On the Achievable Rates of FDD Massive MIMO Systems with Spatial Channel Correlation

Zhiyuan Jiang, SIEEE,
Tsinghua University,
Beijing, 100084, China.
jiang-zy10@mails.tsinghua.edu.cn

Andreas F. Molisch, FIEEE, Giuseppe Caire, FIEEE,
University of Southern California,
Los Angeles, CA 90089-2565, USA.
molisch@usc.edu; caire@usc.edu

Zhisheng Niu, FIEEE,
Tsinghua University,
Beijing, 100084, China.
niuzhs@tsinghua.edu.cn

***Abstract*—In this paper, we study the optimization of the achievable rates of frequency-division-duplex (FDD) massive multiple-input-multiple-output (MIMO) systems with *spatially correlated channels*, by designing the downlink channel training sequences and the uplink channel feedback codebooks. In particular, the optimal channel training sequences and a Karhunen-Loeve transform followed by entropy coded scalar quantization codebook are proposed to optimize the achievable rates. We compare our achievable rates with time-division-duplex (TDD) massive MIMO systems, i.i.d. FDD systems, and the joint spatial division and multiplexing (JSDM) scheme. It is shown that, the rate-gap between FDD systems and TDD systems is significantly narrowed. Compared to the JSDM scheme, our proposal achieves dimensionality-reduction channel estimation without channel pre-projection, and higher throughput in general, though at higher computational complexity.**

I. INTRODUCTION

Since the inception of massive multiple-input-multiple-output (MIMO) systems [1], most existing work has focused on the time-division-duplex (TDD) mode. This is due to the fact that acquiring channel state information at the transmitter (CSIT) using the closed-loop channel estimation method, which the frequency-division-duplex (FDD) and the uncalibrated TDD systems must deal with, generally entails prohibitively large overhead. Fortunately, it is found that the channel correlation matrices (CCMs) of the channel coefficients, whose estimation cost is drastically lower than instantaneous CSIT due to their slower variation, can greatly help to reduce this overhead [2]–[4].

In this work, utilizing the CCMs, we design a novel approach to optimize the achievable rates of FDD massive MIMO systems. If the CCMs are known at both the BS side and the user side, and if they are low-rank (which commonly occurs in practical deployments), they can be exploited to enable *dimensionality-reduced* channel training and feedback. In particular, we propose an iterative algorithm to find the optimal channel training sequences in this scenario and a Karhunen-Loeve (KL) transform followed by entropy coded scalar quantization (SQ) with reverse water filling bit-loading feedback codebook design (KLSQ). We derive deterministic equivalents of the achievable rates for our schemes with a regularized-zero-forcing (RZF) precoder, considering distinct CCMs of different users, the dimensionality loss due to channel estimation, and the imperfection of acquired CSIT. The proposed approach requires minimal modifications of the widely-adopted pilot-assisted scheme, thus making it desirable to implement in practice. In fact, the proposed channel training and feedback schemes can be seen as an alternative to the pre-projection and effective channel approach in JSDM. Due to the limited space, additional detailed results, as well as proofs of our theorems, can be found in [5].

The remainder of the paper is organized as follows. In Section II, the system model is characterized. In Section III, we specify the proposed eigenspace channel training and feedback schemes, and derive the deterministic equivalents of the achievable rates. Section IV gives the simulation results, including the comparison with TDD and i.i.d. FDD systems, and the JSDM scheme, under various system parameters. Finally, in Section V, we conclude our work.

Notations : Throughout the paper, we use boldface uppercase letters, boldface lowercase letters and lowercase letters to designate matrices, column vectors and scalars, respectively. $\boldsymbol{X}^\dagger$ denotes the complex conjugate transpose of matrix $\boldsymbol{X}$. $\boldsymbol{X}(:,i)$ denotes the i-th column of $\boldsymbol{X}$. x_i denotes the i-th element of vector $\boldsymbol{x}$. diag$[x_1, x_2, ..., x_n]$ denotes a diagonal matrix with $x_1, x_2, ..., x_n$ on its diagonal. $\det(\boldsymbol{X})$ and tr$(\boldsymbol{X})$ denote the determinant and the trace of matrix $\boldsymbol{X}$, respectively. $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes a circularly symmetric complex Gaussian random vector of mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The logarithm $\log(x)$ denotes the binary logarithm. We use Cov$(\cdot)$ to denote the covariance matrix of a random vector.

