Progress In Electromagnetics Research C, Vol. 102, 93–107, 2020

Pilot Contamination Mitigation Based on Interfering User's Angle of Arrival in Massive MIMO Systems

Parfait I. Tebe*, Guangjun Wen, and Kwadwo Ntiamoah-Sarpong

Abstract—A new approach to mitigate pilot contamination in massive multiple-input multiple-output (MIMO) systems is proposed in this paper. We consider two cells from the first tier of copilot cells of a cellular network where the base stations (BSs) are equipped with uniform linear arrays with hybrid beamforming adopted. We consider one cell as the cell of interest containing a typical desired user, and the other cell contains an interfering user sending data and contaminating pilot signals to the BS of the cell of interest. We derive a closed-form expression for the desired user's achievable rate as a function of the interfering user's angle of arrival (AoA). We model the ray propagation from the interfering user to the BS of the cell of interest and its related AoA as Gaussian distribution. Based on the model, we derive closed-form expressions for the pilot contamination level in the cell of interest and for the desired user's data path gain estimation error due to pilot contamination. A perfect agreement is found between theoretical and Monte Carlo simulation results which show that when the interfering user's AoA is increased the pilot contamination level is significantly minimized, the desired user's data path gain estimation error also minimized, and hence its data rate is significantly increased. Moreover, we show in our analysis that the interfering user's AoA can be effectively controlled and increased by reducing the copilot cells' radius.

1. INTRODUCTION

The fifth generation wireless communication networks are expected to significantly improve spectral efficiency, energy efficiency, and user' quality of services. To meet such needs, massive MIMO has been approved as one of the technology candidates and has then recently received significant attention from researchers. The technology involves many single-antenna user terminals served by a large number of antennas deployed at the base stations [1–4]. It has been proven to provide a more significant gain in spectral and energy efficiency than conventional MIMO systems, thanks to the large number of antennas and the application of multiuser detection and beamforming techniques [5–9].

However, massive MIMO huge potential is limited by some factors, mainly pilot contamination which is shown to have a severe effect on the system performance [10–13]. Indeed, the number of orthogonal pilot sequences within a cell is limited in a practical massive MIMO system, and the BSs then receive pilots from other co-channel cells. Such situation generates inter-cell interference and causes pilot contamination in the cells. Pilot contamination makes the channel state information (CSI) imperfect, and imperfect CSI affects the system performance [10–14]. The problem of pilot contamination in massive MIMO technology has been so far addressed by several approaches from researchers. Authors in [15] have done a survey in which they have categorized the commonly proposed techniques in two different approaches: pilot-based and subspace-based. The pilot-based approach includes techniques to transmit pilot and to exploit users' covariance information [16, 17]. Blind and semi-blind techniques

Received 15 April 2020, Accepted 6 May 2020, Scheduled 21 May 2020
* Corresponding author: Parfait I. Tebe (pariteb@yahoo.fr).
The authors are with the Center for RFIC and System Technology, University of Electronic Science and Technology of China, No. 2006, Xiyuan Avenue, West High-Tech Zone, Chengdu 611731, China.

combined with algorithms for channel estimation and prediction-based techniques for channel estimation are used in the subspace-based approach [18–21]. However, the channel estimation methods used in those approaches are not linear and make the system more complex. Moreover, the proposed approaches in which mitigation methods involve dynamic channels are just limited by certain structures, but they need to be exploited for the practicability of the system. Other approaches have also been proposed in [22, 23] and deal with inter-cell interference represented by pilot contamination in millimeter-wave massive MIMO systems. The approach in [22] focuses on antenna subset transmission technique where Walsh codes are used to generate random sequence to distort the side lobes signal from antennas, and a combination in a destructive manner occurs at the reception side. A 2 D-unitary estimation of the signal parameters through rotational invariance techniques algorithm is used in the approach in [23]. However, inter-cell interference represented by pilot contamination in massive MIMO is less critical at millimeter-wave frequencies due to millimeter-waves characteristics. Moreover, those approaches involve more signal processing at the BSs.

In this paper, we propose a new approach to mitigate pilot contamination. We consider a uniform linear array (ULA) where hybrid beamforming with the users' AoA information is adopted. Our approach is based on maximizing the interfering user's AoA at the BS of a cell of interest to minimize the pilot contamination level, a typical desired user's data path gain estimation error, and hence increase the user's achievable data rate, without further signal processing at the BS.

The main contributions of our work are as follows. First, we derive a closed-form expression for the desired user's data rate as a function of an interfering user's AoA. Second, we model the ray propagation from the interfering user to the BS of the cell of interest and its related AoA as Gaussian distribution. Based on the model, we derive closed-form expressions for the pilot contamination level in the cell of interest and for the desired user's data path gain estimation error due to pilot contamination. We show analytically and by simulations that as the interfering user's AoA increases, the pilot contamination level decreases, the user's data path gain estimation error also minimized, and hence its achievable data rate increases. Third, we show in our analysis that the interfering user's AoA can be effectively controlled and increased by reducing the copilot cells' radius. This supports existing work in [24] that, reducing the cell size in a massive MIMO system can mitigate pilot contamination. However, the main difference between this work and [24] is that in [24] the cell size reduction is used to increase the pilot sequence length in a desired cell whereas in this work it is used to maximize the interfering user's AoA which in turn helps to mitigate the effect of pilot contamination. Moreover, results from [24] are valid only when users are located at the edge of the cells, whereas in this work, users are randomly located within their cells, which is the most likely situation in a practical scenario.

The rest of the paper is organized as follows. Sections 2 and 3 present the system model and performance analysis, respectively. Our proposed model and approach are described in Section 4. Sections 5 deals with simulation results and Section 6 concludes the paper.

For readers' convenience, the acronyms used throughout the paper are summarized in Table 1.

Table 1. List of acronyms.

Acronyms	**Full names**
MIMO	Multiple-Input Multiple-Output
BS	Base Station
AoA	Angle of Arrival
CSI	Channel State Information
ULA	Uniform Linear Array
SINR	Signal-to-Interference-plus-Noise Ratio
SIR	Signal-to-Interference Ratio

2. SYSTEM DESCRIPTION

This section presents the channel model with pilot contamination in a massive MIMO system and provides the received signal under a uniform linear array configuration.

2.1. Channel Model with Pilot Contamination

We consider the first layer of a cellular network consisting of seven hexagonal cells sharing the same pilots, using the same frequency (copilot cells) and including the central network cell. Those seven cells are known as interfering cells, and the interference of copilots beyond that layer is assumed to be neglected [25]. This is an approximation of the exact configuration of the first tier of co-channel cells for a cluster size equal to 7 [25], and it is illustrated in Figure 1. All the cells are assumed to have the same size with radius R, and the copilot distance (distance separating the BSs of two copilot cells) is denoted by D. We consider an uplink transmission scenario where each BS is equipped with M antennas receiving signals from K single-antenna users. Ideally, to estimate the channel, each BS is supposed to receive orthogonal pilot signal**s** only from each of its users. However, in practice, the number of orthogonal pilot sequences within a cell is limited in a massive MIMO system. The BSs then receive pilots from users from other co-channel cells, and the channel estimator suffers from a lack of orthogonality between the desired pilots (pilots from the home cell) and the interfering pilots (pilots from other cells). It then causes inter-cell interference with the effect known as pilot contamination.

