Estimation of Sparse Channels in Millimeter-Wave MU-MIMO Systems

Anzhong Hu

School of Communication Engineering, Hangzhou Dianzi University

Hangzhou, China, 310018

[e-mail: huaz@hdu.edu.cn]

*Corresponding author: Anzhong Hu

Received December 4, 2015; revised February 14, 2016; accepted April 6, 2016; published May 31, 2016

Abstract

This paper considers a channel estimation scheme for millimeter-wave multiuser multiple-input multiple-output systems. According to the proposed method, parts of the beams are selected and the channel parameters are estimated according to the sparsity of channels and the orthogonality of the beams. Since the beams for each channel become distinct and the signal power increases with the increased number of antennas, the proposed approach is able to achieve good estimation performance. As a result, the sum rate can be increased in comparison with traditional approaches, and channels can be estimated with fewer pilot symbols. Numerical results verify that the proposed approach outperforms traditional approaches in cases with large numbers of antennas.

Keywords: Millimeter-wave (mm-wave), multiple-input multiple-output (MIMO), channel estimation, sparse, beamspace

This research was supported by Zhejiang Provincial Natural Science Foundation of China under Grant No. LQ16F010007 and Scientific Research Starting Foundation of Hangzhou Dianzi University under Grant No. KYS085614054.The author thanks the two anonymous reviewers and Editor Prof. Zdenek Becvar for very helpful feedback and suggestions on the paper.

http://dx.doi.org/10.3837/tiis.2016.05.009 **ISSN : 1976-7277**

1. Introduction

Millimeter-wave (mm-wave) systems operating from 30 GHz to 300 GHz can provide multi-GHz bandwidth, and are expected to meet the capacity demands of future wireless networks [1], [2]. Meanwhile, the small wavelengths in mm-wave systems enable placing multiple antennas with half-wavelength distance in constrained spaces [3],[4]. These arrangements of large numbers of antennas constitute massive multiple-input multiple-output (MIMO) systems, which are expected to achieve high spectral efficiency in future wireless systems [5],[6].

In mm-wave systems, their performance relies on the employment of efficient channel estimation schemes. Since the wireless channels tend to exhibit sparse-scattering structure as the signal space dimension gets large, the channel sparsity is utilized to improve system performance, as demonstrated in [7-13]. It is well known that mm-wave channels are sparse, thus existing channel estimation approaches for mm-wave systems usually resort to compressed sensing (CS) theory to reduce the estimation overhead [11-15]. In [12], a CS-based channel estimation approach was proposed to estimate the angles and path gains with a limited number of RF chains. Furthermore, a technique for improving the resolution of the angle estimation and reducing the pilot overhead was presented in [13].However, these CS-based approaches are highly complex computationally. For example, even the computational complexity of the least squares solution of the CS-based approaches is proportional to M^2, where M is the number of the base station (BS) antennas. Moreover, all these methods were only proposed for point-to-point transmission scenarios, i.e., they are not suitable for mm-wave multiuser MIMO (MU-MIMO) systems.

Recently, in order to exploit spatial channel sparsity in mm-wave systems, the beamspace MIMO concept was proposed [16-19]. A beamspace MIMO transmitter selects a number of orthogonal beams for base-band precoding, and converts the low-dimensional baseband signals to high-dimensional radio-frequency (RF) signals with an analog beamforming front-end. As the complexity of the base-band precoding and the number RF chains are dramatically reduced, beamspace MIMO systems are capable of achieving near-optimal performance with low complexity [16]. Since the beamspace MIMO concept also utilizes channel sparsity, it is feasible to estimate mm-wave channels in beamspace. For example, the approach in [20] utilized partial spatial beam selection for transmission. As the beams are orthogonal and are not reused among the mobile stations (MSs), the interference can be mitigated during the downlink channel estimation process. However, the uplink channel estimation is assumed to be perfect, which requires traditional orthogonal pilots, just as an approach utilized in [21]. Hence, the pilot overhead is extensive in cases where there is a large number of MSs.

In this paper, a channel estimation approach for mm-wave MU-MIMO systems is proposed. By exploiting the beamspace concept, the correlations of the received pilot vector and the beams are used to select the partial beam. Then, the direction-of-arrivals (DOAs) of the MSs are estimated by matching these correlation values with the theoretical values that are known *a priori*. Finally, path gains are estimated based on the received pilots and the estimated DOAs. The main contributions of this paper are two-fold.

1) By utilizing the orthogonality of the beams, the whole spatial dimension is divided for the MSs. Thus, the number of effective antennas is less than the number of antennas equipped, which means the pilot overhead is reduced in comparison with the modified approach proposed in [20] and [21]. Additionally, the computational complexity of the beamspace processing is proportional to M, which is much lower than that of the CS-based approach. Besides, the proposed approach is suitable for multiuser scenarios, while existing CS approaches are for single user cases.
2) The estimation accuracy of the proposed approach is analyzed and the analysis demonstrates that the proposed approach allows estimates with no error in systems with large numbers of antennas. Thus, the proposed approach performs better than the traditional modified approach in [20] and [21].

Notations: Lower-case (upper-case) boldface symbols denote vectors (matrices); $(\cdot)^H$ denotes the conjugate transpose; $[\cdot]_j$ is the j th element of a vector or the j th column of a matrix; $\mathcal{CN}(\mu,\sigma^2)$ is a circularly-symmetric complex Gaussian random variable with mean μ and variance σ^2; $\mathbb{E}\{\cdot\}$ denotes the expectation, and i is the imaginary unit.

2. System Model

Consider the following system model. A BS equipped with N antennas serves K MSs simultaneously, where each MS has only one antenna[1]. In addition, it is assumed that the number of RF chains at the BS is the same as the number of antennas at the BS[2].

[1]The single antenna assumption is used to simplify the discussion, and is also used in [16] and [18]. The extension of the considered system to the multiple antenna scenario is straightforward, as shown in [12] and [13].
[2]It is known that the number of RF chains is limited in mm-wave systems. The assumption of unconstrained number of RF chains is for simplification. As shown in [12] and [22], limited numbers of RF chains are used to form multiple beamforming vectors on the antennas in multiple time slots. Similarly, the limited RF chains can also be switched to the partial antennas in each time slot, and all the antennas can be accessed with the RF chains in multiple time slots. Thus, the case of unconstrained number of RF chains can be reduced to the case of constrained numbers of RF chains, as the only expense is the reduction of the time for data transmission.

The BS is equipped with a uniform rectangular array (URA) of dimension $N = N_{az}N_{el}$, where N_{az} is the number of antennas in the *y*-axis and N_{el} is the number of antennas in the *z*-axis, as shown in **Fig. 1**.

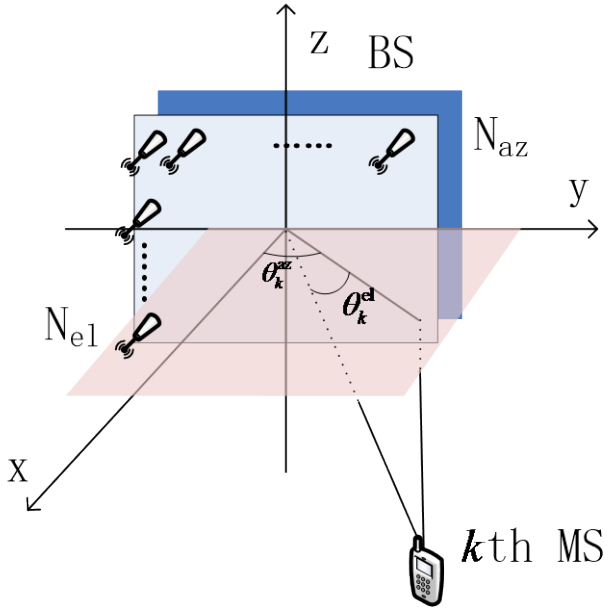

Fig. 1. The geometry of a cell. The URA is placed in the *y*-*z* plane, the MSs are distributed below the BS. The antennas are distributed uniformly in the URA.