II. SYSTEM MODEL

We consider a downlink broadcast channel (BC), where an M-antenna BS serves N single-antenna users. The receive signal of the n-th user is expressed as

$$y_n = \boldsymbol{h}_n^\dagger \boldsymbol{W}\boldsymbol{s} + n_n, \qquad (1)$$

where $\boldsymbol{h}_n$ is the channel vector of user-n, $\boldsymbol{s} \in \mathcal{C}^N$ is the data transmitted to the users, $\boldsymbol{x} = \boldsymbol{W}\boldsymbol{s}$ denotes the precoded downlink signals, $\boldsymbol{W} \in \mathcal{C}^{M\times N}$ denotes the precoding matrix,

This work is sponsored in part by the National Basic Research Program of China (973 Program: 2012CB316001), the National Science Foundation of China (NSFC) under grant No. 61201191 and No. 61321061, Hitachi R&D Headquarter, and Intel Research under the 5G Program.

and $\boldsymbol{y} \in \mathcal{C}^N$ are the received signals of users. The downlink total transmit power constraint is

$$\text{tr}\left\{\mathbb{E}[\boldsymbol{W}\boldsymbol{s}\boldsymbol{s}^\dagger\boldsymbol{W}^\dagger]\right\} \leq P, \tag{2}$$

and $\boldsymbol{n} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_N)$ is the Gaussian distributed uncorrelated noise.

A. Spatial Correlated Channel Matrix

Define the compound downlink channel matrix $\boldsymbol{H} = [\boldsymbol{h}_1, \boldsymbol{h}_2, ..., \boldsymbol{h}_N]^\dagger$, where $\boldsymbol{h}_n \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{R}_n)$. The CCM of user n is

$$\boldsymbol{R}_n = \mathbb{E}\left[\boldsymbol{h}_n\boldsymbol{h}_n^\dagger\right], \tag{3}$$

where by the Karhunen-Loeve representation,

$$\boldsymbol{h}_n = \boldsymbol{R}_n^{\frac{1}{2}}\boldsymbol{z}_n, \tag{4}$$

where $\boldsymbol{z}_n \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_M)$. It is assumed that the channel vectors of users are mutually *independent*, since users are usually well separated. Denote the singular-value-decomposition (SVD) of the CCM as $\boldsymbol{R}_n = \boldsymbol{U}_n\boldsymbol{\Sigma}_n\boldsymbol{U}_n^\dagger$, and $\boldsymbol{\Sigma}_n = \text{diag}[\lambda_1^{(n)}, \lambda_2^{(n)}, ..., \lambda_M^{(n)}]$. It is worthwhile to mention that in this work, we assume the BS and the users have *perfect* knowledge of the CCMs.

III. FDD Massive MIMO Achievable Rates

We now propose our pilot-assisted FDD transmission scheme, which consists of three steps:

- Downlink channel training.
- Uplink CSIT feedback.
- Data transmission.

In the following, we will investigate the aforementioned steps in order, namely design of optimal channel training sequences, a KLSQ feedback codebook design, and the deterministic equivalents of the data transmission rates.

