Optimal Pilot Beam Pattern Design for Massive MIMO Systems

Song Noh†, Michael D. Zoltowski†, Youngchul Sung‡, and David J. Love†
†School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907, USA
Email: songnoh@purdue.edu, mikedz@ecn.purdue.edu, djlove@ecn.purdue.edu
‡Department of Electrical Engineering, KAIST, Daejeon, Korea 305-701
Email: ysung@ee.kaist.ac.kr

***Abstract*—In this paper, channel estimation for massive multiple-input multiple-output (MIMO) systems with a large number of transmit antennas at the base station is considered, and a new algorithm for pilot beam pattern design for optimal channel estimation under the assumption of Gauss-Markov channel processes is proposed. The proposed algorithm designs the optimal pilot beam pattern sequentially by exploiting the statistics of the channel, antenna correlation, and temporal correlation. The algorithm provides a sequentially optimal sequence of pilot beam patterns for a given set of system parameters. Numerical results show the effectiveness of the proposed algorithm.**

I. Introduction

MIMO systems using very large antenna arrays, so called *massive MIMO* systems, have been an active research area to achieve high spectral efficiency [1]–[3]. Massive MIMO can provide performance scaling with the number of transmit antennas using simple signal processing [1]. Such benefits, in practice, can be limited by channel estimation accuracy [4]. Since the available orthogonal training sequences for channel estimation are limited by either the channel coherence time or the interference from multiple users in neighboring cells, near perfect channel estimation can be infeasible for massive MIMO systems.

To tackle the challenges of channel estimation in massive MIMO systems, most recent works consider methods using the reciprocity benefits of time-division duplexing (TDD) to exploit channel reciprocity [4], [5]. In this case, the pilot overhead related to channel estimation can be reduced via uplink channel sounding because the required orthogonal training sequences become independent of the number of transmit antenna at the base station and dependent on the number of serviced users. In the frequency-division duplexing (FDD) case, channel estimation becomes more challenging because traditional small array (e.g., two, four, or eight antenna) MIMO channel sounding approaches require far too much time overhead. There has only been limited worn on massive MIMO channel estimation techniques with FDD [6], [7]. These techniques exploit spatial correlation or closed-loop training to get improved estimation performance. If transmit channel adaptation is needed, FDD systems also require potentially substantial feedback overhead [8]–[10].

This paper considers the problem of downlink channel estimation in FDD massive MIMO systems. We develop a new pilot beam pattern design in which orthogonal pilot sequences are bounded by the channel coherence time. We propose an efficient algorithm which provides the sequentially optimal pilot beam pattern to minimize the channel estimation mean square error (MSE). The key idea behind the new algorithm is the use of the second-order statistics of the channel, the temporal correlation, and the signal-to-noise ratio (SNR) jointly to derive the pilot beam pattern at each training instance.

A. Notation

Vectors and matrices are written in boldface with matrices in capitals. All vectors are column vectors. $\mathbf{A}^T$, $\mathbf{A}^H$, and $\mathbf{A}^*$ indicate the transpose, Hermitian transpose, and the complex conjugate of $\mathbf{A}$, respectively. $[\mathbf{A}]_{i,j}$ denotes the element of $\mathbf{A}$ at the i-th row and j-th column. diag$(d_1, \cdots, d_n)$ is the diagonal matrix composed of elements $d_1, \cdots, d_n$ and diag$(\mathbf{A})$ gives a vector containing the diagonal elements of matrix $\mathbf{A}$. For a vector $\mathbf{a}$, we use $\|\mathbf{a}\|_2$ for 2-norm. For a matrix $\mathbf{A}$, tr$(\mathbf{A})$ and var$(\mathbf{A})$ denote the trace of $\mathbf{A}$, and variance operator, respectively. The Kronecker product is $\otimes$, and vec$(\mathbf{A})$ operator creates a column vector by stacking the elements of $\mathbf{A}$ columnwise. $E\{\mathbf{x}\}$ represents the expectation of $\mathbf{x}$, and $\mathbf{I}_n$ stands for the identity matrix of size n.

