

TRAINING SIGNAL DESIGN FOR CHANNEL ESTIMATION IN 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

[‡]Department of Electrical Engineering, KAIST, Daejeon, Korea 305-701

ABSTRACT

In this paper, the design of training signals for channel estimation in massive multiple-input multiple-output (MIMO) systems is considered. Under a stationary, block Gauss-Markov channel model, a method for optimal pilot beam pattern design for enhanced channel estimation is proposed, exploiting both the properties of Kalman filtering and the spatio-temporal channel correlation. First, pilot beam pattern design is considered under the assumption of orthogonal beam patterns within a block. The orthogonality assumption is subsequently relaxed and the design problem is solved via a greedy approach. Numerical results show the efficacy of the proposed algorithm.

Index Terms— Channel estimation, Kalman filtering, Gauss-Markov model, massive MIMO.

1. INTRODUCTION

MIMO systems using very large transmit antenna arrays, so called *massive MIMO* systems, have drawn interest because of their ability to achieve high spectral efficiency [3, 4]. Based on the law of large numbers, massive MIMO systems can alleviate the effect of thermal noise and fast channel fading plus interference, and an energy efficient system can be realized with cheap, low-power amplifiers for each antenna [3]. However, the benefits of massive MIMO in practical systems are limited by channel estimation accuracy [5]. In particular, when full frequency reuse across neighboring cells is adopted, imperfect channel estimation severely degrades performance, a condition called pilot contamination [2]. Further, the overhead required for channel estimation in massive MIMO can become overwhelming since the available training resources are limited by the channel coherence time and/or the amount of interference induced by neighboring cells.

To tackle the challenge of channel estimation, most prior work is based on time-division duplex (TDD) systems and

This research was funded in part by the MSIP (Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program 2013. A preliminary version of this work dealing with a symbol-fading MISO case was presented in [1]. A complete version of this work was submitted to [2] with an extension to the MIMO case with power allocation and to a block-fading channel model. However, the block-fading case was handled with the orthogonality assumption for the pilot sequence for each transmission slot. In this paper, the sequential design in [2] is extended by relaxing the orthogonality condition on the pilot signals within a training period.

relies on channel reciprocity under the assumption of time-invariant channels within the coherence time and reciprocity calibration [3, 4, 6]. However, frequency-division duplex (FDD) systems are employed in most current wireless cellular systems, and in this case the problem of channel estimation becomes more challenging because MIMO channel sounding requires substantial overhead that scales with the number of antennas. Recently, there has been some work on channel estimation and channel state information (CSI) feedback techniques for FDD massive MIMO [7, 8, 9].

In this paper, we consider the problem of pilot beam design for downlink channel estimation in FDD massive MIMO systems, where the number of symbol times for channel estimation within a channel coherence time is typically much less than the number of transmit antennas. To design an efficient sequence of pilot beam patterns, we exploit the underlying channel statistics in massive MIMO such as channel temporal dynamics [10, 11], channel spatial correlation [12, 13], and the structure of error covariance matrices associated with optimal Kalman filtering. We develop a low-complexity pilot beam pattern design method that sequentially minimizes the channel estimation mean square error (MSE) at each training instant.

2. SYSTEM MODEL

2.1. System Setup

We consider a massive MIMO system with N_t transmit antennas and a single receive antenna ($N_t \gg 1$) over flat Rayleigh-fading channels. The received signal at the k -th symbol time is given by

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

where \mathbf{s}_k is the $N_t \times 1$ transmitted symbol vector, \mathbf{h}_k is the $N_t \times 1$ MISO channel vector, and w_k is a zero-mean independent and identically distributed (i.i.d.) complex Gaussian noise with covariance σ_w^2 . We assume slotted transmission with M consecutive symbols for one slot composed of a pilot transmission period of M_p symbols and a data transmission period of M_d symbols ($M = M_p + M_d$).

We assume that the channel is block-fading, i.e., the channel remains constant during each transmission slot, i.e., $\mathbf{h}_k = \mathbf{h}_l$ for $k = lM + m$, $l = 0, 1, 2, \dots$, and $1 \leq m \leq M$.

