Pilot-Decontamination in Massive MIMO Systems via Network Pilot–Data Alignment

Majid Nasiri Khormuji
Huawei Technologies Sweden AB, Stockholm
Email: majid.n.k@ieee.org

***Abstract*—We propose a pilot decontamination method for noncooperative cellular massive MIMO systems using pilot and data alignment in the time–frequency plane. In this method, pilot and data channels are assigned such that the terminals in any two adjacent cells operate over non-overlapping pilot channels but the terminals may share data channels. The former condition is imposed to circumvent pilot contamination in order to boost the performance of cell-edge terminals and the latter condition is put to enhance the cell-average capacity. The pilot–data channel assignment is blindly done with respect to the channel state information of the terminals. Simulation results, using Marzetta's cellular setup, show significant network capacity gains over the conventional methods at virtually no additional cost.**

I. Introduction

The concept of Massive Multiple-Input Multiple-Output (mMIMO) systems is crystallized in [1] wherein a Time-Division Duplex (TDD) protocol is used for channel acquisition and data transmission. With TDD, uplink pilot signals are used to learn both the uplink and downlink channel using the reciprocity of radio channels. Having learned the channels (a.k.a. spatial signatures), the mMIMO receiver can perform Space-Division Multiple-Access (SDMA) [1], [2]. The channel estimation is of crucial importance as it enables separating the superimposed data streams associated to different users. When the number of antennas in the array increases, the corresponding channel vectors of the terminals asymptotically become *mutually orthogonal* for example for i.i.d. channels as discussed in [1]. This hence ensures concurrent reception or transmission of multiple users' data with a negligible inter-user interference where the interference asymptotically vanishes as long as the known channels are *not* contaminated. However, since there are *limited* numbers of orthogonal sequences that can be accommodated in the coherence windows of radio channels (i.e. a portion of the time–frequency plane over which the radio channels remain almost unchanged), the pilot sequences should be shared by multiple users. This causes pilot contamination which is the main bottleneck of the mMIMO systems. See also [3]–[5] for a more elaborate background and some recent results for mMIMO systems.

In this paper, we introduce a novel transmission subframe with an associated time–frequency mapping for multi-cell scenarios, which significantly reduces the pilot contamination. We assume that there is no cooperation among the cells. In the proposed subframe, some of the frequency tones (i.e. subcarriers) are exclusively used for pilot transmission and the other frequency tones in the subframe are reserved for data transmission. The proposed subframe is then used to construct a multi-cell blind pilot–data alignment without the need for channel state information (neither fast nor slow fading coefficients). To alleviate the pilot contamination, different subcarriers (i.e. non-overlapping frequency tones) are used for pilot transmission from terminals located in adjacent cells. However the subcarriers reserved for data transmission may be shared among the terminals in adjacent cells to increase the system capacity. In other words, our method relies on non-uninform frequency reuse for pilot and data where a higher reuse factor is used for pilot as the effect of the pilot contamination is the main bottleneck of the system. Simulation results using the Marzetta's cellular setup [1] indicate encouraging gains for both cell-edge and cell-average capacities compared to the conventional frequency reuse [1] and partial pilot sequence reuse [6], [7].

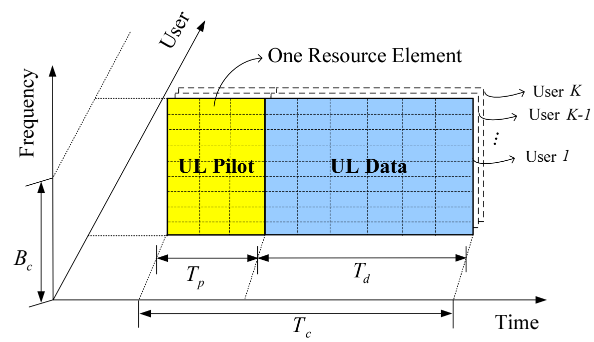


Fig. 1. Transmission frame adapted to the characteristics of radio channels.

