IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 26, 2016, accepted October 20, 2016, date of publication November 7, 2016,
date of current version November 28, 2016.

Digital Object Identifier 10.1109/ACCESS.2016.2625306

Location-Aware Channel Estimation Enhanced TDD Based Massive MIMO

ZHAOCHENG WANG[1], (Senior Member, IEEE), PEIYAO ZHAO[1], (Student Member, IEEE), CHEN QIAN[1], AND SHENG CHEN[2,3] (Fellow, IEEE)

[1]Tsinghua National Laboratory for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing 100084, China
[2]Electronics and Computer Science Department, University of Southampton, Southampton, SO17 1BJ, U.K.
[3]King Abdulaziz University, Jeddah 21589, Saudi Arabia

Corresponding author: P. Zhao (zhaopy10@gmail.com)

This work was supported in part by the National High Technology Research and Development Program of China under Grant 2014AA01A704, in part by the National Natural Science Foundation of China under Grant 61571267, in part by the Beijing Natural Science Foundation under Grant 4142027, in part by the Shenzhen Peacock Plan under Grant 1108170036003286, and in part by the Shenzhen Subject Arrangements under Grant JCYJ20160331184124954. This work was presented at the ICC 2015, Workshop 5G and Beyond-Enabling Technologies and Application.

ABSTRACT Pilot contamination (PC) is a stumbling block in the way of realizing massive multi-input multi-output (MIMO) systems. This contribution proposes a location-aware channel estimation-enhanced massive MIMO system employing time-division duplexing protocol, which is capable of significantly reducing the inter-cell interference caused by PC and, therefore, improving the achievable system performance. Specifically, we present a novel location-aware channel estimation algorithm, which utilizes the property of the steering vector to carry out a fast Fourier transform-based post-processing after the conventional pilot-aided channel estimation for mitigating PC. Our asymptotic analysis proves that this post-processing is capable of removing PC from the interfering users with different angle-of-arrivals (AOAs). Since in practice the AOAs of some users may be similar, we further present a location-aware pilot assignment method to ensure that users utilizing the same pilot have distinguishable AOAs, in order to fully benefit from the location-aware channel estimation. Simulation results demonstrate that the proposed scheme can dramatically reduce the inter-cell interference caused by the re-use of the pilot sequence and improve the overall system performance significantly, while only imposing a modest extra computational cost, in comparison with the conventional pilot-aided channel estimation.

INDEX TERMS Massive multi-input multi-output, time-division duplexing, pilot contamination, inter-cell interference, location-aware channel estimation, pilot assignment.

I. INTRODUCTION

Massive multiple-input multiple-output (MIMO), also known as large-scale MIMO, has gained lots of attentions both from academia and industry, due to its potential to significantly improve the spectral efficiency as well as the energy efficiency, as reported in [1]–[4]. More specifically, based on the property of asymptotic orthogonality of the channel propagation vector, theoretical analysis [5] shows that compared with the conventional MIMO system, massive MIMO can greatly enhance the achievable system's performance by just using simple linear signal processing for uplink (UL) reception and downlink (DL) transmission. Therefore, the performance of massive MIMO systems is critically dependent on the accuracy of channel state information (CSI) estimation.

In time-division duplexing (TDD) based systems, the CSI can be estimated at the base station (BS) with the aid of UL training, which can then be applied for DL transmission by exploiting the reciprocity between the UL and the DL channels. However, since the number of orthogonal pilot sequences is limited by the length of the channel coherence time, non-orthogonal pilot sequences have to be applied among different cells. Consequently, the channel estimation at a BS will be contaminated by the pilots sent from other cells, which is known as the pilot contamination (PC). As demonstrated by Marzetta [6], the inter-cell interference caused by the PC effect severely degrades the achievable performance of TDD based massive MIMO systems.

The current solutions for combating PC can be divided into two categories. The first category of solutions focus on pilot

design and pilot allocation. For example, in the work [7], a pilot design method for TDD based massive MIMO is proposed. Specifically, by maximizing the signal-to-interference ratio (SIR), a pilot design criterion is introduced in [7]. This type of solutions cannot completely eliminate the inter-cell interference caused by PC, since the re-use of the same pilot group is still required.

The second category of solutions include the 'smart' signal processing algorithms. As pointed out in [8], PC may be regarded as a shortage of linear processing algorithms and a sub-space based blind channel estimation algorithm is proposed to reduce the inter-cell interference. In the study [9], a Bayesian estimator is applied to TDD based massive MIMO whereby the second-order statistics of the channel coefficients are required at BS. To further mitigate the inter-cell interference, a pilot assignment is also proposed in [9] which requires the cooperation between BSs. In the work [10], the PC avoidance precoding is proposed which assumes that the large-scale fading coefficients of the users in different cells are known and can be utilized to eliminate the inter-cell interference. Thus, cooperation between BSs is also necessary for the scheme of [10]. In the study [11], the re-use of pilot sequences in single cell is investigated and it is demonstrated that the minimum mean square error (MMSE) detection can reduce the inter-cell interference when the angle-of-arrivals (AOAs) of the users with re-used pilots are non-overlapping. However, the channel covariance matrix is still required for this scheme. Although these smart non-linear signal processing algorithms can eliminate or significantly reduce the inter-cell interference caused by PC, their complexity is usually very high and they all require the knowledge of the second-order statistics, i.e., channel covariance matrices, of all the UL channels. However, the acquisition of such a large amount of second-order statistics at the BSs is extremely time-consuming and, moreover, sharing them among the BSs requires a huge amount of back-haul transmissions.

It is also worth pointing out that two PC eliminating schemes based on 'smart' processing have been proposed [12], [13], which do not require any knowledge of the second-order channel statistics. The scheme of [12] consists of an amalgam of DL and UL training phases, more specifically, a total of $(L+3)$ training phases for an L-cell system. Similarly, the scheme of [13] requires a total of $(L+1)$ UL training phases. Specifically, it consists of a conventional simultaneous UL training phase in which all the MSs simultaneously transmit UL pilots to their BSs, followed by the L consecutive pilot transmission phases in which each cell stays idle at one phase and repeatedly transmits pilot sequences in other phases. Although these two schemes are capable of eliminating PC, they require excessive long channel coherence time which is unlikely to be met in practice.

Recent study and experiments [9], [14], [15] have shown that the channel angle spread observed at the BS is small, since the BS is usually elevated high with few scatterings around. As a consequence, the users with sufficiently large line-of-sight AOA differences tends to have non-overlapping AOAs, where the line-of-sight AOA can be acquired from the location information. Based on the above observation, we propose a novel location-aware channel estimation enhanced TDD based massive MIMO system, which does not require the knowledge of second-order channel statistics and does not impose long training duration. Specifically, to reduce the inter-cell interference caused by the re-use of pilots, an efficient location-aware channel estimation algorithm is derived which utilizes the property of the steering vector to carry out a fast Fourier transform (FFT) based post-processing after the conventional pilot-aided channel estimation. The proposed algorithm can effectively distinguish the users with different AOAs and, therefore, can be applied for efficient inter-cell interference cancellation. Moreover, asymptotically, we prove that the inter-cell interference caused by the users with non-overlapping AOAs can be perfectly eliminated when the number of BS antennas tends to infinity. Furthermore, to improve the feasibility of our proposed scheme, we propose to divide each hexagonal cell into several sectors. For the users inside the same sector, the same post-processing parameters are adopted when performing the location-aware channel estimation algorithm. Thus only the sector information of users rather than the accurate location information is required at the BS, leading to much reduced overhead for locationing. In addition, we also propose a location-aware pilot assignment scheme to maximally benefit from the location-aware channel estimation algorithm, where the inter-cell interference in the system is minimized by jointly considering the effect of AOA differences and distance.

In summary, our solution differs from the previous works of [11] and [13]–[15] in that our algorithm aims at separating different users by exploiting the degrees of freedom in spatial domain. In particular, our solution does not require the knowledge of second-order channel statistics and, consequently, costly acquisition of a large amount of second-order statistics and distributing them among the BSs are avoided. Moreover, our scheme does not impose a long training session. In fact, it requires the exact training duration as the conventional pilot-aided channel estimation. Therefore, our proposed scheme is more suitable for practical use. Simulation results show that our scheme can reduce the inter-cell interference significantly, and the overall system performance is enhanced dramatically with only a slightly increase in complexity, compared with the conventional pilot-aided channel estimation design.

The rest of this paper is organized as follows. The system model is described in Section II. Section III is devoted to the development of our proposed location-aware channel estimation algorithm, while Section IV introduces the location-aware pilot assignment for ensuring that the users with same pilot have non-overlapping AOAs or are far apart. The simulation results are included in Section V to demonstrate the effectiveness of our proposed scheme for enhancing the performance of massive MIMO systems, and our conclusions are drawn in Section VI.

Throughout our discussions, the following notations are adopted. Time-domain and frequency-domain scalars are denoted by normal-face lower-case and upper-case letters, respectively. Boldface lower-case and upper-case symbols denote time-domain column vectors and matrices, respectively, but frequency-domain vectors are also represented by boldface upper-case symbols. The transpose, conjugate, conjugate transpose and inverse operators are denoted by $(\cdot)^{\mathrm{T}}$, $(\cdot)^{*}$, $(\cdot)^{\mathrm{H}}$ and $(\cdot)^{-1}$, respectively. The imaginary axis is denoted by $\mathrm{j} = \sqrt{-1}$, while $|\cdot|$ is the magnitude operator and $\|\cdot\|$ is the norm operator. $\boldsymbol{I}_K$ and $\boldsymbol{0}_K$ represent the $K \times K$ identity matrix and zero matrix, respectively, while $\mathrm{tr}(\cdot)$ denotes the matrix trace operation and $E\{\cdot\}$ stands for the expectation operator. Additionally, $\lfloor\cdot\rceil$ denotes the integer rounding operation.

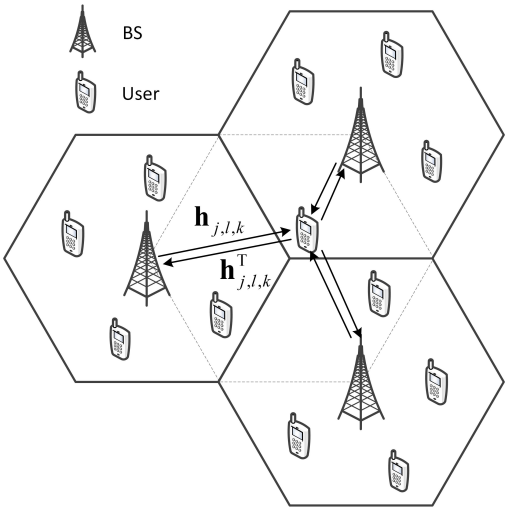


FIGURE 1. An illustrative example of multi-cell multi-user TDD massive MIMO system comprising $L = 3$ cells each with $K = 4$ single antenna users, where the uplink channel and downlink channel are reciprocal.

