Performance Analysis of Massive MIMO Networks with Random Unitary Pilot Matrices

Rusdha Muharar
Department of Electrical and Computer Engineering
Universitas Syiah Kuala, Indonesia

Jamie Evans
Department of Electrical and Electronic Engineering,
The University of Melbourne, Australia

Abstract—A common approach to obtain channel state information for massive MIMO networks is to use the same orthogonal training sequences in each cell. We call this the full-pilot reuse (FPR) scheme. In this paper, we study an alternative approach where each cell uses different sets of orthogonal pilot (DOP) sequences. Considering uplink communications with matched filter (MF) receivers, we first derive the SINR in the large system regime where the number of antennas at the base station, the number of users in each cell, and training duration grow large with fixed ratios. For tractability in the analysis, the orthogonal pilots are drawn from Haar distributed random unitary matrices. The resulting expression is simple and easy to compute. As shown by the numerical simulations, the asymptotic SINR approximates the finite-size systems accurately. Secondly, we derive the user capacity of the DOP scheme under a simple power control and show that it is generally better than that of the FPR scheme.

I. INTRODUCTION

One of the candidates for 5G technology is the massive Multiple-Input and Multiple-Output (MIMO) system (see e.g., [1], [2]) introduced by Marzetta in [3]. In massive MIMO cellular networks, a large number of small and low-cost antennas, in the order of hundreds, is employed at base stations (BSs). This enables an aggressive spatial multiplexing which can lead to a ten times capacity increase compared to conventional MIMO systems [2], [3].

A common approach in the uplink training of massive MIMO systems is the full-pilot reuse (FPR) where the same orthogonal training sequences are used in each cell, see e.g., [3], [4], [5]. The other approach proposed in [3] is to use different orthogonal pilots in different cells and we denote this approach as the DOP scheme. To the best of our knowledge, this scheme is largely unexplored.

It is argued in [3] that this scheme gives little difference in terms of the achieved SINRs compared to the FPR scheme. However, in the DOP, we get a small amount of contamination from all inter-cell users rather than a (potentially) large amount from a few users (those with the same pilot sequence) and this can lead to a better user capacity (see [6]). Note that the analysis in [3], [6] is performed in the regime where the number of antennas \( N \) tends to infinity and the number users \( K \) is finite \( K \ll N \). This implies that the cell-loading \( K/N \) is close to zero. It should be noted that the analysis in this regime gives a loose approximation for finite-size systems and can also converge slowly [4], [7].

In this paper, we generalize the performance analysis of the DOP scheme for arbitrary numbers of cell-loading. We obtain the approximation of the SINR by performing the analysis in the large system regime where the number of antennas at the base station, the number of users in each cell, and the number of training symbols go to infinity with fixed ratios. For analytical tractability, we choose the training pilots from Haar-distributed random unitary matrices. This approach has been used previously in CDMA systems, see for example [8] and [9]. Our numerical simulations show that the asymptotic results approximate the finite-size systems accurately. In the analysis, we show that the pilot contamination in the asymptotic SINR expression is the average of the square of received powers of all users from the interfering cells (see also [3] for a similar conclusion). This result differs from that obtained in the FPR case [3], [4], [5], where the pilot contamination is the sum of the received power of users from the interfering cells that use the same training sequence.

In this paper, we also consider another performance criterion, i.e., the user capacity. In the downlink with maximum ratio transmission (MRT) precoders, the user capacities of massive MIMO networks for single and multi-cell scenarios have been characterized in [10], [11], respectively. Recent work [6] studies the uplink user capacity when the cell-loading approaches zero. Here, we derive the uplink user capacity for arbitrary numbers of cell-loading under a simple power control where the uplink transmit power of a user is the inverse of the slow path-gain of that user (see also [12]). Our numerical simulations show that even though all users from the interfering cells contribute to the pilot contamination in the DOP scheme, its user capacity can be larger compared to that of FPR. Other related work is [12] that investigates the optimal number of users that maximizes the spectral efficiency of massive MIMO networks.

The following notations are used in this paper. The boldface lower and upper case letters denote vectors and matrices, respectively. \( I_N \) denotes an \( N \times N \) identity matrix, \( E[\cdot] \) and \( a.s. \) denote respectively the statistical expectation and the almost sure convergence. The circularly symmetric complex Gaussian (CSCG) vector with zero mean and covariance matrix \( \Sigma \) is denoted by \( \mathcal{CN}(0, \Sigma) \). \(|a|\) and \( \Re[a] \) denote the magnitude and the real part of the complex variable \( a \), respectively. \( \| \cdot \| \) represents the Euclidean norm. \( \text{Tr}(\cdot) \), \( (\cdot)^T \) and \( (\cdot)^H \) refer to the trace, transpose and Hermitian transpose, of a matrix respectively.
II. SYSTEM MODEL

We consider a multi-cell communication system with $L$ cells. Each cell has a base station equipped with $N$ antennas and the number of users in cell $i$ is denoted by $K_i$. The channel between user $k$ in cell $j$ and the BS in cell $i$ is denoted by the column vector $g_{kji}$ that can be modeled as

$$ g_{kji} = \sqrt{\ell_{kji} h_{kji}} $$

where $h_{kji} \sim \mathcal{CN}(0, I_N)$ represents the fast-fading channel coefficients. It is assumed that the slow-fading coefficient $\ell_{kji}$ is distance-dependent and the shadowing effect is ignored. The channel variations follow the block channel fading model and one block duration is equal to the channel coherence time.

