CONSTRAINED ML ALGORITHMS FOR SEMI-BLIND MIMO CHANNEL ESTIMATION

Aditya K. Jagannatham and Bhaskar D. Rao

Center for Wireless Communication
University of California, SanDiego
9500 Gilman Drive, La Jolla, CA 92093-0407
e-mail: {ajaganna,brao}@ucsd.edu

ABSTRACT

We propose and study algorithms for constrained maximum-likelihood (ML) estimation of a unitary matrix in the context of semi-blind Multi-Input Multi-Output (MIMO) channel estimation. The flat-fading $r \times t$ MIMO channel matrix H for $r \geq t$ can be decomposed as the matrix product $H = WQ^H$, where W is a whitening matrix and Q is a unitary rotation matrix. Exclusive estimation of Q from pilot symbols has been shown to potentially achieve a 3 dB or greater improvement in terms of channel estimation accuracy. We develop and present the OPML, IGML and ROML algorithms for the constrained estimation of the unitary matrix Q that are appropriate for a variety of scenarios, e.g. orthogonal pilots, low complexity etc. Simulation results are provided to demonstrate the efficacy of the algorithms.

Key Words: MIMO, ML, Constrained ML, Unitary, Semi-blind, Channel Estimation.

1. INTRODUCTION

MIMO and smart antenna systems are widely being studied for employment in current and upcoming wireless communication systems. Smart antenna systems, which are built with multiple antennas on receive or transmit side, offer a variety of gains such as improved SNR due to diversity of reception or transmission and also enhanced signal quality from interference suppression. In addition to these, MIMO systems also provide the additional advantage of increased data communication rates for the same SNR by using the multiple spatial multiplexing modes available for communication.

As the number of data channels increases in MIMO systems, the number of associated training streams for the estimation of the channel coefficients increases proportionately which results in reduced spectral efficiency. Moreover, such pilot based techniques tend not to use the statistical information available in unknown data symbols to improve channel estimates. The MIMO channel estimation problem is further complicated because, as the diversity of the MIMO system increases, the SNR (per bit) required to achieve the same system performance (in terms of BER) decreases. The SNR at each antenna is even lower. For instance, employing binary orthogonal FSK modulation and at an operation BER of 2×10^{-3}, while an SNR of 25 dB is required with a single receive antenna, an SNR of 12dB suffices with 4 antennas [1]. Such low SNR environments call for robust channel estimation techniques which use both training and blind data completely.

Semi-blind techniques can potentially enhance the quality of such estimates by making a more complete use of available data. Overhead costs can be reduced by achieving pilot based estimation quality for smaller training symbol pay loads. With a few known training symbols along with blind statistical information, such techniques can avoid the convergence problems associated with completely blind techniques. Early research on semi-blind techniques has been reported in [2]. Extensive work has been done later by Slock et. al. [3], [4] where several semi-blind techniques have been reported.

We utilize the fact that the $r \times t$ MIMO channel matrix H for $r \geq t$ can be decomposed as the product $H = WQ^H$, where W is a whitening matrix and Q is unitary such that $QQ^H = \mathbf{I}$. It is well known that W can be computed blind from the second order statistics of received output data. Training data can then be utilized to estimate only the unitary matrix Q. Significant estimation gains can then be achieved by estimation of such orthogonal matrices which are parameterized by a much fewer number of parameters. However, since Q is a unitary constrained matrix, optimal estimation of Q necessitates the construction of constrained estimators. Such an estimator can be found in [4] for an orthogonal pilot sequence. We refer to this as the OPML estimator and examine its properties. Another salient feature of this work is the development of a novel IGML algorithm for the constrained estimation of Q employing any (not necessarily orthogonal) pilot sequence. Finally we present the ROML algorithm for low complexity implementation of a constrained estimator. Simulation results are presented that

This work was supported by CoRe research grant Cor00-10074.

demonstrate the usefulness of the algorithms developed.