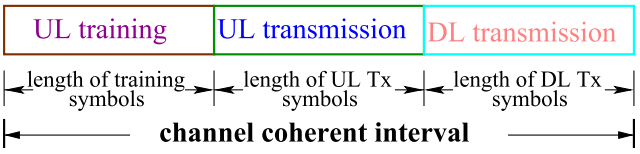


FIGURE 2. TDD protocol frame structure and the channel coherent interval.

II. SYSTEM MODEL

Consider a homogeneous multi-cell multi-user massive MIMO system employing TDD protocol with L hexagonal cells as depicted in Fig. 1. In each cell, K single-antenna users are served at the same time/frequency resource. The BS within each cell is equipped with M antennas, where $M \gg K$. We assume that the frequency reuse factor is 1 and the same frequency band is used by all cells. The data transmission is divided into three stages, as illustrated in Fig. 2.

A. UL TRAINING

In the first stage, the users from all cells transmit pilot signals to their corresponding BSs. Denote the complex pilot sequence for the k-th user of a cell as $\boldsymbol{\phi}_k = \left[\phi_{k,1}\ \phi_{k,2} \cdots \phi_{k,\tau}\right]^{\mathrm{T}} \in \mathbb{C}^{\tau\times 1}$ with length τ and the normalized power $\boldsymbol{\phi}_k^{\mathrm{H}}\boldsymbol{\phi}_k = 1$, and let $\boldsymbol{\Phi} = \left[\boldsymbol{\phi}_1\ \boldsymbol{\phi}_2 \cdots \boldsymbol{\phi}_K\right]^{\mathrm{T}} \in \mathbb{C}^{K\times\tau}$ be the pilot matrix for all the K users of this cell. The pilot sequences used by the users in the same cell are assumed to be orthogonal, i.e., $\boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{H}} = \boldsymbol{I}_K$. Clearly, to obtain this desired orthogonal property, the length of pilot sequences must satisfy $\tau \geq K$. In practical systems, the pilot sequences are generally selected from predefined finite alphabet, such as the Zadoff-Chu (ZC) sequences [16] in Long Term Evolution (LTE).

On the other hand, the length of the pilot sequences is limited by the channel coherence interval (CHI). Specifically, let N_{UL} and N_{DL} be the numbers of UL and DL data symbols transmitted in one frame, respectively. Then the UL training duration or τ must meet the following condition

$$\tau \leq \mathrm{CHI} - N_{\mathrm{UL}} - N_{\mathrm{DL}}. \tag{1}$$

The shortest CHI that the system can cope with while maintaining the orthogonal pilot set is $\mathrm{CHI} = N_{\mathrm{UL}} + N_{\mathrm{DL}} + \tau$, i.e., the shortest pilot sequence length is $\tau = K$. This can only yield the K orthogonal pilots. Hence this same set of pilot sequences has to be re-used for every cell.

In the first stage, therefore, the received signal of the BS in the l-th cell is given by

$$\boldsymbol{Y}_l = \sum_{j=1}^{L} \boldsymbol{H}_{j,l}\boldsymbol{\Phi} + \boldsymbol{N}_l, \tag{2}$$

where $\boldsymbol{Y}_l \in \mathbb{C}^{M\times\tau}$ and $\boldsymbol{N}_l \in \mathbb{C}^{M\times\tau}$ are the received signal matrix and the UL channel additive white Gaussian noise (AWGN) matrix, respectively, while $\boldsymbol{H}_{j,l} = \left[\boldsymbol{h}_{j,l,1}\ \boldsymbol{h}_{j,l,2} \cdots \boldsymbol{h}_{j,l,K}\right] \in \mathbb{C}^{M\times K}$ denotes the channel coefficient matrix linking the K users in the j-th cell to the M antennas of the BS in the l-th cell. Hence, the k-th column of $\boldsymbol{H}_{j,l}$ is the channel vector $\boldsymbol{h}_{j,l,k} = \left[h_{j,l,k,1}\ h_{j,l,k,2} \cdots h_{j,l,k,M}\right]^{\mathrm{T}}$ whose element $h_{j,l,k,i}$ is the channel coefficient between the k-th user in the j-th cell and the i-th antenna of the BS in the l-th cell. At the BS, the orthogonality of the pilot sequences is utilized for channel estimation. Right multiplying $\boldsymbol{Y}_l$ by $\boldsymbol{\Phi}^{\mathrm{H}}$ yields the estimate of $\boldsymbol{H}_{l,l}$ as [6]

$$\widehat{\boldsymbol{H}}_{l,l} = \boldsymbol{Y}_l\boldsymbol{\Phi}^{\mathrm{H}} = \boldsymbol{H}_{l,l} + \sum_{j\neq l} \boldsymbol{H}_{j,l} + \bar{\boldsymbol{N}}_l, \tag{3}$$

where $\bar{\boldsymbol{N}}_l = \boldsymbol{N}_l\boldsymbol{\Phi}^{\mathrm{H}}$ is the equivalent noise matrix at the channel estimate. Thus, for the k-th user in the l-th cell, its channel estimation is given as

$$\widehat{\boldsymbol{h}}_{l,l,k} = \boldsymbol{h}_{l,l,k} + \sum_{j\neq l} \boldsymbol{h}_{j,l,k} + \bar{\boldsymbol{n}}_{l,k}, \tag{4}$$

where $\bar{\boldsymbol{n}}_{l,k}$ is the k-th column of $\bar{\boldsymbol{N}}_l$.

Remark 1: Obviously, the conventional channel estimation given by (4) suffers from serious PC. But this approach imposes the shortest possible training duration of $\tau = K$, which ensures that the system can still operate for the CHI as short as $N_{\mathrm{UL}} + N_{DL} + K$. As mentioned previously, there are other much more complicated channel estimation algorithms [9], [11]–[13], capable of eliminating PC, but they all impose some unrealistic requirements. For example, the PC elimination scheme of [13] can only work for the system with the CHI no shorter than $N_{\mathrm{UL}} + N_{\mathrm{DL}} + (L+1)K$. Since we assume more realistically that the BS has no knowledge of the channel covariance matrices and we do not want to impose excessively long training session, the same shortest possible training duration of $\tau = K$ is applied for our proposed scheme. Note that given the CHI, it is vital to keep the training duration as short as possible in order to maintain the effective spectral efficiency of the massive MIMO system. Moreover, if the training duration is not shorter than the CHI, then the massive MIMO system of Fig. 2 cannot be realized.

Remark 2: The work [7] proposes a pilot design method by maximizing an approximate SIR subject to the requirement that the training sequence length τ satisfies $K \leq \tau < 2K$. Clearly, this scheme cannot eliminate completely PC, as the design cannot provide the LK orthogonal pilots and the reuse of the same pilot group is required. Moreover, when the shortest training duration of $\tau = K$ is imposed, the design of [7] becomes exactly the conventional channel estimation scheme described here, and it suffers from the same amount of PC with the same poor estimation performance.

B. UL TRANSMISSION

In the second stage, the users transmit data to their corresponding BSs. Let the symbol vector transmitted by the K users in the j-th cell be denoted as $\boldsymbol{s}_j = \left[s_{j,1}\ s_{j,2} \cdots s_{j,K}\right]^{\mathrm{T}}$ with $E\left\{s_{j,k}^* s_{j,k}\right\} = 1$ for $1 \leq k \leq K$. Then the received signal $\boldsymbol{y}_l \in \mathbb{C}^{M\times 1}$ at the l-th BS can be written as

$$\boldsymbol{y}_l = \sum_{j=1}^{L} \boldsymbol{H}_{j,l}\boldsymbol{s}_j + \boldsymbol{n}_l^u, \tag{5}$$

where $\boldsymbol{n}_l^u \in \mathbb{C}^{M\times 1}$ is the corresponding UL channel AWGN vector. The channel estimation $\widehat{\boldsymbol{H}}_{l,l}$ obtained in the first stage is used for data detection, and the detected symbol vector for the users in the l-th cell is computed according to

$$\widehat{\boldsymbol{s}}_l = \boldsymbol{D}_l \boldsymbol{y}_l, \tag{6}$$

where $\boldsymbol{D}_l = \left[\boldsymbol{d}_{l,1}\ \boldsymbol{d}_{l,2} \cdots \boldsymbol{d}_{l,K}\right]^{\mathrm{T}} \in \mathbb{C}^{K\times M}$ is the detector coefficient matrix calculated based on $\widehat{\boldsymbol{H}}_{l,l}$. Here, the zero-forcing (ZF) detection algorithm is applied, and $\boldsymbol{D}_l = \left(\widehat{\boldsymbol{H}}_{l,l}^{\mathrm{H}}\widehat{\boldsymbol{H}}_{l,l}\right)^{-1}\widehat{\boldsymbol{H}}_{l,l}^{\mathrm{H}}$.

C. DL TRANSMISSION

In the third stage, the BSs transmit data to their corresponding users. Denote $\boldsymbol{x}_l = \left[x_{l,1}\ x_{l,2} \cdots x_{l,K}\right]^{\mathrm{T}}$ as the symbol vector transmitted by the l-th BS to its corresponding K users with $E\left\{\boldsymbol{x}_l^{\mathrm{H}}\boldsymbol{x}_l\right\} = 1$. By utilizing the channel reciprocity of TDD systems, the signals received by the K users in the l-th cell can be expressed in a vector form as

$$\boldsymbol{r}_l = \sum_{j=1}^{L} \boldsymbol{H}_{l,j}^{\mathrm{T}}\boldsymbol{W}_j\boldsymbol{x}_j + \boldsymbol{n}_l^d, \tag{7}$$

where $\boldsymbol{W}_j = \left[\boldsymbol{w}_{j,1}\ \boldsymbol{w}_{j,2} \cdots \boldsymbol{w}_{j,K}\right] \in \mathbb{C}^{M\times K}$ is the precoding coefficient matrix for the K users in the j-th cell, which is calculated based on the channel estimation obtained in the first stage, and $\boldsymbol{n}_l^d \in \mathbb{C}^{K\times 1}$ denotes the DL channel AWGN vector. As the ZF-precoding can approach the optimal performance in the asymptotic case of $M \to \infty$, we apply the ZF-precoding. Therefore, the precoding matrix is computed according to $\boldsymbol{W}_l = \beta_l \widehat{\boldsymbol{H}}_{l,l}^*\left(\widehat{\boldsymbol{H}}_{l,l}^{\mathrm{T}}\widehat{\boldsymbol{H}}_{l,l}^*\right)^{-1}\boldsymbol{\Gamma}$, in which $\boldsymbol{\Gamma} \in \mathbb{C}^{K\times K}$ is a diagonal matrix for power allocation among different users with $\mathrm{tr}\left(\boldsymbol{\Gamma}^{\mathrm{H}}\boldsymbol{\Gamma}\right) = 1$ and β_l is used for the normalization to ensure the total transmission power constraint.

