PRECODING ON THE BROADCAST MIMO CHANNEL WITH DELAYED CSIT: THE FINITE SNR CASE

Xinping Yi and David Gesbert

EURECOM
Mobile Communications Department
06560, Sophia Antipolis, France
Email: {xinping.yi, david.gesbert}@eurecom.fr

ABSTRACT

Recent information theoretic results suggest that precoding on the multi-user downlink MIMO channel with delayed channel state information at the transmitter (CSIT) could lead to data rates much beyond the ones obtained without any CSIT, even in extreme situations when the delayed channel feedback is made totally obsolete by a feedback delay exceeding the channel coherence time. This surprising result is based on the ideas of interference forwarding and alignment which allow the receivers to reconstruct an information allowing them to cancel out the interference completely, making it an optimal scheme in the infinite SNR regime. In this paper, we formulate a similar problem, yet at finite SNR. We propose a new construction for the precoder which matches the previous results at infinite SNR yet reaches a useful trade-off between interference alignment and signal enhancement at finite SNR, allowing for significant performance improvements in practical settings.

Index Terms— MIMO, delayed feedback, precoding, multi-user

I. INTRODUCTION

Multi-user MIMO systems (or their information-theoretic counterparts “broadcast MIMO channels”), have recently attracted considerable attention from the research community and industry alike. Success is due to their ability to enhance the wireless spectrum efficiency by a factor equal to the number N of antennas installed at the base station, with little restriction imposed on the richness of the multipath channel, the presence or absence of a strong line of sight channel component, and the fact it can easily accommodate single antenna mobile devices. On the downlink of such systems, the ability to beamform (i.e. linearly precode) multiple data streams simultaneously to several users (up to N) comes nevertheless at a price in terms of requiring the base station transmitter to be informed of the channel coefficients of all served users [1]. In frequency division duplex scenarios (the bulk of available wireless standards today), this implies establishing a feedback link from the mobiles to the base station which can carry CSI related information, in quantized format. A common limitation of such an approach, perceived by many to be a key hurdle toward a more widespread use of MU-MIMO methods in real-life networks, lies in the fact that the feedback information typically arrives back to the transmitter with a delay which may cause a severe degradation when comparing the obtained feedback CSIT with the actual current channel state information. Pushed to the extreme, and considering a feedback delay with the same order of magnitude as the coherence period of the channel, the available CSIT feedback becomes completely obsolete (uncorrelated with the current true channel information) and, seemingly non exploitable in view of designing the precoding coefficients.

Recently, this commonly accepted viewpoint was challenged by an interesting information-theoretic work which established the usefulness of stale channel state information in designing precoders achieving significantly better rate performance than what is obtained without any CSIT [2]. The premise of this work [2] is a time-slotted MIMO broadcast channel with a common transmitter serving multiple users and having a delayed version of the correct CSIT, where the delay causes the CSIT to be fully uncorrelated with the current channel vector information. In this situation, it is shown that the transmitter can still exploit the stale channel information: The transmitter tries to reproduce the interference generated to the users in the previous time slots, a strategy we refer in this paper as interference forwarding, while at the same time making sure the forwarded interference occupies a subspace of limited dimension, compatible with its cancelation at the user’s side, a method commonly referred to as interference alignment [3], [4]. Building on such ideas, [2] constructs a transmission protocol which was shown to achieve the maximum Degrees-of-Freedom (DoF) for the delayed CSIT broadcast MIMO channel. Precoding on delayed CSIT MIMO channels have recently attracted more interesting work, dealing with DoF analysis on extended channels, like the X channel and interference channels [5], [6], [7], but also performance analysis including effects of feedback [8] and training [9]. The DoF is a popular information theoretic performance metric indicating the number of interference-free simultaneous data streams which can be communicated over this delayed CSIT channel at infinite SNR, also coinciding with the notion of pre-log factor in the channel capacity expression. In the example of the two antenna transmitter, two user channel, the maximum DoF is shown to be $\frac{4}{3}$, less than the value of 2 which would be obtained with perfect CSIT, but strictly larger than the single DoF obtained in the absence of any CSIT.

Although fascinating from a conceptual point of view, these results are intrinsically focussed on the asymptotic SNR behavior, leaving in particular aside the question of how shall precoding be done practically using stale CSIT at finite SNR. This paper precisely tackles this question. In what follows we obtain the following key results:

- We show finite SNR precoding using delayed CSIT can be achieved by a combination of interference forwarding, alignment together with a signal enhancement strategy.