2. PROBLEM FORMULATION

Consider a flat-fading MIMO channel matrix $H \in \mathbb{C}^{r \times t}$ where t is the number of transmit antennas and r is the number of receive antennas in the system, and each h_{ij} represents the flat-fading channel coefficient between the i^{th} receiver and j^{th} transmitter. Denoting the complex received data by $\mathbf{y} \in \mathbb{C}^{r \times 1}$, the equivalent base-band system can be modelled as

$$\mathbf{y}(k) = H\mathbf{x}(k) + \eta(k), \tag{1}$$

where k represents the time instant, $\mathbf{x} \in \mathbb{C}^{t \times 1}$ is the complex transmitted symbol vector. η is additive white Gaussian noise such that $\mathrm{E}\{\eta(k)\eta(l)\} = \delta(k,l)\sigma_n^2\mathbf{I}$ where $\delta(k,l) = 1$ if $k = l$ and 0 otherwise. Also, the sources are assumed to be spatially and temporally independent with identical source power σ_s^2 i.e. $\mathrm{E}\{\mathbf{x}(k)\mathbf{x}(l)\} = \delta(k,l)\sigma_s^2\mathbf{I}$. The signal to noise ratio (SNR) of operation is defined as $SNR \triangleq \frac{\sigma_s^2}{\sigma_n^2}$. Assume that the channel has been used for a total of N symbol transmissions. Out of these N transmissions, the initial L symbols are known training symbols and the observed outputs are thus training outputs. Let $X_p \in \mathbb{C}^{t \times L}$ be defined as $X_p \triangleq [\mathbf{x}(1), \mathbf{x}(2), \ldots, \mathbf{x}(L)]$, by stacking the training symbols. $Y_p \in \mathbb{C}^{r \times L}$ is given by similarly stacking the received training outputs. The remaining $N - L$ information symbols transmitted are termed as 'blind symbols' and their corresponding outputs as 'blind outputs'. $X_b \in \mathbb{C}^{t \times N-L}, Y_b \in \mathbb{C}^{r \times N-L}$ can be defined analogously for the blind symbols. $\{[X_p, Y_p], Y_b\}$ is the complete available data.

3. ESTIMATION STRATEGIES

3.1. Training Based Estimation

H can be estimated exclusively using the pilot X_p given as

$$\hat{H}_{TS} = Y_p X_p^{\dagger}, \tag{2}$$

where $X_p^{\dagger}$ denotes the Moore-Penrose pseudo-inverse of X_p. This qualifies as training based estimation and is simple to implement. However, it results in poor usage of available bandwidth since the pilot itself conveys no source information.

3.2. Semi-Blind Estimation

Now consider a MIMO channel $H \in \mathbb{C}^{r \times t}$ which has at least as many receive antennas as transmit antennas i.e. $r \geq t$. Then, the channel matrix H can be decomposed as $H = WQ^H$ where $W \in \mathbb{C}^{r \times t}$ is also known as the 'whitening' matrix and $Q \in \mathbb{C}^{t \times t}$, termed as the 'rotation' matrix, is unitary i.e. $Q^H Q = QQ^H = \mathbf{I}$. As shown in [5], the matrix W can be estimated from the received data Y_b alone. We therefore employ the pilot information $\{X_p, Y_p\}$ to exclusively estimate the rotation matrix Q. This semi-blind estimation procedure is termed as a *Whitening-Rotation* (WR) scheme. Let the SVD of H be given as $P\Sigma Q^H$. A possible choice for W is given by $W = P\Sigma$ and we assume this specific choice in the rest of the work. We present next a list of potential assumptions which are employed as appropriate in subsequent parts of the work.

A.1 $W \in \mathbb{C}^{r \times t}$ is perfectly known at the output.

A.2 $X_p \in \mathbb{C}^{t \times L}$ is orthogonal i.e. $X_p X_p^H = \sigma_s^2 L\, \mathbf{I}_{t \times t}$.