We also consider the time-division duplex (TDD) protocol with perfect channel reciprocity between the uplink and downlink channels. In this paper, we focus on the uplink transmission where all scheduled users transmit simultaneously to their base station (BS). In the pilot-based TDD training, the BS estimates each user channel from the pilot symbols sent by each user in the uplink transmission. The BS uses the channel information to decode the transmitted symbols from its users. We should note that the uplink training and the uplink data transmissions should occur in one coherence block time.

A. Uplink Training

Let $T$ be the length of the channel coherence time (in symbols). Let $\tau \leq T$ be the uplink training interval or the number of training symbols. In this training phase, each user in each cell sends the pilot symbols to their BS. We assume a synchronized training, where all cells perform the training at the same time and with the same training period $\tau$. Let $h_{kji}^{[n]}$ be the $n$-th element of $h_{kji}$ that represents the channel from the corresponding user to the $n$-th antenna of BS $i$. Also, let $q_{ki}^{(t)}$ be the pilot symbol sent by user $k$ in cell $i$ at time $t$ and $\check{q}_{ki}$ be the corresponding average training power. Note that, in the uplink training, BS $i$ will estimate the channels of its users $h_{kji}$. Since our assumption that the elements of $h_{kji}$ are independent, therefore we can estimate each element of the channel vector $h_{kji}$ independently. The received signal vector at the $n$-th antenna of BS for $\tau$ training symbols is

$$ y_i^{[n]} = \sum_{j=1}^{L} \sum_{k=1}^{K_j} \sqrt{\rho_{kji}} h_{kji}^{[n]} q_{ki} + n_i^{[n]} $$

where $y_i^{[n]} = [y_{i1}^{[n]} \ldots , y_{\tau i}^{[n]}]^T$, $q_{kj} = [q_{kj}^{(1)} \ldots , q_{kj}^{(\tau)}]^T$, and $n_i^{[n]} = [n_{i1}^{[n]} \ldots , n_{\tau i}^{[n]}]^T \sim \mathcal{CN}(0, \sigma^2 I_{\tau})$ is the receiver noise vector. We also denote $\check{q}_{kji} = \sqrt{\rho_{kji}} g_{kji}$ as the received power of user $k$ from cell $j$ at BS $i$. From (6), we can see that BS $i$ also receives the training transmissions from other cells. It is assumed that the BS knows all the path-gains and training sequences of all users perfectly. We also assume that the training matrix for cell $j$, $Q_j = [q_{j1}, q_{j2}, \ldots , q_{JK_j}]$ is obtained by extracting $K_j \leq \tau$ columns of a $\tau \times \tau$ Haar-distributed random unitary matrix $U_j$ (see also [3], [9]). Thus, $Q_j^TQ_j = I_K$ which implies that the training sequences are orthogonal (orthonormal) across the users in the same cell. It is assumed that $\mathbf{U}_j, j = 1, \ldots , L$ are independent. In other words, different cells employs different sets (independent) of orthogonal training sequences. It is in contrast to the majority of works in massive MIMO where the same orthogonal training sequences are used in each cell.

Let us focus on obtaining the estimate for $h_{kji}^{[n]}$. By correlating the observation vector $y_i^{[n]}$ with the training vector $q_{ki}$, we have

$$ \hat{h}_{kji}^{[n]} = \frac{q_{ki}^H y_i^{[n]}}{\sqrt{\rho_{kji}}}. $$

where $\hat{h}_{kji}^{[n]} = q_{ki}^H y_i^{[n]} \sim \mathcal{CN}(0, \sigma^2)$. The minimum mean-square estimation (MMSE) is employed based on the observation (3). Note that since BS $i$ knows $\rho_{kji}$, and $q_{ki}, \forall k,j$, the vector $y_i^{[n]}$ is Gaussian. Moreover, the scalars $q_{ki}^H y_i^{[n]}$ and $\hat{h}_{kji}^{[n]}$ are jointly Gaussian. The MMSE estimate for $h_{kji}^{[n]}$ is given by

$$ \hat{h}_{kji}^{[n]} = \frac{v_{ki}}{\sqrt{\rho_{kji}}} q_{ki} y_i^{[n]} $$

where

$$ v_{ki} = \frac{\rho_{kji}}{\rho_{kji} + \sum_{j=1}^{L} \sum_{m=1}^{K_j} \rho_{mji} q_{mi}^H q_{mj}^H + \sigma^2} $$

Note that $\frac{v_{ki}}{\sqrt{\rho_{kji}}}$ is the scalar estimator for $\hat{h}_{kji}^{[n]}$ based on the observation (3) and $v_{ki}$ is the variance of the channel estimate. It can be checked that the channel estimates for different users (in the same cell) are correlated. In contrast, they are uncorrelated in the FPR scheme. From (4), we can model $\hat{h}_{kji}^{[n]} \sim \mathcal{CN}(0, v_{ki})$. Note that we have removed the index $n$ in the notation for the channel estimation variance because it does not depend on $n$. The channel estimate vector, $\hat{h}_{kji} = [\hat{h}_{kji}^{[1]} \hat{h}_{kji}^{[2]} \ldots \hat{h}_{kji}^{[N]}]$ can be expressed as follows

$$ \hat{h}_{kji} = v_{ki} \left( h_{kji} + \sum_{j=1}^{L} \sum_{m=1}^{K_j} \frac{\rho_{mji}}{\rho_{kji}} h_{mji} q_{mi}^H q_{mj} + \frac{n_{ki}}{\sqrt{\rho_{kji}}} \right) $$