The system adopts time division duplex (TDD) mode and thus channel reciprocity holds for two-way transmission. Furthermore, it is assumed that the channel undergoes block fading, which means the channel remains constant during the transmission of each coherent block of T_c symbols. The details of the transmission process are as follows.

1) The K MSs transmit the same pilot symbol to the BS simultaneously[3]. The BS estimates the uplink channels and selects the beams.

2) The K MSs transmit N_u data symbols to the BS at the same time. The BS processes the uplink data symbols using beamspace detection.

[3]Note that the MSs can also transmit a larger number of pilot symbols. For example, orthogonal pilots can also be employed, where the number of pilot symbols is K. A detailed description of the extension of the proposed approach for a larger number of pilot symbols will be presented in the simulation part.

3) The BS transmits N_{d} data symbols to the MSs with beamspace precoding and the MSs detect the downlink data symbols.

As can be seen, coherent intervals are used for the three phases, which results into $T_{\mathrm{c}} = N_{\mathrm{u}} + N_{\mathrm{d}} + 1$.

For the uplink transmission, the received pilot vector at the BS is given by

$$\mathbf{r}_{\mathrm{p}} = x_{\mathrm{p}} \sum_{k=1}^{K} [\mathbf{H}]_k + \mathbf{n} \in \mathbb{C}^{N\times 1}, \tag{1}$$

where $\mathbf{H} \in \mathbb{C}^{N\times K}$ is the channel matrix from the MSs to the BS, x_{p} is the pilot symbol transmitted by the MSs, and $\mathbf{n} \in \mathbb{C}^{N\times 1}$ is the noise vector, which is composed of independent and identically distributed (i.i.d.) $\mathcal{CN}(0,1)$ variables. Similarly, at the BS the received data matrix is given by

$$\mathbf{R}_{\mathrm{a}} = \mathbf{H}\mathbf{X}_{\mathrm{a}} + \mathbf{N} \in \mathbb{C}^{N\times N_{\mathrm{u}}}, \tag{2}$$

where $\mathbf{X}_{\mathrm{a}} \in \mathbb{C}^{K\times N_{\mathrm{u}}}$ is composed of the data symbols transmitted by the MSs, $\mathbf{N} \in \mathbb{C}^{N\times N_{\mathrm{u}}}$ is the noise matrix, which is also composed of i.i.d. $\mathcal{CN}(0,1)$ variables. For the downlink transmission, the received data matrix at the MSs is given by

$$\mathbf{T}_{\mathrm{a}} = \mathbf{H}^{H}\mathbf{Y}_{\mathrm{a}} + \mathbf{W} \in \mathbb{C}^{K\times N_{\mathrm{d}}}, \tag{3}$$

where $\mathbf{Y}_{\mathrm{a}} \in \mathbb{C}^{N\times N_{\mathrm{d}}}$ is composed of the BS precoded data symbols, and $\mathbf{W} \in \mathbb{C}^{K\times N_{\mathrm{d}}}$ is the noise matrix, composed of i.i.d. $\mathcal{CN}(0,1)$ variables.

2.1 Channel Model

According to the sparse-scattering property of the mm-wave channels and the cell geometry presented in Fig. 2, the uplink channel of the k-th MS is given by [18]

$$[\mathbf{H}]_k = \beta_k \mathbf{a}(\theta_k^{\mathrm{az}}, \theta_k^{\mathrm{el}}) + \sum_{m=1}^{N_P} \beta_{k,m} \mathbf{a}(\theta_{k,m}^{\mathrm{az}}, \theta_{k,m}^{\mathrm{el}}) \in \mathbb{C}^{N\times 1}, \tag{4}$$

where β_k is the uplink complex path loss and $\mathbf{a}(\theta_k^{\text{az}}, \theta_k^{\text{el}}) \in \mathbb{C}^{N\times 1}$ is the array steering vector for the azimuth, θ_k^{az}, and the elevation, θ_k^{el}, components of the DOA, which are also shown in Fig. 2. In addition, N_P is the number of multipaths, $\beta_{k,m}$ is the uplink complex path loss of the m-th multipath of the k-th MS, and θ_k^{az} and θ_k^{az} are the azimuth and elevation components of the DOA of the m-th multipath of the k-th MS, respectively. Similarly to [18], this paper focuses on pure line-of-sight (LOS) transmission[4], and the uplink channel can be re-expressed as

$$[\mathbf{H}]_k = \beta_k^{\text{u}} \mathbf{a}(\theta_k^{\text{az}}, \theta_k^{\text{el}}). \tag{5}$$

In addition, the uplink path loss is written as [23]:

$$\beta_k = \frac{D(\theta_k^{\text{az}}, \theta_k^{\text{el}})\lambda^2 e^{i\phi_k}}{16\pi^2 d_k^2}, \tag{6}$$

where λ is the transmission wavelength, ϕ_k is a random phase of the path loss of the k-th MS, d_k is the distance between the k-th MS and the BS, and $D(\theta_k^{\text{az}}, \theta_k^{\text{el}})$ is the array directivity as defined in [24]. According to the geometry of the URA, the array steering vector can be expressed as

$$[\mathbf{a}(\theta^{\text{az}}, \theta^{\text{el}})]_n = e^{i\pi(n_{\text{az}} - 0.5(N_{\text{az}}-1))\cos\theta^{\text{el}}\sin\theta^{\text{az}}} \times e^{i\pi(n_{\text{el}} - 0.5(N_{\text{el}}-1))\sin\theta^{\text{el}}}, \tag{7}$$

where $n_{\text{az}} = 0,1,\cdots,N_{\text{az}}-1$, $n_{\text{el}} = 0,1,\cdots,N_{\text{el}}-1$, and $n = n_{\text{el}}N_{\text{az}} + n_{\text{az}} + 1$.

2.2 Beamspace Processing

According to the principle of beamspace MIMO, the N orthogonal vectors corresponding to the beams form the unitary matrix, $\mathbf{U} \in \mathbb{C}^{N\times N}$, in which the m-th orthogonal vector or beam is

$$[\mathbf{U}]_m = \frac{1}{\sqrt{N}} \mathbf{a}(\tilde{\theta}_m^{\text{az}}, \tilde{\theta}_m^{\text{el}}), \tag{8}$$

[4]The assumption of the LOS transmission is also for simplicity of analysis. The extension of the LOS scenario to the multipath scenario is straightforward, because the channel responses of the multipaths can be estimated in the same manner as that of the LOS path at the expense of increased computational complexity.

where $\tilde{\theta}_m^{\text{az}}$ and $\tilde{\theta}_m^{\text{el}}$ satisfy $\cos\tilde{\theta}_m^{\text{el}}\sin\tilde{\theta}_m^{\text{az}} = -1+2m_{\text{az}}/N_{\text{az}}$, $\sin\tilde{\theta}_m^{\text{el}} = -1+2m_{\text{el}}/N_{\text{el}}$, $m=(m_{\text{el}}-1)N_{\text{az}}+m_{\text{az}}$, $m_{\text{az}}=1,2,\cdots,N_{\text{az}}$, and $m_{\text{el}}=1,2,\cdots,N_{\text{el}}$.

