Enabling massive MIMO systems in the FDD mode thanks to D2D communications

Haifan Yin, Laura Cottatellucci, and David Gesbert
EURECOM
Mobile Communications Department
450 Route des Chappes, 06904 Biot, France
Email: {yin, gesbert, cottatel}@eurecom.fr

***Abstract*—We propose novel approaches to design feedback in FDD massive MIMO systems. We exploit synergies between massive MIMO systems and inter-user communications based on D2D. The exchange of local CSI among users, enabled by D2D communications, makes available global CSI at the terminals. Thus, we can construct more *informative* forms of feedback based on this shared knowledge. Two feedback variants are highlighted: 1) cooperative CSI feedback, and 2) cooperative precoder index feedback. For a given feedback overhead, the sum-rate performance is assessed and the gains compared with a conventional massive MIMO setup without D2D are shown.**

I. Introduction

Very large antenna array or massive MIMO networks have impressive potentials to combat interference [1]–[3] based on simple beamforming techniques without requiring complex inter-cell coordination approaches. The challenge of this kind of systems concerns the channel state information (CSI) acquisition at the access point, which is crucial for downlink transmission. In [2], Marzetta limited the applicability of massive MIMO networks to time division duplex (TDD) mode. By appealing to the reciprocity principle, TDD mode enables the acquisition of the CSI for downlink by channel estimation in the uplink via an open-loop feedback scheme that avoids costly feedback. Although the reuse of the same set of pilot sequences in adjacent cells, usually referred to as pilot contamination, seems to have severe detrimental effects on the spectral [2] and power efficiency [4] of the massive MIMO networks compared with ideal CSI knowledge. Nevertheless the promised gains are still of several orders of magnitude [4] and fueled intensive research activities on massive MIMO networks in TDD mode.

In setups where channel reciprocity does not hold such as massive MIMO systems in frequency division duplex (FDD) mode, closed-loop feedback is required which consists of a preliminary phase that we refer to as channel sounding, where the access point transmits training sequences for channel estimation or local CSI acquisition at each user terminal (UT). In a second phase, shortly after CSI feedback, each UT retransmits its local CSI such that the global CSI is available at the access point. Using traditional closed-loop approaches, both the length of training sequences and the necessary feedback for large antenna arrays can become prohibitive. Very recently, building on the reduced rankness of UT channel covariance matrices in massive MIMO systems [5]–[7], promising schemes for FDD mode have been proposed in [6], [8]–[11]. In [6], [8], [9], the authors refer to the subspace actually spanned by the UT channel as the *effective channel* and they cluster UTs with almost completely overlapping effective channels. A pre-beamforming designed using only second-order channel statistics and projecting signals of each cluster onto the effective channel enables a drastic reduction of the training sequence length. Further reductions are also possible by restricting the projection to *reduced-dimension effective channel*. Then, a reduced amount of feedback is required to design the precoder on this latter subspace. In [10], [11], the authors exploit the hidden joint sparsity of the channel for clusters of UTs to reduce both training and feedback by applying compressed sensing techniques.

In [9], a drastic reduction of the required CSI is obtained in the ideal case as the access point is equipped with a large - theoretically infinite - uniform linear antenna array and each cluster consists of UTs located on a ray with origin at the access point. In those ideal conditions the selection of a reduced effective channel with negligible performance loss is strongly simplified and the required training length and CSI is drastically reduced.

The low-rankness of UT channel covariance matrices for more general antenna array settings has been studied in [5], [7]: under more realistic conditions, with UTs of a cluster randomly located in a given sector and arbitrary topology of large antenna array, the dimensions of the effective channel might be still too large.

In order to reduce the amount of feedback in non-reciprocal setups, we propose a three-phase *cooperative* closed-loop feedback as an alternative to the traditional per-user feedback loop. The new scheme exploits a novel synergy between multiuser networks with access points equipped with multiple antenna arrays and device to device (D2D) communications. More specifically, we introduce an intermediate phase in the classical closed-loop feedback: once estimated the channel parameters in the channel sounding phase, the UTs in a cluster exchange the acquired local CSI such that the global CSI is available to a master receiver in the case of centralized processing or to all the receivers in the case of distributed processing. The availability of the global CSI enables a *joint optimized design of the feedback*. We propose two methods to design the feedback under a constraint on the total amount of bits available for the feedback. The first method performs an optimal selection of the reduced effective subspace based on instantaneous knowledge of the global CSI. Then, the coefficients of the reduced effective subspace are quantized and retransmitted. The second approach benefits from the knowledge of the global CSI by selecting the best precoder from a predefined codebook. In both cases we adopt as optimality criterion the maximization of the sum-rate. In the first approach a zero-forcing precoder is implemented at the access point. Both schemes show clear performance improvements compared with the reference massive MIMO system without D2D.