where $n_{ki} = [n_{i1}^{[1]} \ldots , n_{iN}^{[N]}]^T q_{ki} \sim \mathcal{CN}(0, \sigma^2 I_N)$. Note that $\hat{h}_{kji} \sim \mathcal{CN}(0, v_{ki} I_N)$. By the property of MMSE, we can model $h_{kji} = \hat{h}_{kji} + \tilde{h}_{kji}$, where $\tilde{h}_{kji} \sim \mathcal{CN}(0, (1 - v_{ki}) I_N)$ is the channel estimation error vector. Moreover, $\tilde{h}_{kji}$ and $\hat{h}_{kji}$ are independent. Note that the second term of (6) is the pilot contamination term which includes the channels of all users from the interfering cells. In the FPR scheme, only users that have the same training sequence contribute to this term.
B. Uplink Data Transmission

The uplink data transmission will take the duration of \( T - \tau \) symbols. Let \( s_{kj} \) be the transmitted symbol from user \( k \) in cell \( j \) with the average power \( p_{kj} \). The received signal at BS \( i \) is

\[
\mathbf{u}_i = \sum_{k=1}^{L} \sum_{j=1}^{K_j} \sqrt{p_{kj}} \mathbf{f}_{kj} \mathbf{h}_{kj} s_{kj} + \mathbf{n}_i
\]  

(7)

where \( \mathbf{n}_i \sim \mathcal{CN}(0, \sigma^2 I_{N_i}) \) is the receiver noise at BS \( i \). It is assumed that \( s_{kj} \sim \mathcal{CN}(0, 1) \) and is independent of other users’ data symbols. For the rest of the paper, we assume \( p_{kj} = p_{k} \) and thus, \( \rho_{kj} = p_{kj} f_{kj} \).

Let us consider user \( k \) at cell \( i \). A linear receiver, denoted by \( \mathbf{c}_{ki} \), is used by BS \( i \) to decode the received data symbols. The estimate for symbol \( s_{ki} \) is

\[
\hat{s}_{ki} = \mathbf{c}_{ki}^H \mathbf{u}_i = \mathbf{c}_{ki}^H \sum_{j=1}^{K_j} \sqrt{p_{kj}} \mathbf{f}_{kj} \mathbf{h}_{kj} s_{kj} + \mathbf{c}_{ki}^H \mathbf{n}_i.
\]

In this paper, we consider the MF receiver due to its low complexity. A more sophisticated LMMSE receiver is left for future works. The receiver is constructed by using the channel estimate, i.e., \( \mathbf{c}_{ki} = \hat{\mathbf{h}}_{ki} \). The resulting SINR, denoted by \( \gamma_{ki} \), can be written as follows

\[
\gamma_{ki} = \frac{\rho_{ki} \| \hat{\mathbf{h}}_{ki} \|^4}{\xi_{ki} + \sum_{(j,m) \neq (i,k)} \rho_{mji} \| \hat{\mathbf{h}}_{mji} \|^2 + \sigma^2 \| \hat{\mathbf{h}}_{ki} \|^2}
\]  

(8)

where \( \xi_{ki} = \rho_{ki} \| \hat{\mathbf{h}}_{ki} \|^2 \| \mathbf{n}_i \|^2 \) is the self interference noise.

III. Asymptotic Analysis

In this section, we derive the asymptotic uplink SINR \( \tilde{\gamma}_{ki} \) by analyzing it in the large system regime, i.e., when \( K_j, N \) and \( \tau \) tend to infinity with \( K_j / N \) and \( K_j / \tau \) being finite constants. The derivation relies on results in random matrix theory \([13, 14]\), as well as the results on Haar-distributed random matrices (see e.g., \([8, 9, 14]\)).

**Theorem 1**: Let \( \Gamma_{ji} = \text{diag}(\rho_{1j}, \rho_{2j}, \ldots, \rho_{K_jj}) \). Suppose that the empirical spectral distribution (e.s.d.) of \( \Gamma_{ji} \) converges to a non-random distribution \( F_{\Gamma_{ji}} \). As \( K_j, N, \tau \to \infty \) with \( K_j / N \to \alpha_j \) and \( K_j / \tau \to \kappa_j \), the uplink SINR \( \gamma_{ki} \) converges almost surely to \( \tilde{\gamma}_{ki} \) which is given by

\[
\tilde{\gamma}_{ki} = \frac{\rho_{ki}^2 \bar{\nu}_{ki}}{\rho_{ki} \sum_{j \neq i} \kappa_j \mathbb{E}[\Gamma_{ji}] + \bar{\nu}_{ki} \sum_{j \neq i} \kappa_j \mathbb{E}[\Gamma_{ji}^2]} \]  

(9)

where

\[
\bar{\nu}_{ki} = \frac{\rho_{ki}^2}{\rho_{ki} \sum_{j \neq i} \kappa_j \mathbb{E}[\Gamma_{ji}] + \sigma^2}
\]  

(10)

and \( \Gamma_{ji} \) is a random variable with distribution \( F_{\Gamma_{ji}} \).

**Proof**: See Appendix [A]

**Remark 1**: In practice, i.e. when the distribution of \( F_{\Gamma_{ji}} \) is difficult to obtain or \( K_j \) is relatively small, we can replace \( \mathbb{E}[\Gamma_{ji}] \) with its empirical value \( \frac{1}{L} \sum_{l=1}^{L} \rho_{kj} \).