II. System Model

A. System Set-Up

We consider a massive MIMO system with N_t transmit antennas and a single receive antenna ($N_t \gg 1$), as shown in Fig. 1. At the k-th symbol time, the received signal is given by

$$y_k = \mathbf{h}_k^H \mathbf{s}_k + w_k, \quad k = 1, 2, \dots \tag{1}$$

where $\mathbf{s}_k$ is the $N_t \times 1$ transmitted symbol vector at time k, $\mathbf{h}_k$ is the $N_t \times 1$ MISO channel vector at time k, and w_k is a zero-mean independent and identically distributed (i.i.d.) complex Gaussian noise at time k with covariance σ_w^2. We assume that the channel is time-varying and Rayleigh-faded under a state-space model, i.e., the channel dynamic is given by the first-order stationary Gauss-Markov process

$$\mathbf{h}_{k+1} = a\mathbf{h}_k + \sqrt{1-a^2}\mathbf{b}_k, \tag{2}$$

satisfying the Lyapunov equation

$$\mathbf{R}_\mathbf{h} = a^2\mathbf{R}_\mathbf{h} + (1-a^2)\mathbf{R}_\mathbf{b}, \tag{3}$$

where a is the temporal fading correlation coefficient,[1] $\mathbf{b}_k$ is a zero-mean and temporally independent plant Gaussian vector.

[1]Under Jakes' model, $a = J_0(2\pi f_D T_s)$ [11], where $J_0(\cdot)$ is the zero-order Bessel function, f_D is the Doppler frequency shift, and T_s is the transmit symbol interval.

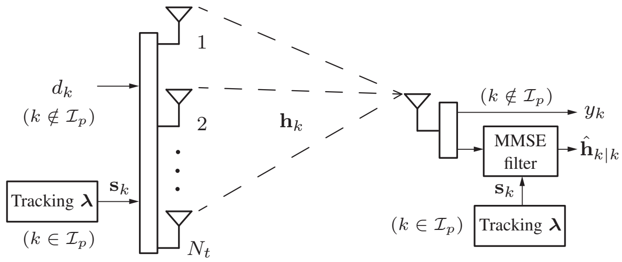


Fig. 1. Massive MIMO system model, $k = lM + m$.

For stationarity, $\mathbf{R_b} = E\{\mathbf{b}_k\mathbf{b}_k^H\} = \mathbf{R_h} = E\{\mathbf{h}_k\mathbf{h}_k^H\}$ for all k.

We assume that the transmission is continuously slotted with M consecutive symbols as one slot and that each slot is composed of a training period of M_p symbols and a data transmission period of M_d symbols ($M = M_p + M_d$). During training periods, a sequence of properly designed known pilot transmit vectors $\mathbf{s}_k$, $k \in \mathcal{I}_p := \{k = lM + m | l = 0, 1, 2, \cdots, m = 1, 2, \cdots, M_p\}$ is transmitted and the channel is estimated. (Note that $\mathbf{s}_k$ at training symbol time k is the pilot beam pattern at time k.) During data transmission periods, on the other hand, unknown data is transmitted. Here, transmit beamforming can be applied based on the estimated channel during the training period.[2]

B. Channel Estimation

We consider the minimum mean square error (MMSE) channel estimation based on the current and all previous observations during training periods, i.e., $\hat{\mathbf{h}}_{k|k} := E\{\mathbf{h}_k | y_p^{(k)}\}$ where $y_p^{(k)}$ is all observations during the pilot transmission up to symbol time k, given by

$$y_p^{(k)} = \{y_{k'} | k' \le k, k' \in \mathcal{I}_p\}.$$

Note that the system equation (1) can be rewritten as

$$y_k = \mathbf{s}_k^H \mathbf{h}_k + w_k, \quad k = lM + m \in \mathcal{I}_p. \tag{4}$$

Then, (2) and (4) form a state-space model and the optimal channel estimation is given by Kalman filtering and prediction applied to this state-space model [12]. During the training period, a measurement update step at each symbol time is available due to the known pilot pattern, and the Kalman channel estimate and the error covariance matrix are given by [12]

$$\hat{\mathbf{h}}_{k|k} = \hat{\mathbf{h}}_{k|k-1} + \mathbf{K}_k(y_k - \mathbf{s}_k^H \hat{\mathbf{h}}_{k|k-1}), \tag{5}$$

$$\mathbf{P}_{k|k-1} = a^2 \mathbf{P}_{k-1|k-1} + (1 - a^2)\mathbf{R_h}, \tag{6}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{s}_k^H \mathbf{P}_{k|k-1}, \tag{7}$$

with $\hat{\mathbf{h}}_{1|0} = \mathbf{0}$ and $\mathbf{P}_{1|0} = \mathbf{R_h}$, where $\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{s}_k(\mathbf{s}_k^H \mathbf{P}_{k|k-1}\mathbf{s}_k + \sigma_w^2)^{-1}$, and $\mathbf{P}_{k|k}$ and $\mathbf{P}_{k|k-1}$ are the estimation and prediction error covariance matrices, respectively, defined as