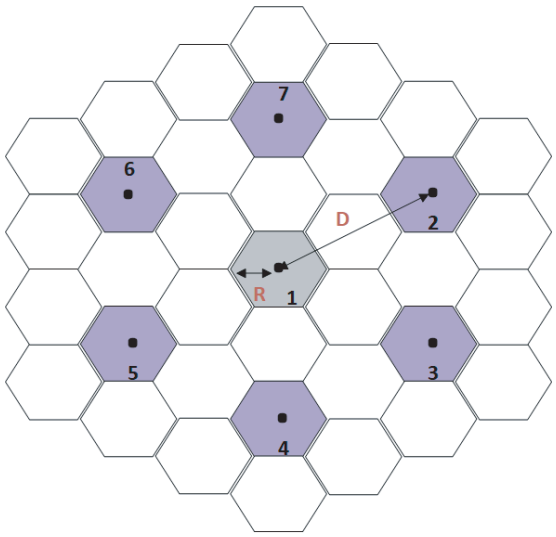

Figure 1. Illustration of first tier of copilot cells.

To simplify our illustration of pilot contamination phenomenon, we consider a system with two co-channel cells as illustrated in Figure 2, and one user from a total number of K is considered in each cell just for illustration purpose. Cell 1 and Cell 2 contain BS1 and BS 2, respectively. During the uplink transmission, BS 1 receives signal and pilot sequence from User 1 of Cell 1 and from User 1 of Cell 2. BS 2 also receives signal and pilot sequence from User 1 of Cell 2 and from User 1 of Cell 1. User 1 of Cell 1 and User 1 of Cell 2 are desired and interfering users for BS 1, respectively. Similarly, User 1 of Cell 2 and User 1 of Cell 1 are desired and interfering users for BS 2, respectively.

Similar to [22] and [26], we consider a hybrid beamforming architecture consisting of a precoder $\mathbf{F}$ used to transmit signal $\mathbf{x}$ and a combiner $\mathbf{W}$ used to extract the transmitted data from the received signal. The signal received at any of the two BSs is given as [22]

$$\mathbf{y} = \mathbf{w}^H(\sqrt{p_d}\mathbf{h}_d(\theta_d, \gamma_d, \phi_d)\mathbf{f}_d\mathbf{s}_d + \sqrt{p_i}\mathbf{h}_i(\theta_i, \gamma_i, \phi_i)\mathbf{f}_i\mathbf{s}_i + \mathbf{n}) \tag{1}$$

where the indices d and i refer to the desired and interfering users of the BS, respectively; p_d and p_i are the transmission powers; $\mathbf{s}_d$ and $\mathbf{s}_i$ are the pilot signals; θ_d and θ_i are the angles of arrival between the

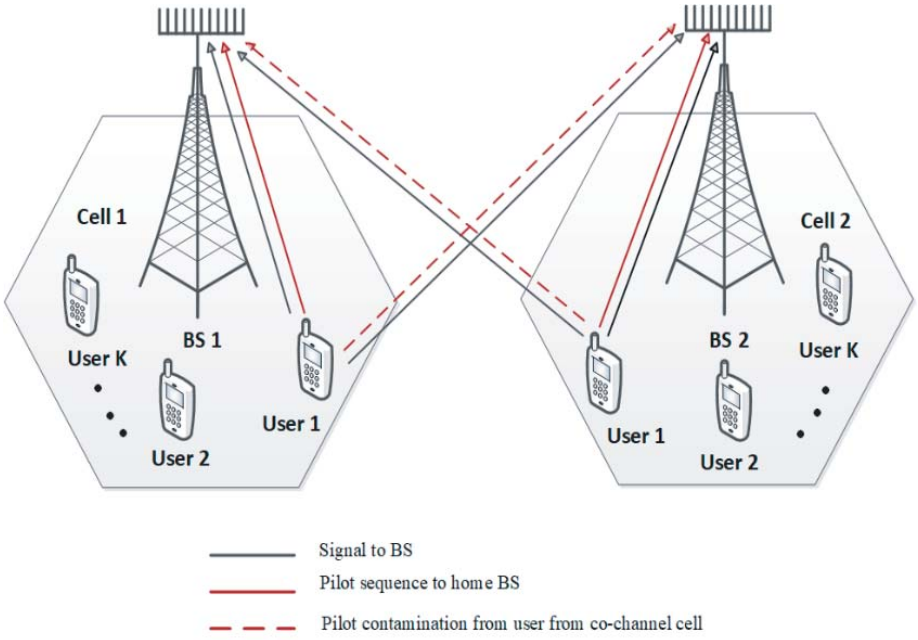

Figure 2. Illustration of pilot contamination.

users and the BSs; ϕ_d and ϕ_i are the angles of departure between the users and the BSs; γ_d and γ_i are the path loss exponents, and $\mathbf{n}$ is the additive white Gaussian noise component with $\mathbf{n} \sim \mathcal{N}(0, \sigma_n^2)$.

2.2. Uniform Linear Array Configuration

A uniform linear array located along the x-y plane is adopted in this work with antennas to capture all the incoming wave information with the assumption that the angles of arrival $\theta \in [0, \pi/2]$ [22]. The received signal in Eq. (1) can be rewritten as [22]

$$\mathbf{y} = \mathbf{w}^H(\sqrt{p_d}\gamma_d \mathbf{h}_d(\theta_d, \phi_d)\mathbf{f}_d \mathbf{s}_d + \sqrt{p_i}\gamma_i \mathbf{h}_i(\theta_i, \phi_i)\mathbf{f}_i \mathbf{s}_i + \mathbf{n}) \tag{2}$$

For maximum reception, perfect alignment is assumed between the receiver and transmitter beams [27–31], and $\mathbf{w} = \mathbf{a}_r(\phi)$, where $\mathbf{a}_r$ is the reception antenna array manifold. Replacing $\mathbf{W}$ in Eq. (2) then gives the received signal as

$$\mathbf{y} = \sqrt{p_d}\gamma_d \mathbf{a}_r^*(\phi_d)\mathbf{h}_d(\theta_d, \phi_d)\mathbf{f}_d \mathbf{s}_d + \sqrt{p_i}\gamma_i \mathbf{a}_r^*(\phi_i)\mathbf{h}_i(\theta_i, \phi_i)\mathbf{f}_i \mathbf{s}_i + \mathbf{v} \tag{3}$$

where $\mathbf{v} = \mathbf{w}^H(\phi)\mathbf{n}$ is the modified noise component, and $(\cdot)^*$ stands for the conjugate operator. Moreover, with the perfect alignment applied, the received signal can also be written as [27–29]

$$\mathbf{y} = (\sqrt{p_d}\gamma_d \mathbf{h}_d(\theta_d)\mathbf{f}_d + \sqrt{p_i}\gamma_i \mathbf{h}_i(\theta_i)\mathbf{f}_i)\mathbf{s} + \mathbf{v} \tag{4}$$

Following the guidelines in [22], the received signal under the ULA configuration is given by

$$\mathbf{y} = \sqrt{p_d M K}\gamma_d \mathbf{s} + \sqrt{\frac{p_i M}{K}}\gamma_i \left[\frac{\sin\left(K\frac{\pi d}{\lambda}(\cos\theta_d - \cos\theta_i)\right)}{\sin\left(\frac{\pi d}{\lambda}(\cos\theta_d - \cos\theta_i)\right)}\right]\mathbf{s} + \mathbf{v} \tag{5}$$

where d is the inter-antenna spacing, and λ is the wavelength of the carrier frequency.