As mentioned previously, the large system analysis is performed to obtain the approximation for the uplink SINR. In practice, the parameters \( K_j, N \) and \( \tau \) are finite and the approximation \((9)\) uses the ratios of those parameters (\( \alpha_j \) and \( \kappa_j \)). As shown in Section [11] \((9)\) can approximate the finite-size systems accurately. Moreover, we can infer from \((9)\) that the uncertainty due to the fast fading has been removed and only that from the large-scale fading (\( \rho_{kj} \)) remains.

Now, let us consider a simple power control scheme where \( p_{kj} = P_{u} \). Thus, the received power \( \rho_{kj} = P_{u} \) is the same for all users in cell \( j \) (see also \([12]\)). Consequently, \( \Gamma_{ji} \) becomes

\[
\Gamma_{ji} = \begin{cases} P_{u} \mathbf{I}_{K_j}, & j = i \\ P_{u} \mathbf{L}_{ji}, & j \neq i \end{cases}
\]

where \( \mathbf{L}_{ji} = \text{diag}(\ell_{ji}, \ell_{j2j}, \ldots, \ell_{K_jji}, \ell_{K_jj}) \). Suppose that the e.s.d. of \( \mathbf{L}_{ji} \) converges almost surely to a deterministic distribution \( F_{\mathbf{L}_{ji}} \). Then, it is straightforward from Theorem [11] to

- - -
show that the limiting SINR under the power control, denoted by $\gamma_{th}^p$, is

$$
\tilde{\gamma}_{ki}^p = \frac{\tilde{v}_i^p}{\sum_{j=1}^L \alpha_j E[L_{ji}] + \tilde{v}_i^p \sum_{j \neq i}^L \kappa_j E[L_{ji}^2]}
$$

(14)

where

$$
\tilde{v}_i^p = \frac{1}{1 + \sum_{j \neq i}^L \kappa_j E[L_{ji}] + \sigma^2/P_u}
$$

(15)

and $L_{ji}$ is a random variable whose distribution $F_{L_{ji}}$.

We can see from (14) and (15) that the limiting SINR and channel estimation variance under power control are the same for all users in cell $i$. Moreover, $\gamma_{th}^p$ is deterministic since it does not depend on the random locations (or particular realizations of the locations) of the users. Under the same power control, the limiting SINR for the FPR scheme is

$$
\tilde{\gamma}_{ki}^{\alpha p} = \frac{\tilde{v}_{ki}^{\alpha p}}{\sum_{j=1}^L \alpha_j E[L_{ji}] + \tilde{v}_{ki}^{\alpha p} \sum_{j \neq i}^L \frac{\bar{\gamma}_{ji}^p}{\bar{\gamma}_{ji}}}
$$

(16)

where $\tilde{v}_{ki}^{\alpha p} = \left(1 + \sum_{j \neq i}^L \tilde{\gamma}_{ji}^p + \sigma^2/P_u\right)^{-1}$. It is obvious that $\gamma_{th}^p$ still depends on the particular realizations of $L_{ji}$ and $L_{ji}^2$. Note that, for a particular realization of $L_{ji}$, $L_{ji}^2$, and $L_{ji}^3$, can be larger or smaller than $\sum_{j \neq i}^L \kappa_j E[L_{ji}]$ and $\sum_{j \neq i}^L \kappa_j E[L_{ji}^2]$, respectively. Consequently, $\gamma_{th}^p$ can be smaller or larger than $\gamma_{th}^p$.

IV. USER CAPACITY UNDER POWER CONTROL

Here we will characterize the user capacity based on the limiting SINR under power control (14). Recall that the limiting SINR, $\gamma_{th}^p$, is deterministic and is the same for all users in cell $i$. Now, let us assume that all users in cell $i$ have the quality of service (QoS) requirement,

$$
\tilde{\gamma}_{th}^p \geq \gamma_{th}^p
$$

(17)

where $\gamma_{th}^p$ is the minimum required SINR for users in cell $i$. Then, the user capacity is defined as the number of users per degree of freedom ($N$), or $\alpha_i$, that satisfies the QoS (17), see also (10), (11). From (17), we have

$$
\alpha_i + \sum_{j \neq i}^L \alpha_j E[L_{ji}] \leq \tilde{v}_i^p \left(\frac{1}{\gamma_{th}^p} - \sum_{j \neq i}^L \kappa_j E[L_{ji}^2]\right).
$$

(18)

To get more insights on the user capacity, let us consider the simplest case where $K_j = K$, so that $\alpha_j = \alpha$, $\kappa_j = \kappa$. Note that $K$ affects both $\alpha$ and $\kappa$ and we can write $\alpha = \frac{K}{\kappa} = \kappa \theta$ with $\theta = \frac{1}{K}$. For a given degree of freedom, there are two ways of defining $\alpha$ that satisfies the QoS in (17), either by fixing $\tau$ or $\kappa$.