Furthermore, for the channel temporal dynamic across slots, we consider a state-space model for a first-order stationary Gauss-Markov process [14], given by

$$\mathbf{h}_{l+1} = a\mathbf{h}_l + \sqrt{1 - a^2}\mathbf{b}_l, \quad (2)$$

that satisfies the Lyapunov equation $\mathbf{R}_h = a^2\mathbf{R}_h + (1 - a^2)\mathbf{R}_b$ and $\mathbf{R}_h = E\{\mathbf{h}_l\mathbf{h}_l^H\} = \mathbf{R}_b = E\{\mathbf{b}_l\mathbf{b}_l^H\}$ for all l , where a is the temporal fading correlation coefficient, \mathbf{b}_l is a zero-mean and temporally independent plant Gaussian vector, and l is the slot index. Since the channel statistics can be estimated [15, 16], we shall assume that a and \mathbf{R}_h are known.

It was observed in [2] that pilot power allocation derived from Kalman filtering is beneficial especially for low SNR regime. Thus, we here consider equal power allocation for the pilot beam pattern at each training symbol time.

2.2. Channel Estimation

We focus on minimum mean square error (MMSE) channel estimation based on the current and all previous received pilot signals given by $\hat{\mathbf{h}}_{l|l} = E\{\mathbf{h}_l|y_p^{(l)}\}$, where $y_p^{(l)} = \{y_{k'}|k' \leq lM + M_p, k' \in \mathcal{I}_p\}$ and $\mathcal{I}_p = \{k = lM + m|l = 0, 1, \dots; 1 \leq m \leq M_p\}$. Here, $y_p^{(l)}$ denotes all received pilot signals up to the l -th training period. During a training period, a sequence of pilot beam vector $\{\mathbf{s}_k\}$ is transmitted for channel estimation. During a data transmission period, the base station transmits unknown data symbols with transmit beamforming based on the estimated channel.

The received signal model (1) for the l -th training period can be rewritten in vector form as

$$\mathbf{y}_l = \mathbf{S}_l^H \mathbf{h}_l + \mathbf{w}_l, \quad (3)$$

where $\mathbf{y}_l = [y_{lM+1}, \dots, y_{lM+M_p}]^T$ and $\mathbf{S}_l = [\mathbf{s}_{lM+1} \dots \mathbf{s}_{lM+M_p}]$. Note that (2) and (3) form a state-space model, and optimal channel estimation is provided by Kalman filtering [17]. During a training period, a measurement update step is available due to the known pilot beam pattern, given by [17]

$$\hat{\mathbf{h}}_{l|l} = \hat{\mathbf{h}}_{l|l-1} + \mathbf{K}_l(\mathbf{y}_l - \mathbf{S}_l^H \hat{\mathbf{h}}_{l|l-1}), \quad (4)$$

$$\mathbf{P}_{l|l} = \mathbf{P}_{l|l-1} - \mathbf{K}_l \mathbf{S}_l^H \mathbf{P}_{l|l-1}, \quad (5)$$

where $\mathbf{K}_l = \mathbf{P}_{l|l-1} \mathbf{S}_l (\mathbf{S}_l^H \mathbf{P}_{l|l-1} \mathbf{S}_l + \sigma_w^2 \mathbf{I}_{M_p})^{-1}$ and $(\mathbf{P}_{l|l}, \mathbf{P}_{l|l-1})$ are the estimation and prediction error covariance matrices, respectively, defined as

$$\mathbf{P}_{l|l'} = E\{(\mathbf{h}_l - \hat{\mathbf{h}}_{l|l'}) (\mathbf{h}_l - \hat{\mathbf{h}}_{l|l'})^H | y_p^{(l')}\} \quad (6)$$

$$\hat{\mathbf{h}}_{0|-1} = \mathbf{0} \text{ and } \mathbf{P}_{0|-1} = \mathbf{R}_h, \quad (7)$$

where $\hat{\mathbf{h}}_{l|l'} = E\{\mathbf{h}_l | y_p^{(l')}\}$ with the initial condition (7).

During the data transmission period, unknown data symbols are transmitted with transmit beamforming based on the current channel estimate $\hat{\mathbf{h}}_{l|l}$. The channel after the data transmission period is predicted for the next slot as [17]

$$\hat{\mathbf{h}}_{l+1|l} = a\hat{\mathbf{h}}_{l|l} \text{ and } \mathbf{P}_{l+1|l} = a^2\mathbf{P}_{l|l} + (1 - a^2)\mathbf{R}_h. \quad (8)$$

3. THE PROPOSED PILOT BEAM PATTERN METHOD: GREEDY SEQUENTIAL DESIGN

In this section, we propose a pilot beam pattern design method that minimizes the channel estimation MSE extracted from Kalman filtering in Section 2.2.