A. Optimal Downlink Training with Per-User CCM

The signal model of the channel training phase is expressed as

$$\begin{aligned} &\boldsymbol{Y}_\tau = \boldsymbol{H}\boldsymbol{X}_\tau + \boldsymbol{N}_\tau \\ &\text{tr}\left[\boldsymbol{X}_\tau\boldsymbol{X}_\tau^\dagger\right] \leq \tau P, \end{aligned} \tag{5}$$

where $\boldsymbol{X}_\tau$ is a $M \times \tau$ training signal matrix, containing the training sequences which is known to the BS and the users. τ is the training length, and $\boldsymbol{Y}_\tau = [\boldsymbol{y}_{\tau,1}, \boldsymbol{y}_{\tau,2}, ..., \boldsymbol{y}_{\tau,N}]^\dagger$ is the corresponding channel output observed by the user, disturbed by Gaussian noise $\boldsymbol{N}_\tau$ with i.i.d. unit variance entries. The n-th user observes

$$\boldsymbol{y}_{\tau,n}^\dagger = \boldsymbol{h}_n^\dagger\boldsymbol{X}_\tau + \boldsymbol{n}_{\tau,n}^\dagger, \tag{6}$$

and applies the minimum-mean-square-error (MMSE) estimation [6, Section 19.5]

$$\hat{\boldsymbol{h}}_n^\dagger = \boldsymbol{R}_n\boldsymbol{X}_\tau(\boldsymbol{X}_\tau^\dagger\boldsymbol{R}_n\boldsymbol{X}_\tau + \boldsymbol{I}_\tau)^{-1}\boldsymbol{y}_{\tau,n}^\dagger. \tag{7}$$

Notice that we assume the CCMs are known to both the users and the BS. Applying the MMSE decomposition, the user channel $\boldsymbol{h}_n$ and the covariance matrix of the channel estimation error due to imperfect channel training are expressed as [7]

$$\begin{aligned} \boldsymbol{h}_n &= \hat{\boldsymbol{h}}_n + \hat{\boldsymbol{e}}_n, \\ \boldsymbol{C}_{\hat{e}_n} &= (\boldsymbol{R}_n^{-1} + \boldsymbol{X}_\tau\boldsymbol{X}_\tau^\dagger)^{-1}, \end{aligned} \tag{8}$$

respectively. By assumption, $\boldsymbol{R}_n$ is the CCM, thus it may be rank-deficient and not invertible. Nonetheless, let $\bar{\boldsymbol{R}}_n = \boldsymbol{R}_n + \epsilon\boldsymbol{I}_M$ such that ϵ is small but $\bar{\boldsymbol{R}}_n$ is invertible. Then (8) holds true if we substitute $\bar{\boldsymbol{R}}_n$ for $\boldsymbol{R}_n$. Then we can let $\epsilon \to 0$ due to the continuity of the function involved.

In [8], the optimal training sequences where users have *identical* CCMs are given, in the sense of minimizing the MSE or the mutual information between the channel coefficients and received signals conditioned on the transmitted block signals. However, the situation is different in the multi-user case, since different users have different CCMs yet share the same downlink training sequences. Thus the training sequence can no longer match one specific CCM, as in the case where user CCMs are identical [8]. We now find the optimal training sequences, in terms of maximizing the conditional mutual information (CMI) between the channel vector and the received signal. The optimization problem, given the training length τ and total transmit power P is first expressed as,

$$\begin{aligned} \text{maximize:} \quad & \sum_{n=1}^{N} \log\det\left(\boldsymbol{I} + \boldsymbol{X}_\tau^\dagger\boldsymbol{R}_n\boldsymbol{X}_\tau\right) \\ \text{s.t.} \quad & \text{tr}\left[\boldsymbol{X}_\tau\boldsymbol{X}_\tau^\dagger\right] \leq \tau P, \end{aligned} \tag{9}$$

and we have the following theorem.

Theorem 1: The training sequences that maximize the CMI satisfy following condition

$$\sum_{n=1}^{N}\left[\boldsymbol{R}_n\boldsymbol{X}_{\text{opt}}\left(\boldsymbol{I}_\tau + \boldsymbol{X}_{\text{opt}}^\dagger\boldsymbol{R}_n\boldsymbol{X}_{\text{opt}}\right)^{-1}\right] = \lambda\boldsymbol{X}_{\text{opt}}, \tag{10}$$

where $\lambda \geq 0$ is a constant chosen to satisfy the power constraint.