3. SYSTEM PERFORMANCE ANALYSIS IN THE PRESENCE OF PILOT CONTAMINATION

This section provides an analysis of a typical desired user's data rate under pilot contamination. We then consider Cell 1 from Figure 2 as the cell of interest and the rate is for User 1 in that cell.

3.1. Achievable Data Rate

For imperfect CSI due to pilot contamination, the user's achievable data rate is given as [32]

$$R_k = (1-\mu)\log_2(1+SINR_k) \tag{6}$$

where $SINR_k$ is the user signal-to-interference-plus-noise ratio; μ is the fraction of an uplink transmission time devoted for channel estimation, and it is related to the total pilot sequence length τ available within the cell and an uplink transmission cycle time T_t by $\mu = \tau/T_t$.

Our investigated system is set to be interference-limited where the noise can be neglected in contrast with the user's interference [22, 24, 33]. Therefore, in our work, we rather consider signal-to-interference ratio (SIR), and Eq. (6) can be rewritten as

$$R_k = (1-\mu)\log_2(1+SIR_k) \tag{7}$$

Moreover, in this work, we are interested in pilot contamination which occurs as an interference where a BS from a given cell receives pilot signals from users from other cells. Therefore, following the guidelines in [23] and [33], and based on Eq. (3), the SIR_k is given as

$$SIR_k = \frac{\left\|\sqrt{p_d}\gamma_d \mathbf{a}_r^*(\phi_d)\mathbf{h}_d\mathbf{f}_d\right\|^2}{\left\|\sqrt{p_i}\gamma_i \mathbf{a}_r^*(\phi_i)\mathbf{h}_i\mathbf{f}_i\right\|^2} \tag{8}$$

where the term in the numerator is only the desired term corresponding to the desired signal in the cell of interest, and the term in the denominator is the interfering term corresponding to the signal from the co-channel cell user. Using Eqs. (5) and (8), the obtained SIR_k becomes

$$SIR_k = \frac{(\sqrt{p_d MK}\gamma_d)^2}{\left[\sqrt{\frac{p_i M}{K}}\gamma_i \left(\frac{\sin\left(K\frac{\pi d}{\lambda}(\cos\theta_d - \cos\theta_i)\right)}{\sin\left(\frac{\pi d}{\lambda}(\cos\theta_d - \cos\theta_i)\right)}\right)\right]^2} \tag{9}$$

Further derivations give

$$SIR_k = \frac{p_d K^2 \gamma_d^2}{p_i\gamma_i^2\left(\frac{\sin\left(K\frac{\pi d}{\lambda}(\cos\theta_d - \cos\theta_i)\right)}{\sin\left(\frac{\pi d}{\lambda}(\cos\theta_d - \cos\theta_i)\right)}\right)^2} \tag{10}$$

The user's achievable data rate in Eq. (7) then becomes

$$R_k = (1-\mu)\log_2\left[1 + \frac{p_d K^2 \gamma_d^2}{p_i\gamma_i^2\left(\frac{\sin\left(K\frac{\pi d}{\lambda}(\cos\theta_d - \cos\theta_i)\right)}{\sin\left(\frac{\pi d}{\lambda}(\cos\theta_d - \cos\theta_i)\right)}\right)^2}\right] \tag{11}$$

3.2. Performance Analysis

For simplicity purpose, we assume that the path loss exponents γ_d and γ_i are equal, and the transmission powers p_d and p_i are also equal [22]. Moreover, in our array configuration, we adopt $d = \lambda/2$ for the

inter-element spacing [33]. The user's data rate then becomes

$$R_k = (1-\mu)\log_2\left[1+\frac{K^2}{\left(\frac{\sin\left(\frac{K\pi}{2}(\cos\theta_d-\cos\theta_i)\right)}{\sin\left(\frac{\pi}{2}(\cos\theta_d-\cos\theta_i)\right)}\right)^2}\right] \tag{12}$$

From Eq. (12), we observe that the user SIR part will increase with the number of users, for all other parameters given and fixed. However, with the presence of pilot contamination, the SIR converges because of the corrupted channel estimate, and it limits the data rate [32]. Our primary objective is then to increase the desired use's achievable data rate despite the presence of pilot contamination. Let's denote

$$(\cos\theta_d-\cos\theta_i) = Z_1 \tag{13}$$

$$\sin\left(\frac{K\pi}{2}(\cos\theta_d-\cos\theta_i)\right) = A \tag{14}$$

$$\sin\left(\frac{\pi}{2}(\cos\theta_d-\cos\theta_i)\right) = B \tag{15}$$

$$\frac{\sin\left(\frac{K\pi}{2}(\cos\theta_d-\cos\theta_i)\right)}{\sin\left(\frac{\pi}{2}(\cos\theta_d-\cos\theta_i)\right)} = f(\theta_i) = Z \tag{16}$$

Z_1 is the angular separation between the desired and interfering users [33]. For K and μ given, R_k will increase if and only if Z decreases; that is, if A and B increase. Numerical results are provided in Table 2 to confirm this theoretical analysis of decrease of Z. From trigonometry properties, given any angle $x \in [0, \pi/2]$, $\sin(x)$ increases as x increases. Therefore, A and B will increase if Z_1increases. It is evident that the values of θ_d and θ_i should be different; otherwise, Z_1 will give zero, A and B will give zero, and Eq. (16) providing Z will be absurd. This can also be justified by the fact that the desired and interfering signals arrive at the BS from two non-overlapping sets of paths because of the different physical locations of the users. The probability that their paths have the same AoA is then equal to zero [16, 26]. Moreover, from trigonometry properties, given any angle $x \in [0, \pi/2]$, $\cos(x)$ decreases as x increases and vice versa. Therefore, increasing θ_d and θ_i at the same time would not guarantee an increase of Z_1. Then, one needs to be fixed while the other is increased or decreased depending on the following two cases.

Case 1: $\theta_d > \theta_i$, i.e., $\cos\theta_d < \cos\theta_i$.

For θ_i fixed, $\cos\theta_i$ is fixed, and increasing Z_1 means increasing $\cos\theta_d$, which means decreasing θ_d.

For θ_d fixed, $\cos\theta_d$ is fixed, and increasing Z_1 means decreasing $\cos\theta_i$, which means increasing θ_i.

Conclusion 1: For $\theta_d > \theta_i$, increasing Z_1 means that either θ_i is fixed and θ_d decreased (but not to the value of θ_i) or θ_d is fixed and θ_i increased (but not up to the value of θ_d).

Case 2: $\theta_d < \theta_i$, i.e., $\cos\theta_d > \cos\theta_i$.

For θ_i fixed, increasing Z_1 means increasing $\cos\theta_d$, which means decreasing θ_d.