II. Signal and Channel Models

We consider a massive MIMO base station serving K single-antenna users in the cell. The base station has M antennas and operates in FDD mode. The downlink channel between the base station and the k-th user is denoted by $\mathbf{h}_k^H \in \mathbb{C}^{1\times M}$. The full downlink channel can therefore be represented by

$$\mathbf{H}^H = \begin{bmatrix}\mathbf{h}_1^H \\ \vdots \\ \mathbf{h}_K^H\end{bmatrix}_{\in\mathbb{C}^{K\times M}}. \quad (1)$$

The downlink transmission is modeled by:

$$\mathbf{y} = \mathbf{H}^H\mathbf{B}\mathbf{s} + \mathbf{n}, \quad (2)$$

where $\mathbf{y} \in \mathbb{C}^{K\times 1}$ is the received signal at all users, $\mathbf{s} \in \mathbb{C}^{K\times 1}$ denotes the vector of i.i.d. Gaussian signals with zero-mean and unit-variance, and $\mathbf{n}$ represents the spatially and temporally additive white Gaussian noise (AWGN) with zero-mean and element-wise variance σ_n^2. $\mathbf{B}$ is the downlink beamformer which has the total power P.

We investigate a scenario where a group of K users are located close to each other so they can be assumed to share similar spatial statistics of channels (the study of how this assumption can be relaxed is carried out in a companion full-length paper [12]). Hence we assume the channel covariances of these users are identical, i.e., $\forall k, \mathbb{E}\{\mathbf{h}_k\mathbf{h}_k^H\} = \mathbf{R}$. Due to limited angle spread followed by incoming paths originating from high-level base station, the channel covariance $\mathbf{R}$ typically exhibits low-rank property [5] [9] [7]. We denote the rank of $\mathbf{R}$ as d. Applying eigen value decomposition (EVD) to $\mathbf{R}$:

$$\mathbf{R} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^H \quad (3)$$

Without loss of generality, we assume the eigenvalues in $\mathbf{\Sigma}$ are in descending order, so that the first d eigenvalues are non-negligible while the others can be neglected. We extract the first d columns of $\mathbf{U}$ in order to form a sub-matrix $\mathbf{U}_1 \in \mathbb{C}^{M\times d}$. The columns in $\mathbf{U}_1$ are ranked in descending order according to their average powers (or their corresponding eigenvalues). Now the channel vector $\mathbf{h}_k$ is in the column space of $\mathbf{U}_1$, e.g., $\forall k, \mathbf{h}_k$ is a linear combination of the columns of $\mathbf{U}_1$. We may write:

$$\mathbf{H} = \begin{bmatrix}\mathbf{h}_1 & \mathbf{h}_2 & \cdots & \mathbf{h}_k\end{bmatrix} = \mathbf{U}_1\mathbf{A}, \quad (4)$$

where $\mathbf{A}\in \mathbb{C}^{d\times K}$ is defined as:

$$\mathbf{A} \triangleq \begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_K\end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1K} \\ a_{21} & a_{22} & \cdots & a_{2K} \\ \vdots & \vdots & \dots & \vdots \\ a_{d1} & a_{d2} & \cdots & a_{dK}\end{bmatrix}. \quad (5)$$

