

# Iterative Estimation of Sparse and Doubly-Selective Multi-input Multi-output (MIMO) Channel

Jun Won Choi, Kyeongyeon Kim, Thomas J. Riedl, and Andrew C. Singer

\*University of Illinois at Urbana-Champaign

1308 W. Main, Urbana, 61801, Illinois, USA, Email: jwchoi@uiuc.edu

**Abstract**—The estimation of doubly-selective channels is challenging since long channel impulse response should be estimated with a fast tracking speed. Provided that a structure of the channel response is sparse, i.e., only a few of channel gains are nonzero, a tracking performance of the channel estimator can be improved significantly by avoiding estimation of zero taps. In this paper, we study estimation of fast time-varying and long reverberant channels that have a sparse structure in multi-input multi-output (MIMO) systems. In order to exploit the sparse structure, we parameterize the locations of nonzero taps using a binary vector and incorporate it into the state-space system built upon auto-regressive (AR) modeling of the time-varying channel gains. Then, we derive a joint estimate of the binary vector and channel gains based on maximum likelihood (ML) criterion. Expectation maximization (EM) algorithm is derived to find a sparse structure and channel gains iteratively. According to the simulation study performed over MIMO Rician fading channels, the proposed sparse channel estimation technique outperforms the previous estimation schemes, especially when Doppler rate is high.

## I. INTRODUCTION

In most of wireless communication systems, multi-path channels can be modeled as linear time-variant system. Estimation of such a time-varying channel impulse response is often required for various purposes, e.g., channel equalization and synchronization. An adaptive filtering technique such as recursive least square (RLS) algorithm [9] are a popular method for estimating time-varying channels. However, a tracking performance of such adaptive channel estimation algorithms degrades for high Doppler and large delay spread channels [1], [2]. Fortunately, some of such doubly-spread channels exhibit the *sparse structure*, i.e., only a few taps of channel are non-zero, due to large bandwidth transmission or particular geometrical landscape. Intuitively, by exploiting the sparse structure of channels, both tracking and error performance of channel estimation can be improved by avoiding to estimate zero taps [3].

So far, there have been several techniques that exploit the sparse structure for channel estimation. In [4], a auxiliary adaptive filter is employed to estimate a full-tap channel and zero taps are identified via thresholding. Then, only non-zero channel taps are estimated using the filter with shorter length. In [3], the detection of active tap locations was derived based on a mean square error (MSE) criterion under the assumption of pre-whitened input signal. On the other hand, various basis selection techniques for finding *sparse representation* of signal have been applied to sparse channel

estimation. In [5], *matching pursuit* (MP) has been applied to estimate taps that are maximally correlated with received signal sequentially, and more elaborated sparse signal recovery techniques, called *compressed sensing*, have been applied in [6], [7]. Recently, sparse channel estimation was formulated as a detection problem for on-off keying in [8]. These approaches are based on the assumption that the channel is locally time-invariant so that they often fail to cope with estimation of rapidly-varying channels with high Doppler spread.

In this paper, we propose a new sparse channel estimation technique, called iterative sparse (IS) channel estimator, which detects a sparse channel structure as well as time-varying channel gains in an iterative fashion. In order to model the sparse channels, the nonzero tap locations are parameterized by a binary vector. This parameter is incorporated in the state-space representation which models a dynamic of time-varying channel gains based on an auto regressive (AR) model. The maximum likelihood (ML) estimate of the binary vector is derived using expectation-maximization (EM) formulation. The resulting sparse channel estimation algorithm consists of Kalman smoothing [10] (E-step) which finds the *a posteriori* channel estimate given the sparse structure and binary integer least square search (M-step) which estimates the sparse structure. These two steps are performed iteratively. In the derivation, we assume that the sparse structure is locally time-invariant while channel gains constantly vary according to AR model. As a result, the iteration is performed over a block of data where a sparse structure does not change, and hence, undesired noise and channel variations within the block are averaged out, producing relatively reliable sparse structure estimate.