A.1 is reasonable if we assume the transmission of a long data stream ($N \rightarrow \infty$) from which W can be estimated with considerable accuracy and A.2 can be easily achieved by using a an integer orthogonal structure such as the Hadamard matrix.

The WR technique potentially improves estimation accuracy because the matrix Q by virtue of its unitary constraint is parameterized by a fewer number of parameters and hence can be determined with greater accuracy from the limited pilot data X_p, Y_p. Indeed, it has been shown in [6],[7] that under A.1 and A.2, the gain of the semi-blind algorithm (in dB) in terms of MSE of estimation is $10\log_{10}\left(\frac{2r}{t}\right)$. Thus for a size 8×4 complex channel matrix H, i.e. $H \in \mathbb{C}^{8 \times 4}$, the estimation gain of the semi-blind technique is 6 dB which represents a significant improvement over the conventional technique described in (2). However, estimation of the unitary constrained matrix Q necessitates the development of constrained algorithms which are presented in the next section.

4. ALGORITHMS

4.1. Orthogonal Pilot ML (OPML) estimator

$\hat{Q} : \mathbb{C}^{r \times L} \rightarrow \mathcal{S}$, where $\hat{Q}$ is the constrained ML estimator of Q and $\mathcal{S}$ is the manifold of unitary matrices, is obtained by minimizing the likelihood

$$\left\|Y_p - WQ^H X_p\right\|^2 \quad \text{such that} \quad QQ^H = \mathbf{I}. \tag{3}$$

Let $\mathcal{M} \triangleq W^H Y_p X_p^H$. We then have the following result for the constrained estimation of Q.

Lemma 1. *Under A.1 and A.2, $\hat{Q}$ the constrained OPML estimate of Q that minimizes the cost function in (3) is given by*

$$\hat{Q} = V_{\mathcal{M}} U_{\mathcal{M}}^H \; \textit{where}, \; U_{\mathcal{M}} \Sigma_{\mathcal{M}} V_{\mathcal{M}}^H = \mathrm{SVD}\left(\mathcal{M}\right). \tag{4}$$

Proof. This technique has been proposed and proved in [4]. □

Since this procedure employs A.2 (orthogonal pilot), it is termed as the OPML estimate. The above expression (4) thus yields a closed form expression for the computation of $\hat{Q}$, the ML estimate of Q. The channel matrix H is then estimated as $\hat{H} = W\hat{Q}^H$.

4.1.1. Properties of the OPML Estimator:

In this section we discuss properties of the OPML estimator. We show that the estimator is biased and hence does not achieve the CRB for finite sample length. However, from the properties of ML estimators, it achieves the CRB asymptotically as the sample length increases. Further, it is also shown in this section that the bound is achieved for all sample lengths at high SNR .

P.1 There does not exist a finite length constrained unbiased estimator of the rotation matrix Q and hence $\hat{Q}$, the OPML estimator of Q is biased.

Proof. Let there exist $\hat{Q}$ such that $\hat{Q} : \mathbb{C}^{r\times L} \rightarrow \mathcal{S}$ is a constrained unbiased estimator of Q. $\mathbb{C}^{r\times L}$ is the observation space (Y_b) and $\mathcal{S}$ is the manifold of orthogonal matrices. Then $\hat{Q} = Q + \bar{E}$ where $\bar{E}$ is such that $\mathrm{E}\left\{\bar{E}\right\} = \mathbf{0}$. Now since $\hat{Q}$ is a constrained estimator we have $\hat{Q}\hat{Q}^H = \mathbf{I}$ and therefore,

$$\left(Q + \bar{E}\right)^H \left(Q + \bar{E}\right) = \mathbf{I},$$

which when simplified using the fact that $QQ^H = \mathbf{I}$ yields

$$Q^H\bar{E} + \bar{E}^H Q + \bar{E}\bar{E}^H = \mathbf{0}.$$

Rearranging terms in the above expression and taking the expectation of quantities on both sides (where the expectation is with respect to the distribution of E conditioned on Q) yields

$$\mathrm{tr}\left(Q^H\,\mathrm{E}\left\{\bar{E}\right\} + \mathrm{E}\left\{\bar{E}\right\}^H\,Q\right) = -\mathrm{tr}\left(\mathrm{E}\left\{\left\|\bar{E}\right\|^2\right\}\right). \tag{5}$$