The channel vector $\mathbf{h}_k, 1 \le k \le K$, is assumed to be $M\times 1$ complex Gaussian, undergoing correlation due to the finite multipath angle spread at the base station side [13]:

$$\mathbf{h}_k = \mathbf{R}^{1/2}\mathbf{h}_{\mathrm{W}k} = \mathbf{U}\mathbf{\Sigma}^{1/2}\mathbf{U}^H\mathbf{h}_{\mathrm{W}k}, k = 1, 2, \ldots, K, \quad (6)$$

where $\mathbf{h}_{\mathrm{W}k} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ is the spatially white $M\times 1$ SIMO channel, $\mathbf{I}_M$ is the $M\times M$ identity matrix, and $\mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ denotes zero-mean complex Gaussian distribution with covariance matrix $\mathbf{I}_M$. From (6) and (5) we may readily obtain the distribution of $\mathbf{a}_k$ as:

$$\mathbf{a}_k \sim \mathcal{CN}(0, \mathbf{\Sigma}_1), 1 \le k \le K, \quad (7)$$

where $\mathbf{\Sigma}_1$ is a diagonal matrix with the d greatest eigenvalues of $\mathbf{R}$ on its diagonal in decending order.

III. Feedback Design without D2D

A traditional feedback strategy for multiuser system is to let each user quantize its downlink channel vector and then send the quantized information back to the base station [14]. In the subsequent precoding stage the user signals cannot be made perfectly orthogonal to each other because of the quantization errors [15]. Note, in massive MIMO regime only partial channel information (namely K coefficients per user) is needed to achieve orthogonality between user signals. This gives a first step towards increasing the CSI quality with given amount of feedback overhead. This is highlighted by simple Proposition 1 below.

By extracting N ($K \le N \le d$) rows from $\mathbf{A}$, we form a matrix $\mathbf{A}_s\in \mathbb{C}^{N\times K}$; and by extracting the corresponding N columns of $\mathbf{U}_1$, we form a matrix $\mathbf{U}_s\in \mathbb{C}^{M\times N}$. We can partially reconstruct the channel matrix as follows:

$$\widetilde{\mathbf{H}} \triangleq \mathbf{U}_s\mathbf{A}_s. \quad (8)$$

A zero-forcing (ZF) beamformer based on the incomplete CSI can be written as:

$$\mathbf{B} = \frac{\sqrt{P}\widetilde{\mathbf{H}}^\dagger}{||\widetilde{\mathbf{H}}^\dagger||_{\mathrm{F}}}, \quad (9)$$

where $||\cdot||_{\mathrm{F}}$ denotes the Frobenius norm, and $\widetilde{\mathbf{H}}^\dagger$ is the Moore-Penrose pseudoinverse:

$$\widetilde{\mathbf{H}}^\dagger = \widetilde{\mathbf{H}}^H(\widetilde{\mathbf{H}}\widetilde{\mathbf{H}}^H)^{-1}.$$

Proposition 1 *The ZF beamformer (9) is able to eliminate inter-user interference completely.*

Proof: We may rewrite $\widetilde{\mathbf{H}}^\dagger$ as follows:

$$\widetilde{\mathbf{H}}^\dagger = \mathbf{U}_s\mathbf{A}_s(\mathbf{A}_s^H\mathbf{A}_s)^{-1}. \tag{10}$$

The channel model is now

$$\mathbf{y} = \mathbf{A}^H\mathbf{U}^H\mathbf{U}_s\mathbf{A}_s(\mathbf{A}_s^H\mathbf{A}_s)^{-1}\frac{\sqrt{P}}{||\widetilde{\mathbf{H}}^\dagger||_{\mathrm{F}}}\mathbf{s} + \mathbf{n}. \tag{11}$$

Since the following equation holds:

$$\mathbf{A}^H\mathbf{U}^H\mathbf{U}_s = \mathbf{A}_s^H, \tag{12}$$

the received signal vector is

$$\mathbf{y} = \frac{\sqrt{P}}{||\widetilde{\mathbf{H}}^\dagger||_{\mathrm{F}}}\mathbf{s} + \mathbf{n}. \tag{13}$$

Hence interference is nulled, proving Proposition 1. □

Proposition 1 indicates that we may form an $\mathbf{A}_s$ by extracting *any* linearly independent K out of d rows of $\mathbf{A}$. If the base station knows this incomplete CSI, it will ensure the users receive zero interference after downlink beamforming (9). When CSI exchange between users is not possible, a typical choice of eigenvectors is the first K rows in $\mathbf{U}_1$, as they are the strongest K eigen modes in statistical point of view. Denote the index of the i-th chosen eigenvector as e_i, and define the set of chosen indices as $\mathcal{G} \triangleq \{e_1, \cdots, e_N\}$ so that $\forall i, 1 \le e_i \le d$, and that $\forall i \neq j, e_i \neq e_j$. When CSI exchange between users is not allowed, the users select the following set of rows from $\mathbf{A}$:

$$\mathcal{G}^{(1)} = \{1, \cdots, K\}, \tag{14}$$

which is known by the base station by default. The users will only send back a quantized version of $\mathbf{A}_s$, i.e., the first K rows of $\mathbf{A}$.