Depending on the locations of the MSs, parts of the beams are selected to construct the beamforming matrix $\mathbf{U}_{\text{b}} \in \mathbb{C}^{N\times N_{\text{b}}}$, where $N_{\text{b}} = K\tilde{N}_{\text{b}}$ is the number of all the beams selected, and $\tilde{N}_{\text{b}} \geq 2$ is the number of the beams selected for each MS. Then, beamspace precoding and the beamspace detection for data transmission phases can be derived using the minimum mean square error (MMSE). For the uplink transmission, the beamspace detection matrix for $\mathbf{R}_{\text{a}}$ in (2) can be expressed as

$$\tilde{\mathbf{X}}_{\text{a}} = \left(\hat{\mathbf{H}}_{\text{b}}^H\hat{\mathbf{H}}_{\text{b}} + \left(1/\rho_{MS}\right)\mathbf{I}_K\right)^{-1}\hat{\mathbf{H}}_{\text{b}}^H\mathbf{U}_{\text{b}}^H\mathbf{R}_{\text{a}} \in \mathbb{C}^{N_{\text{b}}\times N_{\text{u}}}, \tag{9}$$

where $\hat{\mathbf{H}}_{\text{b}} = \mathbf{U}_{\text{b}}^H\hat{\mathbf{H}} \in \mathbb{C}^{N_{\text{b}}\times K}$ is the beamspace channel estimate, $\hat{\mathbf{H}} \in \mathbb{C}^{N\times K}$ is the channel estimate, and ρ_{MS} is the transmission power of each MS, i.e., $\mathbb{E}\left\{\mathbf{X}_{\text{a}}\mathbf{X}_{\text{a}}^H\right\} = \rho_{MS}\mathbf{I}_K$. For the downlink transmission, the beamspace precoding for $\mathbf{T}_{\text{a}}$ in (3) can be rewritten as

$$\mathbf{T}_{\text{a}} = \mathbf{H}^H\mathbf{U}_{\text{b}}\hat{\mathbf{H}}_{\text{b}}\left(\hat{\mathbf{H}}_{\text{b}}^H\hat{\mathbf{H}}_{\text{b}} + \left(K/\rho_{\text{BS}}\right)\mathbf{I}_K\right)^{-1}\tilde{\mathbf{Y}}_{\text{a}} + \mathbf{W} \in \mathbb{C}^{K\times N_{\text{d}}}, \tag{10}$$

where $\tilde{\mathbf{Y}}_{\text{a}} \in \mathbb{C}^{K\times N_{\text{d}}}$ is matrix of the transmitted data symbols of the BS and ρ_{BS} is the total transmission power of the BS, i.e., $\mathbb{E}\left\{\mathbf{Y}_{\text{a}}\mathbf{Y}_{\text{a}}^H\right\} = \rho_{\text{BS}}\mathbf{I}_{N_{\text{d}}}$. Comparing (3) and (10), it can be seen that $\mathbf{Y}_{\text{a}} = \mathbf{U}_{\text{b}}\hat{\mathbf{H}}_{\text{b}}\left(\hat{\mathbf{H}}_{\text{b}}^H\hat{\mathbf{H}}_{\text{b}} + \left(K/\rho_{\text{BS}}\right)\mathbf{I}_K\right)^{-1}\tilde{\mathbf{Y}}_{\text{a}}$.

With the system model presented and the principle of beamspace processing given, the proposed channel estimation approach will be presented in the following section.

3. Channel Estimation

The traditional CS-based approach is highly complex computationally in mm-wave systems with large numbers of antennas. In addition, the modified approach in [20] and [21] is based on pilots, so the estimation performance will not benefit from the increased numbers of antennas. In order to avoid these drawbacks, a beamspace channel estimation approach is proposed. In the proposed estimation process, parts of the beams are firstly selected for beamspace transmission. Then, the DOAs of the MSs are and the path losses are estimated.

3.1 Beam Selection

In order to select the beams for transmission, the correlation of the pilot vector and the orthogonal beams is first calculated using

$$\tilde{\mathbf{r}}_{\mathrm{p}} = \mathbf{U}^H \mathbf{r}_{\mathrm{p}} = x_{\mathrm{p}} \sqrt{N} \sum_{k=1}^{K} \left[\tilde{\mathbf{H}}\right]_k + \tilde{\mathbf{n}} \in \mathbb{C}^{N\times 1}, \tag{11}$$

where $\tilde{\mathbf{H}} = \mathbf{U}^H \mathbf{H} / \sqrt{N} \in \mathbb{C}^{N\times K}$ and $\tilde{\mathbf{n}} = \mathbf{U}^H \mathbf{n} \in \mathbb{C}^{N\times 1}$. According to the expressions of $\mathbf{H}$ and $\mathbf{U}$ in (5) and (8), respectively, we have

$$[\tilde{\mathbf{H}}]_{m,k} = \frac{\beta_k}{N} \mathbf{a}^H(\tilde{\theta}_m^{\mathrm{az}}, \tilde{\theta}_m^{\mathrm{el}}) \mathbf{a}(\theta_k^{\mathrm{az}}, \theta_k^{\mathrm{el}}) = \frac{\eta_k \beta_k \lambda_{m,k}}{N}, \tag{12}$$

where

$$\lambda_{m,k} = \frac{\left(1 - e^{i\pi(\tilde{\alpha}_m - \alpha_k)N_{\mathrm{az}}}\right)\left(1 - e^{i\pi(\tilde{\gamma}_m - \gamma_k)N_{\mathrm{el}}}\right)}{\left(1 - e^{i\pi(\tilde{\alpha}_m - \alpha_k)}\right)\left(1 - e^{i\pi(\tilde{\gamma}_m - \gamma_k)}\right)}, \tag{13}$$

$$\eta_k = e^{i0.5\pi((-N_{\mathrm{az}}+3)\alpha_k + (-N_{\mathrm{el}}+3)\gamma_k)}, \tag{14}$$

and the second equation in (12) is derived based on (5). In addition, $\tilde{\alpha}_m = \cos\tilde{\theta}_m^{\mathrm{el}} \sin\tilde{\theta}_m^{\mathrm{az}}$, $\alpha_k = \cos\theta_k^{\mathrm{el}} \sin\theta_k^{\mathrm{az}}$, $\tilde{\gamma}_m = \sin\tilde{\theta}_m^{\mathrm{el}}$,and $\gamma_k = \sin\theta_k^{\mathrm{el}}$, which are also known as the spatial frequencies.

Let $f(x) = \sin(\pi x M) / \sin(\pi x)$, where $|x| < 2$, and M is an integer not smaller than 2. When $f'(x) = 0$, $f(x) = M$, which is a local maximum or minimum. Meanwhile, when $x = \pm 2$, $f(x) = M$. Since $|x| < 2$, it can be seen that $|f(x)|$ achieves its global maximum when $f'(x) = 0$, and that maximum is equal to M. In addition, it is evident that $f'(x) = 0$ only

when $\sin(\pi x/2)=0$, i.e., $x=0$. Therefore, the global maximum of $|f(x)|$ is achieved when $x=0$. Since $|\tilde{\alpha}_m-\alpha_k|<2, |\tilde{\gamma}_m-\gamma_k|<2$, based on the above results and (13), there is

$$|\lambda_{m,k}|=\left|\frac{\sin(\pi(\tilde{\alpha}_m-\alpha_k)N_{\text{az}}/2)}{\sin(\pi(\tilde{\alpha}_m-\alpha_k)/2)}\right|\left|\frac{\sin(\pi(\tilde{\gamma}_m-\gamma_k)N_{\text{el}}/2)}{\sin(\pi(\tilde{\gamma}_m-\gamma_k)/2)}\right|\le N,$$

and the maximum is achieved when $\tilde{\alpha}_m-\alpha_k=0, \tilde{\gamma}_m-\gamma_k=0$.