3.1. Orthogonal Sequential Design (OSD)

As mentioned in [2], an approach such as dynamic programming for the joint design of $\{\mathbf{S}_l, l = 0, 1, 2, \dots\}$ is difficult due to the intertwined structure of the MSE extracted from the Kalman recursion. Hence, we adopt a greedy sequential approach to minimize the current MSE $\text{tr}(\mathbf{P}_{l|l})$ at each block time l . Since a , \mathbf{R}_h , $\mathbf{P}_{l|l-1}$, and $\mathbf{S}_{l'}$, $l' < l$, are given at block time l , the greedy sequential design problem is stated as follows.

Problem 1 For each slot time l starting from 0, given $\mathbf{S}_{l'}$ for all pilot symbol time $l' < l$, design \mathbf{S}_l such that

$$\min_{\mathbf{S}_l} \text{tr}(\mathbf{P}_{l|l}), \text{ s.t. } \|\mathbf{s}_{lM+m}\|^2 = \rho_p, m = 1, \dots, M_p. \quad (9)$$

Problem 1 itself does not yield a closed-form solution, but when we impose the following constraint on \mathbf{S}_l

$$(C.1) \quad \mathbf{S}_l^H \mathbf{S}_l = \rho_p \mathbf{I}_{M_p},$$

the problem with this additional constraint yields a very nice solution, as given in the following proposition.

Proposition 1 [2] Given all previous pilot signals $\mathbf{S}_{l'}$ ($l' < l$), under the condition (C.1) the pilot beam signal \mathbf{S}_l at the l -th training period minimizing $\text{tr}(\mathbf{P}_{l|l})$ is given by a properly scaled version of the M_p dominant eigenvectors of the Kalman prediction error covariance matrix $\mathbf{P}_{l|l-1}$ for the l -th training period.

The main difference between the above result exploiting the channel temporal correlation and the existing result for the i.i.d. block-fading result in [18] is as follows. While the M_p dominant eigenvectors of \mathbf{R}_h are used for training in the i.i.d. block-fading case [18], the M_p dominant eigenvectors of the Kalman prediction error covariance matrix $\mathbf{P}_{l|l-1}$ should be used for training in the Gauss-Markov block-fading case. To obtain the optimal \mathbf{S}_l , we need to perform an eigen-decomposition (ED) of $\mathbf{P}_{l|l-1}$ at each slot, and this can be computationally expensive in massive MIMO systems ($N_t \gg 1$). However, it was shown that all Kalman prediction error covariance matrices combined with the orthogonal pilot beam pattern design have the same set of eigenvectors with \mathbf{R}_h [2]. This is an important practical implication in that the base station sends pilot beam vectors \mathbf{S}_l chosen from a *fixed* set of pilot beam patterns (the eigenvectors of \mathbf{R}_h).

3.2. General Sequential Design

Note that the orthogonality condition (C.1) on the pilot beamforming vectors may limit the channel estimation performance. When a few dominant directions contain most of

the channel gain [13] or the SNR is not high enough to reduce the channel uncertainty in the subspace spanned by the pilot beam vectors, the system may require some specific pilot beam vectors more than once during a training period. Hence, in this subsection, we relax the orthogonality condition (C.1) and address the pilot design problem again. In this case, the solution is not given nicely as in Proposition 1. To tackle the problem, we adopt a *symbol-time-wise* sequential optimization approach to the Gauss-Markov block-fading channel model to design the set of pilot beam patterns $\mathbf{s}_k \in \mathcal{S}_l = \{\mathbf{s}_{k'} : lM + 1 \leq k' \leq lM + M_p\}$. That is, we convert the time-invariant block Gauss-Markov model to a time-varying symbol-wise Gauss-Markov model, and apply symbol-time-wise Kalman filtering to this time-varying symbol-wise Gauss-Markov model.