D. SYSTEM ANALYSIS

The well-known asymptotic orthogonality for massive MIMO is described by [6]

$$\lim_{M\to\infty} \boldsymbol{h}_{l_1,j_1,k_1}^{\mathrm{H}}\boldsymbol{h}_{l_2,j_2,k_2} = 0, \tag{8}$$

for $l_1 \neq l_2$ or $j_1 \neq j_2$ or $k_1 \neq k_2$, which indicates that as the number of antennas at the BS tends to infinity, the channels between different users become orthogonal. With this asymptotic orthogonality, it can be verified that the intra-cell interference is eliminated. Take the DL transmission as an example. As $M \to \infty$, $\frac{1}{M}\boldsymbol{H}_{j,l}^{\mathrm{H}}\boldsymbol{H}_{j,l}$ tends to a diagonal matrix denoted as $\boldsymbol{D}_{j,l}$, while $\frac{1}{M}\boldsymbol{H}_{j_1,l_1}^{\mathrm{H}}\boldsymbol{H}_{j_2,l_2}$ tends to the zero matrix $\boldsymbol{0}_K$ for $j_1 \neq j_2$ or $l_1 \neq l_2$. Similarly, as channel vector and noise are uncorrelated, $\frac{1}{M}\boldsymbol{H}_{j,l_1}^{\mathrm{H}}\bar{\boldsymbol{N}}_{l_2}$ tends to $\boldsymbol{0}_K$ for $\forall l_1, l_2$. In addition, $\frac{1}{M}\bar{\boldsymbol{N}}_l^{\mathrm{T}}\bar{\boldsymbol{N}}_l^*$ tends to $E\left\{\bar{\boldsymbol{N}}_l^{\mathrm{T}}\bar{\boldsymbol{N}}_l^*\right\}$ which is also a diagonal matrix, as the noise is white. Since the power constraint is expressed as

$$\mathrm{tr}\left(\boldsymbol{W}_l^{\mathrm{H}}\boldsymbol{W}_l\right) = \beta_l^2 \mathrm{tr}\left(\boldsymbol{\Gamma}^{\mathrm{H}}\left(\widehat{\boldsymbol{H}}_{l,l}^{\mathrm{T}}\widehat{\boldsymbol{H}}_{l,l}^*\right)^{-1}\boldsymbol{\Gamma}\right) = P_l, \tag{9}$$

where P_l is the total transmission power of the l-th BS, we have

$$\beta_l = \sqrt{\frac{P_l}{\mathrm{tr}\left(\boldsymbol{\Gamma}^{\mathrm{H}}\left(\widehat{\boldsymbol{H}}_{l,l}^{\mathrm{T}}\widehat{\boldsymbol{H}}_{l,l}^*\right)^{-1}\boldsymbol{\Gamma}\right)}}. \tag{10}$$

Further noting $\widehat{\boldsymbol{H}}_{l,l} = \sum_{j=1}^{L}\boldsymbol{H}_{j,l} + \bar{\boldsymbol{N}}_l$, we have

$$\begin{aligned}
&\lim_{M\to\infty} \frac{1}{\sqrt{M}}\beta_l \\
&= \lim_{M\to\infty} \sqrt{\frac{P_l}{\mathrm{tr}\left(\boldsymbol{\Gamma}^{\mathrm{H}}\left(\frac{1}{M}\widehat{\boldsymbol{H}}_{l,l}^{\mathrm{T}}\widehat{\boldsymbol{H}}_{l,l}^*\right)^{-1}\boldsymbol{\Gamma}\right)}} \\
&= \sqrt{\frac{P_l}{\mathrm{tr}\left(\boldsymbol{\Gamma}^{\mathrm{H}}\left(\sum_{j=1}^{L}\boldsymbol{D}_{j,l}^* + E\left\{\bar{\boldsymbol{N}}_l^{\mathrm{T}}\bar{\boldsymbol{N}}_l^*\right\}\right)^{-1}\boldsymbol{\Gamma}\right)}} = \widetilde{\beta}_l.
\end{aligned} \tag{11}$$

As $M \to \infty$, therefore, the precoding matrix for the l-th cell can be rewritten as

$$\begin{aligned}\boldsymbol{W}_l &= \beta_l \widehat{\boldsymbol{H}}_{l,l}^* \Big(\big(\sum_{j_1=1}^{L} \boldsymbol{H}_{j_1,l}^{\mathrm{T}} + \bar{\boldsymbol{N}}_l^{\mathrm{T}} \big) \big(\sum_{j_2=1}^{L} \boldsymbol{H}_{j_2,l}^* + \bar{\boldsymbol{N}}_l^* \big) \Big)^{-1} \boldsymbol{\Gamma} \\ &= \frac{\widetilde{\beta}_l}{\sqrt{M}} \widehat{\boldsymbol{H}}_{l,l}^* \Big(\sum_{j=1}^{L} \boldsymbol{D}_{j,l}^* + E\{ \bar{\boldsymbol{N}}_l^{\mathrm{T}} \bar{\boldsymbol{N}}_l^* \} \Big)^{-1} \boldsymbol{\Gamma} \\ &= \frac{1}{\sqrt{M}} \widehat{\boldsymbol{H}}_{l,l}^* \boldsymbol{\Delta}_l, \end{aligned} \tag{12}$$

where $\boldsymbol{\Delta}_l = \widetilde{\beta}_l \big(\sum_{j=1}^{L} \boldsymbol{D}_{j,l}^* + E\{ \bar{\boldsymbol{N}}_l^{\mathrm{T}} \bar{\boldsymbol{N}}_l^* \} \big)^{-1} \boldsymbol{\Gamma}$ is a diagonal matrix. By substituting (12) into (7) we obtain

$$\begin{aligned}\boldsymbol{r}_l &= \frac{1}{\sqrt{M}} \sum_{j=1}^{L} \boldsymbol{H}_{l,j}^{\mathrm{T}} \widehat{\boldsymbol{H}}_{j,j}^* \boldsymbol{\Delta}_j \boldsymbol{x}_j + \boldsymbol{n}_l^d \\ &= \frac{1}{\sqrt{M}} \sum_{j=1}^{L} \boldsymbol{H}_{l,j}^{\mathrm{T}} \Big(\sum_{m=1}^{L} \boldsymbol{H}_{m,j}^* + \bar{\boldsymbol{N}}_j \Big) \boldsymbol{\Delta}_j \boldsymbol{x}_j + \boldsymbol{n}_l^d \\ &= \sqrt{M} \boldsymbol{D}_{l,l}^* \boldsymbol{\Delta}_l \boldsymbol{x}_l + \sum_{j \neq l} \sqrt{M} \boldsymbol{D}_{l,j}^* \boldsymbol{\Delta}_j \boldsymbol{x}_j + \boldsymbol{n}_l^d. \end{aligned} \tag{13}$$

Since $\boldsymbol{D}_{l,l}^* \boldsymbol{\Delta}_l$ is diagonal, the first term of (13) indicates that the intra-cell interference is eliminated. However, it can be observed from the second term of (13) that the inter-cell interference still exists even with an infinitely large M. The asymptotic orthogonality allows us to concentrate on inter-cell interference elimination/reduction.

According to [9], we adopt the following narrow-band multi-path channel model

$$\boldsymbol{h}_{j,l,k} = \frac{1}{\sqrt{P}} \sum_{p=1}^{P} \alpha_{j,l,k}^p \boldsymbol{a}\big(\theta_{j,l,k}^p\big) \sqrt{\beta_{j,l,k}}, \tag{14}$$

where P denotes the number of paths between the k-th user in the j-th cell and the BS of the l-th cell, $\alpha_{j,l,k}^p$ is the complex gain of the p-th path which obeys the complex Gaussian distribution with zero mean and $E\{|\alpha_{j,l,k}^p|^2\} = 1$, and $\theta_{j,l,k}^p$ is the AOA of the p-th path, while the column vector $\boldsymbol{a}(\theta) \in \mathbb{C}^{M \times 1}$ is the steering vector with the AOA θ, and $\beta_{j,l,k}$ is the distance-related path-loss between the k-th user in the j-th cell and the BS of the l-th cell given by

$$\beta_{j,l,k} = \frac{c}{d_{j,l,k}^{\gamma}}, \tag{15}$$

in which c is a constant dependent on the average signal-to-noise ratio (SNR) at the cell edge, $d_{j,l,k}$ is the geographical distance, and γ is the path-loss exponent. We assume that $\theta \in [0, \pi]$. For the uniformly-spaced linear array (ULA), for example, the steering vector $\boldsymbol{a}(\theta)$ is given by [14]

$$\boldsymbol{a}(\theta) = \Big[1 \; e^{-\mathrm{j}2\pi \frac{D}{\lambda} \cos(\theta)} \cdots e^{-\mathrm{j}2\pi \frac{(M-1)D}{\lambda} \cos(\theta)} \Big]^{\mathrm{T}}, \tag{16}$$

where λ is the wavelength of the received signal, and $D \leq \lambda/2$ denotes the antenna spacing at BS. It should be pointed out that here we use the ULA as an example but our proposed scheme is not restricted to the ULA. Note that according to the central-limit theorem, since $\alpha_{j,l,k}^p$ is a random number, if P is sufficiently large, each element of $\boldsymbol{h}_{j,l,k}$ follows a complex Gaussian distribution with zero mean, which coincides with the standard assumption regarding the flat-fading MIMO channel matrix. Thus the asymptotic orthogonality still holds for the channel model (14). It is also worth mentioning that the narrow-band model (14) can be easily extended to wide-band systems by using orthogonal frequency division multiplexing (OFDM) technology. Thus our proposed scheme can be extended to wideband massive MIMO systems by adopting OFDM.

Recent study and experiments [9], [14], [15] have shown that the channel angle spread observed at the BS is small. Thus we assume that the AOAs of the multi-path components fall in a small angle interval, which can be expressed as $\theta_{j,l,k}^p \in \big[\theta_{\min}^{(j,l,k)}, \; \theta_{\max}^{(j,l,k)}\big]$ for $1 \leq p \leq P$, where $\theta_{\min}^{(j,l,k)} = \theta_{\mathrm{LOS}}^{(j,l,k)} - \delta_\theta$ and $\theta_{\max}^{(j,l,k)} = \theta_{\mathrm{LOS}}^{(j,l,k)} + \delta_\theta$, in which $\theta_{\mathrm{LOS}}^{(j,l,k)}$ is the line-of-sight AOA and δ_θ is the angle spread. As pointed out in [15], this assumption holds when the BS is much higher than the surrounded structures with few scatters around. This scenario corresponds to macro-cell or micro-cell environments. The measurement results in [17] also demonstrate that for rural and sub-urban environment, and even often in urban environment, the angle spread is small.