From the above analysis, we see that, for variable m, $|\lambda_{m,k}|, \forall k$, reaches the maximum when $\alpha_k=\tilde{\alpha}_m$ and $\gamma_k=\tilde{\gamma}_m$. According to (12), it can also be seen that, $\forall k$, $|[\tilde{\mathbf{H}}]_{m,k}|$ for variable m reaches its maximum value when $\alpha_k=\tilde{\alpha}_m$ and $\gamma_k=\tilde{\gamma}_m$. We denote the set of the indices of the maximal $\tilde{N}_{\text{b}}$ elements of $|[\tilde{\mathbf{H}}]_k|$ as $\mathcal{C}_k$. Then, $\forall m\in\mathcal{C}_k$, $\tilde{\alpha}_m$ is close to α_k, and $\forall m\in\mathcal{C}_k$, $\tilde{\gamma}_m$, is close to γ_k. Moreover, the element in $\mathcal{C}_k$ that corresponds to the maximum in $|[\tilde{\mathbf{H}}]_k|$ is denoted as c_k.

In the ideal condition that the sets, $\mathcal{C}_k$ and $\mathcal{C}_{k'}$, are disjoint for all $k\neq k'$, there are K local maxima of the function

$$g(m_{\text{az}},m_{\text{el}})=\left|\sum_{k=1}^{K}\left[\tilde{\mathbf{H}}\right]_{m,k}\right|, \tag{15}$$

where $m, m_{\text{az}}, m_{\text{el}}$ are defined in the same manner as in the case of (8). Let us assume that $g\left(m_{\text{az}}(k),m_{\text{el}}(k)\right)$ is the local maximum of the function $g(m_{\text{az}},m_{\text{el}})$ that corresponds to the k-th MS. For $m_{\text{az}},m_{\text{el}}$ that are close to $m_{\text{az}}(k),m_{\text{el}}(k)$, the corresponding $g(m_{\text{az}},m_{\text{el}})$ is also close to $g\left(m_{\text{az}}(k),m_{\text{el}}(k)\right)$. Hence, for $\tilde{N}_{\text{b}}-1$ such that $m_{\text{az}},m_{\text{el}}$ that are close to $m_{\text{az}}(k),m_{\text{el}}(k)$, the corresponding $\tilde{N}_{\text{b}}-1$ values of $g(m_{\text{az}},m_{\text{el}})$ are also close to this local maximum. As can be seen, the indices of this local maximum, and the corresponding $\tilde{N}_{\text{b}}-1$

values are exactly the elements of $\mathcal{C}_k$. Thus, the indices of the K local maxima of $g(m_{\text{az}}, m_{\text{el}})$ and the corresponding $(\tilde{N}_{\text{b}} - 1)K$ values are the elements in $\mathcal{C}_k, k = 1, 2, \cdots, K$. The element in $\mathcal{C}_k$ that corresponds to one of the local maxima in $g(m_{\text{az}}, m_{\text{el}})$ is c_k.

As an example, **Fig. 2** illustrates the relation between different sets, $\mathcal{C}_k$ and $\mathcal{C}_{k'}, \forall k \neq k'$. The axes represent the indices, $m_{\text{az}}, m_{\text{el}}$, the area rounded by the red line is the range of the beam indices for positions in the cell, and the blue circles represent the sets of selected indices for two MSs. For the indices in set $\mathcal{C}_1$, c_1 is an index that corresponds to a local maximum of $g(m_{\text{az}}, m_{\text{el}})$. The other elements in $\mathcal{C}_1$ correspond to the other $\tilde{N}_{\text{b}} - 1$ indices that are close to c_1, and correspondingly result in close values of $g(m_{\text{az}}, m_{\text{el}})$.

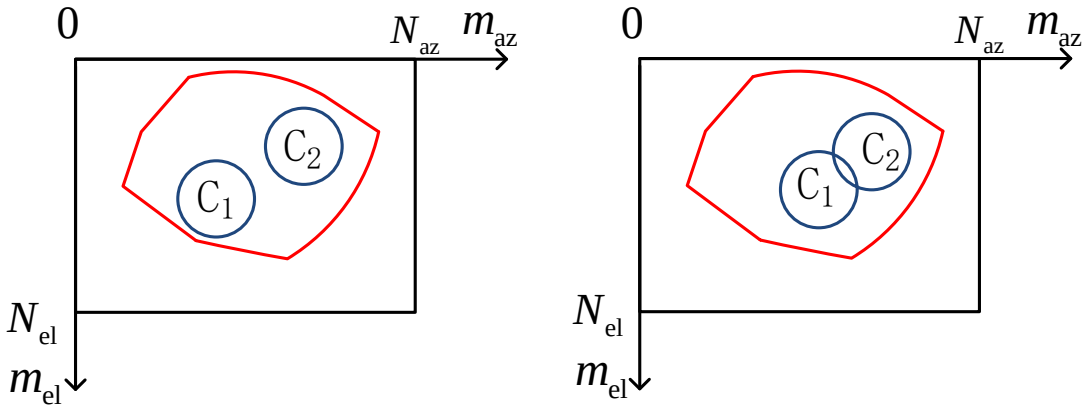

Fig. 2. An example of the relation between different sets of $\mathcal{C}_1$ and $\mathcal{C}_2$.

According to the definition of $\tilde{\mathbf{r}}_{\text{p}}$ in (11), it can be seen that

$$\tilde{g}(m_{\text{az}}, m_{\text{el}}) = |[\tilde{\mathbf{r}}_{\text{p}}]_m| = \left| x_{\text{p}} \sqrt{N} \sum_{k=1}^{K} \left[\tilde{\mathbf{H}} \right]_{m,k} + \tilde{\mathbf{n}} \right| \tag{16}$$

can be used as an estimate of $g(m_{\text{az}}, m_{\text{el}})$ in (15). Therefore, these $N_{\text{b}} = K\tilde{N}_{\text{b}}$ beams can be selected in the same way with the aid of $\tilde{g}(m_{\text{az}}, m_{\text{el}})$. Then, the beamforming matrix, $\mathbf{U}_{\text{b}}$ in

(9) and (10), can be constructed in the following manner. The K columns of $\mathbf{U}$ that have the elements of $\mathcal{C}_k$ as indices are selected to form the $\tilde{N}_{\mathrm{b}}(k-1)+1$th to the $\tilde{N}_{\mathrm{b}}k$th columns of $\mathbf{U}_{\mathrm{b}}$.

When the ideal condition that $\forall k \neq k'$, $\mathcal{C}_k$ and $\mathcal{C}_{k'}$ are disjoint is not satisfied, the same beam may be selected by different MSs, as shown in **Fig. 2**. In [20], various beam selection methods were discussed in to circumvent this problem. Here, a simple approach is employed. In this approach, the beams are selected by the MSs sequentially, i.e., the beams that have been already selected will be excluded in the subsequent selection. For the example in **Fig. 2**, once the elements in $\mathcal{C}_1$ have all been selected, those that are common with $\mathcal{C}_2$, i.e. those in the intersection of $\mathcal{C}_1, \mathcal{C}_2$, will be excluded in the selection of elements in $\mathcal{C}_2$.In addition, there are two principles in the selection process:

1) The beams that correspond to positions outside the cell are not included in the selection of the K local maxima of $\tilde{g}(m_{\mathrm{az}}, m_{\mathrm{el}})$. For the example in **Fig. 2**, the elements outside the red circle are excluded in the selection.

