \chi_{t1}\chi_{1r} & \chi_{t2}\chi_{1r} & \cdots & \chi_{t1}\chi_{2r} & \cdots & \chi_{tr}\chi_{tr} \end{bmatrix}, \quad (\text{A.6})$$

where $\chi \triangleq H^H \Phi$. It can now be seen that $JU = 0$, where U is as defined in (21). For instance, the top t elements of $JU(:, 1)$ (where we employ MATLAB notation and $U(:, 1)$ denotes the first column of U) are given as $[J_{11}^b, J_{12}^b]U(:, 1) = -(\mathbf{h}_1^H \Phi \mathbf{h}_1)\Phi^T \mathbf{h}_2^* + (\mathbf{h}_2^T \Phi)^T (\mathbf{h}_1^H \Phi \mathbf{h}_1) = \mathbf{0}$, and so on. The structure of the FIM for the most general case of arbitrary L_h is complex, but the result can be seen to hold by employing a symbolic manipulation software tool such as the MATLAB symbolic toolbox package.

It can be seen that the top part of the null space basis matrix $U(H)$ is $U(H(0))$. As assumed earlier, $\text{rank}(H(0)) = t$. Now it is easy to see that if $U(H)$ is rank deficient, $U(H(0))$ is rank deficient, and, from its structure, $H(0)$ is rank deficient violating the assumption. Hence, $\text{rank}(U(H)) = t^2$. \square

B. Proof of Theorem 4

Proof. J^b and $J^p(i)$, $i \in \overline{1, k}$ are positive semidefinite (PSD) matrices. We use the following property: if A, B are PSD matrices, $(A + B)\mathbf{v} = \mathbf{0} \Leftrightarrow A\mathbf{v} = B\mathbf{v} = \mathbf{0}$. Therefore, $J_{\theta_{\mathcal{H}}}\mathbf{v} = \mathbf{0} \Leftrightarrow J^b\mathbf{v} = J^p(i)\mathbf{v} = \mathbf{0}_{2rtL_h \times 1}$, for all $i \in 1, k$. In other words,

$$\mathcal{N}(J) = \mathcal{N}(J^b) \bigcap_{i=1}^{L_p} \mathcal{N}(J^p(i)). \quad (\text{B.1})$$

Let $\mathbf{v} \in \mathcal{N}(J)$. Then, from the null space structure of J^b in (21), it follows that $\mathbf{v} = U(\mathbf{H})\mathbf{s}$, where $\mathbf{s} \in \mathbb{C}^{t^2 \times 1}$. Also,

$$J^p(i)\mathbf{v} = J^p(i)U(\mathbf{H})\mathbf{s} = \mathbf{0}, \quad \forall i \in \overline{1, L_p}. \quad (\text{B.2})$$

From Lemma 6, this implies that ${}^iV_i U(H(0))\mathbf{s} = \mathbf{0}$, for all $i \in \overline{1, L_p}$, where iV_i is as defined in (33). Let the matrix $\mathcal{T}(\mathbf{u}) : \mathbb{C}^{t \times 1} \rightarrow \mathbb{C}^{2t \times t^2}$ be defined as

$$\mathcal{T}(\mathbf{u}) \triangleq \begin{bmatrix} 0 & -\mathbf{u}_1^* & -\mathbf{u}_2^* & \cdots \\ -\mathbf{u}_1^* & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ \mathbf{u}_2 & \mathbf{u}_1 & 0 & \cdots \\ 0 & 0 & \mathbf{u}_1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (\text{B.3})$$

Recall that $H(0)$ is assumed to have full column rank. Then, from the structure of $U(H(0))$ given in (21), it can be shown that the relation above holds if and only if, $\mathbb{P}\mathbf{s} = \mathbf{0}$, where the matrix $\mathbb{P} \in \mathbb{C}^{2L_p \times t^2}$ is given as

$$\mathbb{P} \triangleq \begin{bmatrix} \mathcal{T}(\mathbf{x}_p(1)) \\ \mathcal{T}(\mathbf{x}_p(2)) \\ \vdots \\ \mathcal{T}(\mathbf{x}_p(L_p)) \end{bmatrix}. \quad (\text{B.4})$$

It can then be seen that matrix $\mathbb{G} \in \mathbb{C}^{2tL_p \times L_p^2}$ forms a basis for the left nullspace of \mathbb{P} , that is, $\mathbb{G}^T \mathbb{P} = \mathbf{0}$, where

$$\mathbb{G} \triangleq \begin{bmatrix} \mathbf{x}_p(1) & \mathbf{x}_p(2) & \mathbf{0} & \cdots \\ \mathbf{x}_p(1)^* & \mathbf{0} & \mathbf{x}_p(2)^* & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{x}_p(1) & \cdots \\ \mathbf{0} & \mathbf{x}_p(1)^* & \mathbf{0} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (\text{B.5})$$