It can immediately be observed that the right hand side is strictly less than 0 while the left hand side is equal to zero (by virtue of $\mathrm{E}\left\{\bar{E}\right\} = \mathbf{0}$) and hence the contradiction. ☐

The above result then implies that the CRB cannot be achieved in a general scenario as there does not exist an unbiased estimator which is necessary for the achievement of the CRB. However, the properties presented next guarantee the asymptotic achievability of the CRB both in sample length and SNR.

P.2 The OPML estimator achieves the CRB as the pilot sequence length $L \rightarrow \infty$.

Proof. Follows from the asymptotic property of ML estimators, reviewed in [8]. ☐

P.3 The OPML estimator of Q achieves the CRB at high SNR, i.e. as $\frac{\sigma_s^2}{\sigma_n^2} \rightarrow \infty$.

Proof. The above result can be proved using the theory of matrix eigenspace perturbation analysis detailed in [9]. A complete proof can be found in [7]. ☐

4.2. Iterative ML procedure for general pilot - IGML

In this section we present the IGML algorithm to compute the estimate for any given pilot sequence X_p, i.e. when A.2 does not necessarily hold. As it is shown later, the proposed IGML scheme reduces to the OPML under A.2. The ML cost-function to be minimized is given as in (3). Let A.1 hold true and $\tilde{Y}_p \triangleq P^H Y_p$. The Lagrange cost to be minimized can then be formulated as

$$\sum_{i=1}^{t}\left\|\tilde{Y}_p(i) - \sigma_i \mathbf{q}_i^H X_p\right\|^2 + \sum_{i=1}^{t}\mathrm{Re}\left\{\lambda_i\left(\mathbf{q}_i^H\mathbf{q}_i - 1\right)\right\} + \sum_{i=1}^{t}\sum_{j=i+1}^{t}\mathrm{Re}\left\{\mu_{ij}\left(\mathbf{q}_j^H\mathbf{q}_i\right)\right\},$$

where $\lambda_i \in \mathcal{R}$, $\mu_{ij} \in \mathcal{C}$ are the Lagrange multipliers, $\tilde{Y}_p(i) \in \mathbb{C}^{1\times L}$ is the i-th row (output at the i-th receiver) and $\mathbf{q}_i$ is the i-th column of Q for $1 \leq i,j \leq t$. Define the matrix of Lagrange multipliers $S \in \mathbb{C}^{t\times t}$ as $S_{ii} \triangleq \lambda_i$, $S_{ij} \triangleq \mu_{ij}/2$ if $i < j$ and $S_{ij} \triangleq \mu_{ji}^*/2$ if $i > j$. Observe that S is a hermitian symmetric matrix, i.e. $S = S^H$. The above cost function can now be differentiated with respect to $\mathrm{Re}\left\{\mathbf{q}_i\right\}, \mathrm{Im}\left\{\mathbf{q}_i\right\}$ for $1 \leq i \leq t$. These quantities can then be equated to 0 for extrema and after some manipulation, the resulting equations can be represented in terms of complex matrices as

$$X_p\tilde{Y}_p^H\Sigma - X_pX_p^HQ\Sigma^2 = QS, \tag{6}$$

where Q is unitary. We avoid repeated mention of this constraint in the foregoing analysis and it is implicitly assumed to hold. Let $\mathcal{A} \triangleq X_p\tilde{Y}_p^H\Sigma = X_pY_p^HW$. After some manipulations, the resulting equation can be written as