IV. Cooperative Feedback of CSI with D2D

Once CSI exchange is allowed between users, the users can make a joint decision of which set of eigenvectors in $\mathbf{U}_1$ to choose, or equivalently which set of rows of $\mathbf{A}$ to extract in order to form $\mathbf{A}_s$. As a simple example, we may consider the signal-to-noise ratio (SNR) at user side as a criterion:

$$\mathrm{SNR} = \frac{P}{||\widetilde{\mathbf{H}}^\dagger||_{\mathrm{F}}^2} = \frac{P}{\mathrm{tr}\{(\mathbf{A}_s^H\mathbf{A}_s)^{-1}\}}. \tag{15}$$

We choose K out of d eigenvectors from $\mathbf{U}_1$ so that the SNR is maximized.

$$\mathcal{G}^{(2)} = \underset{N=K}{\arg}\min \mathrm{tr}\{(\mathbf{A}_s^H\mathbf{A}_s)^{-1}\}. \tag{16}$$

The optimality of the eigenvector selection decision is achieved by exhaustive search. This problem is loosely reminiscent of the TX antenna selection in conventional MIMO systems. The number of possible candidates is $\binom{d}{K}$, which scales exponentially with K. Despite its optimality, the exhaustive method has a high computational complexity, which necessitates low-complexity selection algorithms. Gorokhov et al. proposed a decremental selection algorithm in [16] [17] for the purpose of antenna selection. In his approach, the rows of the channel matrix are removed one by one, while minimizing the capacity degradation. This approach can be adapted to the selection of eigenvectors. We start with the full effective channel $\mathbf{A}$. Under the condition that the SNR reduction is minimum, the rows of $\mathbf{A}$ are removed one by one, until we have K rows left. The searching space is reduced from $\binom{d}{K}$ down to an order of d^2.

After the joint decision is made, i.e., $\mathcal{G}^{(2)}$ is obtained, the users will send back the corresponding $\mathbf{A}_s$ (quantized), as well as the indices in $\mathcal{G}^{(2)}$ to the base station.

V. Cooperative Feedback of Precoder Index with D2D

We now consider another cooperative feedback design approach based on precoder feedback. Local CSI exchange via D2D communications enables users to jointly choose a precoding matrix, which makes the multi-user system analogous to a point-to-point MIMO system. Some classical precoder selection schemes [18], [19] can directly apply. A precoder codebook is composed of a finite set of precoding matrices predetermined a priori. Such a codebook is used to approximate the precoder, e.g., the normalized Moore-Penrose pseudoinverse of $\mathbf{A}$ for a ZF type precoder. The codewords can be generated according to the distribution of $\mathbf{A}$, which is given in (7). The codebook design and optimization are non-trivial and out of the scope of this paper. Nevertheless a simple random codebook generation method is given in Section VI-B. After local CSI exchange, the users jointly select the best precoding matrix according a certain criterion and feed back the index of the selected precoding matrix. The selection criterion may vary depending on the complexity requirement of the system. However an intuitive choice is the downlink sum-rate. Assume the codebook $\mathcal{X}$ is known at both the base station and the master UT. Any codeword $\mathbf{X} \in \mathcal{X}$ is of the size $d \times K$. The codewords are normalized such that $\forall \mathbf{X} \in \mathcal{X}, ||\mathbf{X}||_F = \sqrt{P}$. If $\mathbf{X}$ is chosen, the downlink precoding matrix is $\mathbf{U}_1\mathbf{X}$. Since the users know the instantaneous channel, they are able to compute the downlink sum-rate when a certain precoder $\mathbf{X}$ is selected. We define $\underline{\mathbf{H}} \triangleq \mathbf{A}^H\mathbf{X}$. The downlink data model is now