Note that the channel is not varying during a slot, and we have symbol-time-wise measurement updates during the training period of the slot. Define $\hat{\mathbf{h}}_k^l \triangleq E\{\mathbf{h}_l | y_p^{(k)}\}$, where $y_p^{(k)} = \{y_{k'} | k' \leq k, k' \in \mathcal{I}_p\}$, and $\mathbf{P}_k^l = E\{(\mathbf{h}_l - \hat{\mathbf{h}}_k^l)(\mathbf{h}_l - \hat{\mathbf{h}}_k^l)^H | y_p^{(k)}\}$. Then, the measurement update steps (4) and (5) can be rewritten as

$$\hat{\mathbf{h}}_k^l = \hat{\mathbf{h}}_{k-1}^l + \mathbf{K}_k(y_k - \mathbf{s}_k^H \hat{\mathbf{h}}_{k-1}^l), \quad (10)$$

$$\mathbf{P}_k^l = \mathbf{P}_{k-1}^l - \mathbf{K}_k \mathbf{s}_k^H \mathbf{P}_{k-1}^l, \quad (11)$$

$lM + 1 \leq k \leq lM + M_p$, where $\hat{\mathbf{h}}_{lM}^l = \hat{\mathbf{h}}_{l|l-1}^l$, $\mathbf{P}_{lM}^l = \mathbf{P}_{l|l-1}^l$, and $\mathbf{K}_k = \mathbf{P}_{k-1}^l \mathbf{s}_k (\mathbf{s}_k^H \mathbf{P}_{k-1}^l \mathbf{s}_k + \sigma_w^2)^{-1}$. At the end of the slot, the channel is updated and thus a prediction step is applied:

$$\hat{\mathbf{h}}_{l+1|l}^l = a \hat{\mathbf{h}}_{M_p}^l \quad \text{and} \quad \mathbf{P}_{l+1|l}^l = a^2 \mathbf{P}_{M_p}^l + (1 - a^2) \mathbf{R}_h. \quad (12)$$

Now the greedy sequential pilot design to this symbol-time-wise model is stated as follows.

Problem 2 For each slot time l starting from 0, given $\mathcal{S}_{l'}$ for all pilot symbol time $l' < l$, design \mathcal{S}_l such that

$$\min_{\mathbf{s}_k} \text{tr}(\mathbf{P}_k^l), \quad \text{s.t.} \quad \|\mathbf{s}_k\|_2^2 = \rho_p, \quad \mathbf{s}_k \in \mathcal{S}_l. \quad (13)$$

The solution to Problem 2 is given by the following proposition.

Proposition 2 Given all previous pilot signals $\mathcal{S}_{l'}$ ($l' < l$), the pilot beam pattern $\mathbf{s}_k \in \mathcal{S}_l$ minimizing $\text{tr}(\mathbf{P}_k^l)$ is given by a scaled dominant eigenvector of \mathbf{P}_{k-1}^l .

The proof is omitted since it can be shown similarly as in the symbol-time-wise, time-invariant case in [1]. Furthermore, employing the sequentially optimal \mathcal{S}_l , it can be shown that $\mathbf{P}_k^l, \mathbf{P}_{k-1}^l, \mathbf{R}_h$, all have the same common set of eigenvectors. That is, we can represent $\mathbf{P}_{k-1}^l = \mathbf{U} \mathbf{\Lambda}_{k-1}^l \mathbf{U}^H$, where $\mathbf{R}_h = \mathbf{U} \mathbf{\Lambda}_0 \mathbf{U}^H$ is the ED of \mathbf{R}_h , and $\mathbf{\Lambda}_{k-1}^l$ is a diagonal matrix composed of the eigenvalues of \mathbf{P}_{k-1}^l . Note that if the selected eigenvectors for $lM + 1 \leq k \leq lM + M_p$ are all different, then the obtained \mathcal{S}_l satisfies the orthogonality condition because $\mathbf{U}^H \mathbf{U} = \mathbf{I}_{N_t}$.