$$Q^H\mathcal{T} = \mathcal{T}^HQ.$$

where $\mathcal{T} \triangleq \mathcal{A} + \left(L\,\sigma_s^2\,\mathbf{I}_{t\times t} - X_pX_p^H\right)Q\Sigma^2$. Thus from the above equation, $Q^H\mathcal{T}$ is hermitian symmetric or in other words $Q^H\mathcal{T} = \mathcal{T}^HQ$. Also, if $U_\mathcal{T}\Lambda_\mathcal{T}V_\mathcal{T}^H = \mathrm{SVD}\left(\mathcal{T}\right)$ then, $Q^HU_\mathcal{T}\Lambda_\mathcal{T}V_\mathcal{T}^H = \mathrm{SVD}\left(Q^H\mathcal{T}\right)$. We have then from the symmetry of $Q^H\mathcal{T}$,

$$Q^HU_\mathcal{T} = V_\mathcal{T} \Rightarrow Q = U_\mathcal{T}V_\mathcal{T}^H. \tag{7}$$

(7) gives the critical step in the iterative algorithm which is succinctly presented below. Some of the definitions above are repeated for the sake of completeness.

IGML Algorithm: Let A.1 hold, i.e. $\hat{W} = W = P\Sigma$. X_p is the transmitted pilot symbol sequence and not necessarily orthogonal. We then compute the constrained ML estimate of $\hat{Q}$ as follows.

- S.1 Compute $\mathcal{A} = X_p Y_p^H W$, where Y_p is the received output data.
- S.2 Let $\hat{Q}_0$ denote the initial estimate of the unitary matrix Q. Compute $\hat{Q}_0$ by employing X_p , W and Y_p in (4). Let $\mathcal{B} \triangleq \left(L\,\sigma_s^2\,\mathbf{I}_{t\times t} - X_p X_p^H\right)$.
- S.3 Repeat for $\mathcal{N}$ iterations. At the k^{th} iteration i.e.$1 \leq k \leq \mathcal{N}$,
  - S.3.1 Compute $\mathcal{T}_k = \mathcal{A} + \mathcal{B}\hat{Q}_{k-1}\Sigma^2$.
  - S.3.2 Compute refined estimate of $\hat{Q}_k$ from $\mathcal{T}_k$ by employing (7).
- S.4 Compute the estimate of H as $\hat{H} = W\hat{Q}^H$.

$\mathcal{N}$, the number of iterations is small and typically $\mathcal{N} \leq 5$ as found in our simulations. It can now also be noticed that if A.2 holds, $X_p X_p^H = L\,\sigma_s^2\,\mathbf{I}$. Therefore, $\mathcal{T} = \mathcal{A} = X_p Y_p^H W$. The SVD of $\mathcal{T}$ is then given by $U_{\mathcal{T}}\Lambda_{\mathcal{T}}V_{\mathcal{T}}^H = V_{\mathcal{M}}\Sigma_{\mathcal{M}}U_{\mathcal{M}}^H$. It follows that the IGML solution given as

$$\hat{Q} = U_{\mathcal{T}}V_{\mathcal{T}}^H = V_{\mathcal{M}}U_{\mathcal{M}}^H \tag{8}$$

is similar to the solution given in (4). Thus, when X_p is orthogonal, the IGML algorithm converges in a single iteration to the OPML solution.