For θ_d fixed, increasing Z_1 means decreasing $\cos\theta_i$, which means increasing θ_i.

Conclusion 2: For $\theta_d < \theta_i$, increasing Z_1 means that either θ_i is fixed and θ_d decreased, or θ_d is fixed and θ_i increased.

Conclusion 1 and conclusion 2 are the same. However, there are some restrictions in conclusion 1 that θ_d should not reach the value of θ_i if it is decreased, and θ_i should not reach the value of θ_d if it is increased. This is to avoid $\theta_d = \theta_i$ case as justified earlier in this work. Moreover, as θ_d belongs to the desired user in the cell of interest and θ_i to the interfering user causing pilot contamination, we will choose to keep θ_d fixed and vary θ_i in a way that will help in achieving our main objective which is to

minimize the effect of pilot contamination in the cell of interest. That is, we keep θ_d fixed and increase θ_i to comply with conclusion 1 and conclusion 2. Numerical results are also provided in Table 2 to confirm the theoretical analysis of increase of Z_1. We set $K = 30$, $\theta_d = 15°$, and we vary θ_i from 60° to 75°.

To accomplish our objective, we propose, in the next section, a model for the interfering user signal path and AoA based on which the results from this section can be achieved.

4. INTERFERING PATH AOA MODEL

This section provides a model for the interfering path and the associated AoA, and based on the model we show how an increase of the AoA can effectively minimize the pilot contamination, reduce the desired user's data path gain estimation error, and hence improve its achievable rate.

We consider that the response between the antennas at the BS of the cell of interest and the antennas of the desired and interfering users is given by a superposition of rays having distinct AoAs. The spatial response of each ray is determined by the AoA together with the spatial distribution of the antenna elements, their patterns and the carrier frequency [26]. We, therefore, suppose that the BS in the cell of interest estimates each channel separately from the signals received from the desired and interfering users. Based on that, we consider propagation according to the Gaussian distribution model where the ray channel from the interfering user to the BS of the cell of interest is described by m dimensional vector $\mathbf{g}_i$. The elements of $\mathbf{g}_i$ are then assumed to be Gaussian random variables with covariance matrix given by

$$\mathbf{R}_i = \mathrm{E}\{\mathbf{g}_i,\mathbf{g}^*\} \tag{17}$$

The power received from the interfering user is also assumed Gaussian distributed in azimuthal angle with the AoA θ_i being the mean and σ_i the standard deviation of the distribution also known as the angular spread. Under such assumption, the covariance matrix then becomes

$$\mathbf{R}_i = G_i\mathbf{R}(\theta_i, \sigma_i) \tag{18}$$

where G_i is the real-valued scalar referred to as the path gain between the BS and the interfering user.

4.1. Effect of the Interfering User's AoA on the Pilot Contamination Level

In this subsection, we show how an increase of θ_i under our model can effectively minimize the effect of pilot contamination. Following the guidelines in [26], the pilot contamination level at the desired signal reception in our case can be given to be bounded as

$$\eta \leq \frac{\kappa}{\sqrt{\tau}M \min\limits_{n,l} |\psi_{in} - \psi_{dl}|} \tag{19}$$

where κ is a constant; n and l are the n^{th} and l^{th} interference path and data path per user, respectively; ψ_{in} and ψ_{dl} are the inter-antenna phase difference and the constant phase difference between the observations by two adjacent antennas, respectively [16, 26], and $\min\limits_{n,l} |\psi_{in} - \psi_{dl}|$ is the minimum spatial separation between data and interfering paths. Assuming the worst-case where the pilot contamination level is the highest, Eq. (19) can be rewritten as

$$\eta = \frac{\kappa}{\sqrt{\tau}M \min\limits_{n,l} |\psi_{in} - \psi_{dl}|} \tag{20}$$

The interfering and desired user' channels are represented as

$$\mathbf{h}_i = \beta_{in}\mathbf{v}(\psi_{in}) = \beta_{in}\ [1, e^{j\psi_{in}}, ..., e^{j(M-1)\psi_{in}}]^2 \tag{21}$$

and

$$\mathbf{h}_d = \beta_{dl}\mathbf{v}(\psi_{dl}) = \beta_{dl}\ [1, e^{j\psi_{dl}}, ..., e^{j(M-1)\psi_{dl}}]^2 \tag{22}$$

respectively. Here, β_{in} and β_{dl} are the channel coefficients. $\mathbf{v}(\psi_{in})$ and $\mathbf{v}(\psi_{dl})$ are the interference and data path vectors, respectively. Moreover, the inter-antenna and constant phase differences are given by

$$\psi_{in} = \frac{2\pi d}{\lambda} \cos\theta_i \tag{23}$$

and

$$\psi_{dl} = \frac{2\pi d}{\lambda} \cos\theta_d \tag{24}$$

Substituting Eqs. (23), (24) and $d = \lambda/2$ into Eq. (20) gives

$$\eta = \frac{\kappa}{\sqrt{\tau}\pi M \min\limits_{n,l} |\cos\theta_i - \cos\theta_d|} \tag{25}$$

To conform to the analysis done in the previous section, we consider that the values of θ_i and θ_d in Eq. (25) are the minimum possible values from n^{th} and l^{th} paths. Let's denote

$$|\cos\theta_i - \cos\theta_d| = Z_2 \tag{26}$$

Considering Eq. (25) and given κ, τ, and M, the pilot contamination level will decrease if and only if Z_2 increases. In the previous section, we considered fixed θ_d and two cases in the analysis. We then follow the same guidelines and show that for $\theta_d > \theta_i$ (case 1 in the previous section), increasing θ_i will decrease the value of Z_2. However, for $\theta_d < \theta_i$ (case 2 in the previous section), increasing θ_i will increase the value of Z_2, and hence reduce the pilot contamination level. The results from this analysis then exclude case 1 and agree with case 2, for pilot contamination minimization, and numerical results are provided in Table 2 to confirm the theoretical analysis of increase of Z_2. From these results and the ones from the previous section we deduce and propose the following theorem.

Theorem 1: With a uniform linear array deployed in a massive MIMO system subject to pilot contamination and where the users' AoAs are distributed in $[0, \pi/2]$, for fixed AoA of any desired user (from a cell of interest) and less than that of any interfering user, an increase of the interfering user's AoA can minimize the pilot contamination level.

Table 2. Theoretical values of A, B, Z, Z_1, Z_2 and Z_3 for $K = 30$ and $\theta_d = 15°$.