- We propose a precoder construction generalizing the ideas of [2] where a compromise between interference alignment and orthogonality within the desired signal channel matrix is striken.
- The precoder coefficients are interpreted as beamforming vector coefficients in a dual interference channel scenario, which can be optimized in a number of ways, including using an MMSE metric, or virtual SINR metric.

Numerical evaluation reveal a substantial performance benefit in terms of data rate in the low to moderate SNR region, but coinciding with the performance of [2] when the SNR grows to infinity.

Notation: Matrices and vectors are represented as uppercase and lowercase letters, and transpose and conjugate transpose of a matrix are denoted as $(\cdot)^T$ and $(\cdot)^H$, respectively. Further, $tr(\cdot)$ and $\|\cdot\|$ represent the trace of a matrix and the norm of a vector.

II. SYSTEM MODEL

We consider 2-user MU-MIMO downlink systems with a base station equipped with 2 antennas and two single-antenna users. Note that the limitation to a small number of users/antennas allows for a greater clarity of explanation. However, in the same spirit of the original work of [2] being extendable to arbitrary number of antennas and users, our proposed precoding concepts can be also extended, although this point will be addressed in details in the journal version of this paper. The system model and notations are as follows.

We consider a time slotted transmission protocol in the downlink direction where the multi-antenna channel vector from the transmitter to i-th user ($i = A, B$), in the j-th time slot, is denoted by $\mathbf{h}_i(j) = [h_{i1}(j)\ h_{i2}(j)]^T$ and is an independent and identically distributed (i.i.d.) complex random vector with zero-mean and unit-variance (i.e., $\mathcal{CN}(0,1)$). It is assumed that three time slots are used to send a total of four data symbols (two symbols for each user), yielding an average rate efficiency of $4/3$ symbols/channel use. The 2×1 symbol vectors intended to user A and B are respectively denoted by $\mathbf{S}_A$ and $\mathbf{S}_B$.

In the following we briefly review the basic transmission and decoding protocol proposed in [2], referred to in the literature as the MAT algorithm.

We assume that at time j, the transmitter is informed perfectly about channels $\mathbf{h}_i(k)$, $k < j$. We make no assumption about any correlation between the channel vectors across multiple time slots (could be fully uncorrelated), making it impossible for the transmitter to use classical MU-MIMO precoding to serve the users. The key point of the MAT algorithm is to establish the feasibility of transmitting and successfully detecting (at least in the high SNR regime) $\mathbf{S}_A$ and $\mathbf{S}_B$ over a three-time-slot protocol. Note that this provides a multiplexing gain (MG) of $4/3$ strictly over what is obtained without any CSIT (MG of 1), although less than the MG of 2 obtained with non delayed CSIT and classical ZF precoding.

The time-slotted protocol goes as follows: At time slot 1, the transmitter sends $\mathbf{S}_A$ over the two transmit antennas, without precoding. At time slot 2, it sends $\mathbf{S}_B$ over the two transmit antennas, also without precoding. At time slot 3, the transmitter makes use of the knowledge of $\mathbf{h}_i(k)$, $k = 1, 2, i = A, B$ in order to transmit to the users a signal $\mathbf{h}_B^T(1)\mathbf{S}_A + \mathbf{h}_A^T(2)\mathbf{S}_B$, which reconstructs the interference they have seen in the previous two slots, enabling them to do interference suppression. The signal vector received over the three time slots at user A, for instance, (a similar model is obtained at user B but not written out for lack of space) is given by:

$$\bar{\mathbf{y}}_A = \sqrt{\frac{P}{2}}\bar{\mathbf{H}}_{A1}\mathbf{S}_A + \sqrt{\frac{P}{2}}\bar{\mathbf{H}}_{A2}\mathbf{S}_B + \mathbf{n}_A \tag{1}$$

where $\bar{\mathbf{y}}_A = (y_A(1), y_A(2), y_A(3))^T$ and $y_A(k)$ is the received signal at user A in time slot k, $\mathbf{n}_A = (n_A(1), n_A(2), n_A(3))^T$ is the Gaussian noise vector with zero-mean and variance σ^2, P is the total transmit power at each time slot, and the channel matrices are

$$\bar{\mathbf{H}}_{A1} = \begin{bmatrix} \mathbf{h}_A^T(1) \\ \mathbf{0} \\ h_{A1}(3)\mathbf{h}_B^T(1) \end{bmatrix}, \ \bar{\mathbf{H}}_{A2} = \begin{bmatrix} \mathbf{0} \\ \mathbf{h}_A^T(2) \\ h_{A1}(3)\mathbf{h}_A^T(2) \end{bmatrix}. \tag{2}$$