The remaining part of the paper is organized as follows. Section II provides a background on the pilot contamination and discusses the throughput of two users with interfering pilot channels. Section III briefly reviews previously proposed methods for pilot decontamination in the literature. Section IV presents our blind pilot–data alignment for pilot decontamination in cellular mMIMO systems. Section V provides the asymptotic aggregate rate of the system for the our proposed method in Section IV. Section VI discusses numerical evaluations and compares the performance of the proposed scheme to that of the prior art. Section VII finally concludes the paper.

II. Background on Pilot Contamination

Fig. 1 illustrates a transmission subframe adapted to the characteristics of the radio channel, designed for mMIMO systems [1], [8]. To learn the radio channel in a cell with K single-antenna users, K orthogonal pilot sequences, each associated to a user, are required over a time–frequency grid of the size $T_c \times B_c$ where T_c denotes the coherence time in number of symbols and B_c denotes the coherence bandwidth of the channel in number of subcarriers. The number of users are given by $K = T_p B_c$ [symbols×tones] where T_p is the time consumed for uplink pilot transmission. The optimal number of scheduled users to maximize the *asymptotic* aggregate-rate in the conventional TDD is given by [1], [8]

$$K^* = \tfrac{1}{2} T_c B_c \tag{1}$$

That is the rate increases by scheduling one additional user as long as the number of users is less than $0.5 T_c B_c$. Scheduling a larger number of users, results to a larger overhead in order to ensure orthogonal pilot transmission for channel estimation and hence the transmission becomes *payload-limited* up to $K = T_c B_c - 1$ users.

To schedule the users in multi-cell scenarios, the pilot sequences for channel estimation should be reused because there are not enough orthogonal pilot sequences when the number of users is above $T_c B_c$. Therefore, the phenomenon known as *pilot contamination* causes a severe degradation in the performance. To illustrate the pilot contamination affect; consider two cell-edge users that transmit the same pilot symbol followed by data symbol sequences over shared uplink resources. Then the mMIMO access node receives the signal

$$\boldsymbol{y}_p = \boldsymbol{h}_1 x_p + \boldsymbol{h}_2 x_p + \boldsymbol{z}_p \tag{2}$$

where $\boldsymbol{y}_p$ denotes the received noisy signal vector, $\boldsymbol{x}_p$ denotes the transmitted pilot symbol from both users (i.e. pilot reuse), and $\boldsymbol{h}_i$ denotes the channel vector between user i and the antenna array at the access node, $\boldsymbol{z}_p$ denotes AWGN. All vectors have the dimension $n_t \times 1$. Then, using the MMSE channel estimation, the estimated channel will be

$$\hat{\boldsymbol{h}} = \boldsymbol{h}_1 + \boldsymbol{h}_2 + \boldsymbol{z}_e \tag{3}$$

where $\boldsymbol{z}_e$ denotes the channel estimation error. The received noisy superimposed data signal can be then written as

$$\boldsymbol{y}_d = \boldsymbol{h}_1 x_{d_1} + \boldsymbol{h}_2 x_{d_2} + \boldsymbol{z}_d \tag{4}$$

where $\boldsymbol{y}_d$ denotes the received noisy signal vector, x_{d_i} denotes the transmitted data symbol from user i and $\boldsymbol{z}_d$ denotes AWGN. The access node can perform spatial filtering to separate the data stream of the first user. Using the matched filtering (MF) (a.k.a. maximum-ratio combining (MRC)) and treating interference as noise, using the approach in [1] for i.i.d. channels with unit variance the rate

$$R = \frac{T_c - T_p}{T_c} \log\left(1 + \frac{P_1}{P_2}\right) \tag{5}$$

is achievable when $n_t \to \infty$ and the average transmit power of each user is set to P_i. From (5), it is seen that a large antenna array under pilot contamination helps to remove the noise and small-scale fading but the inter-user interference remains. Thus for $P_1 = P_2$ and $T_p = 0.5 T_c$, the rate $\frac{1}{2}$ [bits/s/Hz] is achievable for the two contaminated users. Therefore pilot reuse causes a notable rate-loss in spite of the fact that there exist many active antenna elements at the access node.