## II. ITERATIVE SPARSE (IS) CHANNEL ESTIMATION

### A. Sparse Channel Model

In the discrete-time and base-band system model, the  $n$ th signal sample transmitted through the time-variant channel can be expressed as

$$y_n = \sum_{k=-K_f}^{K_p} h_{n,k} x_{n-k} + w_n, \quad (1)$$

where  $x_n$  and  $w_n$  are the  $n$ th input and noise sample and  $h_{n,k}$  is the complex channel impulse response that includes the effect of pulse-shaping, channel, and front-end processing.

This model can be extended to the multi-input multi-output (MIMO) case with  $n_t$  transmit and  $n_r$  receive antennas, i.e.,

$$\begin{bmatrix} y_n^1 \\ \vdots \\ y_n^{n_r} \end{bmatrix} = \sum_{k=-K_f}^{K_p} \begin{bmatrix} h_{n,k}^{1,1} & \cdots & h_{n,k}^{1,n_t} \\ \vdots & \ddots & \vdots \\ h_{n,k}^{n_r,1} & \cdots & h_{n,k}^{n_r,n_t} \end{bmatrix} \begin{bmatrix} x_{n-k}^1 \\ \vdots \\ x_{n-k}^{n_t} \end{bmatrix} + \begin{bmatrix} w_n^1 \\ \vdots \\ w_n^{n_r} \end{bmatrix}, \quad (2)$$

where the superscript used in this notation represents antenna index. Define  $\mathbf{h}_n^{i,j} = [h_{n,-K_f}^{i,j} \cdots h_{n,K_p}^{i,j}]^T$ , that is, a channel impulse response (CIR) from the  $j$ th transmit to the  $i$ th receive antenna. In order to describe a sparse structure of the channel, we represent a channel tap as a product of binary value  $a_{n,k}^{i,j}$  and complex channel gain  $g_{n,k}^{i,j}$ , i.e.,  $h_{n,k}^{i,j} = a_{n,k}^{i,j} g_{n,k}^{i,j}$ . Note that  $a_{n,k}^{i,j} = 1$  for non-zero taps and  $a_{n,k}^{i,j} = 0$  for zero taps. Hence, the CIR from the  $j$ th transmit to the  $i$ th receive antenna is expressed as

$$\mathbf{h}_n^{i,j} = \text{diag}(\mathbf{a}_n^{i,j}) \mathbf{g}_n^{i,j}, \quad (3)$$

where  $\mathbf{a}_n^{i,j} = [a_{n,-K_f}^{i,j} \cdots a_{n,K_p}^{i,j}]^T$  and  $\mathbf{g}_n^{i,j} = [g_{n,-K_f}^{i,j} \cdots g_{n,K_p}^{i,j}]^T$ . We assume that the number of non-zero taps for  $\mathbf{h}_n^{i,j}$  is less than  $K_{i,j}$ , i.e.,  $\|\mathbf{a}_n^{i,j}\|^2 \leq K_{i,j}$ .

The sparsity vector  $\mathbf{a}_n^{i,j}$  are assumed to be locally time-invariant so that during the period of  $T$  symbols,  $n = Tt + 1, \dots, Tt + T$ , it holds that  $\mathbf{a}_{Tt+1}^{i,j} = \cdots = \mathbf{a}_{Tt+T}^{i,j} = \mathbf{a}_t^{i,j}$ , where  $t$  denotes the index of block. From (2) and (3), the received signal is expressed in vector form as

$$\mathbf{y}_n = \mathbf{A}_t \text{diag}(\mathbf{x}_n) \mathbf{g}_n + \mathbf{w}_n, \quad (4)$$