Interestingly, it appears from the above protocol that the interference $\mathbf{S}_B$ seen by user A arrives with an effective channel matrix $\bar{\mathbf{H}}_{A2}$ which is of rank one, making it possible for user A to combine the three received signals in order to retrieve $\mathbf{S}_A$ while canceling out $\mathbf{S}_B$ completely. This process is referred to as alignment of interference signal $\mathbf{S}_B$, as it mimics the approach taken in interference channel in e.g. [3]. A similar property is exploited at user B as well.

III. THE GENERALIZED MAT ALGORITHM (GMAT)

Although optimal in terms of the MG, at infinite SNR, the above approach can be substantially improved at finite SNR. The key reason is that, at finite SNR, a good receiver filter at user A or B will not attempt to use all degrees of freedom to eliminate the interference but will try to strike a compromise between interference canceling and enhancing the detectability of the desired signal in the presence of noise. Taking into account this property of basic receivers leads us to revisit the design of the protocol and in particular the design of the precoding coefficients as function of the knowledge of past channel vectors.

The idea of the Generalized MAT approach (GMAT) consists in enabling the use of arbitrary precoding vectors in the last phase of the protocol. Without loss of generality, we propose that in the third time slot, the transmitter now sends

$$x(3) = \mathbf{w}_1^T\mathbf{S}_A + \mathbf{w}_2^T\mathbf{S}_B \tag{3}$$

from the first antenna alone with the power constraint $\|\mathbf{w}_1\|^2 + \|\mathbf{w}_2\|^2 \leq 2$. This power constraint balances the transmit and received power used over the three time slots. Consequently, the effective channel matrices introduced in eq-(1) are now generalized and given by

$$\bar{\mathbf{H}}_{A1} = \begin{bmatrix} \mathbf{h}_A^T(1) \\ \mathbf{0} \\ h_{A1}(3)\mathbf{w}_1^T \end{bmatrix}, \ \bar{\mathbf{H}}_{A2} = \begin{bmatrix} \mathbf{0} \\ \mathbf{h}_A^T(2) \\ h_{A1}(3)\mathbf{w}_2^T \end{bmatrix} \tag{4}$$

and, by analogy, for user B,

$$\bar{\mathbf{H}}_{B1} = \begin{bmatrix} \mathbf{h}_B^T(1) \\ \mathbf{0} \\ h_{B1}(3)\mathbf{w}_1^T \end{bmatrix}, \ \bar{\mathbf{H}}_{B2} = \begin{bmatrix} \mathbf{0} \\ \mathbf{h}_B^T(2) \\ h_{B1}(3)\mathbf{w}_2^T \end{bmatrix}. \tag{5}$$

Interestingly, we can interpret the role of $\mathbf{w}_1$ as trying to strike a balance between aligning the interference channel of $\mathbf{S}_A$ at user B and enhancing the detectability of $\mathbf{S}_A$ at user A. In algebraic terms this can be done having a compromise between obtaining a rank deficient $\bar{\mathbf{H}}_{B1}$ and an orthogonal matrix for $\bar{\mathbf{H}}_{A1}$. When it comes to $\mathbf{w}_2$, the compromise is between obtaining a rank deficient $\bar{\mathbf{H}}_{A2}$ and an orthogonal matrix for $\bar{\mathbf{H}}_{B2}$.

IV. COMPUTATION OF THE GMAT PRECODERS

The computation of generalized precoding vectors in the proposed GMAT method can use several options. Two of them are briefly described here. The first is based on the optimization of a virtual MMSE metric, yielding an iterative algorithm to find $\mathbf{w}_i$, $i = 1, 2$, while the second considers the SINR maximization in a dual interference channel, yielding suboptimal yet closed-form solutions. Note that none of these approaches have anything in common with finite SNR interference alignment methods with CSIT, such as, e.g., [10], [11], [12], since the nature of our problem is conditioned by the delayed CSIT scenario.