Remark 1 By Proposition 2 combined with the simultaneous diagonalizability, the measurement update step (11) during the training period can be rewritten as

$$\mathbf{P}_k^l = \mathbf{U} \mathbf{\Lambda}_{k-1}^l \mathbf{U}^H - \mathbf{U} \left(\frac{\rho_p \lambda_{k-1,i}^2}{\rho_p \lambda_{k-1,i} + \sigma_w^2} \mathbf{e}_i \mathbf{e}_i^T \right) \mathbf{U}^H, \quad (14)$$

where $\mathbf{\Lambda}_{k-1}^l = \text{diag}(\lambda_{k-1,1}, \dots, \lambda_{k-1,N_t})$ and \mathbf{e}_i is the i -th unit vector for $i = \text{argmax}_j \lambda_{k-1,j}$. The optimal \mathbf{s}_k is given by $\mathbf{s}_k = \sqrt{\rho_p} \mathbf{u}_i$ by Proposition 2, where $\mathbf{U} = [\mathbf{u}_1 \dots \mathbf{u}_{N_t}]$. The prediction step (12) at $k = lM + M$ can be rewritten as

$$\mathbf{P}_{l+1|l}^l = a^2 \mathbf{U} \mathbf{\Lambda}_{M_p}^l \mathbf{U}^H + (1 - a^2) \mathbf{R}_h. \quad (15)$$

Summarizing the above result, we provide the symbol-time-wise sequentially optimal pilot design method for the Gauss-Markov block-fading channel model that removes the orthogonality limitation of the design in Proposition 1, and thus allows the same pilot pattern to be used more than once during a training period, depending on the channel statistics (as in Algorithm 1). In addition, the proposed design can be leveraged to reduce the complexity of a feedback-based pilot beam pattern design [7, 19] with some modified assumptions. That is, the proposed method requires the system to store only N_t eigenvectors of \mathbf{R}_h , instead of all possible M_p combinations of all eigenvectors of \mathbf{R}_h . Thus, the receiver feeds back the index of the selected pilot beam pattern among $\{1, \dots, N_t\}$ for the next pilot symbol time. Although we only consider the MISO case, the result derived here can be extended to the MIMO case by using approaches similar to those in [2]. Furthermore, by the Toeplitz distribution theorem (TDT) [20, 21], the large Toeplitz covariance matrix \mathbf{R}_h can be eigen-decomposed by a discrete Fourier transform (DFT) matrix. As a result, the pilot beam patterns are approximated by DFT vectors without much loss in performance and simulation will be presented in Section 4.

Algorithm 1 Sequentially Optimal Pilot Beam Pattern Design

Require: Perform the ED of $\mathbf{R}_h = \mathbf{U} \mathbf{\Lambda}_0 \mathbf{U}^H$. Store $\lambda_0 = \text{diag}(\mathbf{\Lambda}_0)$ and $\mathbf{U} = [\mathbf{u}_1 \dots \mathbf{u}_{N_t}]$.
 $\lambda = \lambda_0$ where $\lambda = [\lambda_1, \dots, \lambda_{N_t}]^T$
while $l = 0, 1, \dots$ **do**
 for $m = 1$ to M **do**
 if $m \leq M_p$ **then**
 $i = \text{argmax}_j \lambda_j$
 $\mathbf{s}_k = \sqrt{\rho_p} \mathbf{u}_i$
 $\lambda_i \leftarrow \frac{\lambda_i \sigma_w^2}{\rho_p \lambda_i + \sigma_w^2}$ (See (14).)
 end if
 end for
 $\lambda \leftarrow a^2 \lambda + (1 - a^2) \lambda_0$ (See (15).)
end while

4. NUMERICAL RESULTS

In this section, we provide some numerical results to evaluate the performance of the proposed algorithm. We adopt