where

$$\mathbf{g}_n = \begin{bmatrix} \mathbf{g}_n^{1,1} \\ \vdots \\ \mathbf{g}_n^{1,n_t} \\ \vdots \\ \mathbf{g}_n^{n_r,1} \\ \vdots \\ \mathbf{g}_n^{n_r,n_t} \end{bmatrix}, \quad \mathbf{x}_n = \mathbf{1}_{n_r} \otimes \begin{bmatrix} x_{n-K_f}^1 \\ \vdots \\ x_{n+K_p}^1 \\ \vdots \\ x_{n-K_f}^{n_t} \\ \vdots \\ x_{n+K_p}^{n_t} \end{bmatrix}$$

$$\mathbf{y}_n = \begin{bmatrix} y_n^1 \\ \vdots \\ y_n^{n_r} \end{bmatrix}, \quad \mathbf{w}_n = \begin{bmatrix} w_n^1 \\ \vdots \\ w_n^{n_r} \end{bmatrix},$$

and  $\mathbf{1}_{n_r}$  is the  $n_r \times n_r$  matrix whose elements are one. We assume that  $\mathbf{w}_n$  is uncorrelated Gaussian, i.e.,  $E[\mathbf{w}_n \mathbf{w}_n^H] = N_0 \mathbf{I}_{n_r}$ . For convenience, we let  $N = n_r n_t (K_f + K_p + 1)$  and  $N_t = n_t (K_f + K_p + 1)$ . For the  $t$ th block, the  $n_r$ -by- $N$  matrix  $\mathbf{A}_t$  is written

$$\mathbf{A}_t = \begin{bmatrix} \left( \mathbf{a}_t^{1,1} \right)^T \cdots \left( \mathbf{a}_t^{1,n_t} \right)^T & \mathbf{0} & \mathbf{0}_{1 \times N_t} \\ \mathbf{0}_{n_r \times N_t} & \ddots & \mathbf{0}_{n_r \times N_t} \\ \mathbf{0}_{1 \times N_t} & \mathbf{0} & \left( \mathbf{a}_t^{n_r,1} \right)^T \cdots \left( \mathbf{a}_t^{n_r,n_t} \right)^T \end{bmatrix}. \quad (5)$$

From now on, we will use the notation  $\mathbf{A}_t(\theta_t)$  by introducing the parameter

$$\theta_t = \left[ \left( \mathbf{a}_t^{1,1} \right)^T \cdots \left( \mathbf{a}_t^{1,n_t} \right)^T, \left( \mathbf{a}_t^{2,1} \right)^T \cdots \left( \mathbf{a}_t^{2,n_t} \right)^T, \right. \quad (6)$$

$$\left. \cdots \left( \mathbf{a}_t^{n_r,1} \right)^T \cdots \left( \mathbf{a}_t^{n_r,n_t} \right)^T \right]^T. \quad (7)$$

The channel gain vector  $\mathbf{g}_n$  is modeled by a  $p$ -th order autoregressive (AR) process, i.e.,

$$\mathbf{g}_n = \sum_{i=1}^p \mathbf{F}_i \mathbf{g}_{n-i} + \mathbf{v}_n. \quad (8)$$

Then, from (4) and (8), we can provide the state-space model for the  $t$ th block as

$$\mathbf{z}_n = \mathbf{F} \mathbf{z}_{n-1} + \begin{bmatrix} \mathbf{v}_n \\ \mathbf{0}_{(p-1)N \times 1} \end{bmatrix} \quad (9)$$

$$\mathbf{y}_n = \mathbf{C}(\theta_t) \mathbf{z}_n + \mathbf{w}_n, \quad (10)$$

where  $\mathbf{z}_n = [\mathbf{g}_n^T, \dots, \mathbf{g}_{n-p+1}^T]^T$ ,

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_1 & \cdots & \mathbf{F}_{p-1} & \mathbf{F}_p \\ & & \mathbf{I}_{(p-1)N} & \mathbf{0}_{(p-1)N \times N} \end{bmatrix}, \quad (11)$$

$$\mathbf{C}(\theta_t) = \mathbf{A}_t(\theta_t) \text{diag}(\mathbf{x}_n) \begin{bmatrix} \mathbf{I}_N & \mathbf{0}_{N \times (p-1)N} \end{bmatrix}, \quad (12)$$

and  $\mathbf{v}_n$  is zero-mean Gaussian process with  $E[\mathbf{v}_n \mathbf{v}_n^H] = \mathbf{V}_n$  and  $E[\mathbf{v}_n \mathbf{v}_m^H] = \mathbf{0}$  for  $n \neq m$ . Notice that the system matrices  $\mathbf{F}_1, \dots, \mathbf{F}_p$  for the  $p$ -th order AR model account for correlations in time and over different transmit and receive antennas.