IV-A. Optimization based on virtual MMSE

Since the transmitter does not know $\mathbf{h}_i(3)$ at time slot 3, the optimization of the precoder can't involve such information. Fortunately, we point out that the trade-off between interference alignment and signal matrix orthogonalization presented above can be formulated in a way that is fully independent of $\mathbf{h}_i(3)$. In fact, we introduce the virtual received signal given below, where $\mathbf{h}_i(3)$ is ignored (deterministic fading is assumed over the third time slot).

$$\mathbf{y}_i = \sqrt{\frac{P}{2}}\mathbf{H}_{i1}\mathbf{S}_A + \sqrt{\frac{P}{2}}\mathbf{H}_{i2}\mathbf{S}_B + \mathbf{n}_i, i = A, B \tag{6}$$

where the virtual channel matrices are now given by:

$$\mathbf{H}_{i1} = \begin{bmatrix} \mathbf{h}_i^T(1) \\ \mathbf{0} \\ \mathbf{w}_1^T \end{bmatrix}, \ \mathbf{H}_{i2} = \begin{bmatrix} \mathbf{0} \\ \mathbf{h}_i^T(2) \\ \mathbf{w}_2^T \end{bmatrix}, \ i = A, B \tag{7}$$

Now consider the MMSE optimum RX filters at user A and B respectively over this channel:

$$\mathbf{V}_A = \sqrt{\frac{2}{P}}\left(\mathbf{H}_{A1}\mathbf{H}_{A1}^H + \mathbf{H}_{A2}\mathbf{H}_{A2}^H + \gamma\mathbf{I}\right)^{-1}\mathbf{H}_{A1} \tag{8}$$

$$\mathbf{V}_B = \sqrt{\frac{2}{P}}\left(\mathbf{H}_{B1}\mathbf{H}_{B1}^H + \mathbf{H}_{B2}\mathbf{H}_{B2}^H + \gamma\mathbf{I}\right)^{-1}\mathbf{H}_{B2} \tag{9}$$

where $\gamma = \frac{2\sigma^2}{P}$, and the corresponding optimal MSEs are

$$\begin{aligned} & J_A(\mathbf{w}_1, \mathbf{w}_2) \\ = \ & tr\left(\mathbf{I} - \mathbf{H}_{A1}^H(\mathbf{H}_{A1}\mathbf{H}_{A1}^H + \mathbf{H}_{A2}\mathbf{H}_{A2}^H + \gamma\mathbf{I})^{-1}\mathbf{H}_{A1}\right) \end{aligned} \tag{10}$$

$$\begin{aligned} & J_B(\mathbf{w}_1, \mathbf{w}_2) \\ = \ & tr\left(\mathbf{I} - \mathbf{H}_{B2}^H(\mathbf{H}_{B1}\mathbf{H}_{B1}^H + \mathbf{H}_{B2}\mathbf{H}_{B2}^H + \gamma\mathbf{I})^{-1}\mathbf{H}_{B2}\right) \end{aligned} \tag{11}$$

The GMAT-MMSE precoding solutions are now given by:

$$\min_{\mathbf{w}_1, \mathbf{w}_2} \quad J = J_A(\mathbf{w}_1, \mathbf{w}_2) + J_B(\mathbf{w}_1, \mathbf{w}_2) \tag{12}$$

$$s.t. \quad \|\mathbf{w}_1\|^2 + \|\mathbf{w}_2\|^2 \leq 2. \tag{13}$$

As the above optimization does not lend itself easily to a closed form solution, we first propose an iterative alternating optimization procedure, based on a gradient descent of the cost function $J_A(\mathbf{w}_1, \mathbf{w}_2) + J_B(\mathbf{w}_1, \mathbf{w}_2)$. This solution is referred later as the GMAT-MMSE solution. Convexity properties of this cost function are investigated in [13]. The iterative procedure is based on:

$$\hat{\mathbf{w}}_i(k+1) = \hat{\mathbf{w}}_i(k) - \beta\frac{\partial(J)}{\partial \mathbf{w}_i}, \ i = 1, 2$$

where k is the iteration index and β is a small step size. The partial derivation is tedious but straightforward and is given in [13] for lack of space. Instead, we focus here on presenting an alternative solution.

IV-B. Optimization based on dual interference channels

To avoid the need for an iterative algorithm, we propose an alternative approach based on maximizing a SINR metric in a dual interference channel. This dual channel model is obtained from building orthogonal subspaces to the actual channel vectors. This will lead us to a convenient, albeit suboptimal, closed-form solutions for $\mathbf{w}_1$, $\mathbf{w}_2$. This technique is referred to as GMAT-DSINR (Dual SINR).