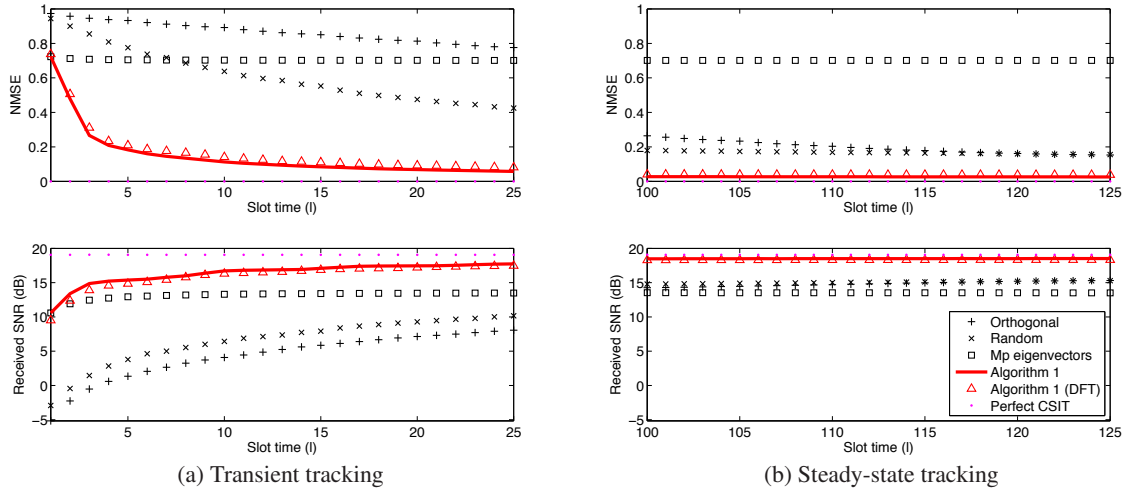


Fig. 2. NMSE and SNR versus the slot time index l where $M_p = 2$, $M = 5$, $N_t = 250$ (10×25), and $\sigma_w^2 = 10^{-1.5}$

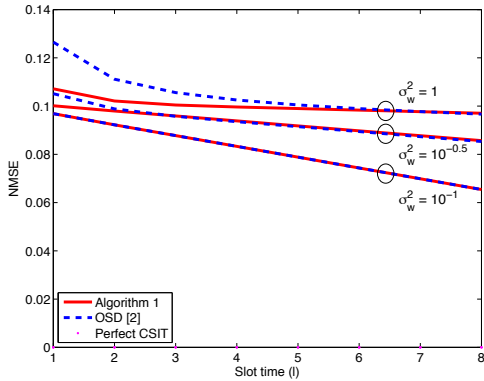


Fig. 1. NMSE performance comparison for $M_p = 3$, $r = 0.9$

2.5GHz carrier frequency and the terminal velocity $v = 3\text{km/h}$ with $T_s = 0.5\text{ms}$.¹ The channel estimation performance was measured by the normalized mean square error (NMSE), i.e., $\frac{1}{\text{tr}(\mathbf{R}_h)} \text{tr}(\mathbf{P}_{lM+M_p|lM+M_p})$ at the end of the l -th training period. The pilot symbol SNR was defined as ρ_p/σ_w^2 with $\rho_p = 1$ and we assumed that the pilot-to-data power ratio is the same. The received SNR was defined as $|\mathbf{s}_k^H \hat{\mathbf{h}}_{k|k}|^2 / (\mathbf{s}_k^H \mathbf{P}_{k|k} \mathbf{s}_k + \sigma_w^2)$.

First, we considered the *uniform correlation model* [19, 23], given by $[\mathbf{R}_h]_{i,j} = (1-r)\delta_{ij} + r$, where r is the correlation coefficient between two adjacent antenna elements and δ_{ij} is the Kronecker delta. Fig. 1 shows the performance of the proposed methods for a uniform linear array (ULA) with $N_t = 32$. Since this correlated channel is mainly distributed around some dominant directions, Algorithm 1 benefits by relaxing the orthogonality condition and using a suitable pilot beam pattern more than once in a slot and thus enhances the MSE performance especially at low SNR.

Next, we considered the *one-ring* channel model (which models typical cellular configurations [12, 13]) for a 10×25

uniform planar array (UPA).² We compared the performance of the proposed method to several methods [24, 18] and the results are plotted in Fig. 2. Under optimal Kalman filtering, we considered a round-robin selection for the orthogonal and random pilot patterns that were initialized at the start of the simulation. Fig. 2(a) shows that the proposed method learns the channel state well, and also provides a considerable (received) SNR gain. Note that the DFT/TDT-based method yields almost the same performance. Orthogonal and random pilot beam patterns are ineffective because such patterns span all of the N_t -dimensional space instead of capturing the dominant channel uncertainty in space. The fixed M_p dominant eigenvectors of \mathbf{R}_h [18] can only minimize the channel MSE along the M_p eigen-directions, and thus the corresponding channel estimation performance is saturated quickly because it leaves the $N_t - M_p$ eigen-directions intact. (See [2] for a detailed discussion.)