Proof: The proof is straightforward by deriving the KKT conditions of the Lagrangian dual problem of (9). ■

Based on this theorem, the optimal channel training sequences can be found by an iterative algorithm, which is specified in our journal paper [5] and is omitted for brevity.

B. Uplink CSIT Feedback

After the users estimate their respective channel coefficients based on received channel training signals, they feed back their estimates using predefined codebooks. In this subsection, we propose the entropy encoded scalar quantization after KL transform, which is a simple way to universally approach the optimal vector quantization (VQ) performance. The comparison with two near-optimal VQ approaches can be found in [5].

1) Entropy Coded Scalar Quantization: We consider a scalar quantization (component by component) of the transformed channel vector. Specifically, denote

$$\hat{\boldsymbol{h}}_n^{\text{KL}} = \boldsymbol{U}_n^{\dagger}\hat{\boldsymbol{h}}_n = \boldsymbol{\Sigma}_n^{\frac{1}{2}}\boldsymbol{z}_n - \boldsymbol{U}_n^{\dagger}\hat{\boldsymbol{e}}_n \tag{11}$$

as the KL-transform of the channel vector of user-n, after channel traning. Putting aside the channel training error $\hat{\boldsymbol{e}}_n$, this yields M mutually independent Gaussian variables with non-identical variances. The reverse water-filling approach (RWF) [9] can be implemented to achieve the rate-distortion function in this scenario, i.e., we allocate the quantization bits according to the following conditions

$$\begin{aligned} &\sum_{i=1}^{M} \min\left[\gamma, \lambda_i^{(n)}\right] = D \\ &R_i = \log\left(\frac{\lambda_i^{(n)}}{\gamma}\right) \\ &\sum_{i=1}^{M} R_i = B_n, \end{aligned} \tag{12}$$

where D is the total MSE distortion, R_i denotes the number of bits allocated to the i-th component of $\hat{\boldsymbol{h}}_n^{\text{KL}}$, B_n is the total number of feedback bits for user-n, and γ denotes the water level. The MSE distortion for the i-th component is

$$D_i = \min\left[\gamma, \lambda_i^{(n)}\right]. \tag{13}$$

After the BS recovers the KL-transformed channel vector from the user feedback, it can reconstruct the channel vector by the inverse KL-transform. By this scheme, we obtain the relationship between the channel estimation at the BS side and the real channel, i.e.,

$$\boldsymbol{h_n} = \hat{\hat{\boldsymbol{h}}}_n + \underbrace{\hat{\boldsymbol{e}}_n + \boldsymbol{U}_n\hat{\hat{\boldsymbol{e}}}_n}_{\boldsymbol{\varepsilon}_n}, \tag{14}$$

$$\text{Cov}(\boldsymbol{\varepsilon}_n) = \underbrace{\boldsymbol{C}_{\hat{e}_n}}_{\mathcal{M}_1} + \underbrace{\boldsymbol{U}_n\boldsymbol{D}_n\boldsymbol{U}_n^{\dagger}}_{\mathcal{M}_2} \tag{15}$$

$$\hat{\hat{\boldsymbol{R}}}_n \triangleq \text{Cov}(\hat{\hat{\boldsymbol{h}}}_n) = \boldsymbol{R}_n - \text{Cov}(\boldsymbol{\varepsilon}_n), \tag{16}$$

where $\boldsymbol{C}_{\hat{e}_n}$ is defined in (8), and $\boldsymbol{D}_n \triangleq \text{diag}\left[D_1, D_2, ..., D_M\right]$. Observing the error covariance matrix in (15), $\mathcal{M}_1$ and $\mathcal{M}_2$ represent channel estimation error due to imperfect channel training and CSI quantization error respectively.

Remark 1: There are several approaches to mimic such behavior using a scalar quantizer, e.g., apply a Huffman code on each of the components with $\lambda_i^{(n)} > \gamma$, based on the fact that the component is Gaussian distributed with variance $\lambda_i^{(n)}$. The advantage of this quantizer is that it does not involve any VQ, thus can be implemented very efficiently in parallel. Notice also that when $\boldsymbol{U}_n$ is a slice of a DFT matrix (as in large linear antenna arrays), the KL-transform can be well approximated by an FFT, therefore the overall quantization can be made extremely computationally efficient.