## B. Algorithm Derivation

In this section, we aim to derive an ML estimate of the parameter  $\theta_t$  given the sequence of the  $t$ th block data, i.e.,  $\mathbf{y}_{tT+1}, \dots, \mathbf{y}_{tT+T}$ . The ML estimate  $\theta_t^{\text{ML}}$  is given by

$$\theta_t^{\text{ML}} = \arg \max_{\theta_t \in \Omega} \ln Pr(\mathbf{y}_{1:T}; \theta_t) \quad (13)$$

where the subscript  $1 : T$  denotes the set of symbol indices within the  $t$ th block and the set  $\Omega$  is defined as  $\{\theta_t : \|\mathbf{a}_t^{i,j}\|^2 \leq K_{i,j}\}$ . By considering  $(\mathbf{y}_{1:T}, \mathbf{g}_{1:T})$  as a *complete data*, we can formulate the EM algorithm for finding the solution to (13) as

- Expectation step

$$Q(\theta_t; \theta_t^{(k)}) = E_{\mathbf{g}_{1:T}} \left[ \ln Pr(\mathbf{y}_{1:T}, \mathbf{g}_{1:T} | \theta_t; \theta_t^{(k)}) \right] \quad (14)$$

- Maximization step

$$\theta_t^{(k+1)} = \arg \max_{\theta_t \in \Omega} Q(\theta_t; \theta_t^{(k)}), \quad (15)$$

where  $\theta_t^{(k)}$  is the value of  $\theta_t$  after the  $k$ th iteration. The steps in (14) and (15) are repeated until the estimate  $\theta_t^{(k)}$  reaches a stationary point.

1) *Expectation Step*: By invoking Bayes formula, the probability  $Pr(\mathbf{g}_{1:T}, \mathbf{y}_{1:T} | \theta_t)$  is expressed as

$$\ln Pr(\mathbf{g}_{1:T}, \mathbf{y}_{1:T}; \theta_t) \quad (16)$$

$$= \ln Pr(\mathbf{y}_{1:T} | \mathbf{g}_{1:T}; \theta_t) + \ln Pr(\mathbf{g}_{1:T}) \quad (17)$$

$$= \sum_{n=tT+1}^{tT+T} -\frac{1}{N_0} \|\mathbf{y}_n - \mathbf{A}_t(\theta_t) \mathbf{X}_n \mathbf{g}_n\|^2 + \ln Pr(\mathbf{g}_{1:T}) + C, \quad (18)$$

where  $\mathbf{X}_n = \text{diag}(\mathbf{x}_n)$ . Since the second and third terms in (18) do not relate to  $\theta_t$ , we have

$$Q(\theta_t; \theta_t^{(k)}) \quad (19)$$

$$= \sum_{n=tT+1}^{tT+T} E \left[ -\frac{1}{N_0} \|\mathbf{y}_n - \mathbf{A}_t(\theta_t) \mathbf{X}_n \mathbf{g}_n\|^2 | \mathbf{y}_{1:T}; \theta_t^{(k)} \right] + C' \quad (20)$$

$$= \text{tr} \left( 2\Re \left( \sum_{n=tT+1}^{tT+T} (\mathbf{y}_n \mathbf{m}_n^H \mathbf{X}_n^H) \mathbf{A}_t(\theta_t)^H \right) \right) - \text{tr} \left( \mathbf{A}_t(\theta_t) \left( \sum_{n=tT+1}^{tT+T} \mathbf{X}_n (\boldsymbol{\Sigma}_n + \mathbf{m}_n \mathbf{m}_n^H) \mathbf{X}_n^H \right) \mathbf{A}_t(\theta_t)^H \right) + C', \quad (21)$$