$\theta_i(°)$	A	B	Z	Z_1	Z_2	Z_3
60	0.373898	0.012773	29.2725	0.4659	0.4659	0.4659
61	0.385456	0.013189	29.2255	0.4811	0.4811	0.4811
62	0.397065	0.013610	29.1745	0.4964	0.4964	0.4964
63	0.408718	0.014034	29.1234	0.5119	0.5119	0.5119
64	0.420409	0.014462	29.0699	0.5275	0.5275	0.5275
65	0.432129	0.014894	29.0136	0.5433	0.5433	0.5433
66	0.443871	0.015329	28.9562	0.5591	0.5591	0.5591
67	0.455628	0.015768	28.8957	0.5751	0.5751	0.5751
68	0.467393	0.016210	28.8336	0.5913	0.5913	0.5913
69	0.479159	0.016655	28.7696	0.6075	0.6075	0.6075
70	0.490916	0.017103	28.7035	0.6239	0.6239	0.6239
71	0.502660	0.017554	28.6350	0.6403	0.6403	0.6403
72	0.514381	0.018008	28.5640	0.6569	0.6569	0.6569
73	0.526072	0.018464	28.4917	0.6735	0.6735	0.6735
74	0.537727	0.018923	28.4165	0.6902	0.6902	0.6902
75	0.549338	0.019384	28.3397	0.7071	0.7071	0.7071

4.2. Effect of the Interfering User's AoA on the Desired User's Data Path Gain Estimation

The interfering path gain G_i between the BS and the interfering user under our model affects the desired user's data path gain by introducing an estimation error given as [26]

$$\varepsilon = \left|\hat{G}_d - G_d\right| \leq \frac{2\alpha_{\max}(L+N-1)}{\psi_{\min} M} + o\left(\frac{1}{M}\right) \tag{27}$$

where G_d and $\hat{G}_d$ are the desired user's data path gain and its estimation, respectively; $\alpha_{\max}$ is the maximum of the magnitudes of all path gains; L and N are the number of data paths and interference paths per user, respectively; $\circ(\frac{1}{M})$ is the computational complexity associated to the estimation, and $\psi_{\min}$ is the minimum AoA separation between the desired and interfering users pair of paths. For our case study $\psi_{\min} = |\psi_{dl} - \psi_{in}|$. For the worst case gain estimation error due to worst case pilot contamination level assumed in the previous subsection, Eq. (27) can be rewritten as

$$\varepsilon = \left|\hat{G}_d - G_d\right| = \frac{2\alpha_{\max}(L+N-1)}{|\psi_{dl} - \psi_{in}|\, M} + o\left(\frac{1}{M}\right) \tag{28}$$

Substituting Eqs. (23), (24) and $d = \lambda/2$ into Eq. (28) gives

$$\varepsilon = \left|\hat{G}_d - G_d\right| = \frac{2\alpha_{\max}(L+N-1)}{M\pi\,|\cos\theta_d - \cos\theta_i|} + o\left(\frac{1}{M}\right) \tag{29}$$

Our objective here is to show how the desired user's data path gain estimation error is minimized as a result of pilot contamination level minimization. Let's denote

$$|\cos\theta_d - \cos\theta_i| = Z_3 \tag{30}$$

Following the same analysis guidelines of Z_1 and Z_2, it is inferred that increasing θ_i will increase Z_3, and hence minimize ε. Numerical results are also provided in Table 2 to confirm the theoretical analysis of increase of Z_3. The result from this analysis added to Theorem 1 gives the following Theorem 2.

Theorem 2: With a uniform linear array deployed in a massive MIMO system subject to pilot contamination and where the users' AoAs are distributed in $[0, \pi/2]$, for fixed AoA of any desired user (from a cell of interest) and less than that of any interfering user, an increase of the interfering user's AoA can minimize the pilot contamination level, and hence minimize the desired user's data path gain estimation error. When the data path gain estimation error is minimized, the user *SIR* is maximized, and hence its achievable rate increases.

4.3. Effect of the Cell Size on the Proposed Model

The performance gain of an antenna array system depends on the angular distribution of the energy received by the BS [26, 33]. Following the guidelines in [35], the standard deviation σ_i of the Gaussian angle distribution with respect to the AoA θ_i which is the mean in our model is given as

$$\sigma_i = \frac{90^\circ \times l_0}{\pi r} \tag{31}$$

where r is the distance between the interfering user and the BS receiving signal from it, which is the BS of the cell of interest in our case study; l_0 is a propagation distance related constant.

Equation (31) shows that as the distance between the interfering user and the BS of the cell of interest decreases, the standard deviation of the Gaussian angle distribution with respect to its AoA will increase. In statistics, the mean and standard deviation often go together because they both describe different but complementary things about data distribution. The mean measures where the data are centered, and the standard deviation measures how spread out the data are. A larger standard deviation always means that the data are more spread out than a smaller one. Given that in this work the standard deviation is the angular spread related to the mean (interfering user's AoA), the larger its value means the more it spreads out and then the larger the AoA is. From this analysis, we can conclude that when the distance between the interfering user and the BS of the cell of interest decreases, the interfering user's AoA increases. Since the interfering user's location is random, and the BS is assumed

to be located in the center of the cell, from Figure 1 and Figure 2, the most effective way to reduce the distance between the two is to reduce the copilot distance D. Such a distance is related to the cell radius R as [24]

$$D = \sqrt{3}R \tag{32}$$

Equation (32) shows that D is reduced when R is reduced, and we then deduce and propose the following theorem.

Theorem 3: When the ray propagation from the interfering user to the BS of the cell of interest and the related AoA θ_i are modelled as Gaussian distribution, reducing the copilot cells radius will increase θ_i.

From Theorem 1, Theorem 2, and Theorem 3, we conclude that in a massive MIMO system where a uniform linear array is adopted at the BSs with user's AoAs distributed in $[0, \pi/2]$, the effect of pilot contamination can be minimized when the AoA (modeled as Gaussian distribution) from any interfering user from a copilot cell is maximized as a result of cell size reduction. When the effect of pilot contamination is minimized, the error from the estimation of the data path gain of any desired user is minimized, and hence the use's achievable data rate can be maximized.

5. SIMULATION RESULTS

In this section, analytical and simulated results are provided to verify our analysis and evaluate our proposed approach. We performed the simulations over 10000 channel realizations using Monte Carlo. We considered an uplink transmission system where the BSs are equipped with ULAs with adjacent antennas separated by $d = \lambda/2$ and the users'AoAs $\theta \in [0, \pi/2]$. We fixed the AoA θ_d of the desired user at 15° and varied the interfering user's AoA θ_i from 60° to 75°. For generality, we used some simulation parameters from [25, 26, 32, 36–38], and the major ones are presented in Table 3.

Table 3. Simulation parameters.

Parameter	Value	Parameter	Value
System bandwidth, BW	20 MHz	Fraction of uplink transmission time devoted for channel estimation, μ	0.3
Carrier frequency, f_c	2 GHz	Pilot sequence length available within the cell of interest, τ	15
Cells radius, R	2000 m	Constant κ	200
Path loss exponent, γ	3.5	Maximum of the magnitudes of all paths gain, $\alpha_{\max}$	0.1
User uplink transmit power, p	23 dBm	Number of data paths per user, L	2
		Number of interfering paths per user, N	2

Our analysis is mainly based on varying θ_i. However, throughout the work, only its cosine appears in the final analytical results. Hence, plotting those results against θ_i will give oscillating graphs which will not be easy and appropriate to interpret, especially in the cases of data rate, pilot contamination level and data path gain estimation error, which are expected to be either increasing or decreasing functions in a smooth manner. Therefore, for simplicity and clarity of results analysis and interpretation, we first plotted the angular separation functions involving θ_i in the simulations. That is, we plotted Z, Z_1, Z_2, and Z_3 all against θ_i to confirm the theoretical analysis. The results are depicted in Figure 3. Even though the results from this plotting do not give the exact values provided in Table 2 due to the sinusoidal nature of the various functions, the theoretical analysis is validated. The results show that as the interfering user's AoA increases, the values of Z_1, Z_2, and Z_3 are the same. Moreover, it is shown that Z decreases, and Z_1, Z_2, and Z_3 increase as θ_i increases. This confirms the theoretical analysis in Section 3 and Section 4.