III. LOCATION-AWARE CHANNEL ESTIMATION

The basic idea is as follows. By introducing a post-processing after the conventional pilot-aided channel estimation (3), which utilizes the properties of the steering vector to distinguish the users with different AOAs, the interference from the users in other cells with the same pilot sequence but having different AOAs can be significantly reduced. We now detail how this is achieved.

A. EXPLOITING PROPERTY OF THE STEERING VECTOR

For the purpose of clearly illustrating the main concept and without loss of generality, consider the steering vector $\boldsymbol{a}(\theta)$ given in (16). We observe that $\boldsymbol{a}(\theta)$ can be regarded as a single-frequency signal with frequency $f_{\boldsymbol{a}} = \frac{D}{\lambda} \cos(\theta)$. Thus, as the number of the antennas at the BS tends to infinity, the Fourier transform of $\boldsymbol{a}(\theta)$ tends to a δ-function. The basic idea is to utilize this property to design a filter to distinguish the users with different AOAs.

The N-points discrete Fourier transform (DFT) of the steering vector in (16) is defined by

$$\begin{aligned} A(n) &= \sum_{m=0}^{M-1} a(m) e^{-\mathrm{j}\frac{2\pi}{N} nm} \\ &= \frac{1 - e^{-\mathrm{j}2\pi M \left(\frac{D}{\lambda} \cos(\theta) + \frac{n}{N} \right)}}{1 - e^{-\mathrm{j}2\pi \left(\frac{D}{\lambda} \cos(\theta) + \frac{n}{N} \right)}}, \quad 0 \leq n \leq N-1, \end{aligned} \tag{17}$$

in which $N \geq M$ and $a(m) = e^{-\mathrm{j}2\pi \frac{mD}{\lambda} \cos(\theta)}$ is the m-th element of $\boldsymbol{a}(\theta)$, where $0 \leq m \leq M-1$.

TABLE 1. The FFT-aided location-aware channel estimation

Parameters	the thresholds of the AOAs of the users k in the l-th cell, $\theta_{\min}^{(l,k)}$ and $\theta_{\max}^{(l,k)}$, and the size of FFT N
Inputs	the received signal $\boldsymbol{Y}_l$, and the pilot sequence $\boldsymbol{\Phi}$
Step 1	Conventional pilot-aided channel estimation:
	multiply $\boldsymbol{Y}_l$ by $\boldsymbol{\Phi}^{\mathrm{H}}$ to obtain $\widehat{\boldsymbol{h}}_{l,l,k}$ of (4)
Step 2	FFT-aided post-processing:
2.1	perform the N-point FFT on $\widehat{\boldsymbol{h}}_{l,l,k}$ to obtain $\mathbf{F}_{l,l,k}$
2.2	compute $n_{\min}^{(l,k)}$ and $n_{\max}^{(l,k)}$ of (23)
2.3	compute $\widehat{\mathbf{F}}_{l,l,k}$ by (25)
Step 3	Output the estimation:
3.1	perform the N-point IFFT on $\widehat{\mathbf{F}}_{l,l,k}$ to obtain $\widehat{\boldsymbol{f}}_{l,l,k}$
3.2	obtain the estimation result $\bar{\boldsymbol{h}}_{l,l,k}$ of (26)

Noting $1-x^M=(1-x)(1+x+x^2+\cdots+x^{M-1})$ and letting $q_n=\frac{D}{\lambda}\cos(\theta)+\frac{n}{N}$, we can simplify (17) as

$$A(n)=\frac{1-e^{-\mathrm{j}2\pi q_n M}}{1-e^{-\mathrm{j}2\pi q_n}}=1+e^{-\mathrm{j}2\pi q_n}+e^{-\mathrm{j}2\pi 2q_n}+\cdots+e^{-\mathrm{j}2\pi(M-1)q_n}. \quad (18)$$

Considering that each term in (18) has unit norm, the following expression holds

$$|A(n)|=\left|\sum_{i=0}^{M-1}e^{-\mathrm{j}2\pi iq_n}\right|\le M,\quad 0\le n\le N-1. \quad (19)$$

If and only if $e^{-\mathrm{j}2\pi iq_n}=1$ for every $0\le i\le M-1$ in (19), the magnitude of $A(n)$ attains its maximal value M. This means that when $q_n=\frac{D}{\lambda}\cos(\theta)+\frac{n}{N}\in\mathbb{Z}$, i.e., q_n is an integer, the maximal value M is achieved. By letting

$$n_{lim}=\arg\max_{0\le n\le N-1}|A(n)|, \quad (20)$$

we have

$$n_{lim}=\lfloor g_N(\theta)\rceil, \quad (21)$$

where the function $g_N(\theta)$ is given by

$$g_N(\theta)=\begin{cases}N-N\dfrac{D}{\lambda}\cos(\theta), & \theta\in[0,\ \pi/2),\\ -N\dfrac{D}{\lambda}\cos(\theta), & \theta\in[\pi/2,\ \pi].\end{cases} \quad (22)$$

If $N\frac{D}{\lambda}\cos(\theta)\notin\mathbb{Z}$, the maximal value M cannot be reached, but $\left|A\left(n_{lim}\right)\right|$ is still close to M since each term $e^{-\mathrm{j}2\pi iq_{n_{lim}}}$ in (19) is close to 1. We will show later in an asymptotic analysis that as M tends to infinity with $N=M$, the peak $\left|A\left(n_{lim}\right)\right|$ will account for most of the signal power and the power that leakages to other $|A(n)|$ is negligible.

B. PROPOSED CHANNEL ESTIMATION ALGORITHM

By utilizing the property shown in Section III-A, we can design a filter to distinguish the users with different AOAs. Denote the N-point FFT of the estimated channel coefficient vector $\widehat{\boldsymbol{h}}_{l,l,k}$ as $\mathbf{F}_{l,l,k}$. Under the assumption that the AOA of the k-th user in the l-th cell to its corresponding BS is limited within the interval $\left[\theta_{\min}^{(l,k)},\ \theta_{\max}^{(l,k)}\right]$, most energy of $\mathbf{F}_{l,l,k}$ falls in the interval $\mathbb{I}\left(n_{\min}^{(l,k)},n_{\max}^{(l,k)}\right)$ according to the above mentioned property. The operator $\mathbb{I}\left(n_{\min}^{(l,k)},n_{\max}^{(l,k)}\right)$ specifies an interval with the two boundary points $n_{\min}^{(l,k)}$ and $n_{\max}^{(l,k)}$ that are computed according to

$$\begin{cases}n_{\min}^{(l,k)}=\lfloor g_N\left(\theta_{\min}^{(l,k)}\right)\rceil,\\ n_{\max}^{(l,k)}=\lfloor g_N\left(\theta_{\max}^{(l,k)}\right)\rceil.\end{cases} \quad (23)$$

Here we have draped the index j from $\theta_{\min}^{(j,l,k)}$ and $\theta_{\max}^{(j,l,k)}$, as $j=l$. The interval $\mathbb{I}\left[\theta_{\min}^{(l,k)},\ \theta_{\max}^{(l,k)}\right]$ satisfies the condition that for $\forall\theta\in\left[\theta_{\min}^{(l,k)},\ \theta_{\max}^{(l,k)}\right]$, $\lfloor g_N(\theta)\rceil\in\mathbb{I}\left(n_{\min}^{(l,k)},n_{\max}^{(l,k)}\right)$. From this condition, we can obtain that if $0\le\theta_{\min}^{(l,k)}<\theta_{\max}^{(l,k)}<\pi/2$ or $\pi/2\le\theta_{\min}^{(l,k)}<\theta_{\max}^{(l,k)}\le\pi$, $\mathbb{I}\left(n_{\min}^{(l,k)},n_{\max}^{(l,k)}\right)=\left[n_{\min}^{(l,k)},\ n_{\max}^{(l,k)}\right]$, while if $0\le\theta_{\min}^{(l,k)}<\pi/2<\theta_{\max}^{(l,k)}\le\pi$, $\mathbb{I}\left(n_{\min}^{(l,k)},n_{\max}^{(l,k)}\right)=\left[0,\ n_{\max}^{(l,k)}\right]\cup\left[n_{\min}^{(l,k)},\ N\right]$. This is in fact equivalent to

$$\mathbb{I}\left(n_{\min}^{(l,k)},n_{\max}^{(l,k)}\right)=\begin{cases}\left[n_{\min}^{(l,k)},n_{\max}^{(l,k)}\right] & \text{for } n_{\min}^{(l,k)}\le n_{\max}^{(l,k)},\\ \left[0,\ n_{\max}^{(l,k)}\right]\cup\left[n_{\min}^{(l,k)},\ N\right] & \text{for } n_{\min}^{(l,k)}>n_{\max}^{(l,k)}.\end{cases} \quad (24)$$

Thus we can force the values of those elements of $\mathbf{F}_{l,l,k}$ outside the interval $\mathbb{I}\left(n_{\min}^{(l,k)},n_{\max}^{(l,k)}\right)$ to be zero to cancel out the signals with the AOAs outside the interval $\left[\theta_{\min}^{(l,k)},\ \theta_{\max}^{(l,k)}\right]$. In addition, the effect of the noise is also reduced by this spatial filtering.

After the FFT-based filtering, we have

$$\widehat{F}_{l,l,k}(n)=\begin{cases}F_{l,l,k}(n), & n\in\mathbb{I}\left(n_{\min}^{(l,k)},n_{\max}^{(l,k)}\right),\\ 0, & n\notin\mathbb{I}\left(n_{\min}^{(l,k)},n_{\max}^{(l,k)}\right),\end{cases} \quad (25)$$

where $F_{l,l,k}(n)$ denotes the n-th element of $\mathbf{F}_{l,l,k}$. Then perform the inverse FFT (IFFT) on $\widehat{\mathbf{F}}_{l,l,k}=\left[\widehat{F}_{l,l,k}(0)\,\widehat{F}_{l,l,k}(1)\cdots\widehat{F}_{l,l,k}(N-1)\right]^{\mathrm{T}}$ and denote the result as $\widehat{\boldsymbol{f}}_{l,l,k}=\left[\widehat{f}_{l,l,k}(0)\,\widehat{f}_{l,l,k}(1)\cdots\widehat{f}_{l,l,k}(N-1)\right]^{\mathrm{T}}$. The estimation of the channel coefficient vector is thus obtained as

$$\bar{\boldsymbol{h}}_{l,l,k}=\left[\widehat{f}_{l,l,k}(0)\,\widehat{f}_{l,l,k}(1)\cdots\widehat{f}_{l,l,k}(M-1)\right]^{\mathrm{T}}. \quad (26)$$

Noted that (25) is actually the operation that places a one-dimensional rectangular window on $\mathbf{F}_{l,l,k}$ to perform a filtering in spatial domain. Different selections of windows may lead to different performance of channel estimation. This FFT-aided location-aware channel estimation is summarized in Table 1.