where  $\hat{\boldsymbol{\Sigma}}_n = \text{Cov}[\mathbf{g}_n \mathbf{g}_n^H | \mathbf{y}_{1:T}; \theta_t^{(k)}]$  and  $\mathbf{m}_n = E[\mathbf{g}_n | \mathbf{y}_{1:T}; \theta_t^{(k)}]$ . The conditional mean and variance,  $\mathbf{m}_n$  and  $\boldsymbol{\Sigma}_n$  can be obtained by fixed-interval Kalman smoothing algorithm. The forward and backward recursion rules of Rauch-Tung-Streibel (RTS) Kalman smoother [10] are summarized as follows;

- Forward recursion rule

$$\hat{\mathbf{m}}_{n+1} = \mathbf{F} \hat{\mathbf{m}}_n + \mathbf{K}_{n+1} (\mathbf{y}_{n+1} - \mathbf{A}_t(\theta_t^{(k)}) \mathbf{X}_{n+1} \hat{\mathbf{m}}_n) \quad (22)$$

$$\hat{\boldsymbol{\Sigma}}_{n+1} = (\mathbf{I} - \mathbf{K}_{n+1} \mathbf{A}_t(\theta_t^{(k)}) \mathbf{X}_{n+1}) \mathbf{Q}_{n+1}, \quad (23)$$

where  $\hat{\boldsymbol{\Sigma}}_n = \text{Cov}[\mathbf{g}_n \mathbf{g}_n^H | \mathbf{y}_{1:n}; \theta_t^{(k)}]$ ,  $\hat{\mathbf{m}}_n = E[\mathbf{g}_n | \mathbf{y}_{1:n}; \theta_t^{(k)}]$ ,  $\mathbf{Q}_{n+1} = (\mathbf{F} \hat{\boldsymbol{\Sigma}}_n \mathbf{F}^H + \mathbf{V}_n)$ , and the Kalman gain  $\mathbf{K}_n$  is expressed as

$$\mathbf{K}_n = \mathbf{Q}_n \mathbf{X}_n^H \mathbf{A}_t(\theta_t^{(k)})^H \quad (24)$$

$$\cdot \left[ \mathbf{A}_t(\theta_t^{(k)}) \mathbf{X}_n \mathbf{Q}_n \mathbf{X}_n^H \mathbf{A}_t(\theta_t^{(k)})^H + N_0 \mathbf{I} \right]^{-1}. \quad (25)$$

- Backward recursion rule

$$\mathbf{m}_n = \hat{\mathbf{m}}_n + \mathbf{S}_n (\mathbf{m}_{n+1} - \mathbf{F} \hat{\mathbf{m}}_n) \quad (26)$$

$$\boldsymbol{\Sigma}_n = \hat{\boldsymbol{\Sigma}}_n + \mathbf{S}_n (\boldsymbol{\Sigma}_{n+1} - \mathbf{Q}_{n+1})^H \mathbf{S}_n^H, \quad (27)$$

where  $\mathbf{S}_n = \hat{\boldsymbol{\Sigma}}_n \mathbf{F} \mathbf{Q}_{n+1}^{-1}$ .

We first compute  $\hat{\mathbf{m}}_n$  and  $\hat{\boldsymbol{\Sigma}}_n$  from  $n = tT + 1$  to  $tT + T$  recursively and then obtain  $\mathbf{m}_n$  and  $\boldsymbol{\Sigma}_n$  using them in a reverse

order. Notice that the initial value of  $\hat{\mathbf{m}}_n$  and  $\hat{\boldsymbol{\Sigma}}_n$  can be taken from the last value in the previous block.