Thus, $\text{rank}(\mathbb{P}) \leq 2tL_p - L_p^2$, and, therefore, right nullity (or nullity) of \mathbb{P} is $\dim(\mathcal{N}(\mathbb{P})) \geq t^2 - (2tL_p - L_p^2)$. And, therefore, $\text{rank}(J_{\theta_{\mathcal{H}}}) \leq 2rtL_h - \dim(\mathcal{N}(\mathbb{P})) = 2rtL_h - t^2 + (2tL_p - L_p^2)$. \square

Lemma 6. Let $J^p(i)\mathbf{v} = \mathbf{0}$, for all $i \in \overline{1, L_p}$, where $\mathbf{v} = U(\mathbf{H})\mathbf{s}$. Then, ${}^iV_iU(H(0))\mathbf{s} = \mathbf{0}$, for all $i \in \overline{1, L_p}$, where iV_i is as defined in (33).

Proof. Consider $J^p(1)$, the FIM contribution of the first transmitted pilot symbol. It can be seen clearly that $J^p(1)$ is given as

$$J^p(1) = \begin{bmatrix} {}^1V_1 & \mathbf{0}_{2rt \times (2L_h-2)rt} \\ \mathbf{0}_{(2L_h-2)rt \times 2rt} & \mathbf{0}_{(2L_h-2)rt \times (2L_h-2)rt} \end{bmatrix}. \quad (\text{B.6})$$

Hence, $J^p(1)U(\mathbf{H})\mathbf{s} = \mathbf{0}$ implies that ${}^1V_1U(H(0))\mathbf{s} = \mathbf{0}$. Further, from the properties of the matrix Kronecker product, one has $AB \otimes CD = (A \otimes C)(C \otimes D)$. Substituting $A = \mathbf{x}_p(i)$, $B = \mathbf{x}_p(i)^H$, $C = D = \mathbf{I}_r$, one can then obtain that $K(\mathbf{x}_p(1))U(H(0))\mathbf{s} = \mathbf{0}$

$$K(\mathbf{x}_p(1)) \triangleq \begin{bmatrix} \mathbf{x}_p(i)^H \otimes \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{x}_p(i)^T \otimes \mathbf{I}_r \end{bmatrix}. \quad (\text{B.7})$$

Since $H(0)$ (and hence $H(0)^*$) is full rank, after some manipulation it can be shown that the above condition implies $\mathcal{T}(\mathbf{x}_p(1))\mathbf{s} = \mathbf{0}$. Now consider the contribution of the second pilot transmission $J^p(2)$. $J^p(2)U(\mathbf{H})\mathbf{s} = \mathbf{0}$ implies that

$$K(\mathbf{x}_p(2))U(H(0))\mathbf{s} + K(\mathbf{x}_p(1))U(H(1))\mathbf{s} = \mathbf{0}. \quad (\text{B.8})$$

Since $\mathcal{T}(\mathbf{x}_p(1))\mathbf{s} = \mathbf{0}$ and $U(H(0)), U(H(1))$ have the same structure, it can be shown that $K(\mathbf{x}_p(1))U(H(1))\mathbf{s} = \mathbf{0}$, and, hence, it follows from the above equation that $K(\mathbf{x}_p(2))U(H(0))\mathbf{s} = \mathbf{0}$ which in turn implies that $\mathcal{T}(\mathbf{x}_p(2))\mathbf{s} = \mathbf{0}$ and hence ${}^2V_2U(H(0)) = \mathbf{0}$, and so on. This proves the lemma. \square

Acknowledgment

This work was supported by UC Discovery Grant com04-10176.