A. Fixed $\tau$

Recall that $\alpha = \kappa \theta$. By substituting the expression for $\tilde{v}_i^p$ in (15) into (18), we obtain

$$
A_i \kappa^2 + B_i \kappa - \frac{1}{\gamma_{th}^p} \leq 0
$$

(19)

where

$$
A_i = \theta \left(\sum_{j=1}^L E[L_{ji}]\right) \left(\sum_{j \neq i}^L E[L_{ji}]\right)
$$

$$
B_i = \theta \left(1 + \sigma^2/P_u\right) \sum_{j=1}^L E[L_{ji}] + \sum_{j \neq i}^L E[L_{ji}^2].
$$

Now we need to find the range of $\kappa$ (or $\alpha$) such that (19) holds with $A_i, B_i, \gamma_{th}^p > 0$ and $0 \leq \kappa \leq 1$. The left hand side of (19) is a quadratic equation with one positive root and one negative root. The positive root (solution) is

$$
\kappa = \frac{-B_i + \sqrt{B_i^2 + \frac{4A_i}{\gamma_{th}^p}}}{2A_i}
$$

(20)

It is easy to check that the left hand side of (19) is a convex function and has a negative value at $\kappa = 0$. Thus, (19) will hold when

$$
0 \leq \kappa \leq \kappa_{\tau}.
$$

Equivalently, the cell loading that satisfies (17) is

$$
0 \leq \alpha \leq \alpha_{\tau}, \quad \text{with} \quad \alpha_{\tau} = \kappa_{\tau} \theta.
$$

(21)

Now, we will check the condition where $\kappa_{\tau} \leq 1$ holds. It is equivalent to $\sqrt{B_i^2 + \frac{4A_i}{\gamma_{th}^p}} \leq 2A_i + B_i$. By squaring both sides and performing some algebraic manipulations, $\kappa_{\tau} \leq 1$ holds when

$$
\frac{1}{\gamma_{th}^p} \leq A_i + B_i.
$$

Otherwise, we set $\kappa_{\tau} = 1$ (or $\alpha_{\tau} = \theta$).

Considering (20) and (21), it can be easily checked that $\alpha_{\tau}$ is an increasing function of $\tau$. Thus, a larger $\tau$ will increase the user capacity.

B. Fixed $\kappa$

In this case, $\kappa$ is fixed and thus $\tau$ varies as $K$ varies. The corresponding user capacity is simply given by

$$
\alpha \leq \alpha_{\kappa} = \tilde{v}_i^p \frac{\gamma_{th}^p}{\gamma_{th}^p - \kappa \sum_{j \neq i}^L E[L_{ji}^2]}.
$$

(22)

Since $\alpha_{\kappa} \geq 0$, the right-hand side of the above equality must satisfy

$$
\frac{1}{\gamma_{th}^p} \geq \kappa \sum_{j \neq i}^L E[L_{ji}^2].
$$

It is obvious that the choice of $\kappa$ affects the user capacity. It can be checked that $\tilde{v}_i^p \left(\frac{1}{\gamma_{th}^p} - \kappa \sum_{j \neq i}^L E[L_{ji}^2]\right)$ is increasing as $\kappa$ decreases. Note that by decreasing $\kappa$, we have a better
the FPR scheme.

Remark 3: As discussed in both subsections above, increasing \( \tau \) (or decreasing \( \kappa \)) will result in a higher user capacity. On the other side, this will reduce the uplink transmission time and consequently the cell sum rate. This trade-off is also a concern raised in [3]. One way to handle this trade-off is to find an optimal \( \tau \), for \( K \leq \tau \leq T \), that maximizes the cell sum rate. It is obvious that \( \tau = T \) is never an optimal solution since it gives a zero sum-rate. Thus, the optimal \( \tau \) is either at \( \tau = K \) or at \( \tau \in (K, T) \). Let us consider the worst case in terms of the user capacity, i.e., \( \tau = K \) (or \( \kappa = 1 \)). The user capacity in this case is easily obtained from (22) (or (21)) and is given by

\[
\alpha_K = \frac{1}{1 - \sum_{j \neq i} \mathbb{E}[L_{ji}]} \left( 1 + \sum_{j \neq i} \mathbb{E}[L_{ji}] + \bar{s}_x^2 \right).
\]

The numerical simulation for this case is shown in Figure 4 in the following section. We can see that the user capacity of the DOP scheme is considerably larger compared to that of the FPR scheme.

V. NUMERICAL SIMULATIONS

In this section, we will validate the approximation of our large system results for finite \( K_j, N \), and \( \tau \). We consider a 7-cell hexagonal layout where the inner cell radius is normalized to 0.9 and the distance between the base stations is normalized to 2. The users are uniformly distributed across the cell. We adopt the bounded path-loss model where the slow path-gain is modeled as \( \ell_{kji} = (1 + d_{kji}^{-\kappa})^{-1} \) where \( \kappa \) is the path-loss exponent and \( d_{kji} \) is the distance between user \( k \) in cell \( j \) and BS \( i \). We set \( \kappa = 3.7 \) and ignore the shadowing effect in the model. The experiment takes parameters \( P_a = 0 \) dB, \( K_j = K = 20 \), \( \kappa = \frac{2}{3} \) and the number of antennas varies from 50 to 500 with interval 50. We generated 500 channel realizations according to the Rayleigh distribution. Figure 1 shows the SINR (in decibels) as the number of antennas varies from 50 to 500. We see that there is some probability that smaller \( \kappa \) gives a better performance (SINR). The figure also reveals that there is some probability that \( \bar{\gamma}_i^p \) is larger than \( \bar{\gamma}_k^p \). This gap reduces as \( N \) increases and it is approximately 2.3 dB when \( N = 500 \).