2) *Maximization step*: In the maximization step, We maximize  $Q(\theta_t; \theta_t^{(k)})$  over  $\theta_t$ . For the sake of brevity, we let

$$\mathbf{D} = \sum_{n=tT+1}^{tT+T} (\mathbf{y}_n \mathbf{m}_n^H \mathbf{X}_n^H) \quad (28)$$

$$\mathbf{B} = \sum_{n=tT+1}^{tT+T} (\mathbf{X}_n (\boldsymbol{\Sigma}_n + \mathbf{m}_n \mathbf{m}_n^H) \mathbf{X}_n^H). \quad (29)$$

Then, (21) becomes

$$Q(\theta_t; \theta_t^{(k)}) = \text{tr} \left( 2\Re \left( \mathbf{D}^H \mathbf{A}_t(\theta_t)^H \right) \right) \quad (30)$$

$$- \mathbf{A}_t(\theta_t) \mathbf{B} \mathbf{A}_t(\theta_t)^H + C'. \quad (31)$$

Assume that the conditional covariance matrix  $\boldsymbol{\Sigma}_n$  is positive-definite. Then,  $\mathbf{B}$  is also positive-definite. Since  $\mathbf{B} = \mathbf{B}^H$ , we can obtain the square root of  $\mathbf{B}$ , or  $\mathbf{L}$ , i.e.,  $\mathbf{B} = \mathbf{L} \mathbf{L}^H$ . Then, we have

$$Q(\theta_t; \theta_t^{(k)}) = \text{tr} \left( -(\mathbf{D} \mathbf{L}^\dagger - \mathbf{A}_t(\theta_t) \mathbf{L}) (\mathbf{D} \mathbf{L}^\dagger - \mathbf{A}_t(\theta_t) \mathbf{L})^H \right) + C', \quad (32)$$

where  $\mathbf{L}^\dagger$  is the Pseudo-inverse of  $\mathbf{L}$ . Let  $\mathbf{d}_i^H$  be the  $i$ th row vector of  $\mathbf{D}$  and  $\theta_{t,i}$  be  $[(\mathbf{a}_n^{i,1})^T \cdots (\mathbf{a}_n^{i,n_t})^T]^T$ , i.e.,  $\theta_t = [\theta_{t,1}^T, \cdots, \theta_{t,n_r}^T]^T$ . Then, we can write (32) as

$$Q(\theta_t; \theta_t^{(k)}) = \sum_{i=1}^{n_r} -\|\mathbf{L}^\dagger \mathbf{d}_i - \mathbf{L}_i \theta_{t,i}\|^2 + C', \quad (33)$$

where  $\mathbf{L}_i$  is the sub-matrix of  $\mathbf{L}$  with the  $N_t i + 1$ -th to  $N_t i + N_t$ -th columns. Since the parameters in  $\theta_{t,1}, \cdots, \theta_{t,n_r}$  are not overlapped, we can maximize  $Q(\theta_t; \theta_t^{(k)})$  over each of  $\theta_{t,1}, \cdots, \theta_{t,n_r}$ , separately. Therefore,  $\theta_{t,1}, \cdots, \theta_{t,n_r}$  that minimize (32) become

$$\theta_t^{(k+1)} = \begin{bmatrix} \theta_{t,1}^{(k+1)} \\ \vdots \\ \theta_{t,n_r}^{(k+1)} \end{bmatrix}, \quad (34)$$

where

$$\theta_{t,i}^{(k+1)} = \arg \min_{\theta_{t,i} \in \Theta} \|\mathbf{p} - \mathbf{L}_i \theta_{t,i}\|^2, \quad (35)$$

where  $\mathbf{p} = \mathbf{L}^\dagger \mathbf{d}_i$ . In order to obtain  $\theta_t^{(k+1)}$ , the squared norm of (35) should be minimized over all combination of binary vectors. This kind of optimization problem is referred to as an *integer least square* (ILS) problem that has been actively studied in recent years. To solve this problem efficiently, we apply a tree search algorithm, called *M algorithm* [11] which finds best  $M$  candidates for each element of  $\theta_{t,i}$  successively. Notice that the *M algorithm* finds the sub-optimal solution to (35) due to its greedy nature, but it provides significant reduction in computational complexity compared to an exhaustive