Remark 3: In the above analysis, we have assumed a ULA whose steering vector is defined in (16). However, with some

appropriate modifications on the FFT-aided post-processing, the proposed algorithm can be easily extended to other array geometries. For example, for the BS equipped with a two-dimensional uniformly-spaced rectangle array (URA), the two-dimensional FFT should be applied instead of the one-dimensional FFT applied in the ULA case. Moreover, the position of the window is determined by not only the AOA in azimuth dimension, but also the AOA in elevation dimension. The elements of the two-dimensional frequency-domain transformed channel estimate that are outside the azimuth and elevation window can be forced to be zero to cancel out the signals with the azimuth and elevation AOAs outside the window, where the azimuth and elevation window is calculated based on the azimuth and elevation AOAs in a similar way as given in (23) and (24). In fact, in the generic three-dimensional MIMO [18]–[20], we have more degrees of freedom in spatial domain to exploit, and the condition that the users with the same pilot sequence have non-overlapping AOAs is more likely to be met.

The line-of-sight AOA of the k-th user in the l-th cell $\theta_{\text{LOS}}^{(l,k)}$ can be obtained from the location information and a preset small angle spread δ_θ can then be applied to calculate the interval $\left[\theta_{\min}^{(l,k)},\ \theta_{\max}^{(l,k)}\right]$ with $\theta_{\min}^{(l,k)} = \theta_{\text{LOS}}^{(l,k)} - \delta_\theta$ and $\theta_{\max}^{(l,k)} = \theta_{\text{LOS}}^{(l,k)} + \delta_\theta$. To eliminate the need for continuously tracking users' locations, we propose to divide each hexagonal cell into several sectors and to apply a constant interval $\left[\theta_{\min}^{(l,k)},\ \theta_{\max}^{(l,k)}\right]$ for all the users in the same sector. Specifically, for a user located in the s-th sector whose range covers the angle interval $[\Theta_{\min}^{(s)},\ \Theta_{\max}^{(s)}]$, we have $\theta_{\min}^{(l,k)} = \Theta_{\min}^{(s)} - \delta_\theta$ and $\theta_{\max}^{(l,k)} = \Theta_{\max}^{(s)} + \delta_\theta$. Consequently, only the sector information rather than the accurate location information is required, which can be easily estimated by existing positioning techniques. In fact, in most of the current mobile networks, each cell is naturally divided into sectors, and the sector information is already available at the BS. Since a user may take at least several or several tens of seconds to cross a sector, mobile scenarios are supported. Furthermore, we propose a location-aware pilot assignment scheme to maximize the performance of the location-aware channel estimation algorithm by guaranteeing non-overlapping AOAs for the users with the same pilot, as will be shown in Section IV.

Remark 4: The conventional channel estimation scheme (3) enjoys a well-known low computational complexity, which is on the order of $\mathsf{O}(\tau M)$, where $\tau = K$. Our proposed location-aware channel estimation algorithm introduces a FFT-based filtering operation, which imposes a further computational complexity on the order of $\mathsf{O}\left(\log_2(N)N\right)$. Hence, the total complexity of our location-aware channel estimation algorithm is $\mathsf{O}\left(\tau M + \log_2(N)N\right)$. As will be shown in the next section, the FFT size N is on the same order of M. Therefore, our proposed location-aware channel estimation algorithm requires approximately twice complexity of the conventional channel estimation scheme. Like the conventional channel estimation scheme, our proposed scheme can operate under the fastest changing environment that the system of Fig. 2 can cope with, as it also only requires the shortest training duration of K.

The practical advantage of our algorithm becomes apparent by examining the requirements of the existing "smart" algorithms. For the "smart" schemes which require the knowledge of the channel covariance matrices [9]–[11], a total of LK $M \times M$ channel covariance matrices must be acquired, and a huge amount of backhaul transmissions must be spent to distribute these channel covariance matrices. This consumes huge amount of computational complexity and takes a long period of time to complete. Since it is impractical to periodically do this acquisition, effectively, these schemes may only operate under a static network environment. For the PC elimination scheme of [13], signal cancellation also imposes further additional complexity on the order of $\mathsf{O}(KM)$, and this signal cancellation amplifies the noise significantly. Moreover, this scheme requires a training duration of $(L+1)K$ and, therefore, it cannot be implemented for the system with the CHI shorter than $(L+1)K + N_{\text{UL}} + N_{\text{DL}}$.

C. ASYMPTOTIC ANALYSIS

The asymptotic performance of our proposed location-aware channel estimation algorithm when the number of BS antennas tends to infinity is given in the following proposition.

Proposition 1: For sufficiently large N, the ratio of the power $P\left(A\left(n_{lim}\right)\right)$ to the total signal power $\sum_{n=0}^{N-1} P(A(n))$ can be expressed by

$$\frac{P\left(A\left(n_{lim}\right)\right)}{\sum_{n=0}^{N-1} P(A(n))} \approx \frac{M^2}{MN} = \frac{M}{N}. \tag{27}$$

Proof: Since the Fourier transform is unitary, we have

$$\sum_{m=0}^{M-1} |a(m)|^2 = \frac{1}{N} \sum_{n=0}^{N-1} |A(n)|^2, \tag{28}$$

where $N \geq M$. Because every $|a(m)| = 1$, $\sum_{m=0}^{M-1} |a(m)|^2 = M$, and we have $\frac{1}{N}\sum_{n=0}^{N-1} |A(n)|^2 = M$. Note that the instantaneous power of $A(n)$ is given by

$$\begin{aligned} P(A(n)) &= A^*(n)A(n) = \sum_{m=0}^{M-1} e^{\mathrm{j}2\pi m q_n} \sum_{i=0}^{M-1} e^{-\mathrm{j}2\pi i q_n} \\ &= \sum_{m=0}^{M-1}\sum_{i=0}^{M-1} e^{\mathrm{j}2\pi (m-i) q_n} \\ &= M + \sum_{m=1}^{M-1} (M-m) e^{\mathrm{j}2\pi m q_n} + \sum_{m=1}^{M-1} (M-m) e^{-\mathrm{j}2\pi m q_n} \\ &= M + 2\sum_{m=1}^{M-1} (M-m)\cos\left(2\pi m q_n\right), \end{aligned} \tag{29}$$

where the last equation comes from Euler equation. Denote $n_{lim} = Z\left(n_{lim}\right) + R\left(n_{lim}\right)$, where $Z\left(n_{lim}\right) = g_N(\theta)$, and $R\left(n_{lim}\right) = n_{lim} - g_N(\theta)$ represents the rounding error of n_{lim}.

In this way, we can express $q_{n_{lim}}$ as

$$q_{n_{lim}} = \frac{D}{\lambda}\cos(\theta) + \frac{Z(n_{lim}) + R(n_{lim})}{N} = b + \frac{1}{N}R(n_{lim}) = b + \varepsilon(n_{lim}), \quad (30)$$

where $b = \frac{D}{\lambda}\cos(\theta) + \frac{Z(n_{lim})}{N} = \frac{D}{\lambda}\cos(\theta) + \frac{g_N(\theta)}{N} \in \{0, 1\}$, and $\varepsilon(n_{lim}) = \frac{1}{N}R(n_{lim})$ is the rounding error of $q_{n_{lim}}$. Although $R(n_{lim})$ can be as large as 0.5, $\varepsilon(n_{lim})$ is very close to 0 if N is sufficiently large. This means that

$$\lim_{N\to\infty} q_{n_{lim}} = b. \quad (31)$$

Thus as N tends to infinity, we have

$$\lim_{N\to\infty} P(A(n_{lim})) = M + 2\sum_{m=1}^{M-1}(M-m)\cos(2\pi mb) = M + 2\sum_{m=1}^{M-1}(M-m) = M^2. \quad (32)$$

Consequently, for sufficiently large N, the ratio of the power $P(A(n_{lim}))$ to the total signal power can be expressed by

$$\frac{P(A(n_{lim}))}{\sum_{n=0}^{N-1} P(A(n))} \approx \frac{M^2}{MN} = \frac{M}{N}. \quad (33)$$

The approximation becomes exact, as $N \to \infty$. ■

Proposition 1 indicates that for a fixed and sufficiently large N, increasing M ensures that most of the signal power is accounted by the maximal value $A(n_{lim})$. Furthermore, as $N \geq M$, it can be concluded that by choosing $M = N$, the power ratio (27) reaches the maximum value of 1 asymptotically, which means that all the signal power is allocated on $A(n_{lim})$ and there exists no leakage power. This analysis also shows that with large M, by using the proposed channel estimation algorithm, the users using the same pilot sequence in different cells can be distinguished if they have non-overlapping AOAs.

Thus if M tends to infinity with $N = M$, our FFT-based post-processing can distinguish the users with different AOAs without any loss of energy and the interference caused by the users with different AOAs can be eliminated perfectly. However, for finite M and N, according to (30), increasing N can reduce the rounding error of $q_{n_{lim}}$ and consequently increases the resolution for different AOAs. Therefore, choosing $N > M$ is still more appropriate in practice.

IV. LOCATION-AWARE PILOT ASSIGNMENT

The effectiveness of the location-aware channel estimation method relies on the non-overlapping AOAs of different users with the same pilot. To maximize the performance of this algorithm, a location-aware pilot assignment scheme is introduced to make sure that the users in different cells with the same pilot sequence have different AOAs at the target BS. More specifically, each hexagonal cell is divided into S sectors. Denote $\Theta_{\max}^{(s)}$ and $\Theta_{\min}^{(s)}$ as the maximal and minimal AOAs of the users in the s-th sector, respectively. That is, the angle interval of the s-th sector is $[\Theta_{\min}^{(s)}, \Theta_{\max}^{(s)}]$. For convenience and without loss of generality, we set $\Delta_{\Theta}^{(s)} = \Theta_{\max}^{(s)} - \Theta_{\min}^{(s)} = \Delta_{\Theta} = \frac{2\pi}{S}$, i.e., all the sectors have the same angular range. Each sector is assigned with one pilot sequence and all the users within one sector use the same pilot. The pilot sequences assigned to different sectors in the same cell are orthogonal. To serve multiple users at the same time/frequency resource, the BS selects one user from each sector. As pointed out previously, with the aid of GPS, WiFi or other existing positioning technologies, the user's location information is much easier to obtain, compared to the channel statistical information, such as the channel covariance matrix. In fact, we only need to know which sector a user is in. Since each BS has the sector information of the users within its coverage area, it can send its users' sector information to the pilot assignment process. The amount of backhaul transmission required is minimal, in comparison with transmitting a number of huge-size channel covariance matrices. We propose that two principles should be considered in pilot assignment.