C. Data Transmission

For fair comparison, also in line with the work in [3] and [10], we consider the RZF linear precoder schemes. The precoder treats the channel estimates as the real channel coefficients. Corresponding achievable rates on account of the imperfect channel estimations are computed in the following section. The RZF precoding matrix is expressed as

$$\boldsymbol{W}_{\text{rzf}} = \zeta\left(\hat{\hat{\boldsymbol{H}}}^{\dagger}\hat{\hat{\boldsymbol{H}}} + M\alpha\boldsymbol{I}_M\right)^{-1}\hat{\hat{\boldsymbol{H}}}^{\dagger}, \tag{17}$$

where $\hat{\hat{\boldsymbol{H}}} = \left[\hat{\hat{\boldsymbol{h}}}_1, \hat{\hat{\boldsymbol{h}}}_2, ..., \hat{\hat{\boldsymbol{h}}}_N\right]^{\dagger}$, ζ is a normalization scalar to fulfill the power constraint in (2), and α is the regularization factor. Based on (2), we obtain

$$\zeta^2 = \frac{N}{\text{tr}\left[\hat{\hat{\boldsymbol{H}}}\left(\hat{\hat{\boldsymbol{H}}}^{\dagger}\hat{\hat{\boldsymbol{H}}} + M\alpha\boldsymbol{I}_M\right)^{-2}\hat{\hat{\boldsymbol{H}}}^{\dagger}\right]}, \tag{18}$$

where equal power allocation is assumed, i.e., $\left[\mathbb{E}[\boldsymbol{s}\boldsymbol{s}^{\dagger}]\right]_{i,i} = \frac{P}{N}$. Define

$$\boldsymbol{K}_{\text{rzf}} = \left(\hat{\hat{\boldsymbol{H}}}^{\dagger}\hat{\hat{\boldsymbol{H}}} + M\alpha\boldsymbol{I}_M\right)^{-1}, \tag{19}$$

the signal-to-interference-and-noise-ratio (SINR) of user n is

$$\gamma_{n,\text{rzf}} = \frac{\left|\hat{\hat{\boldsymbol{h}}}_n^{\dagger}\boldsymbol{K}_{\text{rzf}}\hat{\hat{\boldsymbol{h}}}_n\right|^2}{\frac{N}{P\zeta^2} + \left|\boldsymbol{\varepsilon}_n^{\dagger}\boldsymbol{K}_{\text{rzf}}\hat{\hat{\boldsymbol{h}}}_n\right|^2 + \boldsymbol{h}_n^{\dagger}\boldsymbol{K}_{\text{rzf}}\hat{\hat{\boldsymbol{H}}}_{[n]}^{\dagger}\hat{\hat{\boldsymbol{H}}}_{[n]}\boldsymbol{K}_{\text{rzf}}\boldsymbol{h}_n}, \tag{20}$$

where $\hat{\hat{\boldsymbol{H}}}_{[n]} = \left[\hat{\hat{\boldsymbol{h}}}_1, ..., \hat{\hat{\boldsymbol{h}}}_{n-1}, \hat{\hat{\boldsymbol{h}}}_{n+1}, ..., \hat{\hat{\boldsymbol{h}}}_N\right]^{\dagger}$. The training dimensionality loss is the length of the training sequence τ. Assuming the feedback information is transmitted over the uplink MIMO-multiple-access-channel (MIMO-MAC), and based on [7], the total feedback dimensionality loss is computed as

$$\delta = \frac{\sum_{n=1}^{N} B_n}{C_{\text{MIMO-MAC}}}. \tag{21}$$

For the ease of exposition, we assume $B_n = B$, $\forall n$, and

$$C_{\text{MIMO-MAC}} = \kappa \min\left[M, N\right]\log(M\mathsf{SNR}_{\text{ul}}), \tag{22}$$