$$\mathbf{P}_{k|k'} = E\left\{(\mathbf{h}_k - \hat{\mathbf{h}}_{k|k'})(\mathbf{h}_k - \hat{\mathbf{h}}_{k|k'})^H | y_p^{(k')}\right\}, \tag{8}$$

where $\hat{\mathbf{h}}_{k|k'} := E\{\mathbf{h}_k | y_p^{(k')}\}$. During the data transmission period, the channel is predicted without the measurement update step, based on the last channel estimate of the previous training period as [12]

$$\begin{aligned} \hat{\mathbf{h}}_{lM+M_p+i|lM+M_p} &= a^i \hat{\mathbf{h}}_{lM+M_p|lM+M_p} \\ \mathbf{P}_{lM+M_p+i|lM+M_p} &= a^{2i}\mathbf{P}_{lM+M_p|lM+M_p} + (1 - a^{2i})\mathbf{R_h}, \end{aligned} \tag{9}$$

where $i = 1, \ldots, M_d$. During the data transmission period, transmit beamforming can be applied based on the current channel estimate $\hat{\mathbf{h}}_{lM+M_p+i|lM+M_p}$, $i = 1, \ldots, M_d$. For maximum rate transmission, eigen-beamforming [1], [13] can be applied. For maximal ratio transmit beamforming based on the current channel estimate, the beamforming weight vector is given by

$$\mathbf{s}_k = \frac{\hat{\mathbf{h}}_{lM+M_p+i|lM+M_p}}{\|\hat{\mathbf{h}}_{lM+M_p+i|lM+M_p}\|_2} d_k, \tag{10}$$

where d_k is the k-th data symbol with signal power ρ_d ($k = lM + M_p + i$).

III. The Proposed Pilot Beam Pattern Design

In this section, we propose a pilot beam pattern design method for the channel estimation considered in the previous section under an estimation MSE criterion. The channel estimation MSE is directly related to the effective signal-to-noise ratio (SNR) and the training-based channel capacity [14].

Note from (9) that during the l-th data transmission period, the channel estimation error depends only on a, $\mathbf{R_h}$ and the estimation error covariance matrix $\mathbf{P}_{lM+M_p|lM+M_p}$ at the last pilot symbol time. Since a and $\mathbf{R_h}$ are given, we need to minimize the estimation MSE, tr$(\mathbf{P}_{lM+M_p|lM+M_p})$, at the last pilot symbol time by properly designing the pilot beam pattern sequence $\{\mathbf{s}_k,\ k = l'M + m,\ l' \le l,\ m = 1, \cdots, M_p\}$. Note that $\mathbf{P}_{lM+M_p|lM+M_p}$ is a function of $\mathcal{S} := \{\mathbf{s}_j | j = l'M + m, m = 1, \cdots, M_p, j \le lM + M_p\}$. Thus, to minimize the MSE at time $k = lM + M_p$, $\mathcal{S}$ should be jointly optimized. However, such joint optimization is too complicated since the impact of $\mathcal{S}$ on $\mathbf{P}_{lM+M_p|lM+M_p}$ is intertwined. Furthermore, optimal channel estimation at $k = lM + M_p$ for some l is not the only optimization goal since the MSE at $k = l'M + M_p$ for each and every $l' < l$ should be optimized for the l'-th data transmission period. Hence, we adopt a greedy sequential optimization approach to design the pilot beam pattern sequence. That is, we optimize pilot $\mathbf{s}_k$ at time k to minimize tr$(\mathbf{P}_{k|k})$, given $\mathbf{s}_{k'}$ at all pilot symbol time $k' < k$, starting from $k = 1$.

Problem 1: For each pilot symbol time k starting from 1, given $\mathbf{s}_{k'}$ for all pilot symbol time $k' < k$, design $\mathbf{s}_k$ such that

$$\min_{\mathbf{s}_k} \text{tr}\left(\mathbf{P}_{k|k}\right) \tag{11}$$

$$\text{s.t. } \|\mathbf{s}_k\|_2^2 = \rho_p. \tag{12}$$

[2] For transmit beamforming in FDD, some form of channel state information (CSI) should be fed back to the transmitter from the receiver. Here, a quantized $\mathbf{y}_k$ can be fed back and channel estimation is performed at the transmitter, or channel estimation is performed at the receiver and the quantized channel estimate can be fed back. The focus of the paper is not feedback quantization but optimal design of the pilot beam pattern for channel estimation.