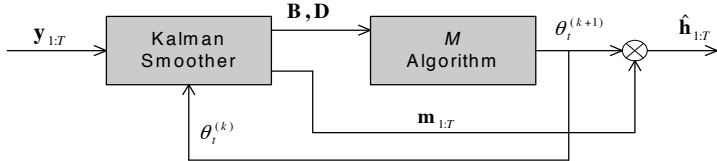


Fig. 1. Block diagram of the IS channel estimator.

search. In addition, by exploiting the constraint  $\theta_{t,i} \in \Theta$ , we can reduce the search candidates further.

The block diagram of the IS channel estimator is provided in Fig. 1. The Kalman smoother and  $M$ -algorithm produce the quantities  $\mathbf{B}$  and  $\mathbf{D}$ , and the sparse structure estimate  $\theta_t^{(k+1)}$  and exchange them in an iterative manner. The iterations repeat until there is no change in the sparse structure estimate further.

### III. SIMULATIONS

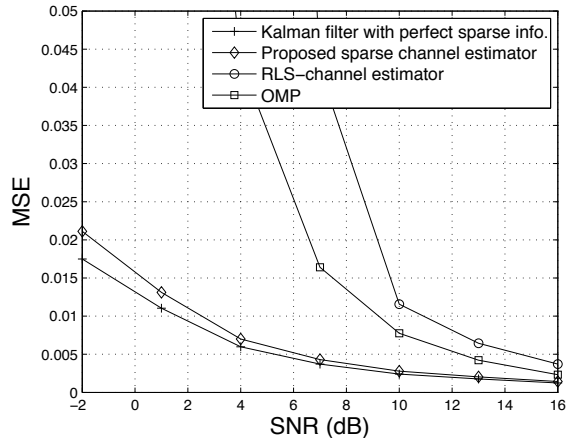
In this section, we provide the simulation results for Rician fading MIMO channels.

#### A. Simulation Setup

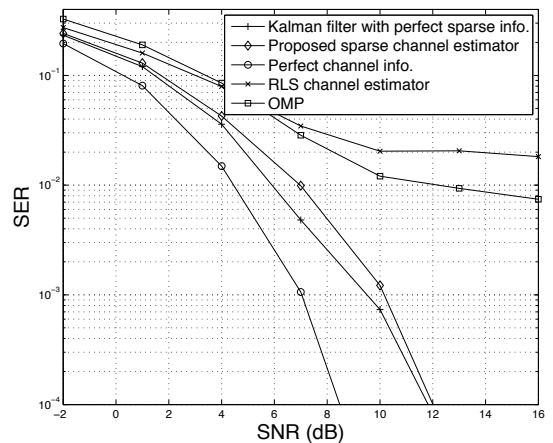
A total of  $10^5$  symbols are generated. The modulation order is set to QPSK and the number of training symbols and data symbols are set to 200 and 1200, respectively. The  $2 \times 2$  MIMO system is considered. The number of non-zero taps is set to 6 out of a total of 20 taps for each channel link. The parameters  $K_{i,j}$  for the sparse constraints are equally set to 8. The sparse structure of channels does not change during the block but randomly generated across the blocks. The 3rd-order AR model is used to generate a fading channel gain. The AR parameters are chosen such that an autocorrelation of fading process derived from the Jake's model [14] is closely approximated by an AR process (see [13] for details) and the parameters for the Kalman smoother are chosen accordingly. We generate channel gains based on wide sense stationary and uncorrelated scattering (WSSUS) assumption [12] and set the power of main tap to 0dB and those of the next taps to  $-6$ dB,  $-4$ dB,  $-2$ dB,  $-7$ dB,  $-10$ dB,  $-6$ dB,  $\dots$  in the ascending order of distance from the main tap. In addition, the  $\kappa$  value of each tap for Rician fading is set to 20, 10, 6, 4, 1, 1  $\dots$  in the same order. We measure the normalized mean square error between the real channel and its estimate and symbol error rate (SER) when the IS channel estimator is combined with MMSE decision-feedback equalizer. We compare our IS channel estimator with the RLS channel estimator [9], orthogonal matching pursuit (OMP) method [5], and Kalman channel estimator which has perfect knowledge on sparse structure. Hence, the performance of this genie-aided Kalman channel estimator provides the performance bound that the proposed sparse channel estimator can achieve.