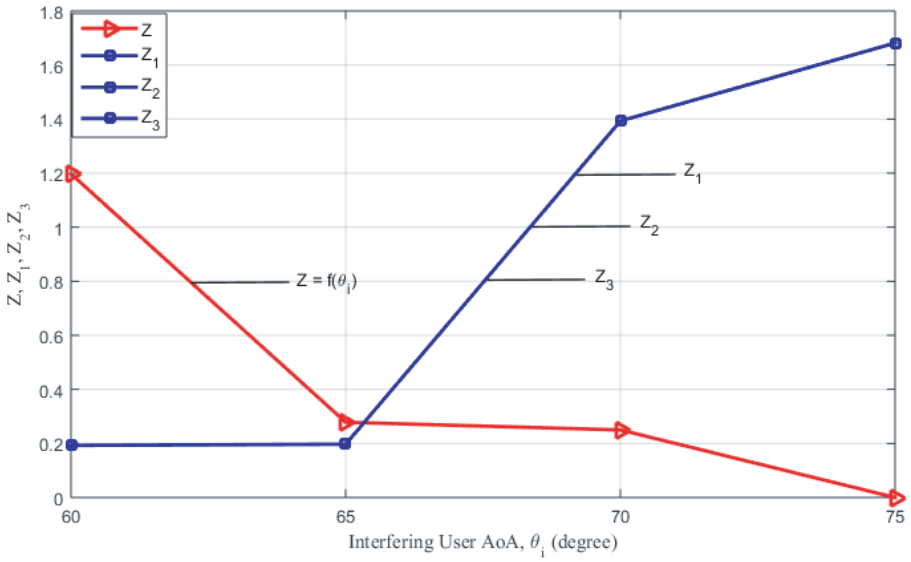

Figure 3. Angular separations versus interfering user's angle of arrival θ_i.

Next, we performed simulations and plotted the final analytical results as functions of Z, Z_2, and Z_3, and all the results show a perfect agreement between our simulation and theory.

In Figure 4, the pilot contamination level is depicted against Z_2 using Eq. (25), and we performed simulations for 200, 250, and 300 BS antennas. The results show that the pilot contamination level significantly decreases as Z_2 increases. Based on these results and the ones from Figure 3 (Z_2 increases as θ_i increases), it can be deduced that the pilot contamination level significantly decreases as the interfering user's AoA increases, which confirms Theorem 1. Moreover, the results show lower pilot contamination level for a higher number of BS antennas. This is justified by the role the number of BS antennas which plays in the performance of a massive MIMO system. That is, a higher number of BS antennas can achieve higher data rate when there is no or weak pilot contamination but better CSI

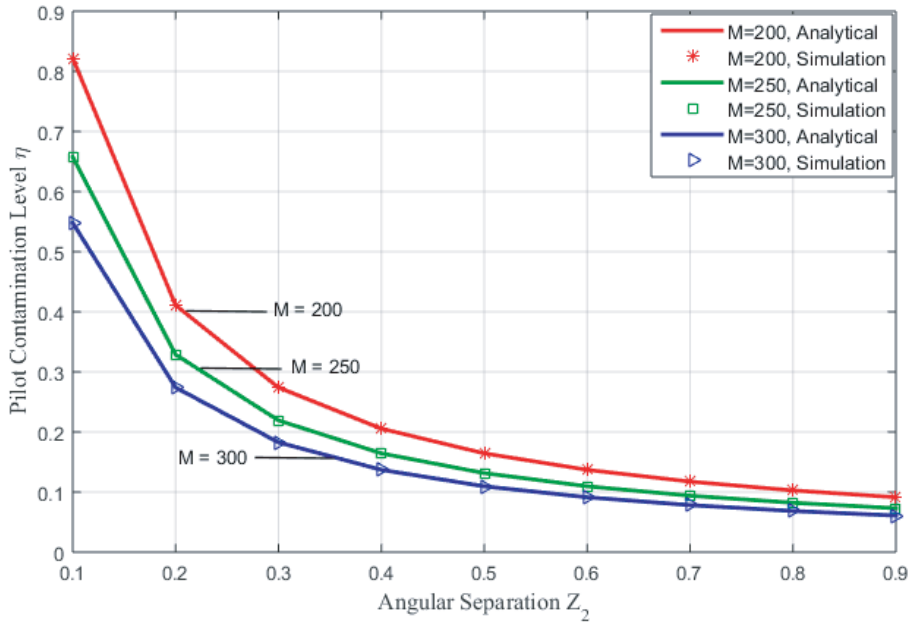

Figure 4. Pilot contamination versus angular separation Z_2.

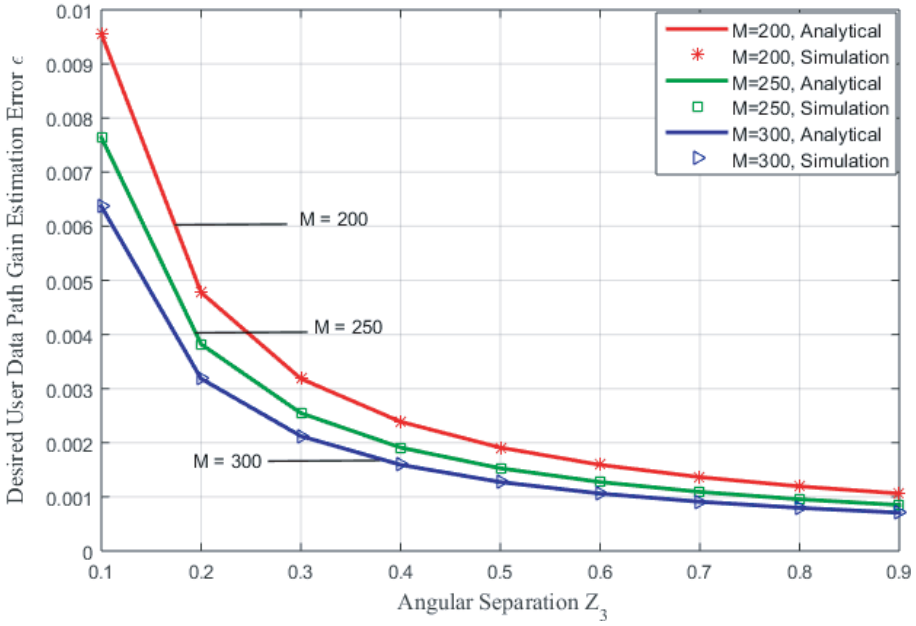

Figure 5. Desired user's data path gain estimation error versus angular separation Z_3.