A. PILOT ASSIGNMENT PRINCIPLE AND METRIC

The first one is that the pilot assignment should only assign the same pilot sequence to the sectors in different cells which have different AOAs at BSs. This principle ensures that the users with the same pilot have non-overlapping AOAs and, therefore, the performance of the proposed location-aware channel estimation can be enhanced significantly. In most cases, although the angle spread of the main lobe is small, there may be some side lobes. These side lobes may still affect the system performance. To reduce this side lobe effect, the difference between the AOAs of the users with the same pilot should be as large as possible.

The second principle is that if it is difficult to ensure the non-overlapping AOAs for the users with the same pilot, the distance between the interference sector and the BS of the target cell should be sufficiently large. According to the channel model (15), the interference strength decreases exponentially fast with the increase of distance. Thus, if the non-overlapping condition cannot be satisfied, the path loss can be utilized to reduce the potential inter-cell interference.

Without loss of generality, we assume that the number of sectors is $S = K$. Let $\theta_{l,m,s}$ be the AOA of the "representative" user in the s-th sector of the l-th cell at the BS in the m-th cell and $d_{l,m,s}$ be the distance between this representative user and the BS in the m-th cell. Fig. 3 illustrates this representative user information $(\theta_{l,m,s}, d_{l,m,s})$. Any user in s-th sector of the l-th cell is represented by this location information $(\theta_{l,m,s}, d_{l,m,s})$ with respect to the BS in the m-th cell. In this way, we avoid the need for accurate user location information. To investigate the inter-cell interference of the interference cells to the m-th cell, the following metric is proposed for the

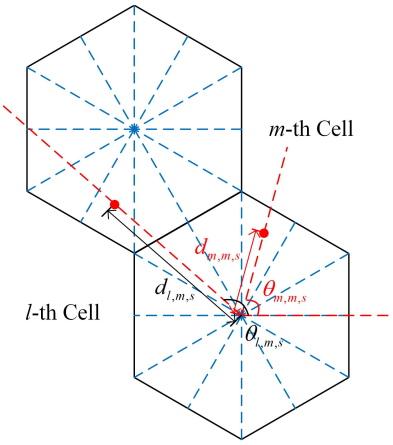


FIGURE 3. Illustration of the representative user in s-th sector of l-th cell with respect to BS in m-th cell, $(\theta_{l,m,s}, d_{l,m,s})$.

pilot assignment:

$$R_{l,m,s} = \frac{\left|\boldsymbol{t}^{\mathrm{T}}(\theta_{m,m,s})\boldsymbol{t}(\theta_{l,m,s})\right|}{d_{l,m,s}^{\gamma}}, \tag{34}$$

where $\boldsymbol{t}(\theta) = [\cos(\theta)\ \sin(\theta)]^{\mathrm{T}}$ is the directional vector with unit length. $R_{l,m,s}$ measures the interference caused by the s-th sector in the l-th cell to the s-th sector in the m-th cell by considering not only the correlation between the two AOAs but also the distance of the interference users. Smaller $R_{l,m,s}$ indicates that the difference between the AOAs of different users is large and/or the interference sector is far away from the target BS, resulting in smaller inter-cell interference and better performance. Thus, it is a good metric for the inter-cell interference, and assigning pilots based on this metric is exactly utilizing the above-mentioned two principles.

By applying the metric $R_{l,m,s}$ in (34), the optimal pilot assignment can be obtained by exhaustively searching all the possible patterns to find the solution that minimizes $\sum_l \sum_m \sum_s R_{l,m,s}$. However, the computational complexity of exhaustive search, given as $(S!)^L$, is prohibitively high in practice, especially when the number of cells L is large. Fortunately, we notice that the interference caused by the sectors in the nearest two adjacent cells is the severest and owing to the path loss, the interference from the sectors in far-away cells is much smaller. Thus we only consider the interference from the nearest two adjacent cells to reduce the computational complexity. Based on these considerations, we propose a three-step simplified method to obtain a sub-optimal pilot assignment pattern.

B. PROPOSED SUBOPTIMAL PILOT ASSIGNMENT ALGORITHM

Initialization. Starting from a randomly selected cell, e.g., the central cell of Fig. 4, randomly assign the orthogonal pilots to the sectors of this cell. Then the pilot assignment for the sectors in the other cells is carried out with the following three-step procedure.

1) In the first step, the sectors in the adjacent cells are assigned by a non-exhaustive search. For a sector that has been assigned with pilot, the adjacent sectors in its nearest two cells are checked and one of them is selected to assign with the same pilot sequence. Specifically, the metric $R_{l,m,s}$ of each sector in the nearest two cells is calculated and the sector with the smallest $R_{l,m,s}$ is chosen.

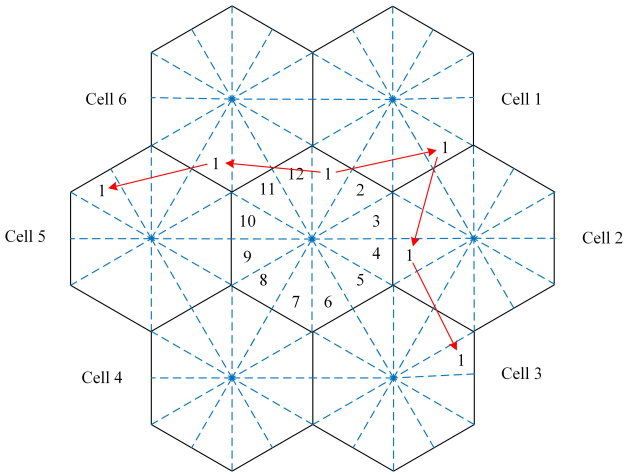


FIGURE 4. An example of the initialization and first step of the proposed pilot assignment method, with the number of cells $L = 7$ and the number of sectors in each cell $S = 12$. Number in each sector is the pilot number assigned.

The first step of the pilot assignment is illustrated in Fig. 4, where the central cell is initially assigned with pilots. Starting from the sector 1 of the central cell, which happens to be assigned with the pilot 1, the sectors in its nearest two adjacent cells, i.e., cell 1 and cell 6, are checked and the sector with smallest $R_{l,m,s}$ is assigned with the same pilot 1. Then we repeat this process to all the possible sectors that are already assigned with pilots. After the first step, there may be still some blank sectors located on the edges of the network that need to be assigned with pilots.

2) In the second step, these remaining sectors are assigned with pilots by an exhaustive search and the assignment is based on the smallest $\sum_s R_{l,m,s}$ for each cell. Note that the summations over m and l have been removed because in the first step, the sectors in the nearest cells have been assigned and the remaining sectors will not affect their interference cells. Normally, the number of the blank sectors left in each cell is small and the complexity of the second-step exhaustive search is reduced significantly. An example of the obtain pilot assignment pattern is shown in Fig. 5.

3) For the cells where each sector contains one user, the pilot assignment is completed. However, for a cell, where some sectors have no users, while some sectors contain multiple users, we need to re-assign the pilots of no-user sectors to the users of multiple-user sectors. For a sector with multiple users, a user is randomly selected to assign with the sector's pilot, while the other users are randomly assigned with the pilots of the no-user sectors.

Remark 5. It can be clearly seen that the aforementioned pilot assignment scheme utilizes the sector information, which is readily available at BSs, rather than the accurate user

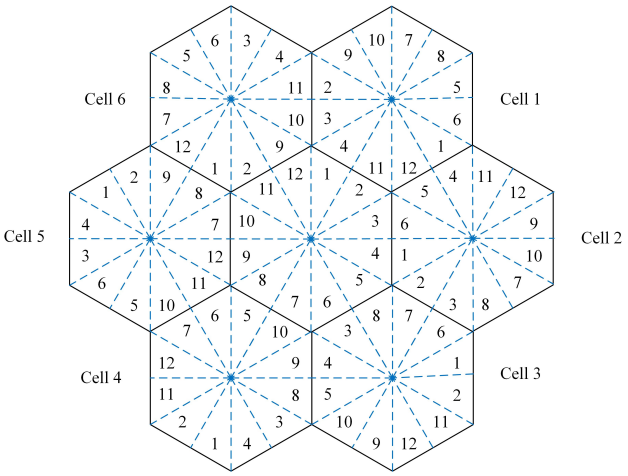


FIGURE 5. Pilot assignment obtained after the first two steps of the proposed method, where the number of cells $L = 7$ and the number of sectors in each cell $S = 12$. Number in each sector is the pilot number assigned.

location information. Therefore, the proposed location-aware channel estimation scheme is very practical as it does not need to acquire accurate user location information. As pointed out previously, it takes at least several or several tens of seconds for a user to cross a sector, which is hundreds or thousands times longer than the typical mobile channel's coherent time. The implication is that the pilot assignment remains valid for the network operation duration over hundreds of frames, with the frame structure of Fig. 2. Only when the users' sector information have changed significantly, the pilots need to re-assign. Moreover, since the BSs have the users' sector information, when to re-start the pilot assignment procedure is automatically determined.

It is also apparent that the proposed location-award pilot assignment scheme is very different from the the location-award pilot assignment scheme of [21].

TABLE 2. Basic simulation parameters

Cell radius	500 m
Number of sectors per cell S	12
Number of users per cell K	12
Path loss exponent γ	3.5
Variance of shadow fading	8 dB
Number of paths per user P	50
Length of pilot sequence	12
Angle spread δ_θ	10 degrees
Cell edge SNR	15 dB

V. NUMERICAL RESULTS

A homogeneous multi-cell multi-user MIMO scenario with $L = 7$ hexagonal cells and $K = 12$ users per cell is investigated to demonstrate the effectiveness of the proposed scheme. The simulation system's parameters are listed in Table 2. Each cell is divided into $S = 12$ sectors. The users in the center cell, which is denoted as cell 0, are the target users considered, and the surrounding cells, which are numbered by cell 1 to cell 6, are the interference cells. For the location-aware channel estimation, the number of FFT is set to $N = 8192$ and is fixed when the number of antennas M at the BS varies. The three-step suboptimal pilot assignment method proposed in Section IV is applied and a typical result obtained is depicted in Fig. 5.