where $\kappa \in (0, 1)$ is a scalar representing the diversity-multiplexing tradeoff in MIMO-MAC as defined in [7]. The achievable sum rate considering imperfect channel training and feedback, $\bar{R}_{\text{rzf}}$, is expressed as the solution of the following optimization problem

$$\begin{aligned} \text{maximize:} \quad & \left(1 - \frac{\tau + \delta}{T}\right)\sum_{n=1}^{N}\log(1 + \gamma_{n,\text{rzf}}) \\ \text{s.t.} \quad & 1 \leq \tau + \delta \leq T, \\ & \tau \geq 1, \quad \delta \geq 1, \end{aligned} \tag{23}$$

where the optimization is over the training and feedback length. Since our focus is on the performance of the downlink

BC achievable rates with correlated channels, we use an exhaustive search to find the optimal training and feedback length.

Leveraging the deterministic equivalent techniques provided in [10], with necessary modifications, we can derive analytic expressions for $\gamma_{n,\text{rzf}}$ in (23). For ease of exposition, we assume the dominant ranks we choose in the feedback schemes are identical, i.e., $r_n = r$, $\forall n$.

Following the approach in [10], when M goes to infinity, the SINR of user n, $\gamma_{n,\text{rzf}}$, satisfies

$$\gamma_{n,\text{rzf}} - \gamma^o_{n,\text{rzf}} \overset{M\rightarrow\infty}{\longrightarrow} 0 \text{ with probability } 1, \tag{24}$$

where $\gamma^o_{n,\text{rzf}}$ is a deterministic quantity that can be computed as

$$\gamma^o_{n,\text{rzf}} = \frac{\frac{(\hat{e}^o_n)^2}{(1+\hat{e}^o_n)^2}}{\frac{\phi^o}{P} + \hat{\hat{E}}^o_n + I^o_n}, \tag{25}$$

where the parameters involved are specified in [5]. The derivation is mostly based upon [10], with generalizations to uncorrelated channel estimation error matrices. The details can be found in [5].

IV. Numerical Results

In our simulations, we evaluate the FDD massive MIMO achievable rates with various spatially correlated channel models, and compare those with the TDD system, the FDD system with i.i.d. channels, and the JSDM scheme. The per-user CCMs are computed according to the one-ring (OR) model [3] and the Laplacian Model [6], with a linear antenna array, antenna spacing D, and angular spread Δ.

A. Comparison with TDD and i.i.d. FDD systems

The achievable rates under the proposed scheme are shown in Fig. 1, in comparison with i.i.d. FDD systems and also TDD systems. The achievable rates of FDD systems with correlated channels are obtained using the training sequences obtained by the iterative algorithm in Section III-A, and the KLSQ feedback codebook design in Section III-B. First, it is noteworthy that in FDD systems, in general, the achievable sum rate is not monotonously increasing with the number of BS antennas, as it does in the TDD system, due to the fact that when the number of BS antennas grows large, the channel estimation dimensionality loss of FDD will become non-negligible. Therefore, there is a large rate gap between the i.i.d. FDD system and the TDD system, rendering the FDD mode unfavorable for massive MIMO transmission.

Nevertheless, when the channel is spatially correlated, the achievable sum rate under per-user CCMs is significantly larger than that in i.i.d. channels for FDD systems, especially when the number of BS antennas is large, thanks to the judiciously designed channel estimation schemes. The rate gap between the TDD mode and the FDD mode is narrowed significantly, especially when M is moderate, which suggests that it is promising to exploit the large-system gain even with FDD.