In this paper, we consider the MISO case only. In the MIMO case in which channel estimation is performed at each receive antenna separately, the MISO result here can directly be applied. The MIMO case with joint processing across the multiple receive antennas for channel estimation is beyond the scope of the current paper.

A. The Proposed Algorithm

In the MISO case, the solution to Problem 1 is given by the following proposition.

Proposition 1: Given all previous pilot $\mathbf{s}_{k'}$ ($k' < k$), the pilot beam pattern $\mathbf{s}_k$ at time k minimizing tr($\mathbf{P}_{k|k}$) is given by a scaled dominant eigenvector of the error covariance matrix of the Kalman *prediction* for time k.

Proof: Case 1) $k \neq lM+1$: From (7), $\arg\min_{\mathbf{s}_k} \text{tr}(\mathbf{P}_{k|k})$ can be written as

$$\arg\max_{\mathbf{s}_k} \ \text{tr}\left(\mathbf{P}_{k|k-1}\mathbf{s}_k(\mathbf{s}_k^H\mathbf{P}_{k|k-1}\mathbf{s}_k+\sigma_w^2)^{-1}\mathbf{s}_k^H\mathbf{P}_{k|k-1}\right). \tag{13}$$

Since tr($\mathbf{ABC}$) = tr($\mathbf{BCA}$), (13) can be rewritten as

$$\arg\max_{\mathbf{s}_k} \ \text{tr}\left((\mathbf{s}_k^H\mathbf{P}_{k|k-1}\mathbf{s}_k+\sigma_w^2)^{-1}\mathbf{s}_k^H\mathbf{P}_{k|k-1}^2\mathbf{s}_k\right). \tag{14}$$

Note that (14) is the *generalized Rayleigh quotient* with respect to (w.r.t.) the pencil $(\mathbf{P}_{k|k-1}^2, \mathbf{P}_{k|k-1}+\sigma_w^2/\rho_p\mathbf{I}_{N_t})$. Thus, for any non-zero vector $\mathbf{s}_k$, (14) satisfies the following bound [15]

$$\frac{\rho_p\lambda_{N_t}^2}{\rho_p\lambda_{N_t}+\sigma_w^2} \leq \frac{(\mathbf{s}_k^H\mathbf{P}_{k|k-1}^2\mathbf{s}_k)}{\mathbf{s}_k^H\left(\mathbf{P}_{k|k-1}+\sigma_w^2/\rho_p\mathbf{I}_{N_t}\right)\mathbf{s}_k} \leq \frac{\rho_p\lambda_1^2}{\rho_p\lambda_1+\sigma_w^2}, \tag{15}$$

where λ_1 and λ_{N_t} are the largest and smallest eigenvalues of $\mathbf{P}_{k|k-1}$, respectively, and optimal $\mathbf{s}_k$ is given by the eigenvector of the Kalman prediction error covariance matrix $\mathbf{P}_{k|k-1}$ corresponding to λ_1 scaled by $\sqrt{\rho_p}$.

Case 2) $k = lM+1$: In this case, we have M_d prediction steps without measurement update steps before the first pilot symbol time k in the l-th slot. The measurement update form (7) at k is valid with $\mathbf{P}_{k|k-1}$ replaced by the Kalman prediction for time k based on all the previous pilot beam patterns given by

$$\mathbf{P}_{k|(l-1)M+M_p} = a^{2M_d}\mathbf{P}_{(l-1)M+M_p|(l-1)M+M_p} + (1-a^{2M_d})\mathbf{R}_{\mathbf{h}}.$$

Thus, the proof in Case 1) is applicable to this case just with $\mathbf{P}_{k|k-1}$ replaced by $\mathbf{P}_{k|(l-1)M+M_p}$. ■

Thus, the optimal $\mathbf{s}_k$ is obtained from the Kalman prediction error covariance matrix which is given for given pilot $\mathbf{s}_{k'}$, $k' < k$. Interestingly, it can be shown that the pilot beam pattern $\mathbf{s}_k$ obtained from (13) is equivalent to the first principal component direction of $\mathbf{P}_{k|k-1}$ given by

$$\arg\max_{\|\mathbf{s}_k\|_2^2=\rho_p} \ \text{var}\left(\mathbf{s}_k^H(\mathbf{h}_k - \hat{\mathbf{h}}_{k|k})\right). \tag{16}$$

Note that to obtain the (sequentially) optimal $\mathbf{s}_k$, we need to perform the eigen-decomposition (ED) of the Kalman prediction error covariance matrix with size $N_t \times N_t$ at each pilot symbol time k and this is computationally expensive since N_t is large for massive MIMO systems. The following proposition provides a useful property of the eigen-space of the Kalman prediction error covariance matrix that can be exploited for constructing an efficient beam pattern design algorithm.