2) When the number of local maxima in $\tilde{g}(m_{\mathrm{az}}, m_{\mathrm{el}})$, $\tilde{K}$, is smaller than K, the maxima remaining after $K-\tilde{K}$ are selected instead. When $\tilde{K}$ is larger than K, the $\tilde{K}-K$ minimum in the $\tilde{K}$ local maxima are excluded in the selection of elements in $\mathcal{C}_k$.

3.2 DOA Estimation

According to (12), it can be seen that

$$\frac{[\tilde{\mathbf{H}}]_{m,k}}{[\tilde{\mathbf{H}}]_{m',k}} = \frac{\lambda_{m,k}}{\lambda_{m',k}}. \tag{17}$$

From the beam selection part, it is known that $\forall m \in \mathcal{C}_k$, the value of α_k is close to $\tilde{\alpha}_m$, and, correspondingly, the value of γ_k is close to $\tilde{\gamma}_m$ and that, additionally, $\sum_{k=1}^{K}\left[\tilde{\mathbf{H}}\right]_{m,k}$ can be used instead of $[\tilde{\mathbf{H}}]_{m,k}$, in the ideal case that $\forall k \neq k'$, $\mathcal{C}_k$ and $\mathcal{C}_{k'}$ are disjoint. Based on this property, the spatial frequencies α_k and γ_k can be estimated as

$$(\hat{\alpha}_k, \hat{\gamma}_k) = \arg\min_{\tilde{\alpha}_k, \tilde{\gamma}_k} \left| \frac{[\tilde{\mathbf{r}}_{\mathrm{p}}]_m}{[\tilde{\mathbf{r}}_{\mathrm{p}}]_{c_k}} - \frac{\breve{\lambda}_{m,k}}{\breve{\lambda}_{c_k,k}} \right|, \tag{18}$$

where m is such that $|\tilde{\alpha}_m - \tilde{\alpha}_{c_k}| \leq 2/N_{\mathrm{az}}$ and $|\tilde{\gamma}_m - \tilde{\gamma}_{c_k}| \leq 2/N_{\mathrm{el}}$, and

$$\breve{\lambda}_{m,k} = \frac{\left(1-e^{i\pi(\tilde{\alpha}_m-\tilde{\alpha}_k)N_{\mathrm{az}}}\right)\left(1-e^{i\pi(\tilde{\gamma}_m-\tilde{\gamma}_k)N_{\mathrm{el}}}\right)}{\left(1-e^{i\pi(\tilde{\alpha}_m-\tilde{\alpha}_k)}\right)\left(1-e^{i\pi(\tilde{\gamma}_m-\tilde{\gamma}_k)}\right)}.$$

In addition, $\tilde{\alpha}_m$ and $\tilde{\gamma}_m$ are defined below (8) and (14), and c_k is defined above (15). This constraint on m means that only the element in $\mathcal{C}_k$ that is close to c_k is used for estimation and that the search ranges for α_k and γ_k, i.e., $[\min(\tilde{\alpha}_m), \max(\tilde{\alpha}_m)]$ and $[\min(\tilde{\gamma}_m), \max(\tilde{\gamma}_m)], \forall m \in \mathcal{C}_k$, can be obtained, respectively. Once these spatial frequencies are estimated, the DOAs can be estimated as $\hat{\theta}_k^{\mathrm{el}} = \arcsin\hat{\gamma}_k$, $\hat{\theta}_k^{\mathrm{az}} = \arcsin(\hat{\alpha}_k / \cos\hat{\theta}_k^{\mathrm{el}})$.

When the ideal condition that $\forall k \neq k'$, $\mathcal{C}_k$ and $\mathcal{C}_{k'}$ are disjoint is not satisfied, the constraint on m mentioned above is not applicable. This means that there is no element in $\mathcal{C}_k$ that is close to c_k. In this case, the above estimation is not suitable, so $\tilde{\alpha}_{c_k}$ and $\tilde{\gamma}_{c_k}$, the meanings of which are explained in the previous paragraph, are used as estimates of α_k and γ_k, respectively.

3.3 Channel Parameter Estimation

With the estimated DOAs, it can be seen that only the uplink path loss β_k remains an unknown in (5). According to (11) and (12), β_k can be estimated as

$$\hat{\beta}_k = \frac{\sqrt{N}[\tilde{\mathbf{r}}_{\mathrm{p}}]_{c_k}}{x_{\mathrm{p}}\hat{\eta}_k\hat{\lambda}_{c_k,k}}, \tag{19}$$

where $\hat{\lambda}_{c_k,k}$ and $\hat{\eta}_k$ are estimated by replacing α_k and γ_k, and m with $\hat{\alpha}_k$, $\hat{\gamma}_k$ and c_k in (13) and (14), respectively. Note that although the above estimation is derived under the ideal condition that $\mathcal{C}_k$ and $\mathcal{C}_{k'}$ are disjoint, it is also applicable when this condition does not hold.

Now the channel parameters have been estimated. In a system with perfect synchronization and power control, the path loss of each MS can be coarsely estimated prior to channel estimation. Thus, the known path losses of the MSs can be utilized to resolve the ambiguity of the MSs, i.e., the corresponding estimated channel vector of each MS can be separated from other estimated channel vectors. A similar assumption and a corresponding method for distinguishing MSs can be seen in [25]. In the case that orthogonal pilots are employed, the pilots can also be used to distinguish the MSs.

With the proposed channel estimation approach presented, the performance of the system will be investigated in the following section.

4. Performance Analysis

In this section, the proposed approach is compared with the modified approach in [20] and [21] and the performance of the proposed approach in systems with large numbers of antennas is analyzed.

4.1 Comparison with Traditional Approaches

Although mm-wave MU-MIMO systems can achieve extremely high data rates, the channel estimation approach employed in these systems is not specific, as evident the traditional pilot-based approach in [20] and [21]. When the traditional pilot-based approach is modified for channel estimation in the mm-wave MU-MIMO system presented here, the number of uplink pilot symbols is K, and the received pilot matrix is $\mathbf{H}\mathbf{X}_{\mathrm{p}} + \mathbf{N}$, where $\mathbf{X}_{\mathrm{p}} \in \mathbb{C}^{K\times K}$ is the matrix of pilots, and $\mathbf{N} \in \mathbb{C}^{N\times K}$ is the noise matrix. Then, the uplink

channel matrix is estimated by multiplying $\mathbf{X}_{\mathrm{p}}^{-1}$ with the received pilot matrix. The beams can be then selected based on the uplink channel estimate.

According to the previous analysis, the number of pilot symbols is K for the traditional approach, which is the dimension of $\mathbf{X}_{\mathrm{p}}$ in the previous paragraph. On the contrary, only one pilot symbol is necessary in the proposed approach, which is the dimension of x_{p} in (1). Thus, the advantage of the proposed approach is that it can transmit more data symbols in one coherence interval and is thus expected to achieve a higher sum rate.