III. Pilot Decontamination: Prior Art

We next briefly review pilot decontamination methods proposed in the literature [1], [6]–[16].

A. Frequency Planning

Pilot contamination is a variant of interference, which can be thereby harnessed by the conventional frequency reuse type of solutions. This approach is considered in [1]. This solution, benefits the cell-edge users but it *reduces* the cell-average capacity. In [1], it is shown that with frequency reuse factor of three, the cell-edge experience is notably improved but the cell-average capacity is reduced by 30% as compared to that with frequency reuse factor of one.

B. Partial Pilot Sequence Reuse

An alternative method to harness the pilot contamination is to use partial (a.k.a. fractional) pilot sequence reuse [6], [7]. That is the terminals located in the adjacent cells are assigned mutually orthogonal pilot sequences. For example for the sequence reuse factor of three, the pilot contamination among any two adjacent cells can be avoided by dividing the available set of pilot sequences into three mutually orthogonal sets. This methods enhances the capacity per terminal. However, the reduction of cell-average capacity due to a lower number of scheduled users per cell is the main shortcoming of this scheme. This solution is therefore not suitable for cases with many users demanding a simultaneous access in the cells.

C. Channel Statistics

Radio channels in some scenarios can be modeled in a way that some long-term or second-order statistics can be used for reduction of the interference on the pilot signals. The work [9] proposed pilot precoding based on large-scale coefficients which reduces the pilot contamination. However, this method requires that modulated symbols as well as the large-scale coefficients to be known at all base stations, which puts a serious drawback on the solution. The work [10] utilizes the knowledge of temporal variation of the radio channels in order to alleviate the pilot contamination. This method is not applicable to narrow band channels. The work [11]–[13] also take advantages of properties of the second order statistics of the channels to harness the interference in channel estimations.

D. Coordinated Transmissions

The work [8] proposes a semi-orthogonal multiple-access to address the pilot contamination by introducing a semi-orthogonal feature in pilot transmission. The work [14], [15] proposes pilot transmission with time-shift across the cells. These methods require a coordination among the users/cells and additional signallings. The work [16] also investigates a coordinated approach to the pilot assignment. However, the results in [7] indicate that the coordination approach of [16] does not improve the performance of the mMIMO network as compared to that with a random pilot assignment.

IV. Proposed Scheme: Blind Pilot–Data Alignment

In this section, we present our proposed scheme. We first discuss a time–frequency transmission subframe that we use throughout and then explain a cellular blind pilot–data alignment scheme using the transmission subframe. The scheme is referred to as *blind* because it does not rely on any type of channel state information. We additionally assume that there is no cooperation among the base stations as that in [1].

A. Transmission Subframe

Fig. 2 depicts the proposed transmission subframe that we employ in the paper. In this subframe, some frequency tones are reserved for pilot signals and the remaining tones are used for data transmission. The coherence bandwidth B_c is divided into two subbands, where the pilot subband is B_p and the data subband is $B_d = B_c - B_p$. The maximum number of users with mutually orthogonal pilot sequences that can be scheduled is $K = B_p T_c$. In contrast to the subframe in Fig. 1, the pilots in the subframe in Fig. 2 are transmitted during the entire transmission interval but on fewer subcarriers.

B. Network Pilot–Data Alignment

Fig. 3 shows the network time–frequency planning with the proposed blind pilot–data alignment using the subframe in Fig. 2. The proposed pilot–data alignment is constructed based on two conditions:

- any two adjacent cells operate over *non-overlapping* frequency tones for pilot transmission; and
- two adjacent cells may operate over *overlapping* frequency tones for data transmission.

The first condition is imposed to avoid the pilot contamination over any two adjacent cells. This hence enhances the capacity of the cell-edge terminals. To ensure the first condition, it is required to assign at least three non-overlapping pilot subbands (i.e. the frequency reuse factor of three only for pilot tones). The second condition allows sharing the frequency tones used for data among the adjacent cells. This creates additional inter-cell interference, however this additional interference can be alleviated using spatial filtering using large-scale arrays since the corresponding channel estimates are not contaminated. Therefore, the second condition boosts the cell-average capacity beyond the conventional solutions.