Proposition 2: The Kalman filtering error covariance matrix $\mathbf{P}_{k|k}$ and the Kalman prediction error covariance matrix $\mathbf{P}_{k|k'}$ generated with sequentially optimal $\mathbf{s}_k$ obtained from Proposition 1 are simultaneously diagonalizable with $\mathbf{R}_{\mathbf{h}}$ for any k and $k'(< k)$, under the assumption of $\mathbf{P}_{1|0} = \mathbf{R}_{\mathbf{h}}$.[3]

Proof: Proof is by induction. Let $\mathbf{R}_{\mathbf{h}} = \mathbf{U}\mathbf{\Lambda}_1\mathbf{U}^H$ be the ED of $\mathbf{R}_{\mathbf{h}}$. First note that $\mathbf{P}_{1|0} = \mathbf{R}_{\mathbf{h}}$. For any pilot symbol time $k = lM + m$ ($m = 1, \cdots, M_p$), suppose that $\mathbf{P}_{k|k-1} = \mathbf{U}\mathbf{\Lambda}_k\mathbf{U}^H$ is the ED of $\mathbf{P}_{k|k-1}$, where $\mathbf{\Lambda}_k := \text{diag}(\lambda_{k,1}, \ldots, \lambda_{k,N_t})$ and $k_m := \arg\max_i \lambda_{k,i}$. By Proposition 1, $\mathbf{s}_k$ is given by a scaled eigenvector of $\mathbf{P}_{k|k-1}$ corresponding to the largest eigenvalue λ_{k,k_m}. Then, from the measurement update (7), we have $\mathbf{P}_{k|k}$, given by

$$\mathbf{P}_{k|k} = \mathbf{U}\text{diag}\left(\lambda_{k,1}, \ldots, \lambda_{k,k_m-1}, \frac{\lambda_{k,k_m}\sigma_w^2}{\rho_p\lambda_{k,k_m}+\sigma_w^2}, \lambda_{k,k_m+1}, \ldots, \lambda_{k,N_t}\right)\mathbf{U}^H \tag{17}$$

$$=: \mathbf{U}\bar{\mathbf{\Lambda}}_k\mathbf{U}^H. \tag{18}$$

Thus, $\mathbf{P}_{k|k}$ and $\mathbf{P}_{k|k-1}$ are simultaneously diagonalizable. Furthermore, it is easy to see that after the prediction step (6), $\mathbf{P}_{k+1|k}$ and $\mathbf{P}_{k|k}$ are simultaneously diagonalizable. Since $\mathbf{P}_{1|0} = \mathbf{R}_{\mathbf{h}} = \mathbf{U}\mathbf{\Lambda}_1\mathbf{U}^H$, $\mathbf{P}_{k|k}$ and $\mathbf{P}_{k|k-1}$ in the first training period have the same set of eigenvectors as $\mathbf{R}_{\mathbf{h}}$.

Now consider a symbol time k during the first data transmission period. In this case, the prediction error covariance matrix is given by

$$\begin{aligned}\mathbf{P}_{M_p+i|M_p} &= a^{2i}\mathbf{P}_{M_p|M_p} + (1-a^{2i})\mathbf{R}_{\mathbf{h}} \\ &= \mathbf{U}\left(\mathbf{\Lambda}_1 - a^{2i}(\mathbf{\Lambda}_1 - \bar{\mathbf{\Lambda}}_{M_p})\right)\mathbf{U}^H,\end{aligned} \tag{19}$$

where $i = 1, \ldots, M_d$. Thus, any prediction error covariance matrix during the first data period is simultaneously diagonalizable with $\mathbf{P}_{k|k}$ for $k \leq M_p$. Since this Kalman recursion repeats, we have proved the claim. ■