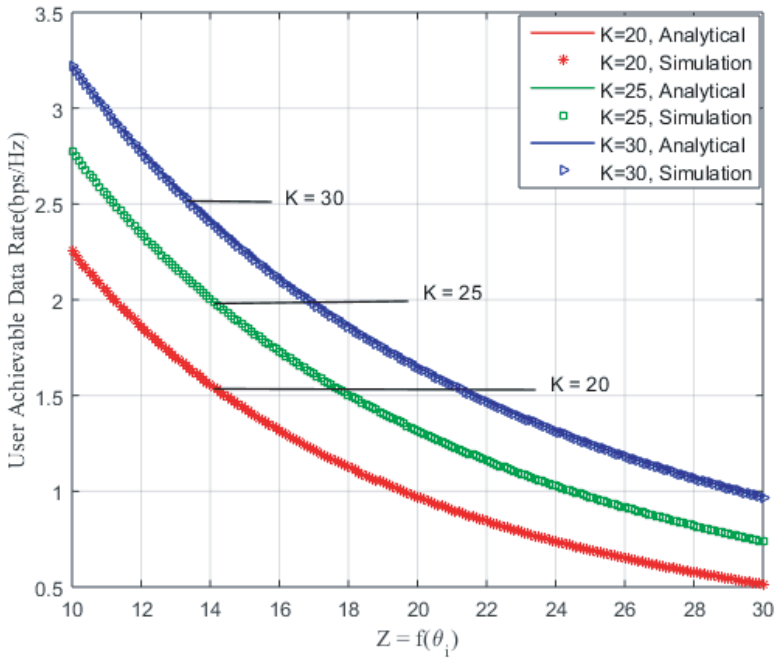

Figure 6. User's achievable data rate versus angular separation function $Z = f(\theta_i)$.

knowledge.

Figure 5 depicts the desired user's data path gain estimation error against Z_3 using Eq. (29), and for 200, 250, and 300 BS antennas. It is shown that the error significantly decreases as Z_3 increases. From these results and the ones from Figure 3 (Z_3 increases as θ_i increases), we deduce that the user's data path gain estimation error is minimized as the interfering user's AoA increases, which is a result of pilot contamination effect minimization, and it then confirms Theorem 2.

In Figure 6, the desired user's achievable data rate is depicted against Z using Eq. (12). We

performed simulations for 20, 25, and 30 number of users. The results show a higher data rate at lower values of Z. Based on these results and the ones from Figure 3 (Z decreases as θ_i increases), it can be claimed that the desired user's data rate increases as the interfering user's AoA increases. This is then a direct effect of data path gain estimation error minimization which is also due to the pilot contamination effect minimization as shown earlier. These results also confirm Theorem 2. Moreover, the results from the figure show higher data rate for a higher number of users, which is in accordance to one of the massive MIMO principles aiming to serve more users with higher data rate.

A performance comparison between the results from our proposed approach and some existing referenced results is shown in Table 4. We consider the data rate with the average values chosen for the comparison. It is realized that our approach offers a competitive advantage than the referenced results in terms of user data rate.

Table 4. Results comparison.

References	Carrier frequency	System bandwidth	Transmit power	Number of users	Spectral efficiency (SE) or User data rate $\mathbf{R_k}$
[17]	2 GHz	20 MHz	23 dBm	10	SE $\approx$ 9 b/s/Hz
[22]	28 GHz	100 MHz	37 dBm	64	SE $\approx$ 3 b/s/Hz
This work	2 GHz	20 MHz	23 dBm	20	$R_k \approx 1.45$ b/s/Hz

6. CONCLUSIONS

In this paper, we propose a new technique to minimize the effect of pilot contamination in massive MIMO technology. The approach is based on increasing any interfering user's AoA in an uplink transmission system where pilot contamination occurs as the effect of inter-cell interference. We considered two copilot cells from the first tier of copilot cells of a cellular network where the BSs are equipped with uniform linear arrays. We modeled the interfering path and the associated AoA as Gaussian distribution; we derived closed-form expressions for a typical desired user's achievable rate under pilot contamination, pilot contamination level, and the desired user's data path gain estimation error due to pilot contamination. Our results show that increasing the interfering user's AoA will minimize the pilot contamination level, reduce the desired user's data path gain estimation error, and hence increase its achievable data rate. In other words, the minimization of the pilot contamination level will significantly improve the reliability of the channel estimation, hence the user's achievable data rate. Finally, we showed in our analysis that the interfering user's AoA can be effectively controlled and increased by reducing the copilot cells' radius. This supports existing work in [24] that reducing the cell size in a massive MIMO system can mitigate pilot contamination. However, this work differs from [24] as follows.

i. In [24] the cell size reduction is used to increase the pilot sequence length in a desired cell whereas in this work it is used to maximize the interfering user's AoA which in turn helps to mitigate the effect of pilot contamination.

ii. Results from [24] are valid only when users are located at the edge of the cells, whereas in this work, users are randomly located within their cells, which is the most likely situation in a practical scenario.

Even though reducing the cell size could require more cell sites deployment in the network, the advantages of small cells deployment recognized as another promising technique for the next generation cellular systems come as a tradeoff. Such advantages include increasing capacity in areas with higher user densities, improving the network energy efficiency, and extending handset battery life by reducing power consumption. The developed results from this work will be further extended to a multi-cell system composed of more than two cells with other types of antenna array configurations deployed at the BSs and also with users' AoA distributed in a broader range.