$$\mathbf{y} = \underline{\mathbf{H}}\mathbf{s} + \mathbf{n}. \tag{17}$$

The rate of user k is:

$$r_k \triangleq \log_2\left(1 + \frac{|\underline{\mathbf{H}}_{k,k}|^2}{\sigma_n^2 + \sum\limits_{l\neq k}|\underline{\mathbf{H}}_{k,l}|^2}\right), \tag{18}$$

where $\underline{\mathbf{H}}_{k,l}$ stands for the (k,l)-th element of $\underline{\mathbf{H}}$. The downlink sum-rate is defined as

$$\mathcal{C} \triangleq \sum_{k=1}^{K} r_k. \tag{19}$$

A description of the algorithm is as follows:

(1) The slave UTs send their measured downlink CSI to the master UT. The master UT now has the effective channel matrix $\mathbf{A}$.

(2) The master UT searches the codebook and finds the index of the precoder that maximizes the sum-rate:

$$i = \arg_{X_i \in \mathcal{X}} \max\{\mathcal{C}\}. \tag{20}$$

(3) The master UT feeds back the index i to the base station.

(4) The base station performs downlink beamforming (17) using the selected precoder.

VI. Numerical Results

This section contains numerical evaluations of different feedback mechanisms. First we introduce our physical channel model. We assume the base station antennas form a uniform linear array (ULA). It is worthwhile to note that the proposed methods of this paper and their results also hold for some other different settings of antenna placement, e.g., the random linear array, two dimensional uniform array, or even the two-dimensional distributed array [7]. For ease of exposition we take a ULA for example. The downlink channel between the base station and the k-th user is obtained by the following model [20]:

$$\mathbf{h}_k^H = \sqrt{\beta_k} \sum_{q=1}^{Q} \left(\mathbf{a}(\theta_{kq})\right)^H e^{j\varphi_{kq}}, \tag{21}$$

where Q is the number of i.i.d. paths, β_k denotes the path loss for channel $\mathbf{h}_k$, and it is dependent on the prescribed average SNR at cell edge. $e^{j\varphi_{kp}}$ is the i.i.d. random phase, which is independent over channel index k and path index q. $\mathbf{a}(\theta)$ is the signature (or phase response) vector by the array to a path originating from the angle θ, as shown in [21]

$$\mathbf{a}(\theta) \triangleq \begin{bmatrix} 1 \\ e^{-j2\pi \frac{D}{\lambda}\cos(\theta)} \\ \vdots \\ e^{-j2\pi \frac{(M-1)D}{\lambda}\cos(\theta)} \end{bmatrix}, \tag{22}$$

where D is the antenna spacing at the base station and λ is the signal wavelength.

In simulations of this paper, we assume the users share the same scattering environment, giving rise to identical channel covariance matrix for all users. In other words, θ_{kq} has an i.i.d. distribution, and $\forall k, \beta_k = \beta$. This is the worst case scenario due to the resemblance of the channel of all users.

We consider a cluster of 3 single-antenna UTs being served by a base station equipped with 50 antennas. The cell radius is 1000 meters and the users are located 800 meters away from the base station. The angle of departure (AoD) of any user channel follows a uniform distribution from 80 degrees to 100 degrees, i.e., $\forall k, \forall q, \theta_{kq} \sim \mathcal{U}(80, 100)$. Due to limited angle spread (20 degrees), the rank of the channel covariance $\mathbf{R}$ is around $d = 15$ (see [5]). Since channel covariance is assumed known by the users, we consider only the reduced-dimension subspace (effective subspace), where the 50×1 channel vector can be effectively represented by a linear combination of 15 eigenvectors of the channel covariance.

When CSI exchange is not allowed between users, the traditional approach is that each user quantizes its own channel and sends back the quantized CSI to the base station. Then the base station designs a precoder based on the quantized CSI. We introduce this method as a reference system.