Algorithm 1 Sequentially Optimal Pilot Beam Pattern Design

Require: Perform the ED of $\mathbf{R}_{\mathbf{h}} = \mathbf{U}\mathbf{\Lambda}_1\mathbf{U}^H$. Store $\boldsymbol{\lambda}_1 = \text{diag}(\mathbf{\Lambda}_1)$ and $\mathbf{U} = [\mathbf{u}_1, \cdots, \mathbf{u}_{N_t}]$.
  $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1$ where $\boldsymbol{\lambda} = [\lambda_1, \cdots, \lambda_{N_t}]^T$
  while $l = 0, 1, \cdots$ **do**
    for $m = 1$ to M **do**
      if $m \leq M_p$ **then**
        $i = \text{argmax}_j \lambda_j$
        $\mathbf{s}_k = \sqrt{\rho_p}\mathbf{u}_i$
        $\lambda_i \leftarrow \frac{\lambda_i\sigma_w^2}{\rho_p\lambda_i+\sigma_w^2}$
      end if
      $\boldsymbol{\lambda} \leftarrow a^2\boldsymbol{\lambda} + (1-a^2)\boldsymbol{\lambda}_1$
    end for
  end while

[3]Such an initial parameter is a typical value for the Kalman filter, there will be no loss.

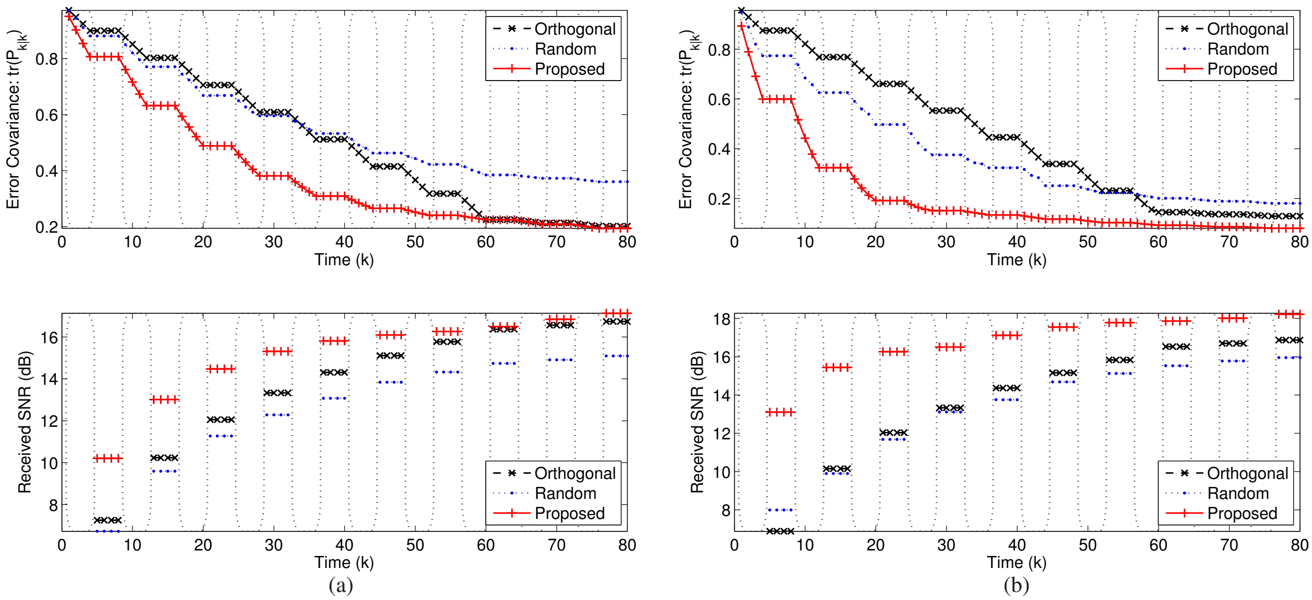


Fig. 2. MSE and SNR versus time index k, SNR=20dB, $v = 3km/h$: (a) $r = 0.4$ and (b) $r = 0.8$. (The dotted rectangles denote the pilot transmission periods.)

Thus, all Kalman prediction error covariance matrices that are used for sequentially optimal pilot beam pattern design, have the same set of eigenvectors as $\mathbf{R_h}$. Note that (17) shows how a sequentially optimal pilot beam pattern at time k reduces the channel estimation error by changing the eigenvalue distribution and (19) shows how the eigenvalues of the channel prediction error covariance matrix change during the pure prediction step. Exploiting these facts, we can construct an efficient algorithm to obtain the sequence of optimal pilot beam patterns that sequentially minimize the channel estimation MSE at each given symbol time. The algorithm is summarized in Algorithm 1.