4.2.1. 'Rotation-Optimization' ML (ROML)

The above suggested IGML scheme to compute $\hat{Q}$ for a general pilot sequence X_p might be computationally complex owing to the SVD computations involved. Thus, to avoid the complexity involved in the full computation of the optimal ML solution, we propose a simplistic ROML procedure for the sub-optimal estimation of Q, thus trading complexity for optimality. The first step of ROML involves construction of a modified cost function as

$$\min_{Q} \left\|\tilde{W}Y_p - Q^H X_p\right\|^2 \quad \text{where} \quad QQ^{\mathrm{H}} = \mathbf{I}. \tag{9}$$

$\bar{Y}_p = \tilde{W}Y_p$ is the whitening pre-equalized data. Several choices can then be considered for the pre-equalization filter $\tilde{W}$. The standard Zero-Forcing (ZF) equalizer is given by $\tilde{W}_{ZF} = W^\dagger$ (where † denotes the Moore-Penrose pseudo-inverse). Alternatively, a robust MMSE pre-filter is given as $\tilde{W}_{MMSE} = \sigma_s^2 W^H\left(\sigma_s^2 WW^H + \sigma_n^2\mathbf{I}\right)^{-1}$. Defining $\mathcal{D} \triangleq \tilde{W}Y_p X_p^H$, the cost minimizing $\hat{Q}$ for the modified cost in (9) is given as

$$\hat{Q} = V_{\mathcal{D}}U_{\mathcal{D}}^H \text{ where } U_{\mathcal{D}}S_{\mathcal{D}}V_{\mathcal{D}}^H = \mathrm{SVD}\left(\mathcal{D}\right). \tag{10}$$

This result for problem (9) follows by noting its similarity to problem (3). However, the resulting estimate does not have any statistical optimality properties as it does not compute the solution to the true cost function given in (3). This estimate of Q can now be employed to initialize the IGML procedure to minimize the true cost. However, to avoid the complexity associated with an SVD computation, a constrained minimization procedure (ex: `'fmincon'` in MATLAB) can now employed to converge to the solution with the t^2 non-linear constraints given by the unit norm and mutual orthogonality of the rows of Q. This procedure then yields $\hat{Q}$ which is close to the optimal ML estimate and the low computational cost of the proposed solution makes it attractive to implement in practical systems.

5. SIMULATION RESULTS

Our simulation set-up consists of a 8×4 MIMO channel H (i.e. $r = 8, t = 4$). H was generated as a matrix of zero-mean circularly symmetric complex Gaussian random entries such that the sum variance of the real and imaginary parts was unity. Orthogonal pilot sequences are constructed using the Hadamard structure. For a general pilot sequence and data vectors, symbols were drawn from a 16-QAM signal constellation. Noise vectors $\eta(k)$ were generated as spatio-temporally uncorrelated complex Gaussian random vectors and with variance of each element equal to σ_n^2.

Experiment 1: We evaluate the MSE performance of the different constrained ML estimators of Q under A.1 and compare it to the training based estimate given by (2). A statistically white pilot $\left(\mathrm{E}\left\{X_p X_p^H\right\} = L\sigma_s^2\mathbf{I}\right)$ was employed for the IGML, ROML and training based schemes while an orthogonal pilot X_p was used for the OPML scheme with $X_p X_p^H = L\sigma_s^2\mathbf{I}$, thus maintaining constant source power. The MSE of estimation of the channel matrix H has been computed for different pilot lengths L in the range $20 \leq L \leq 100$. Figure 1 shows the error for these different schemes and also that for the exclusive training based scheme. It can be seen that the semi-blind schemes are 6dB more efficient than the training scheme.

Experiment 2: Finally, we consider P_e of detection of the transmitted symbol vectors employing $\hat{H}$ estimated from

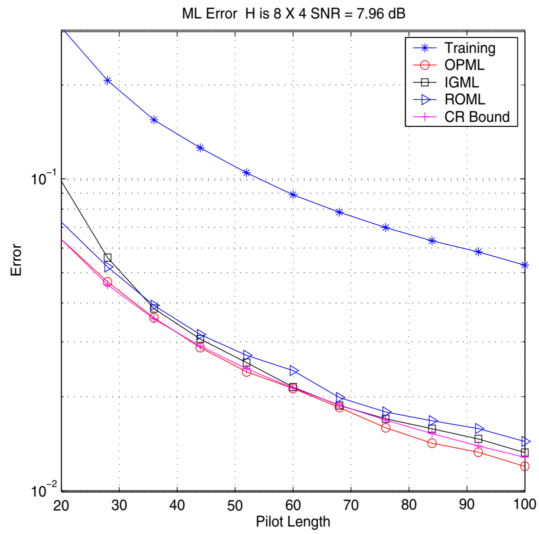

Fig. 1. Computed MSE Vs SNR.