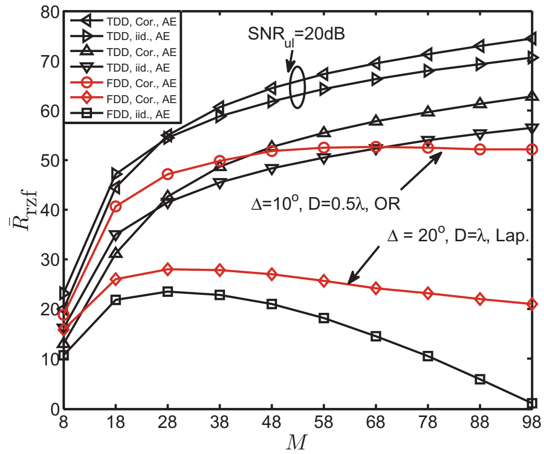

Fig. 1. Achievable sum rates in massive MIMO systems with i.i.d. channels, per-user correlation channels, TDD mode and FDD mode respectively. The downlink and uplink SNR are set to 20 dB and 10 dB, respectively, unless labeled otherwise. The channel block length is $T = 200$. The number of users in the cell is $N = 8$. The regularization factor of the RZF precoder is $\alpha = 0.01$.

Two clarifications should be made. First the achievable rates of FDD systems are even larger than TDD systems under some parameters shown in Fig. 1. The phenomenon is explained by the fact that the uplink SNR is set 10 dB lower than the downlink SNR in the corresponding simulation results, which is typical for a cellular system due to the smaller transmit power of user-terminals, rendering the TDD system performance inferior due to the imperfect *uplink* channel training. Observe that when M becomes larger, the TDD system sum rate will go up unbounded, eventually surpassing the FDD system. Moreover, when the uplink SNR is set to be the same as the downlink SNR, see corresponding curves, the TDD system performs better, which is as expected. Secondly, the performance with correlated channels is slightly worse than the i.i.d. channels when the number of BS antennas is small, due to the fact that the channel capacity with i.i.d. channels is larger than the one with correlated channels, regardless of the estimation overhead, and when the number of BS antennas is small, the estimation overhead is negligible compared with the channel block length.

B. Comparison with JSDM

In Fig. 2, we compare the achievable sum rates obtained by the proposed eigenspace channel estimation to the JSDM scheme [3], [4], which was the first to exploit the spatial correlation to benefit the FDD massive MIMO system. Note that the uplink CSIT feedback is not treated in the previous JSDM papers [3]. To make a fair comparison, we assume that the JSDM scheme uses an isotropical random VQ feedback codebook, since it is unknown whether the JSDM scheme can also benefit from a better-designed codebook for correlated channels after the pre-projection of channel vectors. To get more insights and understand the simulation results better, it is important to first illustrate the merits and demerits of the JSDM scheme compared to our scheme.

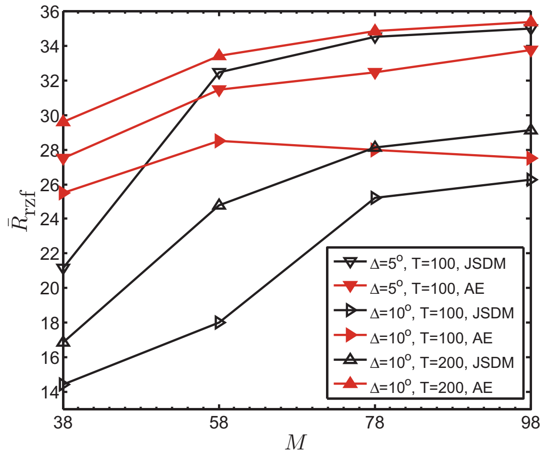

Fig. 2. The achievable sum rates (AE) obtained by the eigenspace channel estimation, compared to the JSDM scheme. The downlink and uplink SNR are both 10 dB. The number of simultaneous users is 8.