A. Cooperative CSI Feedback

Three different CSI feedback regimes are evaluated and the sum-rate performances are given in Fig. 1. For the sake of fairness, we assume the same amount of quantization bits is available in the three regimes. Each user will send back 16 bits of information to the base station. The curve "Quantize full CSI, no D2D" denotes the sum-rate performance of the conventional method when the full effective channel matrix $\mathbf{A}$ is quantized and sent back to base station. In this approach each user quantizes its effective channel vector, i.e., its corresponding column in $\mathbf{A}$, into 16 bits. Upon the reception of users' feedback, the base station constructs the effective channel matrix and performs ZF precoding based on it. The curve "Use strongest K eigen modes, no D2D" shows the performance of each user using K dominant eigenvectors of $\mathbf{R}$, as described in the reference system (14). The curve "Cooperative feedback of CSI, D2D" refers to the proposed feedback regime where the users exchange CSI and search using the decremental selection method for the best 3 eigen modes. The users need to feed back the quantized projections and the indices of the three eigenvectors that are chosen. We omit the result of exhaustive search due to the fact that it has higher complexity yet negligible performance improvement compared with the low-complexity decremental method. Despite the fact that the D2D method has fewer quantization bits available, which results from the requirements of feeding back the indices, we still observe a clear performance gain of the D2D method.

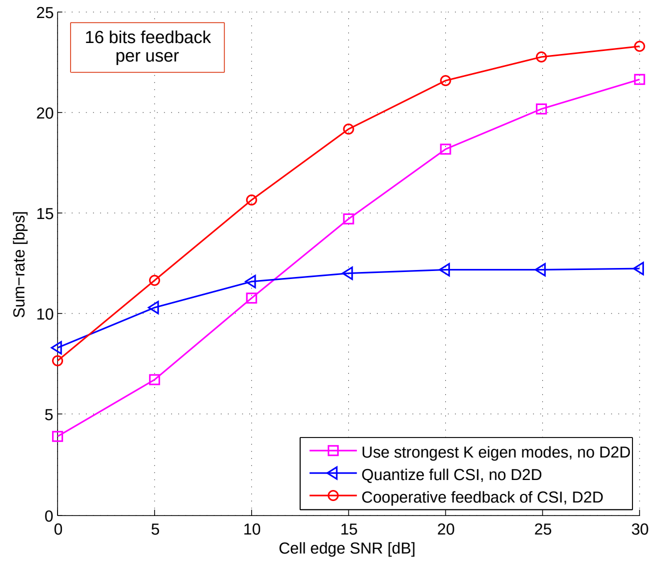

Fig. 1. DL sum-rates with/without feedback cooperation, feedback overhead: 16 bits per user.

B. Cooperative Precoder Index Feedback

In the following we will evaluate the performance of precoder index feedback method. We keep the simulation settings the same as in Section VI-A except that the amount of feedback overhead is now 4 bits per user. We still work on the effective subspace. The curve “Quantize full CSI, no D2D” and “Use strongest K eigen modes, no D2D” in Fig. 2 denote the same non-cooperative methods as shown in Section VI-A. We omit the performance of cooperative CSI feedback here due to the lack of available quantization bits, as the representation of the indices of the chosen eigen modes alone requires $\log_2 \binom{d}{K} \approx 9$ bits.

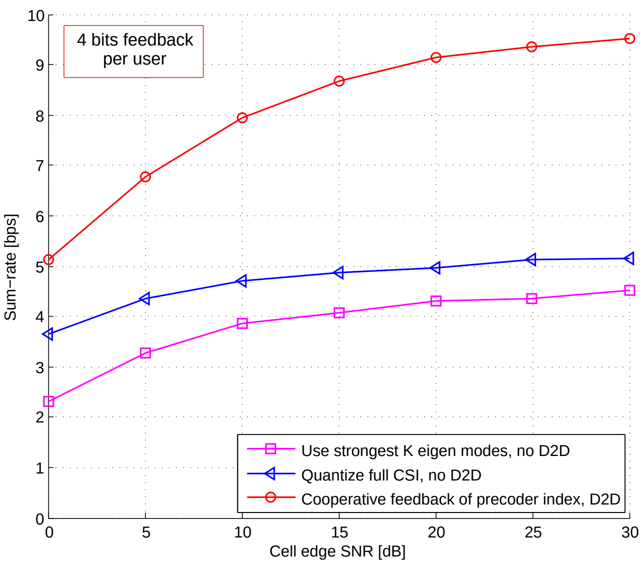

Fig. 2. DL sum-rates of cooperative precoder selection and non-cooperative CSI feedback, feedback overhead: 4 bits per user.