References

- [1] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*, Cambridge University Press, 1st edition, 2003.
- [2] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.
- [3] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Transactions on Signal Processing*, vol. 50, no. 5, pp. 1051–1064, 2002.
- [4] F. A. Dietrich and W. Utschick, "Pilot-assisted channel estimation based on second-order statistics," *IEEE Transactions on Signal Processing*, vol. 53, no. 3, pp. 1178–1193, 2005.
- [5] C. Fragouli, N. Al-Dhahir, and W. Turin, "Training-based channel estimation for multiple-antenna broadband transmissions," *IEEE Transactions on Wireless Communications*, vol. 2, no. 2, pp. 384–391, 2003.
- [6] L. Tong and S. Perreau, "Multichannel blind identification: from subspace to maximum likelihood methods," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 1951–1968, 1998.
- [7] J. K. Tugnait and B. Huang, "Multistep linear predictors-based blind identification and equalization of multiple-input multiple-output channels," *IEEE Transactions on Signal Processing*, vol. 48, no. 1, pp. 26–38, 2000.
- [8] D. Pal, "Fractionally spaced semi-blind equalization of wireless channels," in *Proceedings of the 26th Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 642–645, October 1992.
- [9] D. Pal, "Fractionally spaced equalization of multipath channels: a semi-blind approach," in *Proceedings of the 1993 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 12–9, April 1993.
- [10] A. Medles, D. T. M. Slock, and E. de Carvalho, "Linear prediction based semi-blind estimation of MIMO FIR channels," in *Proceedings of the 3rd IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC '01)*, pp. 58–61, Taoyuan, Taiwan, March 2001.
- [11] E. de Carvalho and D. T. M. Slock, "Blind and semi-blind FIR multichannel estimation: (global) identifiability conditions," *IEEE Transactions on Signal Processing*, vol. 52, no. 4, pp. 1053–1064, 2004.
- [12] S. Shahbazpanahi, A. B. Gershman, and J. H. Manton, "Closed-form blind MIMO channel estimation for orthogonal space-time block codes," *IEEE Transactions on Signal Processing*, vol. 53, no. 12, pp. 4506–4517, 2005.
- [13] T. Wo, A. Scherb, P. A. Hoeher, and K. Kammeyer, "Analysis of semi-blind channel estimation for FIR-MIMO systems," in *Proceedings of the 4th International Symposium on Turbo Codes and Related Topics*, Munich, Germany, 2006.
- [14] Y. Hua, "Fast maximum likelihood for blind identification of multiple FIR channels," *IEEE Transactions on Signal Processing*, vol. 44, no. 3, pp. 661–672, 1996.
- [15] D. T. M. Slock, "Blind joint equalization of multiple synchronous mobile users using oversampling and/or multiple antennas," in *Proceedings of the 28th Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1154–1158, 1994.
- [16] N. Ammar and Z. Ding, "On blind channel identifiability under space-time coded transmission," in *Proceedings of the 36th Asilomar Conference on Signals Systems and Computers*, vol. 1, pp. 664–668, November 2002.
- [17] A. L. Swindlehurst and G. Leus, "Blind and semi-blind equalization for generalized space-time block codes," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2489–2498, 2002.
- [18] T. J. Moore, B. M. Sadler, and R. J. Kozick, "Regularity and strict identifiability in MIMO systems," *IEEE Transactions on Signal Processing*, vol. 50, no. 8, pp. 1831–1842, 2002.
- [19] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 10, pp. 1783–1795, 1990.
- [20] A. Gorokhov and P. Loubaton, "Blind identification of MIMO-FIR systems: a generalized linear prediction approach," *Signal Processing*, vol. 73, no. 1-2, pp. 105–124, 1999.
- [21] A. K. Jagannatham and B. D. Rao, "Whitening-rotation-based semi-blind MIMO channel estimation," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 861–869, 2006.
- [22] N. Al-Dhahir and A. H. Sayed, "The finite-length multi-in-put multi-output MMSE-DFE," *IEEE Transactions on Signal Processing*, vol. 48, no. 10, pp. 2921–2936, 2000.

- [23] N. Al-Dhahir and A. H. Sayed, "A system-theoretic foundation for blind equalization of an FIR MIMO channel system," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 49, no. 4, pp. 425–435, 2002.
- [24] A. van den Bos, "A Cramer-Rao lower bound for complex parameters," *IEEE Transactions on Signal Processing*, vol. 42, no. 10, p. 2859, 1994.
- [25] S. M. Kay, *Fundamentals of Statistical Signal Processing. Vol. I: Estimation Theory*, Prentice Hall PTR, 1st edition, 1993.
- [26] A. Medles and D. T. M. Slock, "Augmenting the training sequence part in semiblind estimation for MIMO channels," in *Proceedings of the 37th Asilomar Conference on Signals, Systems and Computers*, pp. 1825–1829, November 2003.
- [27] E. de Carvalho and Dirk T. M. Slock, "Cramer-Rao bounds for semi-blind, blind and training sequence based channel estimation," in *Proceedings of the 1st IEEE Workshop on Signal Processing Advances in Wireless Communications*, pp. 129–132, 1997.
- [28] K. S. Miller, *Complex Stochastic Processes: An Introduction to Theory and Application*, Addison-Wesley, 1st edition, 1974.
- [29] M. F. Siyau, P. Nobles, and R. F. Ormondroyd, "Channel estimation for layered space-time systems," in *Proceedings of the 4th IEEE Workshop on Signal Processing Advances in Wireless Communications*, pp. 482–186, June 2003.
- [30] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, 2003.
- [31] T. Marzetta, "'BLAST training: estimating channel characteristics for high-capacity space- time wireless," in *Proceedings of the 37th Annual Allerton Conference on Communications, Control, and Computing*, pp. 22–24, Monticello, Ill, USA, September 1999.
- [32] J. G. Proakis, *Digital Communications*, McGraw-Hill Higher Education, New York, NY, USA, 2001.