V. Asymptotic Sum-Rate Characterization

We, in this section, investigate the mMIMO systems with an infinite number of antennas at the access nodes and assume that the channel is varying according to the block fading such that the channel over the coherence window of $T_c \times B_c$ is modeled as a complex number. For terminal k located in cell l and base station antenna m in cell j, the channel for all channel uses in the coherence window is represented as

$$g_{mjkl} = h_{mjkl} \cdot \beta_{jkl}^{1/2} \tag{6}$$

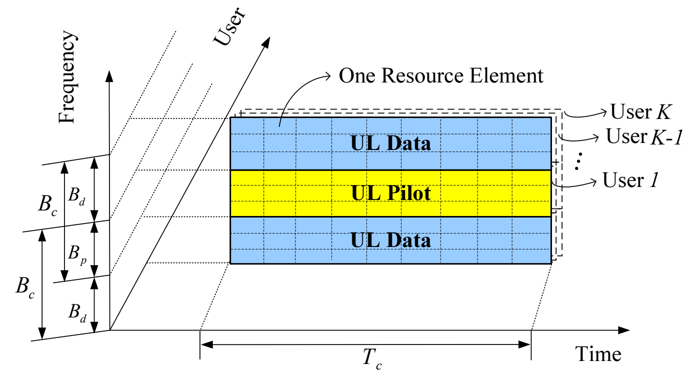


Fig. 2. Proposed transmission frame with disjoint tones for data and pilot.

where h_{mjkl} denotes fast fading modelled as i.i.d. Rayleigh fading with unit variance and β_{jkl} denote large-scale fading which is constant over the antenna array and is given by

$$\beta_{jkl} = \frac{z_{jkl}}{r_{jkl}^{\gamma}} \tag{7}$$

where z_{jkl} denotes the shadow fading which is a log-normal random variable with a standard deviation of σ, r_{jkl} denotes the distance between terminal k in cell l and the base station in cell j, and γ denotes the path-loss exponent.

Each base station using the uplink pilot sequences estimates the channels. The channel estimates are contaminated by the users that are using the same pilot sequence according to

$$\hat{\boldsymbol{g}}_{jkl} = \underbrace{\boldsymbol{g}_{jkl}}_{\text{Desired Channel}} + \underbrace{\sum_{i=1, i\neq l}^{L} \boldsymbol{g}_{jki}}_{\text{Pilot Contamination}} + \boldsymbol{z}_{e,jkl} \tag{8}$$

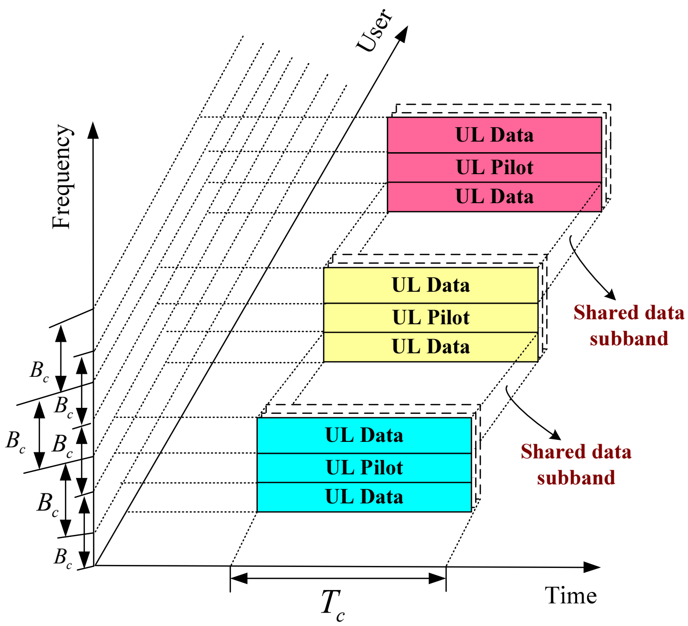
where $\boldsymbol{z}_{e,jkl}$ denotes the channel estimation error. The summation index in (8) runs over the cells with interfering pilot signals (i.e. the cells with the same color in Fig. 3b). Now using the same approach as that in [1], when the number of antennas in the array increases, we can obtain the following sum-rate for the K users in cell l