In the proposed algorithm, the dominant eigenvalue λ_i is tracked at each symbol time during the training period. That is, the maximum eigenvalue index is searched and the corresponding eigenvector $\mathbf{u}_i$ is used as the pilot beam pattern for the corresponding symbol time. After incorporating the reduction of the dominant eigenvalue by the measurement update based on the pilot beam pattern and the eigenvalue change by the prediction step, the dominant eigenvalue index for the next time step is searched again and this operation iterates. The proposed algorithm can be run only if the channel statistics a, $\mathbf{R_h}$, the slot information (M_p, M_d) and SNR are known, and the optimal sequence of pilot beam patterns depends on these parameters. Unless the antenna elements are uncorrelated, we have a nontrivial optimal sequence. The necessary parameters can be shared between the transmitter and the receiver at the beginning of the transmission session, and then the receiver can run the algorithm to know the currently used pilot pattern by itself. Note that the proposed algorithm requires the ED of $\mathbf{R_h}$ only once and all other computation is simple arithmetic.

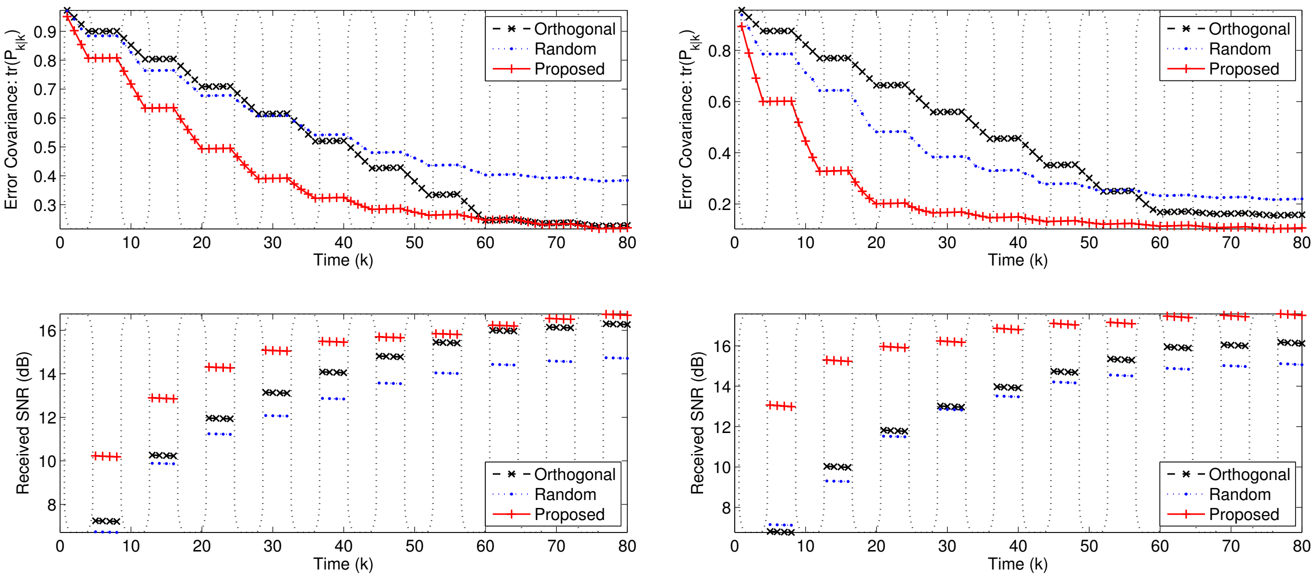
IV. Numerical Results

In this section, we provide numerical results to evaluate the performance of the proposed algorithm. Since massive MIMO systems assume $N_t \gg 1$, we consider $N_t = 32$ transmit antennas and a single receive antenna, and set $M_p = 4$ and $M = 8$ due to the insufficient pilot training. We adopt $2.5GHz$ carrier frequency and a symbol duration of $100\mu s$. We compare two different cases: $v = 3km/h, 30km/h$ in which $a = 0.9999, 0.9995$, respectively. The considered channel spatial correlation model is the *quadratic exponential correlation model* [16], [17], given by

$$[\mathbf{R_h}]_{i,j} = \begin{cases} r^{(j-i)^2} & \text{if } j \geq i \\ \left(r^{(i-j)^2}\right)^* & \text{if } j < i \end{cases}, \qquad (20)$$

where $|r| < 1$ and normalized so that $\text{tr}(\mathbf{R_h}) = 1$. The channel estimation performance was measured by the trace of the Kalman estimation error covariance, averaged under $1,000$ Monte Carlo runs. The noise variance σ_w^2 is determined according to the SNR, i.e., SNR $= \frac{\rho_p}{\sigma_w^2}$ $(\rho_p = \rho_d = 1)$, and the received SNR is defined as $\frac{|\mathbf{s}_k^H \hat{\mathbf{h}}_{k|k'}|^2}{\mathbf{s}_k^H \mathbf{P}_{k|k'} \mathbf{s}_k + \sigma_w^2}$ imposed by imperfect channel estimation.