Figure 2 compares the cdf of \( \bar{s}_{k_i}^p \) and \( \bar{s}_i^p \). Note that \( \bar{s}_i^p \) is a deterministic quantity. Hence, its cdf is a unit impulse located at \( \bar{s}_i^p \). To produce the curves, we set \( \alpha = 0.1 \) and \( \kappa = \{\frac{2}{3}, \frac{5}{3}, 1\} \). Note that, \( \bar{s}_i^p \) depends on \( \kappa \) while \( \bar{s}_{k_i}^p \) does not. It is obvious from the curves for \( \bar{s}_i^p \) that smaller \( \kappa \) gives a better performance (SINR). The figure also reveals that there is some probability that \( \bar{s}_{k_i}^p \) is larger than \( \bar{s}_{k_i}^p \). This gap reduces as \( \kappa \) increases.

The user capacity for the different orthogonal training sequences (DOP scheme) has been presented in Section [1]. For the FPR scheme, the formulation of user capacity is slightly different since \( \bar{s}_{k_i}^p \) is a random quantity contributed by the large-scale fading terms of the pilot contamination. Hence, the user capacity region, for the simplest case (\( K_j = K, \forall j \))...
is defined as the value of $\alpha$ that satisfies

\[
P(\gamma^{s,p}_{ki} \geq \gamma^{th}_i) \geq 1 - \beta
\]

where $\beta$ is the outage probability. The above is equivalent to $P(X_{ki} \leq 0) \geq 1 - \beta$ or $F_{X_{ki}}(0) \geq 1 - \beta$ where

\[X_{ki} = \alpha \sum_{j=1}^{L} \mathbb{E}[L_{ji}] \left(1 + \frac{\sigma^2}{P_u} + \sum_{j\neq i} L_{kji} + \sum_{j\neq i} \frac{L_{kji}^2}{\gamma^{th}_i} \right).
\]

Figure 3 presents the comparisons of the user capacities of the FPR obtained from (23) with $\alpha = 1 + 0.2$ and of the DOP represented by $\alpha$ in (21) with $\theta = \{0.5, 1, 1.5\}$. As expected, higher outages lead to higher user capacities for the FPR while higher $\theta$ (higher $\tau$ with $N$ fixed) gives higher user capacities for DOP. In general, we can see that the FPR has a higher user capacity compared to the DOP for very low $\gamma^{th}_i$ (less than $-6$ dB) and the reverse occurs for low to high $\gamma^{th}_i$. The user capacity of the FPR scheme is zero starting at $\gamma^{th}_i = -2$ dB for outage probability 5% and at $\gamma^{th}_i = 0$ dB. On the other hand, the user capacities for the DOP scheme are about $0.28 - 0.35$ and $0.2$ when $\gamma^{th}_i = -2$ dB and $\gamma^{th}_i = 0$ dB, respectively. This shows the advantage of the DOP scheme over the FPR scheme in terms of user capacity and is consistent with the findings in [6]. A similar conclusion can be drawn by comparing the user capacities of FPR and $\alpha$ of DOP with $\kappa = \{0.5, 1\}$ as shown in Figure 4. Even for the worst case, $\kappa = 1$, the user capacities of the DOP are higher than those of the FPR at all considered values of $\gamma^{th}_i$. For example, at $\gamma^{th}_i = -10$ dB, the user capacities in the FPR scheme and the DOP scheme are about 1.1 and 1.7, respectively. Hence, for example, if the BS has 500 antennas, the DOP scheme can admit 300 more users than the FPR scheme can.

VI. CONCLUSION

In this paper, we consider the performance analysis of uplink massive MIMO systems with different sets of orthogonal training pilots in different cells employed during the uplink training. We obtained a novel expression for the asymptotic SINR in the large system regime that approximates finite-size systems accurately. We show that under a simple power control, the asymptotic SINR is a deterministic quantity and its denominator contains the average interference and mean-squared interference (pilot contamination) of all users from the interfering cells. We also derived the user capacity of the system and compared to that of the FPR scheme. Our simulation results showed that the DOP generally gives a better user capacity even for low values of the required SINRs.

APPENDIX A

PROOF OF THEOREM 1

We begin the proof by establishing the limiting result for the channel estimate variance $\nu_{ki}$. Then, we proceed to obtain the limiting results for the numerator and the denominator of the uplink SINR $\gamma_{ki}$ which are denoted by $A_{ki}$ and $B_{ki}$ respectively. In the proof we use the following definitions and results from large random matrices.

Definition 1 ([13] Definition 2.6]): Two infinite sequences $a_n$ and $b_n$ are defined to be asymptotically equivalent as $n \to \infty$, denoted by $a_n \asymp b_n$, if only if $|a_n - b_n| \xrightarrow{a.s.} 0$.

Definition 2 ([13] Definition 2.7]): Two infinite sequences $a_{k,n}$, $b_{k,n}$, indexed by $k = 1, 2, \cdots, n$, are defined to be uniformly asymptotically equivalent as $n \to \infty$, denoted by $a_{k,n} \asymp_k b_{k,n}$, if only if $\max\{|a_{k,n} - b_{k,n}| \leq N \} \xrightarrow{a.s.} 0$.

Lemma 1 ([13] Lemma 2.5]): If $a_{k,n} \approx_k b_{k,n}$, then

\[
\frac{1}{n} \sum_{k=1}^{N} a_{k,n} \approx \frac{1}{n} \sum_{k=1}^{N} b_{k,n}.
\]

Lemma 2 ([16] Lemma 1]): Let $A \in \mathbb{C}^{N \times N}$ be a deterministic matrix with uniformly bounded spectral radius for
all $N$. Let $x = \frac{1}{\sqrt{N}}[x_1, x_2, \ldots, x_N]^\top$ where the $x_i$’s are i.i.d. with zero mean, unit variance and finite eighth moment. Let $y$ be a similar vector independent of $x$. Then, we have $\mathbf{Ax}^\mathbf{H} - \frac{1}{\sqrt{N}}\text{Tr}(\mathbf{A}) \xrightarrow{a.s.} 0$, and $\mathbf{Ay}^\mathbf{H} \xrightarrow{a.s.} 0$.