$$C_{ul} = B \cdot \frac{B_c - B_p}{1.5 B_c} \cdot \frac{T_u}{T_s} \sum_{k=1}^{K} \log\left(1 + \frac{\beta_{jkl}^2}{\sum_{i=1, i\neq l}^{L} \beta_{jki}^2}\right) \tag{9}$$

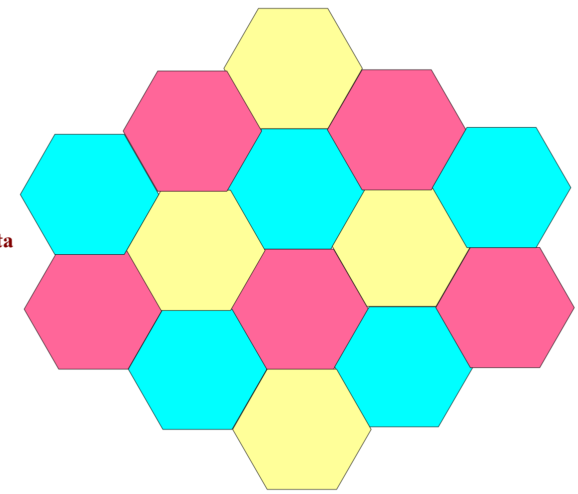
where the rate normalization is done by taking into account the pilot–data alignment in Fig. 3. Here, T_s and $T_u = T_s - T_g$ respectively denote symbol interval and useful symbol interval, where T_g is the length of the guard interval.

VI. Numerical Evaluations

We next consider Marzetta's multi-cell setup for uplink transmission. In our evaluation of the proposed scheme, we consider three groups of cells with pilot–data alignment according to Fig. 3. We choose the simulation parameters as summarized in Table I, which are similar to those in [1]. However, we use the subframe in Fig. 2 where the the channel coherence bandwidth $B_c = 14$ [tones] is divided for pilot and data transmission; i.e. $B_p = 6$, $B_d = 8$ [tones] and



a) Pilot–Data Alignment



b) Cellular System Layout

Fig. 3. The proposed cellular blind pilot–data alignment with the associated cellular mapping using the transmission subframe depicted in Fig. 2.

the coherence time is set to $T_c = 7$ [symbols]. We choose $B_p = 6$ to arrive to the same number of users in each cell (i.e. $K = B_pT_c = 42$ [tones$\times$ symbols]) similar to that in [1].

We consider three reference schemes: frequency reuse factor of one and three [1] and partial pilot sequence reuse factor of three. For the two first cases the number of terminals are 42 and for the third case, the number of terminals is set to 14 such that any two adjacent cells operate using mutually orthogonal pilot sequences. Table II summarizes the number of terminals per cell, cell-edge terminal capacity (i.e. 5% of the cdf), average terminal capacity and cell-average capacity. Fig. 4 shows the cumulative distribution of uplink capacity per terminal in [Mbps]. The scheme with the higher frequency reuse enhances the cell-edge capacity but it notably reduces the average terminal capacity since less bandwidth is allocated per terminal. Using a smaller number of scheduled terminals with the pilot sequence reuse, both cell-edge and average capacities per terminal are enhanced. It however significantly reduces the cell-average capacity since a less number of terminals are scheduled per cell. The cell-average capacity with the pilot sequence reuse and frequency reuse factor of three is the same (see also Appendix for a proof). The proposed blind pilot–data alignment interestingly brings notable gains in both cell-edge and cell-average capacity due to an efficient resource utilization of the adjacent cells such that it alleviates inter-cell pilot contamination over adjacent cells.