Given the channel vectors $\mathbf{h}_i(j)$, $i = A, B$, $j = 1, 2$, define the dual (orthogonal) $\mathbf{h}_i^{\perp}(j)$ to be a unit-norm 2×1 vector orthogonal to $\mathbf{h}_i(j)$. The GMAT-DSINR precoding solutions are now given by

$$\max_{\mathbf{w}_1} \quad \text{DSINR}_1 = \frac{\frac{P}{2}|\mathbf{w}_1^H\mathbf{h}_A^{\perp}(1)|^2}{\frac{P}{2}|\mathbf{w}_1^H\mathbf{h}_B^{\perp}(1)|^2 + \sigma^2\|\mathbf{w}_1\|^2} \tag{14}$$

$$\max_{\mathbf{w}_2} \quad \text{DSINR}_2 = \frac{\frac{P}{2}|\mathbf{w}_2^H\mathbf{h}_B^{\perp}(2)|^2}{\frac{P}{2}|\mathbf{w}_2^H\mathbf{h}_A^{\perp}(2)|^2 + \sigma^2\|\mathbf{w}_2\|^2} \tag{15}$$

$$s.t. \quad \|\mathbf{w}_1\|^2 + \|\mathbf{w}_2\|^2 \leq 2 \tag{16}$$

where DSINR_i is referred to a SINR in a dual domain, where $\mathbf{w}_i$ is interpreted as a receive filter for a system with a desired source with channel $\mathbf{h}_i^{\perp}$ and interference channel $\mathbf{h}_{\bar{i}}^{\perp}$, where $i \neq \bar{i}$. Consequently, the optimal solutions are obtained by

$$\mathbf{w}_1^{opt} = \frac{1}{\alpha}\left(\mathbf{h}_B^{\perp}(1)\mathbf{h}_B^{\perp H}(1) + \frac{2\sigma^2}{P}\mathbf{I}\right)^{-1}\mathbf{h}_A^{\perp}(1) \tag{17}$$

and similarly,

$$\mathbf{w}_2^{opt} = \frac{1}{\alpha}\left(\mathbf{h}_A^{\perp}(2)\mathbf{h}_A^{\perp H}(2) + \frac{2\sigma^2}{P}\mathbf{I}\right)^{-1}\mathbf{h}_B^{\perp}(2) \tag{18}$$

where α is a scalar chosen to satisfy eq-(16).

V. NUMERICAL RESULTS

The effectiveness of the proposed solutions is evaluated in terms of both sum MSE for the reconstructed symbol vectors at the users and in terms of the sum rate per time slot in bps/Hz. The parameters in the simulation are set as follows: 500 gradient-descent iterations for the GMAT-MMSE, $\beta = 0.01$. The performance is averaged over 500 channel realizations. We show in Fig. 1 the sum MSE comparison among GMAT-MMSE with the iteratively updated $\mathbf{w}_1$, $\mathbf{w}_2$, GMAT-DSINR with closed-form solutions in eq-(17) and eq-(18), and the original MAT algorithm with $\mathbf{w}_1 = \mathbf{h}_B(1)$, $\mathbf{w}_2 = \mathbf{h}_A(2)$ and the same power constraint $\|\mathbf{w}_1\|^2 + \|\mathbf{w}_2\|^2 \leq 2$. The gap between GMAT and MAT illustrates improvement of the GMAT-MMSE and GMAT-DSINR algorithms over the initial MAT concept, demonstrating the benefit of trade-off between interference alignment and desired signal orthogonality enhancement. Note that

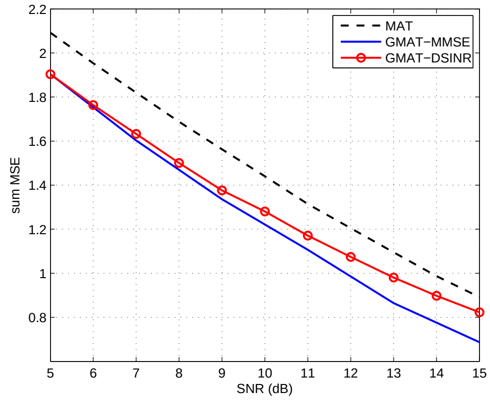

Fig. 1. Sum MSE in 3 time slot vs. SNR.