The angle spread δ_θ is 10 degrees and two distributions of angle spread are considered. The first one is the uniform distribution, where the AOAs are uniformly and randomly distributed in the interval $[\theta_c - \delta_\theta,\ \theta_c + \delta_\theta]$ with θ_c indicating the AOA of line of sight. The second one is the Gaussian distribution with mean θ_c and variance $(\delta_\theta)^2$. The uniform distribution represents the scenario that the condition of non-overlapping AOAs can be satisfied while the Gaussian distribution cannot guarantee this condition.

The mean square error (MSE) of the channel estimation is calculated according to

$$\mathrm{MSE} = 10\log_{10}\left(\frac{E\left\{\|\widehat{\boldsymbol{h}}_{l,l,s} - \boldsymbol{h}_{l,l,s}\|^2\right\}}{E\left\{\|\boldsymbol{h}_{l,l,s}\|^2\right\}}\right)\ [\mathrm{dB}], \quad (35)$$

where $\boldsymbol{h}_{l,l,s}$ is the true UL channel coefficient vector for the user assigned with the s-th sector pilot in the l-th cell, and $\widehat{\boldsymbol{h}}_{l,l,s}$ is its estimate. Fig. 6 compares the MSE of the location-aware channel estimation averaged over all the $S = 12$ sectors with that of the conventional method. With the increase of the antennas at BS, the average MSE of the conventional channel estimation hardly reduces and remains very poor. By contrast, with the uniformly distributed angle spread, the location-aware channel estimation improves the average MSE performance dramatically, and the average MSE reduces significantly with the increase of M. When $M \simeq 100$, the average MSE is reduced to about -13.5 dB, which represents about 8 dB gain compared with its conventional counterpart. The proposed channel estimation is not as efficient for the Gaussian distributed angle spread as for the uniformly distributed angle spread. This is because

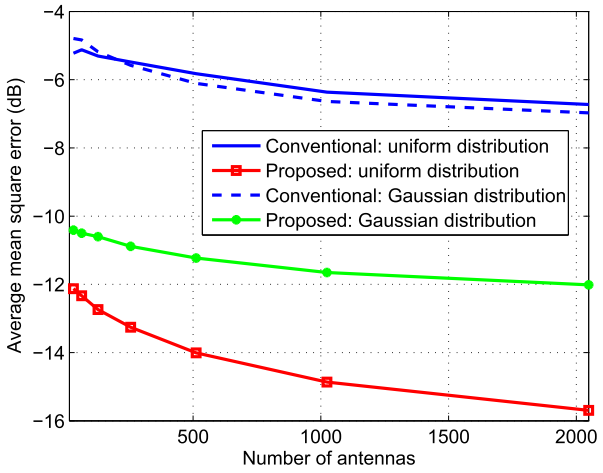


FIGURE 6. Average MSE comparison for the two channel estimation methods under two different angle spread distributions.

for the Gaussian distributed angle spread, the AOAs are not bounded in an interval. As a result, some useful paths are filtered out. However, as we will show later, the improvement on capacity is significant, even with the Gaussian distributed angle spread.

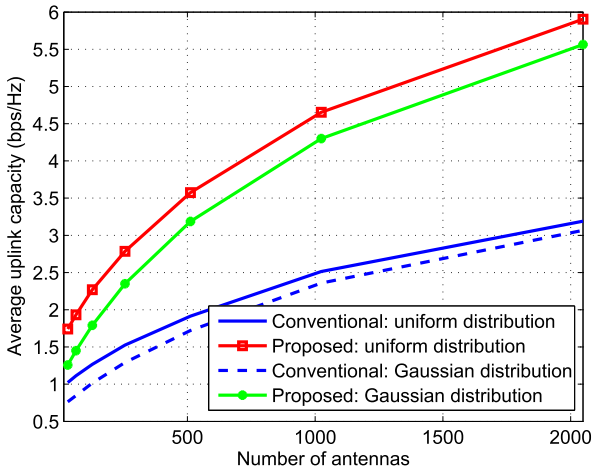


FIGURE 7. Average per user capacity comparison of uplink data transmission for the two channel estimation methods under two different angle spread distributions.

The capacity of UL data transmission is next investigated. The per user UL capacity of the user that adopts the s-th sector pilot is given in (36) at the bottom of this page, where $\boldsymbol{d}_{l,s}^{\mathrm{T}}$ is the detection row vector for the user assigned with the s-th sector pilot of l-th cell, $\boldsymbol{h}_{l_1,l_2,s}$ is the channel coefficient vector from the user assigned with the s-th sector pilot of l_1-th cell to the BS of l_2-th cell, and σ_n^2 denotes the UL noise power during the UL transmission. The ZF detection is applied. The first sum in the denominator of (36) is the intra-cell interference power while the second sum is the inter-cell interference power. Fig. 7 shows the simulation results for the per user UL capacity averaged over all the $S = 12$ sectors, where it can be seen that the UL performance is improved significantly using the proposed location-aware channel estimation, compared with the conventional method. When $M \simeq 100$, the data rates of the proposed scheme are higher than 2.8 bps/Hz and 2.3 bps/Hz with the uniformly distributed and Gaussian distributed angle spreads, respectively, while the capacity of the conventional method is lower than 1.6 bps/Hz.

For DL data transmission, the capacity of the user that adopts the s-th sector pilot of the central cell is computed as given in (37) at the bottom of this page, where $\boldsymbol{w}_{l,s}$ is the ZF precoding vector for the user assigned with the

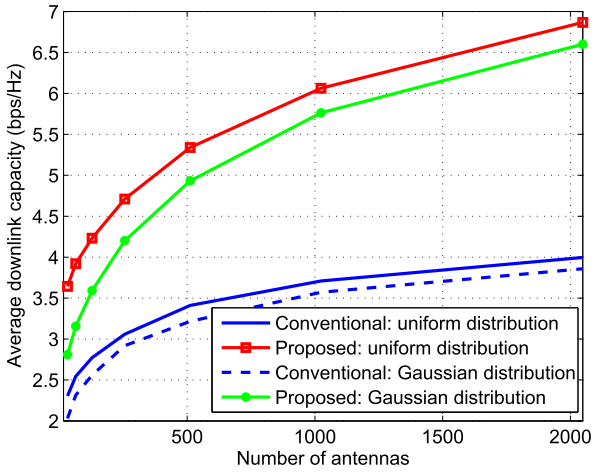


FIGURE 8. Average per user capacity comparison of downlink data transmission for the two channel estimation methods under two different angle spread distributions.

s-th sector pilot of l-th cell, and σ_n^2 denotes the DL noise power during DL transmission. Again it can be seen that the interference comes from the two sources: the first one is the intra-cell interference, denoted by $\sum_{j\neq s}\|\boldsymbol{h}_{0,0,s}^{\mathrm{T}}\boldsymbol{w}_{0,j}\|^2$, and the second one is the inter-cell interference, denoted by $\sum_{i\neq 0}\|\boldsymbol{h}_{0,i,s}^{\mathrm{T}}\boldsymbol{w}_{i,s}\|^2$. Fig. 8 compares the results of the per user capacity for DL transmission achieved by the proposed and conventional schemes, averaged over all the $S = 12$ sectors. As expected, the per user DL capacity is also improved by using the proposed scheme, and the gain over the conventional method increases considerably as M increases for both the uniform and Gaussian distributions of angle spread.

From the above simulation results, we observe that with the Gaussian distributed angle spread, although the MSE of the channel estimation obtained by the proposed location-aware channel estimator only reduces slowly with the increase of M, both the achievable UL and DL capacities still increase quickly with the increase of M, and the gains obtained over the conventional method are significant. This is because according to the analysis of Section III-C, with larger M, the resolution of different AOAs by the proposed method is significantly better. Thus the users with different AOAs can be distinguished much more accurately and the inter-cell interference caused by the reuse of pilot is also reduced significantly, resulting in much better performance.

As explained previously, the recent study and experiments all point to the fact that the channel angle spread observed at BS is small, and typical angle spread value is a few degrees for most practical scenarios. This is the reason that we set

$$C_s^u = E\left\{\log_2\left(1+\frac{\|\boldsymbol{d}_{0,s}^{\mathrm{T}}\boldsymbol{h}_{0,0,s}\|^2}{\sum_{j\neq s}\|\boldsymbol{d}_{0,s}^{\mathrm{T}}\boldsymbol{h}_{0,0,j}\|^2+\sum_{i\neq 0}\|\boldsymbol{d}_{0,s}^{\mathrm{T}}\boldsymbol{h}_{i,0,s}\|^2+\sigma_n^2}\right)\right\}, \tag{36}$$

$$C_s^d = E\left\{\log_2\left(1+\frac{\|\boldsymbol{h}_{0,0,s}^{\mathrm{T}}\boldsymbol{w}_{0,s}\|^2}{\sum_{j\neq s}\|\boldsymbol{h}_{0,0,s}^{\mathrm{T}}\boldsymbol{w}_{0,j}\|^2+\sum_{i\neq 0}\|\boldsymbol{h}_{0,i,s}^{\mathrm{T}}\boldsymbol{w}_{i,s}\|^2+\sigma_n^2}\right)\right\}, \tag{37}$$

the angle spread to 10° in the above simulation investigation. In order to further evaluate the influence of angle spread on the achievable performance of our proposed location-aware channel estimation scheme, we vary the angle spread from 5° to 30°, while fixing the number of antennas at BS to $M = 512$. It is worth emphasizing that an angle spread value of 30° is rare, which is unlikely to occur in practice.

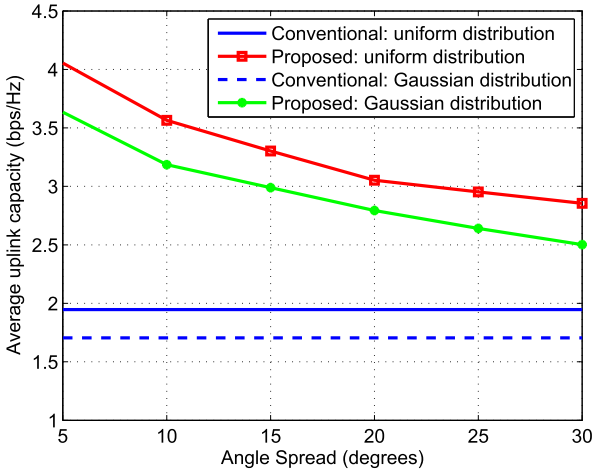


FIGURE 9. Average per user capacity comparison of uplink data transmission for the two channel estimation methods as the function of angle spread, given the BS antenna number $M = 512$.