On the other hand, with the traditional approach in [20] and [21], the channels can be directly estimated with the pilots and the estimation error is only caused by noise. However, with the proposed approach, the channels are estimated through searching the DOAs and the estimation error is not only caused by the noise but also by interference. Therefore, the disadvantage of the proposed approach is that the channel estimation error may be larger than that of the traditional approach. Fortunately, this disadvantage can be effectively mitigated by increasing the number of antennas, as will explained in the next subsection.

4.2 Asymptotic Analysis

In this subsection, the performance of the proposed approach in the asymptotic limit is investigated, where the asymptotic limit means $N_{\mathrm{az}}, N_{\mathrm{el}} \to \infty$. The asymptotic property of $\lambda_{m,k}$ in (13) is presented first.

Theorem 1:1) When $(\tilde{\alpha}_m - \alpha_k)N_{\mathrm{az}}$ and $(\tilde{\gamma}_m - \gamma_k)N_{\mathrm{el}}$ do not tend to zero and one of them tends to infinity, $\lambda_{m,k} / N = 0$; 2) When $(\tilde{\alpha}_m - \alpha_k)N_{\mathrm{az}}$ and $(\tilde{\gamma}_m - \gamma_k)N_{\mathrm{el}}$ are finite, $\lambda_{m,k} / N \neq 0$.

Proof: For the first part of the theorem, there is

$$\frac{\lambda_{m,k}}{N} = \frac{x_0 x_1 x_2}{N(\tilde{\alpha}_m - \alpha_k)(\tilde{\gamma}_m - \gamma_k)} = 0,$$

where $x_0 = \left(1 - e^{i\pi(\tilde{\alpha}_m - \alpha_k)N_{\mathrm{az}}}\right)\left(1 - e^{i\pi(\tilde{\gamma}_m - \gamma_k)N_{\mathrm{el}}}\right)$, $x_1 = \dfrac{\tilde{\alpha}_m - \alpha_k}{1 - e^{i\pi(\tilde{\alpha}_m - \alpha_k)N_{\mathrm{az}}}}$, and $x_2 = \dfrac{\tilde{\gamma}_m - \gamma_k}{1 - e^{i\pi(\tilde{\gamma}_m - \gamma_k)N_{\mathrm{el}}}}$ are finite.

For the second part, there is

$$\frac{\lambda_{m,k}}{N}=\frac{x_0 N^{-1}}{\left(1-e^{x_3/N_{\text{az}}}\right)\left(1-e^{x_4/N_{\text{el}}}\right)}=\frac{x_0}{x_3 x_4}\neq 0,$$

where $x_3=i\pi(\tilde{\alpha}_m-\alpha_k)N_{\text{az}}$ and $x_4=i\pi(\tilde{\gamma}_m-\gamma_k)N_{\text{el}}$ are finite.

For any α_k and γ_k, the second property in the theorem above is satisfied with by a value of $m\in\{1,2,\cdots,N\}$. Thus, for the set $\mathcal{C}_k$,the element $m\in\mathcal{C}_k$ corresponds to the one of the maximal $\tilde{N}_{\text{b}}$ values of $\lambda_{m,k}$. Since the distances between the MSs are non-zero, the differences between the DOAs of two MSs as well as the differences between α_k and γ_k of two MSs are also non-zero, which means the condition of the first part of the theorem above is satisfied for $m\in\mathcal{C}_{k'},\forall k\neq k'$. Then, for the second part of the theorem, it can be seen that $\forall k\neq k'$, $\mathcal{C}_k$ and $\mathcal{C}_{k'}$, are disjoint. In other words, the ideal condition mentioned in the previous derivation is satisfied. Additionally, for $m\in\mathcal{C}_k$, $[\tilde{\mathbf{r}}_{\text{p}}]_m=x_{\text{p}}\sqrt{N}\sum_{k=1}^{K}\left[\tilde{\mathbf{H}}\right]_{m,k}$ is satisfied in the asymptotic limit. Thus, the DOAs as well as the path losses can be estimated with no error.

In reality, the number of antennas is finite, i.e., $N_{\text{az}},N_{\text{el}}$ are finite. In this case, the estimation in (18) can be expressed as

$$(\hat{\alpha}_k,\hat{\gamma}_k)=\arg\min_{\breve{\alpha}_k,\breve{\gamma}_k}\left|\frac{x_{\text{p}}\sqrt{N}\sum_{k=1}^{K}\left[\tilde{\mathbf{H}}\right]_{m,k}+\left[\tilde{\mathbf{n}}\right]_m}{x_{\text{p}}\sqrt{N}\sum_{k=1}^{K}\left[\tilde{\mathbf{H}}\right]_{c_k,k}+\left[\tilde{\mathbf{n}}\right]_{c_k}}-\frac{\breve{\lambda}_{m,k}}{\breve{\lambda}_{c_k,k}}\right|,$$

which is based on (11). It can be seen that the estimation error is inevitable because of the interference $\left[\tilde{\mathbf{H}}\right]_{m,k'}$, $\left[\tilde{\mathbf{H}}\right]_{c_k,k'}$, and the noise $\tilde{\mathbf{n}}$. Therefore, the estimation performance of the proposed channel estimation approach is not as good as its corresponding asymptotic performance.

According to this asymptotic analysis, it can be seen that the proposed approach benefits from an increased number of antennas. However, for the modified approach in [20] and [21],

the channels are directly estimated using the pilots, as shown in the previous subsection. Hence, the channel estimation performance of the traditional approach is invariant with the number of the antennas. Thus, the proposed approach is also expected to achieve a higher sum rate than the modified approach in [20] and [21] even when the same number of pilot symbols is used in the two approaches.

5. Simulation Results

For the simulation, the system was set up as follows. The radius of the cells was in the range [10m, 100m]. The range of the azimuth angle of the cell was $[-60^{\circ}, 60^{\circ}]$. The height of the BS was 10 meters. The frequency of the two-way channels was set to80 GHz and $K = 10, \tilde{N}_{\mathrm{b}} = 2$ .Given the symmetry of two-way transmissions, in the channel coherence interval only uplink channel estimation and uplink data detection were carried out for the simulation. In addition, the transmission SNR in the simulation ranged from 60 dB to 100dB. As explained in [18], the actual transmission SNR ranges from 52dB to 112dB, which roughly corresponds to a received SNR in the range of -40 dB to 20 dB, due to high path loss. In addition, the corresponding transmission power is in the range of -20 dBm to 40 dBm. Thus, the range of the transmission SNR in this simulation was close to the actual range and is reasonable for implementation. In the following simulation, the “traditional approach” mentioned is the modified approach presented in [20] and [21] that was presented in Section 4.1.

The sum rate was averaged over the channel coherence interval, i.e., the effect of the pilot is taken into account in calculating the sum rate. The channel coherence interval is $T_{\mathrm{c}} = 500$, which was regarded as a reasonable value in [26] and the reference therein. The antennas’ configurations were as follows. For $N = 252$, the number of antennas in each of the two dimensions was $N_{\mathrm{az}} = 7, N_{\mathrm{el}} = 36$; for the second case, $N = 2716$ with $N_{\mathrm{az}} = 16, N_{\mathrm{el}} = 76$; and, $N = 4832$, with $N_{\mathrm{az}} = 32, N_{\mathrm{el}} = 151$.