The JSDM scheme has the advantage to better suppress the channel estimation overhead. Specifically, by grouping the users based on their respective CCMs and performing the pre-beamforming, the equivalent number of BS antennas in each *virtual sector*, i.e., b_g in [3], can be optimized to strike a good balance between the power gain, which scales with b_g, and the channel estimation overhead. In an extreme case, b_g can be made as small as the number of users in each virtual sector, thus, the overall channel estimation overhead scales with the number of users in each virtual sector, which drastically decreases the dimensionality loss. However, on the downside, while the JSDM scheme adopts a *divide-and-multiplex* approach, the division is imperfect, in the sense that the JSDM scheme suffers from the inherent residual *inter-group interference (IGI)*, especially when the CCMs of the users in each group are different, rendering that the pre-beamforming cannot counteract the IGI completely. Notice that in our framework, the proposed dominant channel estimations incorporate all the user CCMs into the scheme design, which significantly mitigates the IGI. Moreover, it is noteworthy that the *computational complexity* of the JSDM scheme is smaller compared with our proposed scheme, since our scheme deals with a higher dimensional channel matrix. [1]

Specifically, we follow the parameters used in the simulation in [4, Section IV-C]. It is observed from Fig. 2 that the JSDM scheme achieves better sum rate when the channel coherence time is small, e.g., $T = 100$, and the number of BS antennas M is large. Qualitatively, this is expected since the small channel coherence time and large M both put more weight in the need to suppress the channel estimation overhead, and based on [3], a large M also leads to the fact that the eigenvectors of the channel correlation matrices can be well approximated by the columns of a Discrete Fourier Transform (DFT) matrix, which ensures orthogonality as long as the angular of arrival (AoA) intervals of different users are disjoint. On the other hand, the achievable sum rate of our proposed channel estimation shows evidently better rate when the channel coherence time is larger, which alleviates the urgency to suppress the channel estimation overhead, or when the angular spread of users is larger, which causes larger residual IGI in the JSDM scheme. Notice that large angular spread also decreases the achievable rates of our scheme, due to the increased channel estimation dimensionality, however our scheme turns out to be more resilient in this regard.

[1]Possible operations on the channel matrix include inversion and SVD, depending on the precoding algorithm.

V. Conclusions

In this paper, we have shown that the low-rank CCMs can be exploited to greatly benefit the massive MIMO system with closed-loop channel estimation, by directly modifying the channel training sequences and the feedback codebook without channel pre-projection. We find the optimal channel training sequences and a KLSQ feedback codebook. By comparing the achievable rates through simulations, it is found that the throughput of the FDD massive MIMO is comparable with TDD under currently typical number of BS antennas in massive MIMO systems, which is up to about 100, and that in many practically relevant situations our scheme achieves higher throughput than JSDM, though at the cost of increased computational complexity.

References

[1] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans Commun.*, vol. 9, pp. 3590–3600, Nov 2010.

[2] B. Clerckx, G. Kim, and S. Kim, "Correlated fading in broadcast mimo channels: curse or blessing?," in *IEEE GLOBECOM*, pp. 1–5, 2008.

[3] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing: The large-scale array regime," *IEEE Trans. Inform. Theory*, vol. 59, pp. 6441–6463, Oct 2013.

[4] J. Nam, A. Adhikary, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing: Opportunistic beamforming, user grouping and simplified downlink scheduling," *Selected Topics in Signal Processing, IEEE Journal of*, vol. PP, pp. 1–1, Mar 2014.

[5] Z. Jiang, A. F. Molisch, G. Caire, and Z. Niu, "Achievable rates of fdd massive mimo systems with spatial channel correlation," *submitted to IEEE Trans. Wireless Comm., available online at arXiv:1406.7486.*

[6] A. F. Molisch, *Wireless communications*. 2nd edition, IEEE Press - Wiley, 2011.

[7] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser mimo achievable rates with downlink training and channel state feedback," *IEEE Trans. Inform. Theory*, vol. 56, pp. 2845–2866, Jun 2010.

[8] J. H. Kotecha and A. Sayeed, "Transmit signal design for optimal estimation of correlated mimo channels," *IEEE Trans. Signal Processing*, vol. 52, pp. 546–557, Feb 2004.

[9] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[10] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, "Large system analysis of linear precoding in correlated miso broadcast channels under limited feedback," *IEEE Trans. Inform. Theory*, vol. 58, pp. 4509–4537, Jul 2012.