#### B. Simulation Results

Fig. 2 shows the plot of normalized MSE versus SNR when Doppler rate is 0.002. The MSE of our sparse channel estimator is close to that of genie-aided Kalman estimator.



(a)



(b)

Fig. 2. (a) MSE versus SNR and (b) BER versus SNR of several channel estimators.

This implies that the IS channel estimator is able to detect the sparse structure with reasonably good accuracy. In addition, the proposed IS channel estimator shows lower normalized MSE compared to the RLS and OMP channel estimators. In Fig. 2 (b), the BER performance is compared when combined with the minimum mean square error decision-feedback equalizer (MMSE-DFE). We also include the performance of the MMSE-DFE when the channel is perfectly known. As shown in the figure, the MMSE-DFE with the IS channel estimator outperforms the MMSE-DFE with the RLS and OMP channel estimation schemes and the performance gap is large in the SNR range higher than 10 dB.

### REFERENCES

- [1] T. H. Eggen, A. B. Baggeroer, and J. C. Preisig, "Communication over doppler spread channels - Part I: channel and receiver presentation," *IEEE Journal of Oceanic Engineering*, vol. 25, pp. 62-71, Jan. 2000.
- [2] T. H. Eggen, A. B. Baggeroer, and J. C. Preisig, "Communication over doppler spread channels - Part II: receiver characterization and practical results," *IEEE Journal of Oceanic Engineering*, vol. 26, pp. 612-621, Oct. 2001.

- [3] J. Homer, I. Mareels, R. R. Bitmead, B. Wahlberg, and F. Gustafsson, "LMS estimation via structural detection," *IEEE Trans. on Signal Processing*, vol. 46, pp. 2651-2663, Oct. 1998.
- [4] M. Kocic, D. Brady, and M. Stojanovic, "Sparse equalization for real-time underwater acoustic communications," *Proc. OCEANS*, Oct. 1995, pp. 1417-1422.
- [5] S. F. Cotter and B. D. Rao, "Sparse channel estimation via matching pursuit with application to equalization," *IEEE Trans. Commun.*, vol. 50, pp. 374-377, March 2002.
- [6] J. L. Paredes, G. R. Arce, and Z. Wang, "Ultra-wideband compressed sensing: channel estimation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, pp. 383-395, Oct. 2007.
- [7] W. U. Bajwa, J. Haupt, G. Raz, and R. Nowak, "Compressed channel sensing," *IEEE CISS 08*, March 2008, pp. 1-6.
- [8] C. Carbonelli, S. Vedantam, and U. Mitra, "Sparse channel estimation with zero tap detection," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1743-1753, May 2007.
- [9] S. Haykin, "Adaptive filter theory Third edition," *Prentice Hall*, 1996.
- [10] H. E. Rauch, F. Tung, and C. T. Striebel, "Maximum likelihood estimates of linear dynamic systems," *AIAA Journal*, vol. 3(8) pp. 1445-50, August, 1965.
- [11] J. B. Anderson and S. Mohan, "Sequential coding algorithms: A survey and cost analysis," *IEEE Trans. Commun.*, vol. COM-32, pp. 169-176, Feb. 1984.
- [12] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Commun.*, vol. 11, pp. 360-393, Dec. 1963.
- [13] K. E. Baddour and N. C. Beaulieu, "Autoregressive modeling for fading channel simulation," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 1650-1662, July 2005.
- [14] W. C. Jakes, "Microwave mobile communications," New York: Wiley, 1974.