We first compared the performance of the proposed method to several existing methods [18] in Fig. 2. For the orthogonal and random pilot methods based on the Kalman filter, we considered a round-robin selection for the initialized pilot beam patterns. Fig. 2 shows the proposed algorithm tracks the channel state fast during tracking periods and also guarantees the received SNR gain. Because of our tracking of spectral distribution of the channel MMSE, the proposed method converges more quickly as the antenna spatial correlation increases by comparing Fig. 2(a) with (b). Note that the proposed method has about an SNR loss of 2dB compared to the perfect channel estimation case in Fig. 2(b). For the fast fading process with $v = 30km/h$ in Fig. 3, the channel MMSE and the received SNR shows a repetitive trajectory curve since the channel MMSE increased during data transmission periods, as expected (6). There is some loss of performance in channel



(a) (b)

Fig. 3. MSE and SNR versus time index k, SNR=20dB, $v = 30km/h$: (a) $r = 0.4$ and (b) $r = 0.8$.

estimation and the received SNR due to the increased temporal fading correlation affecting the spectral distribution of the channel MMSE, however, the proposed method still shows good performance.

V. CONCLUSIONS

We have proposed a new algorithm for the optimal pilot beam pattern design in massive MIMO systems using a first-order stationary Gauss-Markov channel process. The proposed algorithm jointly exploits the statistics of channel, antenna correlation, and temporal correlation to provide a sequentially optimal pilot beam pattern for a given set of system parameters with low computational complexity.

REFERENCES

[1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590 – 3600, Nov. 2010.

[2] H. Ngo and E. G. Larsson, "EVD-based channel estimation in multi cell multiuser MIMO systems with very large antenna arrays," in *Proc. IEEE ICASSP*, Kyoto, Japan, Mar. 2012.

[3] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40 – 60, Jan. 2013.

[4] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640 – 2651, Aug 2011.

[5] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160 – 171, Feb. 2013.

[6] P. H. Kuo, H. T. Kung, and P. A. Ting, "Compressive sensing based channel feedback protocols for spatially-correlated massive antenna arrays," in *Proc. IEEE WCNC*, Paris, France, Apr. 2012.

[7] D. J. Love, J. Choi, and P. Bidigare, "A closed-loop training approach for massive MIMO beamforming systems," in *Proc. IEEE CISS*, Johns Hopkins Univ., Maryland, Mar. 2013.

[8] J. Nam, J.-Y. Ahn, A. Adhikary, and G. Caire, "Joint spatial division and multiplexing: Realizing massive MIMO gains with limited channel state information," in *Proc. IEEE CISS*, Johns Hopkins Univ., Maryland, Mar. 2013.

[9] J. Choi, Z. Chance, D. J. Love, and U. Madhow, "Noncoherent trellis coded quantization for massive MIMO limited feedback beamforming," in *Proc. IEEE ITA*, San Diego, CA, Feb. 2013.

[10] J. Choi, D. J. Love, and U. Madhow, "Limited feedback in massive MIMO systems: Exploiting channel correlations via noncoherent trellis-coded quantization," in *Proc. IEEE CISS*, Johns Hopkins Univ., Maryland, Mar. 2013.

[11] W. C. Jakes, *Microwave Mobile Communication*, Wiley, New York, NY, 1974.

[12] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice-Hall, Upper Saddle River, New Jersey, 2000.

[13] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, no. 6, pp. 585 – 595, Nov. 1999.

[14] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951 – 963, Apr. 2003.

[15] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., Johns Hopkins Univ. Press, Baltimore, MD, 1996.

[16] T.-S. Chu and L. J. Greenstein, "A semi-empirical representation of antenna diversity gain at cellular and PCS base stations," *IEEE Trans. Commun.*, vol. 45, no. 6, pp. 644 – 656, Jun. 1997.

[17] V. Erceg *et al.*, "IEEE P802.11 Wireless LANs TGn Channel Models," in *IEEE P802.11-03-0940*, http://www.ieee802.org/11/.

[18] W. Santipach and M. L. Honig, "Optimization of training and feedback overhead for beamforming over block fading channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6103 – 6115, Dec. 2010.