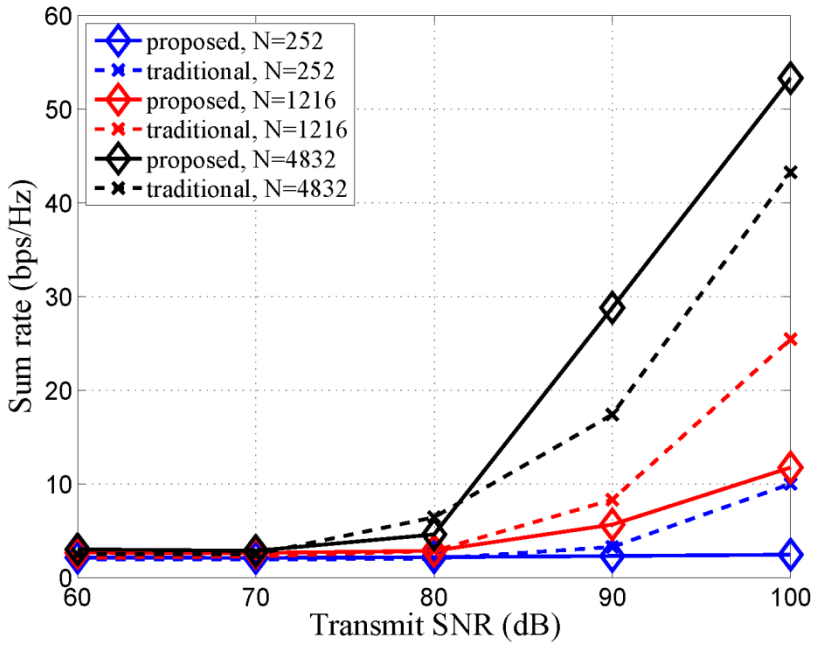

Fig. 3. Illustration of the sum rates versus the SNR.

The sum rates versus the SNR are shown in **Fig. 3**. It can be seen that the sum rates of all approaches increase with the increase of the SNR, while, the increase of the number of antennas also results into an increase of the sum rate. This is because the received signal power is increased with the increase of the number of the transmitting or receiving antennas. In other words, the increase of the sum rate with the number of the antennas reflects the diversity gain, and the diversity gain depends on the number of the antennas rather than the employed approaches. Additionally, the proposed approach tends to achieve a higher sum rate than the existing approach when the number of the antennas increases. This is because the accuracy of the channel estimate of the traditional approach does not vary, while the accuracy of the DOA estimate of the proposed approach improves as the number of the antennas increases. In other words, the increase of the sum rate with the increase of the number of the antennas demonstrates that the estimation performance of the proposed approach improves at the same time. However, the increase of the sum rate with the number of the antennas of the traditional approach only demonstrates the diversity gain. This result verifies the asymptotic analysis of the proposed approach in Theorem 1.

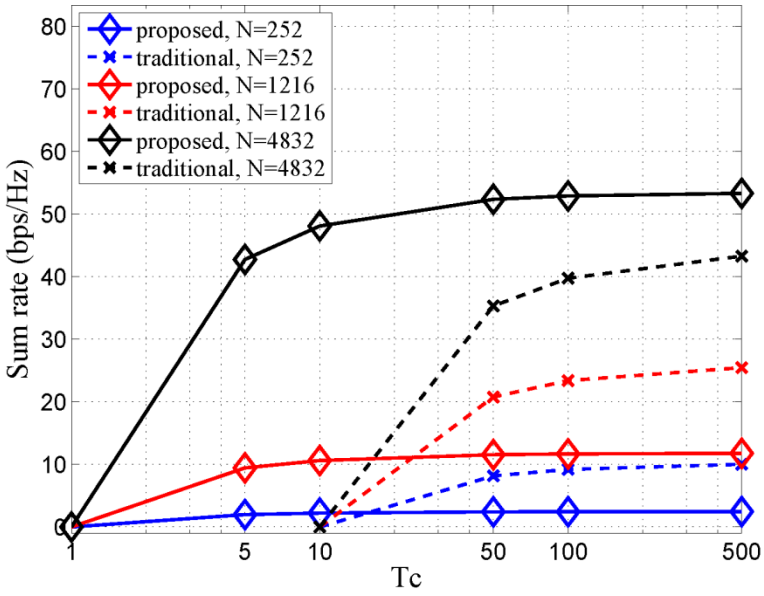

Fig. 4. Illustration of the sum rates versus the channel coherence interval.

The sum rates versus the channel coherence interval are shown in **Fig. 4** for a transmission SNR of 100 dB. It can be seen that the sum rates of both approaches increase with the increase of the channel coherence interval. In addition, the proposed approach is applicable in the case that the channel coherence interval is less than 10, while the traditional approach is not applicable in this case. It can also be seen that the proposed approach has a higher sum rate for the whole range of channel coherence interval values when the number of antennas is $N = 4832$. Note that in the case of long channel coherence intervals, i.e., when T_c = 500 in **Fig. 4**, the proposed approach also performs better than the traditional approach. The ratio of the number of the pilots versus the channel coherence interval is 1/500 for the proposed approach and 10/500 for the traditional approach. Thus, the reduced number of pilot symbols for the proposed approach can save little time for data transmission. In other words, the better performance of the proposed approach in the case of Tc = 500 demonstrates that the proposed approach can estimate the channels more accurately. These results verify that the proposed approach can estimate channels with higher precision and is feasible for a wider range of the channel coherence interval.

However, when the number of antennas is 252 or 1216, the proposed approach tends to perform worse than the traditional approach as the channel coherence interval increases. The reason is as follows. The channel estimation accuracy of the proposed approach relies on the number of antennas, while the traditional approach does not. When the number of antennas is not large enough, e.g., 252 or 1216, the estimation error of the proposed approach is larger than that of the traditional approach. When the channel coherence interval is short, the proposed approach benefits from the reduced number of pilot symbols and thus performs

better than the traditional approach in these cases. When the channel coherence interval is long, the proposed approach does not benefit from the reduced number of pilot symbols, so it is outperformed by the traditional approach.

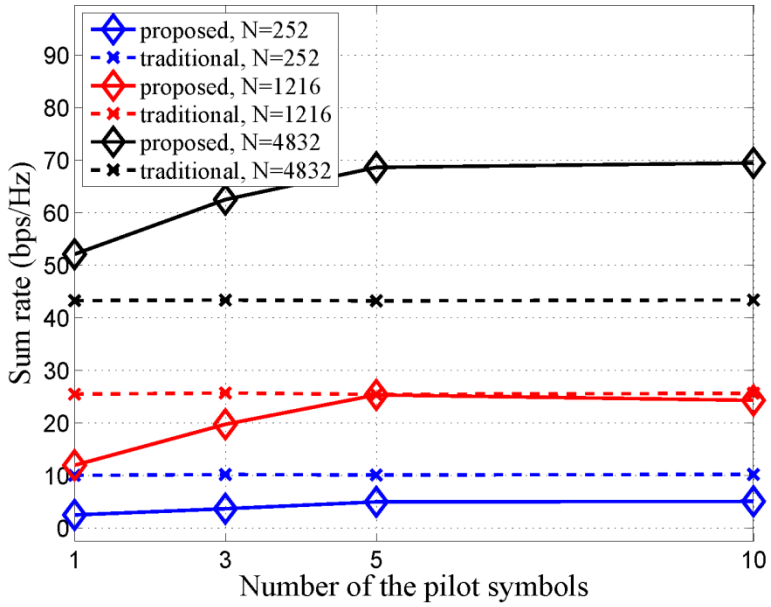

Fig. 5. Illustration of the sum rates versus the number of pilot symbols.