The following lemma which involves the columns of an isometric matrix is analogues to Lemma 2.

**Lemma 3**: Let $w \in \mathbb{C}^N$ be a column of an isometric matrix $\mathbf{W}$ and $z$ be a column of an isometric $\mathbf{Z}$ matrix independent of $\mathbf{W}$. Let $\mathbf{A}$ be as defined in Lemma 2 and independent of $\mathbf{W}$ and $\mathbf{Z}$. Then, $\mathbf{W}^\mathbf{H}\mathbf{A}w = \frac{1}{\sqrt{N}}\text{Tr}(\mathbf{A}) \xrightarrow{a.s.} 0$ and $\mathbf{W}^\mathbf{H}\mathbf{Az} \xrightarrow{a.s.} 0$.

In the following sections, we use the notation $\mathbf{X}_k$ for matrix $\mathbf{X}$ with $k$-th column removed. If $\mathbf{X}$ is a diagonal matrix, $\mathbf{X}_k$ denotes $\mathbf{X}$ with both $k$-th column and $k$-th row deleted.

### A. The large system limit for $v_{ki}$

Here, we only need to find the asymptotic limit for the second term in the denominator of $v_{ki}$ in (5). That term can be written as $\sum_{j \neq i}^N \mathbb{E}Q_{ji}^H \Gamma_{ji} Q_{ji}^H q_{ki}$. Since $q_{ki}$ is independent of $Q_{ji}$ for $j \neq i$, $\mathbb{E}Q_{ji}^H \Gamma_{ji} Q_{ji}^H q_{ki} \xrightarrow{a.s.} \frac{1}{\tau} \text{Tr}(\mathbf{Q}_{ji}^\mathbf{H}\mathbf{Q}_{ji})$ by Lemma 3. Employing the right hand side is equal to $\frac{1}{\tau} \text{Tr} (\Gamma_{ji})$. Suppose that the e.s.d. of $\Gamma_{ji}$ converges almost surely to the non-random limiting spectral distribution (l.s.d.) $F_{\tau,j}$, as $K \tau \to \infty$ with $\frac{1}{\sqrt{K}} \xrightarrow{a.s.} \kappa_j, \frac{1}{\tau} \text{Tr} (\Gamma_{ji}) = \frac{\kappa_j}{\tau} \text{Tr} (\Gamma_{ji})$ converges almost surely to $\kappa_j \mathbb{E} [\Gamma_{ji}]$ where $\Gamma_{ji}$ is a random variable with distribution $F_{\tau,j}$. Consequently, $\mathbb{E}Q_{ji}^H \Gamma_{ji} Q_{ji}^H q_{ki}$, where $\mathbb{E}Q_{ji}^H \Gamma_{ji} Q_{ji}^H q_{ki}$ is defined in (10).

### B. The large system limit for $A_{ki}$

From the channel model described in subsection II-A, $\hat{h}_{kii} \sim \mathcal{CN} (0, v_{ki} I_N)$. By using Lemma 2, $\frac{1}{\mathbb{E}} \mathbb{E} \hat{h}_{kii}^2 \approx v_{ki}$. Since $v_{ki} \xrightarrow{a.s.} \overline{v}_{ki}$, we have $\frac{1}{N} A_{ki} \xrightarrow{a.s.} \rho_{ki} \overline{v}_{ki}^2$.

### C. The large system limit for $B_{ki}$

First let us consider the self interference noise ($\xi_{ki}$). Note that while $\frac{1}{\sqrt{N}} \mathbb{E}h_{kii}^H \Gamma_{ii} \mathbb{E}h_{kii}^H \xi_{ki} \to 0$ by Lemma 2. Consequently, $\frac{1}{\mathbb{E}} \mathbb{E}h_{kii}^H \Gamma_{ii} \mathbb{E}h_{kii}^H \xi_{ki} \approx 0$. For the noise term, while $\frac{1}{\mathbb{E}} \mathbb{E}h_{kii}^H \xi_{ki}^2 \xrightarrow{a.s.} \overline{v}_{ki}$, we have $\frac{1}{\mathbb{E}} \mathbb{E}h_{kii}^H \xi_{ki}^2 \approx 0$. Thus, the noise term converges to zero almost surely.

For the intra-cell interference term, it can be written as $\frac{1}{\mathbb{E}} \mathbb{E}h_{kii}^H \Gamma_{ii} \mathbb{E}h_{kii}^H \xi_{ki}$. Since $\mathbb{E}h_{kii}$ is independent of $\mathbb{E}h_{kii}^H \Gamma_{ii} \mathbb{E}h_{kii}$ and by applying Lemma 2 and the rank-one perturbation lemma [14], the intra-cell interference is asymptotically equivalent to $\frac{1}{\sqrt{N}} \text{Tr} \left( \frac{1}{\mathbb{E}} \mathbb{E}h_{kii} \Gamma_{ii} \mathbb{E}h_{kii}^H \right)$.