VII. Conclusions

We proposed and analyzed a pilot–data alignment scheme in which some of the frequency tones are reserved for pilot and the remaining tones are used for data transmission. The proposed alignment scheme without any need for channel state information, blindly manages the inter-cell pilot contamination, which enhances the cell-edge and cell-average capacities.

TABLE I
Simulation Parameters

Bandwidth	$B = 20$ MHz
Symbol Interval	$T_s = 66.7\mu$s
Guard Interval	$T_g = 4.76\mu$s
Coherence Time	$T_c = 0.5$ ms
Carrier Spacing	$\Delta f = 15$ kHz
Coherence Bandwidth	$B_c = 210$ kHz
Pilot Subband	$B_p = 105$ kHz
Data Subband	$B_d = 105$ kHz
Shadow Fading	log-normal, σ= 8 dB
Path-Loss Exponent	$\gamma = 3.8$
Small-scale fading	i.i.d. Rayleigh fading
Cell Radius	$r_c = 1600$ m
Cell-Hole Radius	$r_h = 100$ m
Number of Terminals	$K = \{42, 14\}$ per cell

Appendix

In this appendix, we prove that the cellular systems with arrays with an infinite number of antennas, the cell-average capacity for frequency reuse and pilot sequence reuse with equal reuse factor is the same. Consider the reuse factor α. For the frequency reuse α, the cell-average capacity is [1]

$$C_{\text{fr}} = \frac{B}{\alpha}\cdot\frac{T_c - T_p}{T_c}\cdot\frac{T_u}{T_s}\sum_{k=1}^{K}\mathbb{E}\log\left(1+\frac{\beta_{jkl}^2}{\sum_{i=1,i\neq l}^{L}\beta_{jki}^2}\right)$$
$$= \frac{KB}{\alpha}\cdot\frac{T_c - T_p}{T_c}\cdot\frac{T_u}{T_s}\mathbb{E}\log\left(1+\frac{\beta_{j1l}^2}{\sum_{i=1,i\neq l}^{L}\beta_{j1i}^2}\right) \quad (10)$$

TABLE II
MMIMO CELLULAR UPLINK CAPACITY

Scheme	Number of Terminals per Cell	Cell-Edge Terminal Capacity [Mbps]	Average Terminal Capacity [Mbps]	Cell-Average Capacity [Gbps]
Frequency Reuse 1	42	0.024	44	1.8
Frequency Reuse 3	42	0.95	28	1.2
Sequence Reuse 3	14	2.6	84	1.2
Proposed Solution	42	2.4	77	**3.2**

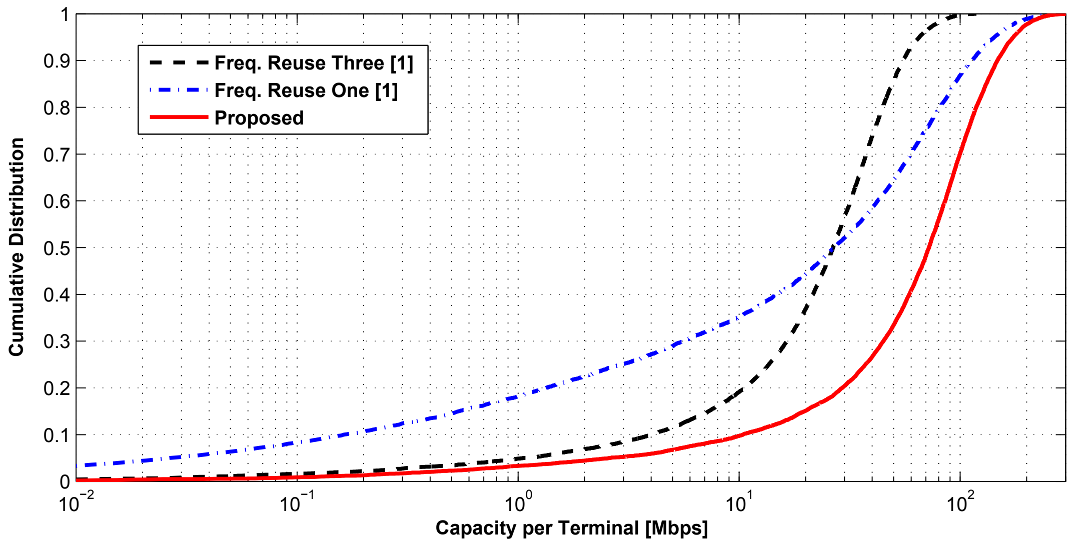


Fig. 4. The cumulative distribution of uplink capacity of the proposed scheme and the conventional solution with frequency reuse factor of one and three.