REFERENCES

1. Larson, E. G., O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, Vol. 52, No. 2, 186–195, Feb. 2014.
2. Rusek, F., 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," *IEEE Sig. Proc. Mag.*, Vol. 30, No. 1, 40–60, Jan. 2013.
3. Ngo, H. Q., E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans Commun.*, Vol. 61, No. 4, 1436–1449, Apr. 2013.
4. Yao, R., T. Li, Y. Liu, X. Zuo, and H. Liu, "Analytical approximation of the channel rate for massive MIMO system with large but finite number of antennas," *IEEE Access*, Vol. 6, 6496–6504, Mar. 2018.
5. Prasad, K. N. R. S. V., E. Hossain, and V. K. Bhargava, "Energy efficiency in massive MIMO-based 5G networks: opportunities and challenges," *IEEE Wirel. Commun.*, Vol. 24, No. 3, 86–94, 2017.
6. Chen, Z., F. Sohrabi, and W. Yu, "Multi-cell sparse activity detection for massive random access: massive MIMO versus cooperative MIMO," *IEEE Trans. Commun.*, Vol. 18, No. 8, 4060–4074, Aug. 2019.
7. Wahhamy, A. A., N. E. Buris, H. Zhu, H. A. Rizzo, and S. Yahya, "An efficient paradigm for evaluating the channel capacity of closed-loop massive MIMO systems," *Progress In Electromagnetics Research C*, Vol. 98, 1–16, 2020.
8. Ge, X., K. Huang, C. X. Wang, X. Hong, and X. Yang, "Capacity analysis of a multi-cell multi-antenna cooperative cellular network with co-channel interference," *IEEE Trans. Wireless Commun.*, Vol. 10, No. 10, 3298–3308, Oct. 2011.
9. Lakshminarayara, S., M. Assaad, and M. Debbah, "Coordinated multicell beamforming for massive MIMO: a random matrix approach," *IEEE Trans. Info. Theory*, Vol. 61, No. 6, 3387–3412, Jun. 2015.
10. Khansefid, A. and H. Minn, "Achievable downlink rates of MRC and ZF precoders in massive MIMO with uplink and downlink pilot contamination," *IEEE Trans. Commun.*, Vol. 63, No. 12, 4849–4864, Dec. 2015
11. Parida, P. and H. S. Dhillon, "Stochastic geometry-based uplink analysis of massive MIMO systems with fractional pilot reuse," *IEEE Trans. Wireless Commun.*, Vol. 18, No. 3, 1651–1668, Mar. 2019.
12. Liu, G., H. Deng, X. Qian, W. Wang, and G. Peng, "Joint pilot allocation and power control to enhance max-min spectral efficiency in TDD massive MIMO systems," *IEEE Access*, Vol. 7, 149191–149201, Oct. 2019.
13. Figueiredo, F. A. P. D., C. F. Dias, F. A. C. M. Cardoso, and G. Fraidenraich, "On the distribution of an effective channel estimator for multi-cell massive MIMO," *IEEE Access*, Vol. 7, 114508–114519, Aug. 2019.
14. Akbar, N., S. Yan, A. M. Khattak, and N. Yang, "On the pilot contamination attack in multi-cell multiuser massive MIMO networks," *IEEE Trans. Commun.*, Vol. 68, No. 4, 2264–2276, Apr. 2020.
15. Elijah, O., C. H. Leow, T. A. Rahman, S. Nunoo, and S. Z. Iliya, "A comprehensive survey of pilot contamination in massive MIMO-5G system," *IEEE Commun Surveys & Tutorials*, Vol. 18, No. 2, 905–923, 2016 (Second Quarter).
16. Yin, H., 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, 264–273, Feb. 2013.
17. Yin, H., D. Gesbert, M. C. Filippou, and Y. Liu, "Decontaminating pilots in massive MIMO systems," *Proc. of IEEE Int. Conf. Commun. (ICC)*, 3170–3175, Jun. 2013.
18. Muller, R., L. Cottatellucci, and M. Vehkapera, "Blind pilot decontamination," *IEEE Journal on Selected Topics in Signal Processing*, Vol. 8, No. 5, 773–786, Oct. 2014.
19. Mazlan, M. H., M. Behjati, R. Nordin, and M. Is mail, "Wiener-based smoother and predictor for massive MIMO downlink system under pilot contamination," *Telecommunication Systems*, Vol. 67, No. 3, 387–399, Mar. 2018.

20. Amiri, E., R. Mueller, and W. Gerstacker, "Blind pilot decontamination in massive MIMO by independent component analysis," *Proc. of IEEEE Global Communications Conference (GLOBECOM)*, 1–5, Dec. 2017.
21. Mazlan, M. H., E. Ali, A. M. Ramly, and R. Nordin, "Pilot decontamination using coordinated wiener predictor in massive-MIMO system," *IEEE Access*, Vol. 6, 73180–73190, Dec. 2018.
22. Smaili, N., M. Djeddou, and A. Azrar, "Pilot contamination mitigation based on antenna subset transmission for mmWave massive MIMO," *International Journal of Communication Systems*, Vol. 31, No. 14, 1–12, Jul. 2018.
23. Ali, E., M. smail, R. Nordin, F. Abdulah, and M. H. Mazlan, "Performance analysis of inter-cell interference reduction using AOA-based beamforming in Mm wave massive MIMO systems," *Proc. of IEEE 13th Malaysia International Conference on Communication (MICC)*, 235–240, Nov. 2017.
24. Tebe, P. I., Y. Kuang, A. E. Ampoma, and K. A. Opare, "Mitigating pilot contamination in massive MIMO using cell size reduction," *IEICE Transactions on Communications*, Vol. E101-B, No. 5, 1280–1290, May 2018.
25. Rappaport, T. S., "Wireless communications: principles and practice," *2nd Edition, Publishing House of Electronics Industry*, 68–71, Prentice Hall, Beijing, China, 2004.
26. Zhu, G., K. Huang, V. K. N. Lau, B. Xia, X. Li, and S. Zhang, "Hybrid beamforming via the Kronecker decomposition for the millimeter-wave massive MIMO systems," *IEEE Journal on Selected Areas in Communications*, Vol.35, No. 9, 2097–2114, Sep. 2017.
27. Maschietti, F., D. Gesbert, P. de Kerret, and H. Wymeersch, "Robust location-aided beam alignment in millimeter wave massive MIMO," *Proc. of IEEE Global Communications Conference (GLOBECOM)*, 1–6, Dec. 2017.
28. Hashemi, M., A. Sabharwal, C. E. Koksal, and N. B. Shroff, "Efficient beam alignment in millimeter wave systems using contextual bandits," *Proc. of IEEE Conference on Computer Communications (INFOCOM)*, 2393–2401, Apr. 2018.
29. Shafin, R., L. Liu, J. Zhang, and Y. C. Wu, "DoA estimation and capacity analysis for 3-D milimeter wave massive MIMO/FD-MIMO OFDM systems," *IEEE Trans. Wireless Commun.*, Vol. 15, No. 10, 6963–6978, Oct. 2016.
30. Cacciola, M., S. Calcagno, G. Megali, F. C. Morabito, D. Pellicano, and M. Versaci, "FEA design and misfit minimization for in-depth flaw characterization in metallic plates with Eddy current nondestructive testing," *IEEE Trans. Magn.*, Vol. 45, No. 3, 1506–1509, Mar. 2009.
31. Versaci, M., "Fuzzy approach and Eddy current NDT/NDE devices in industrial applications," *IEEE Electronics Letters*, Vol. 52, No. 11, 943–945, May 2016.
32. Kazerouni, A., F. J. Lopez-Martinez, and A. Goldsmith, "Increasing capacity in massive MIMO networks via small cells," *Proc. of IEEE Global Communications Conference (GLOBECOM) Workshop*, 358–363, Dec. 2014.
33. Tan, W., X. Feng, G. Liu, W. Tan, M. Zhou, and C. Li, "Spectral efficiency of massive MIMO Systems with multiple sub-arrays antenna," *IEEE Access*, Vol. 6, 31213–31223, Jun. 2018.
34. Marzetta, T. L., "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, Vol. 9, No. 11, 3590–3600, Nov. 2010.
35. Zetterberg, P., "Mobile cellular communications with base station antenna arrays: spectrum e?ciency, algorithms and propagation Models," Ph.D. thesis, Royal Institute of Technology, Sweden, available online on December, 2008.
36. Calcagno, S., F. L. Foresta, and M. Versaci, "Independent component analysis and discrete wavelet transform for artifact removal in biomedical signal processing," *American Journal of Applied Sciences*, Vol. 11, No. 1, 57–68, 2014.
37. Yang, X. and A. O. Fapojuwo, "Performance analysis of hexagonal cellular networks in fading channels," *Wireless Commun. Mob. Computing*, Vol. 2016, No. 16, 850–867, Jan. 2015.
38. Chen, Y., X. Wen, and Z. Lu, "Achievable spectral efficiency of Hybrid beamforming massive MIMO systems with quantized phase shifters, channel non-reciprocity and estimation errors," *IEEE Access*, Vol. 8, 71304–71317, Apr. 2020.