5. CONCLUSIONS

We have considered the problem of pilot beam pattern design for massive MIMO systems under the Gauss-Markov block-fading channel model and have proposed new methods for pilot beam pattern design by exploiting channel statistics and the structure of Kalman filtering with and without an orthogonality condition on the pilot beam vectors within a training period. In the case when an orthogonality condition is imposed on the pilot signals within a training period, the sequentially optimal beam patterns at each training period are given by the dominant eigenvectors of the Kalman prediction error covariance matrix. The proposed algorithm based on a symbol-time-wise sequential approach relaxing the orthogonality condition can further improve the channel estimation performance in highly spatially correlated channels, where a few dominant directions contain the most of the channel gain.

¹For Jakes' model [22], $a = J_0(2\pi f_D T_s)$ where $J_0(\cdot)$ is the zero-order Bessel function, f_D is the Doppler frequency shift, and T_s is the slot interval.

²See [13, 2] for more details about the path loss and the channel model.

6. REFERENCES

- [1] S. Noh, M. D. Zoltowski, Y. Sung, and D. J. Love, "Optimal pilot beam pattern design for massive MIMO systems," in *Proc. IEEE ASILOMAR*, Pacific Grove, CA, Nov. 2013.
- [2] S. Noh, M. D. Zoltowski, Y. Sung, and D. J. Love, "Pilot beam pattern design for channel estimation in massive MIMO systems," *J. Sel. Topics Signal Process.*, submitted for publication. [Online]. Available: <http://arxiv.org/abs/1309.7430>.
- [3] 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.
- [4] 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.
- [5] 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.
- [6] C. Shepard, H. Yu, N. Anand, L. E. Li, T. L. Marzetta, R. Yang, and L. Zhong, "Argos: Practical many-antenna base stations," in *Proc. MobiCom*, Istanbul, Turkey, Aug. 2012.
- [7] 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.
- [8] 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.
- [9] 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 ASILOMAR*, Pacific Grove, CA, Nov. 2013.
- [10] K. E. Baddour and N. C. Beaulieu, "Autoregressive modeling for fading channel simulation," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1650 – 1662, Jul. 2005.
- [11] H. Yu, Y. Sung, H. Kim, and Y. H. Lee, "Beam tracking for interference alignment in slowly-fading MIMO interference channels: Perturbations approach under a linear framework," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1910 – 1926, Apr. 2012.
- [12] D. Shiu and G. J. Foschini and M. J. Gans and J. M. Kahn, "Fading correlation and its effect on the capacity of multi element antenna systems," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502 – 513, Mar. 2000.
- [13] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing: The large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441 – 6463, Oct. 2013.
- [14] L. Tong, B. M. Sadler, and M. Dong, "Pilot-assisted wireless transmissions: General model, design criteria, and signal processing," *IEEE Signal Process. Mag.*, vol. 21, no. 6, pp. 12 – 25, Nov. 2004.
- [15] B. M. Hochwald and T. L. Marzetta, "Adapting a downlink array from uplink measurements," *IEEE Trans. Signal Process.*, vol. 49, no. 3, pp. 642 – 653, Mar. 2001.
- [16] X. Dai, W. Zhang, J. Xu, J. E. Mitchell, and Y. Yang, "Kalman interpolation filter for channel estimation of LTE downlink in high-mobility environments," *EURASIP J. Wireless Commun. and Netw.*, vol. 2012:232, Jul. 2012.
- [17] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice-Hall, Upper Saddle River, New Jersey, 2000.
- [18] J. H. Kotecha and A. M. Sayeed, "Transmit signal design for optimal estimation of correlated MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 546 – 557, Feb. 2004.
- [19] 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.
- [20] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*, University of California Press, Berkeley, CA, 1958.
- [21] Y. Sung, H. V. Poor, and H. Yu, "How much information can one get from a wireless ad hoc sensor network over a correlated random field?," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2827 – 2847, Jun. 2009.
- [22] W. C. Jakes, *Microwave Mobile Communication*, Wiley, New York, NY, 1974.
- [23] R. K. Mallik and J. H. Winters, "Deterministically weighted square-law combining in correlated fading," *IEEE Trans. Commun.*, vol. 61, no. 1, pp. 344 – 356, Jan. 2013.
- [24] 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.