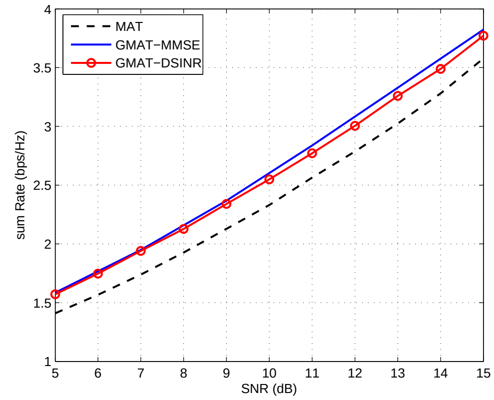

Fig. 2. Sum Rate per time slot vs. SNR.

as the SNR goes to infinity, the advantages reduce as all approaches favor pure interference alignment.

We also show the comparison in terms of sum rate per time slot associated with MMSE receiver in Fig. 2. The SINR of the k-th substream of user A is given by (SINR of user B can be similarly obtained)

$$\eta_{A,k} = \frac{1}{\left[\mathbf{I} - \bar{\mathbf{H}}_{A1}^H(\bar{\mathbf{H}}_{A1}\bar{\mathbf{H}}_{A1}^H + \bar{\mathbf{H}}_{A2}\bar{\mathbf{H}}_{A2}^H + \gamma\mathbf{I})^{-1}\bar{\mathbf{H}}_{A1}\right]_{kk}} - 1.$$

Compared with the initial MAT algorithm, the two GMAT approaches have gained significant improvement at finite SNR and possessed the same slope, which implies the same MG, at high SNR. Interestingly, the closed-form solution performs as well as the iterative GMAT-MMSE.

VI. CONCLUSION

We generalize the concept of precoding over a multi-user MISO channel with delayed CSIT. We proposed a precoder construction algorithm which achieves the same DoF at infinite SNR yet reaches a useful trade-off between interference alignment and signal enhancement at finite SNR. Our proposed precoding concept lends itself to a variety of optimization methods, two of which are shown here (iterative and closed-form). Extensions to more users and antennas can follow similar principles and will presented in the full version of this paper [13].

VII. ACKNOWLEDGEMENTS

Fruitful discussions with Paul de Kerret are thankfully acknowledged.

VIII. REFERENCES

[1] D. Gesbert, M. Kountouris, R. Heath, C-B. Chae, T. Salzer, “From single user to multiuser communications: shifting the MIMO paradigm,” *IEEE Signal Processing Magazine*, Sept. 2007.

[2] M. A. Maddah-Ali and D. Tse, “Completely stale transmitter channel state information is still very useful,” *in Proc. of Allerton Conference*, Monticello, IL, Oct. 2010. [Online].Available: http://arxiv.org/abs/1010.1499.

[3] V. R. Cadambe and S. A. Jafar, “Interference alignment and degrees of freedom of the K-user interference channel,” *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, 2008.

[4] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, “Communication over MIMO X channels: interference alignment, decomposition, and performance analysis,” *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3457–3470, Aug. 2008.

[5] H. Maleki, S. A. Jafar, and S. Shamai, “Retrospective interference alignment,” Arxiv.org, Sept. 2010, avilable at arxiv.org/abs/1009.3593.

[6] A. Ghasemi, A. S. Motahari, and A. K. Khandani, “Interference alignment for the MIMO interference channel with delayed local CSIT,” Feb. 2011, arxiv.org/abs/1102.5673.

[7] C. S. Vaze, M. K. Varanasi, “The degrees of freedom region and interference alignment for the MIMO interference channel with delayed CSI,” Mar. 2011, arxiv.org/abs/1101.5809.

[8] J. Xu, J. G. Andrews, and S. A. Jafar, “Broadcast channels with delayed finite-rate feedback: Predict or observe?,” May 2011, arxiv.org/abs/1105.3686.

[9] A. Adhikary, H. C. Papadopoulos, S. A. Ramprashad and G. Caire, “Multi-user MIMO with outdated CSI: training, feedback and scheduling,” Sept. 2011, arxiv.org/abs/1109.6371.

[10] K. Gomadam, V. R. Cadambe, and S. A. Jafar, “A distributed numerical approach to interference alignment and applications to wireless interference networks,” *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3309–3322, Jun. 2011.

[11] S. W. Peters and R. W. Heath, Jr., “Interference alignment via alternating minimization,” *in Proc. ICASSP 2009*, April 2009, pp. 2445–2448.

[12] F. Negro, S. P. Shenoy, I. Ghauri, and D. Slock, “On the MIMO interference channel,” *in Proc. ITA 2010*, Feb. 2010, pp. 1–9.

[13] X. Yi, D. Gesbert, “Precoding methods for the broadcast MIMO channel with delayed CSIT,” *IEEE Trans. on Signal Processing*, in preparation.