Finally, the sum rates versus the number of the pilot symbols are shown in **Fig. 5**. The transmission SNR is 100 dB, and the channel coherence interval is $T_c = 500$. When the number of the pilot symbols increases, the proposed approach can also be modified to utilize the additional pilots. With the increased number of the pilot symbols, the MSs can be partitioned into different groups, where the MSs in the same group share the same pilot sequence while MSs in different groups employ an orthogonal pilot sequence. Then, the proposed approach can estimate the DOAs and path losses of the MSs group by group. In **Fig. 5**, the number of the pilot symbols for the traditional approach is fixed, and the number of the pilot symbols is $K = 10$. It can be seen that the sum rate of the proposed approach increases with the increase of the number of the pilot symbols. In addition, the proposed approach tends to perform comparably to the traditional approach for $N = 252$ and $N = 1216$, but outperforms it for $N = 4832$. Since the difference between the DOAs of two MSs increases with this partition of the MSs, the proposed approach can distinguish the MSs more easily. Thus, the proposed approach also benefits from an increase of the number of pilot symbols.

6. Conclusion

In this paper, a channel estimation approach is proposed for mm-wave MU-MIMO systems. This method uses fewer pilot symbols than existing approaches and can estimate the channels without interference in the large system limit. Thus, the proposed approach is feasible for mm-wave MU-MIMO systems.

References

[1] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101-107, Jun. 2011. Article (CrossRef Link)

[2] T. S. Rappaport, S. Sun, R. Mayzus, et al., "Millimeter wave mobile communications for 5G cellular: it will work!," *IEEE Access*, vol. 1, pp. 335-349, May 2013. Article (CrossRef Link)

[3] A. M. Sayeed and N. Behdad, "Continuous aperture phased MIMO: basic theory and applications," in *Proc. of Allerton Conf. Commun., Control, and Comput.*, Allerton, USA, pp. 1196-1203, Sep. 2010. Article (CrossRef Link)

[4] A. M. Sayeed and N. Behdad, "Continuous aperture phased MIMO: a new architecture for optimum line-of-sight links," in *Proc. of IEEE Int. Symp. Antennas and Propagation*, pp. 293-296, Jul. 2011. Article (CrossRef Link)

[5] F. Rusek, D. Persson, B. K. Lau, et al., "Scaling up MIMO: opportunities and challenges with very large arrays," *IEEE Signal Processing Mag.*, vol. 30, no. 1, pp. 40-60, Jan. 2013. Article (CrossRef Link)

[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. Article (CrossRef Link)

[7] O. El Ayach, S. Rajagopal, S. Abu-Surra, et al., "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499-1513, Mar. 2014. Article (CrossRef Link)

[8] A. Adhikary, J. Nam, A. J.-Y. Ahn, et al., "Joint spatial division and multiplexing-the large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441-6463, Oct. 2013. Article (CrossRef Link)

[9] A. Adhikary, E. Al Safadi, M. K. Samimi, et al., "Joint spatial division and multiplexing for mm-wave channels," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1239-1255, Jun. 2014. Article (CrossRef Link)

[10] H. Yin, D. Gesbert, M. Filippou, et al., "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. Article (CrossRef Link)

[11] W. U. Bajwa, J. Haupt, A. M. Sayeed, et al., "Compressed channel sensing: a new approach to estimating sparse multipath channels," in *Proc. of IEEE*, vol. 98, no. 6, pp. 1058-1076, Jun. 2010. Article (CrossRef Link)

[12] A. Alkhateeb, O. El Ayach, G. Leus, et al., "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831-846, Oct. 2014. Article (CrossRef Link)

[13] Y. Peng, Y. Li, and P. Wang, "An enhanced channel estimation method for millimeter wave systems with massive antenna arrays," *IEEE Commun. Lett.*, vol. 19, no. 9, pp. 1592-1595, Sep. 2015. Article (CrossRef Link)

[14] D. C. Araujo, A. L. F. de Almeida, J.Axnas, et al., "Channel estimation for millimeter-wave very-large MIMO systems," in *Proc. 22nd European Signal Processing Conference*, Lisbon, Portugal, pp. 81-85, Sep. 2014.

[15] J.Mo, P. Schniter, N. G. Prelcic, et al., "Channel estimation in millimeter wave MIMO systems with one-bit quantization," in *Proc. of 48th Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, USA, pp. 957-961, Nov. 2014. Article (CrossRef Link)

[16] Y. Han, S. Jin, X. Li, et al., "A joint SDMA and interference suppression multiuser transmission scheme for millimeter-wave massive MIMO systems," in *Proc. of Sixth Inter. Conf. Wireless Commun. and Signal Processing*, Hefei, China, pp. 1-5, Oct. 2014. Article (CrossRef Link)

[17] J. Brady, N. Behdad, and A. M. Sayeed, "Beamspace MIMO for millimeter-wave communications: system architecture, modeling, analysis, and measurements," *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3814-3827, Jul. 2013. Article (CrossRef Link)

[18] J. Brady and A. Sayeed, "Beamspace MU-MIMO for high-density gigabit small cell access at millimeter-wave frequencies," in *Proc. of IEEE 15th Int. Workshop Signal Process. Advances Wireless Commun.*, Toronto, Canada, pp. 80-84, Jun. 2014. Article (CrossRef Link)

[19] G. H. Song, J. Brady, and A. Sayeed, "Beamspace MIMO transceivers for low-complexity and near-optimal communication at mm-wave frequencies," in *Proc. of 2013 IEEE Inter. Conf. Acoustics, Speech, Signal Process.*, Vancouver, Canada, pp. 4394-4398, May 2013. Article (CrossRef Link)

[20] P. V. Amadori and C. Masouros, "Low RF-complexity millimeter-wave beamspace-MIMO systems by beam selection," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2212-2223, Jun. 2015. Article (CrossRef Link)

[21] H. Q. Ngo,E. G. Larsson, and T. L. Marzettat, "Uplink power efficiency of multiuser MIMO with very large antenna arrays," in *Proc. of 49th Annual Allerton Conf. Commun., Control, and Computing*, Monticello, USA, pp. 1272-1279, Sep. 2011. Article (CrossRef Link)

[22] S. Payami, M. Shariat, M. Ghoraishi, et al., "Effective RF codebook design and channel estimation for millimeter wave communication systems," in *Proc. of IEEE Inter. Conf. Commun. Workshop*, London, Britain, pp. 1226-1231, Jun. 2015. Article (CrossRef Link)

[23] H. T. Friis, "A note on a simple transmission formula," in *Proc. of the IRE*, vol. 34, no. 5, pp. 254-256, May 1946. Article (CrossRef Link)

[24] C. A. Balanis, Antenna Theory: Analysis and Design, Wiley-Interscience, 2005.

[25] H. Q. Ngo and E. G. Larsson, "EVD-based channel estimation in multicell multiuser MIMO systems with very large antenna arrays," in *Proc. of IEEE Int. Conf. Acoustics, Speech and Signal*

Processing, Kyoto, Japan, pp. 3249–3252, Mar. 2012. Article (CrossRef Link)

[26] P. Schniter and A. Sayeed, "Channel estimation and precoder design for millimeter-wave communications: the sparse way," in *Proc. of 48th Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, USA, pp. 273-277, Nov. 2014. Article (CrossRef Link)

Anzhong Hu received the B.Eng. degree from Zhejiang University of Technology, Hangzhou, China, in 2009, and the Ph.D. degree from Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2014. He is now a lecturer at School of Communication Engineering, Hangzhou Dianzi University. His current research interests include channel estimation and array processing.