Fig. 9 shows the simulation results for the average per user UL capacity achieved by the proposed and conventional schemes as the function of angle spread δ_θ. As expected, the performance gain of our proposed location-aware channel estimation scheme over the conventional one decreases as the increase of angle spread. This is due to the reason that when the angle spread becomes large, the non-overlapping AOAs condition is difficult to be guaranteed, leading to the increase of inter-cell interference. However, even with an angle spread as large as 30°, our proposed scheme still outperforms the conventional channel estimation scheme by about 50%, in terms of achievable UL capacity, as can be clearly seen from Fig. 9. Again, we can also observe that given the same value of angle spread, the UL capacity achieved by the proposed scheme under the uniform AOA distribution is better than that achieved under the Gaussian AOA distribution. Similarly, Fig. 10 compares the average per user DL capacity of our proposed location-aware channel estimator with those of the conventional scheme under different angle spreads, where the similar observations can be drawn as in the UL transmission case. In particular, even at the extreme case of angle spread being 30°, which is unlikely to occur in practice, the proposed scheme still significantly outperforms the conventional one.

The simulation results thus demonstrate that for the angle spread typically encountered in practice, the performance of the proposed location-aware channel estimation with the location-aware pilot assignment is superior. Moreover, the significant performance gain over the low-complexity conventional channel estimation scheme is achieved only at the

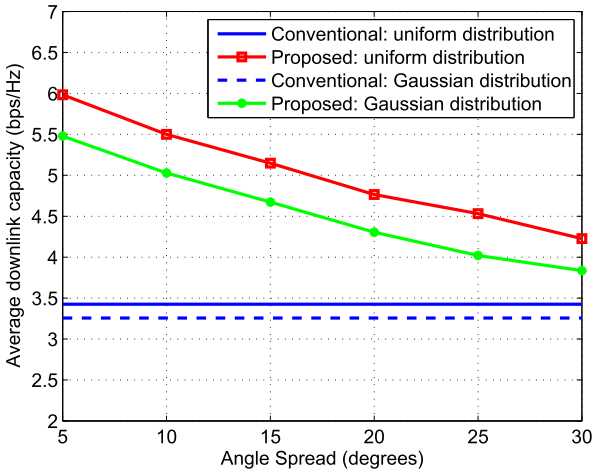


FIGURE 10. Average per user capacity comparison of downlink data transmission for the two channel estimation methods as the function of angle spread, given the BS antenna number $M = 512$.

expense of a slightly increase in complexity, and the complexity of the proposed estimator remains to be on the linear order of M. Like the conventional channel estimation scheme, our proposed location-aware channel estimation and pilot assignment schemes are capable of successfully operating under the shortest CHI condition. Compared to the other existing "smart" schemes [9], [11]–[13] developed for eliminating or reducing inter-cell interference caused by PC, the proposed scheme has much lower implementation complexity and does not impose any unrealistic requirement. In particular, under the operational environment of CHI $= K + N_{\mathrm{UL}} + N_{\mathrm{DL}}$, none of these "smart" schemes can be implemented.

VI. CONCLUSIONS

A novel location-aware channel estimation algorithm has been proposed to reduce the inter-cell interference caused by pilot contamination in TDD based massive MIMO systems. More specifically, an FFT-based post-processing has been introduced after the conventional pilot-aided channel estimation to exploit the spatial selection property of the antenna array's steering vector. Our analysis has shown that this post-processing can effectively distinguish the users with different AOAs. Thus if the users with the same pilots have different AOAs at BSs, the inter-cell interference caused by the re-use of pilots can be reduced considerably. To satisfy the condition of non-overlapping AOAs, a location-aware pilot assignment has further been proposed to improve the achievable system performance, which only utilizes the readily available sector information, rather than the accurate user location information. Simulation results have confirmed that compared with the low-complexity conventional pilot-aided channel estimation method, our location-aware scheme is capable of significantly improves the accuracy of the channel estimate as well as enhancing the uplink and downlink capacities considerably, while only imposing a slightly increased complexity. Moreover, compared with other existing smart

channel estimation schemes for eliminating or reducing inter-cell interference caused by pilot contamination, our proposed scheme has a much lower implementation complexity, imposes no unrealistic condition and does not require extensive coordination between BSs.

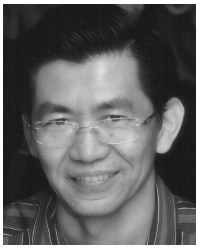
REFERENCES

[1] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
[2] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
[3] C. Shepard *et al.*, "Argos: Practical many-antenna base stations," in *Proc. MobiCom*, Istanbul, Turkey, Aug. 2012, pp. 53–64.
[4] C. Shepard, H. Yu, and L. Zhong, "ArgosV2: A flexible many-antenna research platform," in *Proc. MobiCom*, Miami, FL, USA, Sep./Oct. 2013, pp. 163–166.
[5] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.
[6] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
[7] A. Hu, T. Lv, H. Gao, Y. Lu, and E. Liu, "Pilot design for large-scale multi-cell multiuser MIMO systems," in *Proc. ICC*, Budapest, Hungry, Jun. 2013, pp. 5381–5385.
[8] R. R. Müller, L. Cottatellucci, and M. Vehkaperä, "Blind pilot decontamination," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 773–786, Oct. 2014.
[9] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.
[10] A. Ashikhmin and T. Marzetta, "Pilot contamination precoding in multi-cell large scale antenna systems," in *Proc. ISIT*, Cambridge, MA, USA, Jul. 2012, pp. 1137–1141.
[11] L. You, X. Gao, X.-G. Xia, N. Ma, and Y. Peng, "Massive MIMO transmission with pilot reuse in single cell," in *Proc. ICC*, Sydney, NSW, Australia, Jun. 2014, pp. 4794–4799.
[12] J. Zhang, B. Zhang, S. Chen, X. Mu, M. El-Hajjar, and L. Hanzo, "Pilot contamination elimination for large-scale multiple-antenna aided OFDM systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 759–772, Oct. 2014.
[13] T. X. Vu, T. A. Vu, and T. Q. S. Quek, "Successive pilot contamination elimination in multiantenna multicell networks," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 617–620, Dec. 2014.
[14] J. Nam, A. Adhikary, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing: Opportunistic beamforming, user grouping and simplified downlink scheduling," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 876–890, Oct. 2014.
[15] K. Yu and B. Ottersten, "Models for MIMO propagation channels: A review," *Wireless Commun. Mobile Comput.*, vol. 2, no. 7, pp. 653–666, Nov. 2002.
[16] N. Zhang and S. W. Golomb, "Polyphase sequence with low autocorrelations," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 1085–1089, May 1993.
[17] A. Klein, W. Mohr, R. Thomas, P. Weber, and B. Wirth, "Direction-of-arrival of partial waves in wideband mobile radio channels for intelligent antenna concepts," in *Proc. VTC*, vol. 2. Atlanta, GA, USA, Apr./May 1996, pp. 849–853.
[18] Y.-H. Nam *et al.*, "Full-dimension MIMO (FD-MIMO) for next generation cellular technology," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 172–179, Jun. 2013.
[19] S. Akoum and J. Acharya, "Full-dimensional MIMO for future cellular networks," in *Proc. IEEE RWS*, Newport Beach, CA, USA, Jan. 2014, pp. 1–3.
[20] W. Zhang, Y. Wang, F. Peng, and Y. Yuan, "Interference coordination with vertical beamforming in 3D MIMO-OFDMA networks," *IEEE Commun. Let.*, vol. 18, no. 1, pp. 34–37, Jan. 2014.
[21] Z. Wang, C. Qian, L. Dai, J. Chen, C. Sun, and S. Chen, "Location-based channel estimation and pilot assignment for massive MIMO systems," in *Proc. ICC Workshop*, London, U.K., Jun. 2015, pp. 1264–1268.

ZHAOCHENG WANG (M'09–SM'11) received the B.S., M.S., and Ph.D. degrees from Tsinghua University, Beijing, China, in 1991, 1993, and 1996, respectively. From 1996 to 1997, he was a post-doctoral fellow with Nanyang Technological University, Singapore. From 1997 to 1999, he was with OKI Techno Center (Singapore) Pte. Ltd., Singapore, where he was a Research Engineer and then became a Senior Engineer. From 1999 to 2009, he was with Sony Deutschland GmbH, where he was a Senior Engineer and then became a Principal Engineer. He is currently a Professor of Electronic Engineering with Tsinghua University and serves as the Director of Broadband Communication Key Laboratory, Tsinghua National Laboratory for Information Science and Technology. He has authored or coauthored over 120 journal papers. He holds 34 granted U.S./European patents. He coauthored two books, one of which, *Millimeter Wave Communication Systems*, was selected by the IEEE Series on *Digital and Mobile Communication* (Wiley-IEEE Press). His research interests include wireless communications, visible light communications, millimeterwave communications, and digital broadcasting. He is a fellow of the Institution of Engineering and Technology. He served as an Associate Editor of the IEEE Transactions on Wireless Communications from 2011 to 2015. He serves as an Associate Editor of the IEEE Communications Letters since 2013. He has also served as Technical Program Committee Co-Chair of various international conferences.

PEIYAO ZHAO (S'15) received the B.S. degree (Hons.) from Tsinghua University, Beijing, China, in 2014. He is currently pursuing the Ph.D. degree with the Department of Electronic Engineering, Tsinghua University. His research interests include massive MIMO and millimeterwave communications.

CHEN QIAN received the B.E. and Ph.D. degrees in electronic engineering from Tsinghua University, Beijing, China, in 2010 and 2015, respectively. His research interests include MIMO detection algorithms and channel coding and modulation.



SHENG CHEN (M'90–SM'97–F'08) received the B.Eng. degree in control engineering from the East China Petroleum Institute, Dongying, China, in 1982, and the Ph.D. degree in control engineering from the City, University of London, London, in 1986. In 2005, he received the D.Sc. degree (Hons.) from the University of Southampton, Southampton, U.K. From 1986 to 1999, he held research and academic appointments with the University of Sheffield, U.K, the University of Edinburgh, U.K, and the University of Portsmouth, U.K. Since 1999, he has been with the Department of Electronics and Computer Science, the University of Southampton, UK, where he is currently a Professor in intelligent systems and signal processing. His research interests include adaptive signal processing, wireless communications, modeling and identification of nonlinear systems, neural network and machine learning, intelligent control system design, and evolutionary computation methods and optimization. He has authored over 550 research papers. He is a fellow of the Royal Academy of Engineering, U.K., a fellow of the IET, a Distinguished Adjunct Professor with King Abdulaziz University, Jeddah, Saudi Arabia, and an ISI highly cited researcher in engineering (March 2004).