Note that $\frac{1}{\sqrt{N}} \text{Tr} \left( \frac{1}{\mathbb{E}} \mathbb{E}h_{kii} \Gamma_{ii} \mathbb{E}h_{kii}^H \right)$ is the first moment of the e.s.d. of $\frac{1}{\sqrt{N}} \mathbb{E}h_{kii} \Gamma_{ii} \mathbb{E}h_{kii}^H$. It can be shown that (see also [13, eq. 2.118]), this moment converges almost surely to $\kappa_i \mathbb{E} [\Gamma_{ii}]$. Hence, the intra-cell interference power converges to $\kappa_i \mathbb{E} [\Gamma_{ii}]$ almost surely.

Now, let us consider the inter-cell interference term. The (normalized) interference from cell $j$ can be written as $I_{ki,j} = \frac{1}{\mathbb{E}} \mathbb{E}h_{kii} \Gamma_{ji}^{\frac{1}{2}} \mathbb{Q}_{ji}^H q_{kj} + z_{kj}$. Let us rewrite (6): $\hat{h}_{kii} = \frac{v_{ki}}{\sqrt{\rho_{ki}}} \mathbb{E}h_{kii} \Gamma_{ji}^{\frac{1}{2}} \mathbb{Q}_{ji}^H q_{kj} + z_{kj}$.

1 An $N \times K$ isometric matrix with $K \leq N$ is obtained by taking $K$ columns of an $N \times N$ Haar-distributed random matrix [15, Definition 2.14].
It can be checked that all the summands of the last equation are asymptotically equivalent to zero. Therefore, by using Lemma 1, $T_{kij,3} \xrightarrow{a.s.} 0$.

Summing up over all interfering cells, the limiting result for the total inter-cell interference $\sum_{j \neq i} T_{kij}$ is

$$\frac{\bar{v}_{ki}^2}{\rho_{ki}} \sum_{j \neq i} \kappa_j E \left[ \gamma_{ji}^2 \right] + \frac{v_{ki}^2}{\rho_{ki}} \left( \rho_{ki} + \sum_{l \neq i} \kappa_l E \left[ \Gamma_{li} \right] + \sigma^2 \right) \sum_{j \neq i} \alpha_j E \left[ \Gamma_{ji} \right].$$

Using the definition for $\bar{v}_{ki}$ in (10), the second term can be simplified to $v_{ki} \sum_{j \neq i} \alpha_j E \left[ \Gamma_{ji} \right]$. Thus,

$$\sum_{j \neq i} T_{kij} \xrightarrow{a.s.} v_{ki} \left( \sum_{j \neq i} \alpha_j E \left[ \Gamma_{ji} \right] + \frac{v_{ki}}{\rho_{ki}} \sum_{j \neq i} \kappa_j E \left[ \Gamma_{ji}^2 \right] \right).$$

Bringing up the limiting results for each term of $B_k$, we obtain

$$\frac{1}{N^2} B_{ki} \xrightarrow{a.s.} v_{ki} \left( \sum_{j=1}^{L} \alpha_j E \left[ \Gamma_{ji} \right] + \frac{v_{ki}}{\rho_{ki}} \sum_{j \neq i} \kappa_j E \left[ \Gamma_{ji}^2 \right] \right).$$

By putting the large system results for $A_{ki}$ and $B_{ki}$, (9) follows immediately and this completes the proof.

**References**

[1] F. Boccardi, R. Heath, A. Lozano, T. Marzetta, and P. Popovski, “Five disruptive technology directions for 5G,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, February 2014.

[2] F. Rusek, D. Persson, B. K. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, “Scaling up MIMO: Opportunities and challenges with very large arrays,” *IEEE Commun. Mag.*, vol. 52, no. 11, pp. 69–77, November 2014.

[3] T. Marzetta, “Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas,” *IEEE Transactions on Wireless Communications*, vol. 1, no. 1, 2004.

[4] J. Hoydis, S. ten Brink, and M. Debbah, “Massive MIMO in the UL/DL of Cellular Networks: How Many Antennas Do We Need?” *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, 2013.

[5] N. Krishnan, R. Yates, and N. Mandayam, “Uplink linear receivers for multi-cell multiuser MIMO with pilot contamination: Large system analysis,” *IEEE Transactions on Wireless Communications*, vol. 13, no. 8, 2014.

[6] A. Sivamani and J. Evans, “On uplink user capacity for massive MIMO-cellular networks,” in *Proc. IEEE Glocom*, 2016, to appear.

[7] P. J. Smith, C. Neil, M. Shafi, and P. A. Dmochowski, “On the convergence of massive MIMO systems,” in *Proc. IEEE International Conference on Communications (ICC)*, June 2014, pp. 5191–5196.

[8] M. Debbah, W. Hachem, P. Loubaton, and M. de Courville, “MMSE analysis of certain large isometric random precoded systems,” *IEEE Transactions on Information Theory*, vol. 49, no. 5, 2003.

[9] M. Peacock, I. Collings, and M. Honig, “Asymptotic spectral efficiency of multiuser multisignature CDMA in frequency-selective channels,” *IEEE Transactions on Information Theory*, vol. 52, no. 3, 2006.

[10] J.-C. Shen, J. Zhang, and K. Letaief, “Downlink User Capacity of Massive MIMO Under Pilot Contamination,” *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3183–3193, 2015.

[11] N. Akbar, N. Yang, P. Sadeghi, and R. Kennedy, “Multi-cell multiuser massive MIMO networks: User capacity analysis and pilot design,” *IEEE Transactions on Communications*, vol. PP, no. 99, p. 1, 2016.

[12] E. Björnson, E. G. Larsson, and M. Debbah, “Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?” *IEEE Transactions on Wireless Communications*, vol. 15, no. 2, Feb. 2016.