where the second equality follows since all terminals have i.i.d. β_{jkl}^2 and last equality follows by symmetry and only considering the first user in the target cell.

Next consider pilot sequence reuse factor of α, the cell-average capacity is given by

$$
\begin{aligned}
C_{\text{sr}} &= B \cdot \frac{T_c - T_p}{T_c} \cdot \frac{T_u}{T_s} \sum_{k=1}^{\frac{1}{\alpha}K} \mathbb{E} \log \left(1 + \frac{\beta_{jkl}^2}{\sum_{i=1, i \neq l}^{L} \beta_{jki}^2} \right) \\
&= \frac{KB}{\alpha} \cdot \frac{T_c - T_p}{T_c} \cdot \frac{T_u}{T_s} \mathbb{E} \log \left(1 + \frac{\beta_{jkl}^2}{\sum_{i=1, i \neq l}^{L} \beta_{jki}^2} \right) = C_{\text{fr}}
\end{aligned}
\tag{11}
$$

where the first equality follows since all terminals have i.i.d. β_{jkl}^2 and last equality follows by symmetry and only considering the first user in the target cell. We therefore conclude that the average capacity remains unchanged.

REFERENCES

[1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, 2010.

[2] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1436–1449, 2013.

[3] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, pp. 186–195, 2014.

[4] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *Signal Processing Magazine, IEEE*, vol. 30, no. 1, pp. 40–60, 2013.

[5] 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.

[6] Y. Li, Y.-H. Nam, B. L. Ng, and J. Zhang, "A non-asymptotic throughput for massive MIMO cellular uplink with pilot reuse," in *IEEE Global Communications Conference (GLOBECOM)*, Dec 2012, pp. 4500–4504.

[7] V. Saxena, G. Fodor, and E. Karipidis, "Mitigating pilot contamination by pilot reuse and power control schemes for massive MIMO systems," in *IEEE Vehicular Technology Conference (VTC Spring)*, 2015.

[8] M. N. Khormuji, "Generalized semi-orthogonal multiple-access for massive MIMO," in *Emerging MIMO Technologies and Millimeter-Waves for 5G Networks Workshop*, Glasgow, 2015.

[9] A. Ashikhmin and T. Marzetta, "Pilot contamination precoding in multi-cell large scale antenna systems," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, July 2012.

[10] H. Wang, Z. Pan, J. Ni, S. Wang, and I. Chih-Lin, "A temporal domain based method against pilot contamination for multi-cell massive MIMO systems," in *IEEE 79th VTC-Spring*, 2014.

[11] R. Muller, L. Cottatellucci, and M. Vehkapera, "Blind pilot decontamination," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 8, no. 5, pp. 773–786, Oct 2014.

[12] L. Cottatellucci, R. R. Muller, and M. Vehkapera, "Analysis of pilot decontamination based on power control," in *IEEE 77th Vehicular Technology Conference (VTC Spring)*, 2013.

[13] M. Li, S. Jin, and X. Gao, "Spatial orthogonality-based pilot reuse for multi-cell massive MIMO transmission," in *IEEE International Conference on Wireless Communications & Signal Processing*, 2013.

[14] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. on Wireless Communications,*, vol. 10, no. 8, pp. 2640–2651, 2011.

[15] F. Fernandez, A. Ashikhmin, and T. Marzetta, "Interference reduction on cellular networks with large antenna arrays," in *Proc. of IEEE International Conference on Communications (ICC)*, 2012.

[16] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE Journal on Selected Areas in Communications,*, vol. 31, no. 2, pp. 264–273, 2013.