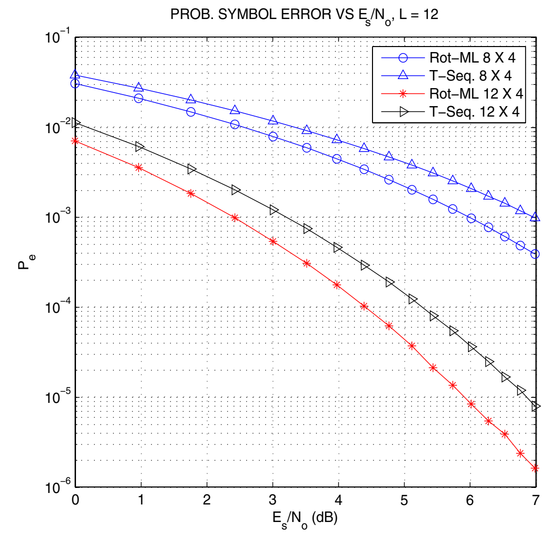

Fig. 2. Probability of Bit Error Vs SNR

different schemes. We compare OPML at the receiver vs the training based estimate of H for a pilot sequence of length $L = 12$ symbols. Figure (5) shows the probability of error detection vs SNR for a linear MMSE receiver at the output for both a 8×4 and 12×4 system H. It can be seen that at an SNR of 6 dB the semi-blind schemes achieves about a 1 dB improvement in probability of bit error detection performance over the exclusive training based estimate.

6. CONCLUSIONS

We have presented algorithms for constrained ML estimation of a unitary matrix in the context of semi-blind MIMO channel estimation. Properties of the OPML algorithm for the constrained ML estimation of Q were examined. The IGML has been presented for the estimation of Q from a general pilot sequence X_p and the ROML as a low complexity alternative.

7. REFERENCES

[1] J.G. Proakis, *Digital Communications*, McGraw-Hill Higher Education, NewYork, 2001.

[2] D. Pal, “Fractionally spaced semi-blind equalization of wireless channels,” *The Twenty-Sixth Asilomar Conference*, vol. 2, pp. 642–645, 1992.

[3] E. De Carvalho and D.T.M.Slock, “Asymptotic performance of ML methods for semi-blind channel estimation,” *Thirty-First Asilomar Conf.*, pp. 1624–8, 1998.

[4] A.Medles, D.T.M. Slock, and E.D.Carvalho, “Linear prediction based semi-blind estimation of MIMO FIR channels,” *Third IEEE SPAWC, Taiwan*, 2001.

[5] A. K. Jagannatham and B. D. Rao, “A semi-blind technique for MIMO channel matrix estimation,” in *Proc. of IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2003) , # 582*, Rome, Italy, 2003.

[6] A. K. Jagannatham and B. D. Rao, “Complex constrained CRB and its application to semi-blind MIMO and OFDM channel estimation,” in *Proc. of the IEEE SAM Workshop, 2004*, Sitges, Barcelona.

[7] A. K. Jagannatham and B. D. Rao, “Whitening rotation based semi-blind MIMO channel estimation,” *Submitted to IEEE Transactions on Signal Processing*.

[8] S. M. Kay, *Fundamentals of Statistical Signal Processing,Vol I: Estimation Theory*, Prentice Hall PTR, first edition, 1993.

[9] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford University Press, Walton St., Oxford, first edition, 1965.