Once D2D is enabled, the users can compute the system performance, i.e., the sum-rates, when different precoders in the codebook are used, and finally pick the best precoder. Note that in simulation, we generate a random codebook as follows: 1) generate a fixed number of realizations of $\mathbf{A}$ according to its distribution (7); 2) compute the normalized Moore-Penrose pseudoinverse of $\mathbf{A}$ for each realization. For the sake of fairness, the cooperative precoder index feedback scheme, marked with “Cooperative feedback of precoder index, D2D”, also has totally 12 bits of feedback overhead available. We can observe significant gain of this method over the traditional ones when D2D is not possible.

Finally, we would like to remark that the cooperative precoder index feedback approach has higher complexity due to exhaustive search. However it has better performance than cooperative CSI feedback, especially when the feedback overhead is small.

VII. Conclusions

We proposed a new cooperative feedback framework for FDD massive MIMO whereby devices rely on local CSI exchange so as to compute a suitable feedback signal. We show two approaches for feedback design, cooperative CSI feedback and cooperative precoder index feedback. These methods help reduce feedback overhead for FDD massive MIMO systems for a given sum-rate performance target, compared with the conventional non-cooperative feedback design.

Acknowledgment

This work was supported by Huawei France Research Center.

References

[1] T. L. Marzetta, “How much training is required for multiuser MIMO?” in *Proc. of the Fortieth Asilomar Conf. on Signals, Systems, & Computers*, Pacific Grove, CA, USA, Oct. 2006.

[2] ——, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” *IEEE Trans. Wireless Commun*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[3] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, “Massive MIMO for next generation wireless systems,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[4] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, “Energy and spectral efficiency of very large multiuser MIMO systems,” *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[5] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, “A coordinated approach to channel estimation in large-scale multiple-antenna systems,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

[6] 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.

[7] H. Yin, D. Gesbert, and L. Cottatellucci, “Dealing with interference in distributed large-scale MIMO systems: A statistical approach,” *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 942–953, Oct 2014.

[8] J. Nam, J. Ahn, and G. Caire, “Joint spatial division and multiplexing: Realizing massive MIMO gains with limited channel state information,” in *46th Annual Conference on Information Sciences and Systems, (CISS 2012)*, Princeton University, NJ, USA, Mar. 2012.

[9] A. Adhikary and G. Caire, “Joint spatial division and multiplexing: Opportunistic beamforming and user grouping,” *arXiv preprint arXiv:1305.7252*, 2013.

[10] X. Rao, V. K. Lau, and X. Kong, “CSIT estimation and feedback for FDD multi-user massive MIMO systems,” in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),*, 2014, pp. 3157–3161.

[11] X. Rao and V. K. Lau, “Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems,” *IEEE Trans Signal Process.*, vol. 62, no. 12, pp. 3261–3271, 2014.

[12] H. Yin, L. Cottatellucci, and D. Gesbert, “Cooperative feedback designs in FDD massive MIMO systems,” in preparation.

[13] A. J. Paulraj, R. Nabar, and D. Gore, *Introduction to space-time wireless communications*. Cambridge University Press, 2003.

[14] N. Jindal, “MIMO broadcast channels with finite-rate feedback,” *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, 2006.

[15] P. Ding, D. J. Love, and M. D. Zoltowski, “Multiple antenna broadcast channels with shape feedback and limited feedback,” *IEEE Trans Signal Process.*, vol. 55, no. 7, pp. 3417–3428, 2007.

[16] A. Gorokhov, “Antenna selection algorithms for MEA transmission systems,” in *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 3, Orlando, FL, USA, May 2002, pp. III–2857–III–2860.

[17] A. Gorokhov, D. A. Gore, and A. J. Paulraj, “Receive antenna selection for MIMO spatial multiplexing: theory and algorithms,” *IEEE Trans Signal Process.*, vol. 51, no. 11, pp. 2796–2807, 2003.

[18] D. J. Love and R. W. Heath, “Limited feedback unitary precoding for spatial multiplexing systems,” *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2967–2976, 2005.

[19] D. J. Love, R. W. Heath, V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, “An overview of limited feedback in wireless communication systems,” *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.

[20] A. F. Molisch, *Wireless communications*. Wiley, 2010.

[21] W. C. Jakes, *Mobile microwave communication*. Wiley, 1974.