Tan *et al. EURASIP Journal on Wireless Communications and Networking* (2017) 2017:181
DOI 10.1186/s13638-017-0966-4

EURASIP Journal on Wireless Communications and Networking

RESEARCH Open Access

Spectral efficiency of multi-user millimeter wave systems under single path with uniform rectangular arrays

Weiqiang Tan[1,2], Shi Jin[2*], Chao-Kai Wen[3] and Tao Jiang[4]

Abstract

In this paper, we investigate the downlink achievable ergodic spectral efficiency (SE) of a single-cell multi-user millimeter wave system, in which a uniform rectangular array is used at the base station (BS) to serve multiple single-antenna users. We adopt a three-dimensional channel model by considering both the azimuth and elevation dimensions under single-path propagation. We derive the achievable ergodic SE for this system in with maximum ratio transmission precoding. This analytical expression enables the accurate and quantitative evaluation of the effect of the number of BS antennas, signal-to-noise ratio (SNR), and the crosstalk (squared inner product between different steering vectors) which is a function of the angles of departure (AoD) of users and the inter-antenna spacing. Results show that the achievable ergodic SE logarithmically increases with the number of BS antennas and converges to a value in the high SNR regime. To improve the achievable ergodic SE, we also propose a user scheduling scheme based on feedback of users' AoD information and obtain the maximum achievable ergodic SE. Furthermore, we consider a dense user scenario where every user's AoD becomes nearly identical and then derive the system's minimum achievable SE.

Keywords: Millimeter wave communications, Multi-user MIMO, Maximum ratio transmission, Spectral efficiency, User scheduling, Uniform rectangular array

1 Introduction

With the proliferation of smart wireless services such as high-definition video streaming, cell broadcasting, and mobile TV, mobile data traffic is envisioned to grow 1000-fold by 2020 [1]. To meet the predicted traffic demand, millimeter wave (mmWave) communications, which operates in the 30–300 GHz band, appear to be a promising candidate for next generation cellular systems that support multiple gigabit-per-second data rates [2]. Some recent results have demonstrated that the use of mmWave communications helps to achieve gigabytes per second (Gbps) data rates in both indoor [3] and outdoor wireless networks [4].

One salient feature of mmWave communications is the significant decrease of carrier wavelength, which allows a large number of antennas to be packed into a base station (BS) or an access point. The excessive use of transmit antennas, known as massive multiple-input multiple-output (MIMO) technology [5, 6], provides substantial array gains to combat severe path loss and establishes reliable communication links. Furthermore, deployment of large antenna arrays at the BS enables efficient precoding for multiple data streams, thus improves the spectral efficiency and energy efficiency [7, 8]. To take advantage of the available space at the BS of a massive MIMO system, it is desirable to arrange the antenna elements as a uniform linear array (ULA), a cubic array, or a uniform rectangular array (URA)[9].

Most importantly, rectangular or cubic arrays enable not only horizontal beam pattern adaptation but also vertical beam pattern adaptation, that is, they enable three-dimensional (3D) beamforming, which provides an additional degree-of-freedom for interference suppression [10]. Therefore, exploitation of 2D antenna array

*Correspondence: jinshi@seu.edu.cn
[2]National Communications Research Laboratory, Southeast University, Nanjing 210096, People's Republic of China
Full list of author information is available at the end of the article

configurations such as uniform circular array (UCA) or URA in mmWave MIMO systems is desirable.

1.1 Related works

Several aspects of mmWave MIMO communications, including channel measurements [11–13], channel estimation [14–16], hybrid analog/digital precoding design [17–20], multiple amplify-and-forward (AF) relaying networks [21–23], and multiple-cell cooperative communication [24], have been investigated in prior works. In particular, various potential techniques for mmWave cellular systems were specified in [11] by considering mmWave channel characteristics. The measurements in [12, 13] demonstrated that mmWave channels have limited light-of-sight (LoS) components and are related to the steering vectors which depend on the antenna array topologies of the transmitter and receiver. To enhance the beamforming gains and enable multiplexing of multiple data streams, hybrid precoding techniques were proposed in [19], where the processing was divided between the analog and digital domains. In [17], a hybrid precoding technique that requires only partial knowledge of mmWave channels was presented. The work in [19] proposed a low-complexity hybrid analog/digital scheme, which effectively relaxes the hardware constraints for downlink multi-user mmWave systems. A similar investigation was conducted in [14]. In [15], a compressed sensing-based channel estimation was proposed by exploiting the sparse nature of large mmWave MIMO channels. By leveraging the static nature of the angle of arrival and angle of departure (AoD), the work in [16] studied the performance of zero-forcing (ZF) beamforming with limited feedback. More recently, authors in [18] studied a multi-user MIMO downlink transmission scheme over Rician fading channels, where the BS exploits statistical channel state information (CSI). From these prior works, we notice that fully utilizing static or statistical information is one of the key requirements for massive mmWave MIMO systems.

Despite these existing body of literature, only few analytical results are available on the achievable ergodic spectral efficiency (SE) of multi-user mmWave MIMO systems, particularly for the 3D channel model. Owing to different propagation characteristics at such high frequencies, the coverage and rate trends differ drastically from the conventional cellular systems. In [25], the coverage and rate performance were studied for mmWave systems. The results of [25] showed that the achievable rate is sensitive to the density of BSs and the blockage distribution. In multi-user mmWave MIMO systems, linear precoding method such as ZF can perform almost as well as dirty paper coding, the multi-user mmWave MIMO system analysis of which is studied in [26]. However, the computation complexity due to the inverse of large dimensional matrix makes it inappropriate for real-time processing. In contrast, maximum ratio transmission (MRT) precoding is simple and exhibit near optimal performance (though worse than ZF) has received significant research interest [5]. Hence, studying the achievable ergodic SE of multi-user mmWave systems using MRT precoding is of special interest.

1.2 Summary of contributions

In this paper, we investigate the achievable ergodic SE with multiple antennas and URA configuration at the BS. In particular, we develop analytical expressions to gain insight into the performance. Our contributions are summarized as follows:

- We present an exact closed-form expression for the achievable ergodic SE by using MRT precoding under the assumption that the BS has perfect channel state information. We also present the achievable ergodic SE that is valid for massive numbers of antenna elements in the high-SNR regime. We make use of these results to discuss the impact of the number of BS antennas, SNR, crosstalk, users' AoD, and the inter-antenna spacing.
- To improve the performance of MRT precoding, a user scheduling scheme is proposed by exploiting statistical users' AoD information. We select a set of user group satisfying the orthogonal criterion and derive the corresponding maximum achievable SE. We also consider a dense user scenario and present an expression for the minimum achievable SE.

The remainder of this paper is organized as follows. Section 2 introduces the system model. Section 3 derives an exact closed-form expression of the achievable ergodic SE for a finite number of users and antenna elements. Based on these analytical results, several practical insights are presented. Numerical and simulation results are also provided in Section 4, and we conclude the paper in Section 5.

Notation—Throughout the paper, matrices and vectors are expressed as upper and lower case boldface letters, respectively. Moveover, $(\cdot)^H$ denotes the conjugate transpose; $\|\cdot\|$ and $|\cdot|$ represent the Euclidean norm and absolute value, respectively; $\mathbb{E}\{\cdot\}$ and $\otimes$ represent the expectation operator and kronecker product, respectively; $\mathbf{I}_M$ denotes an $M \times M$ identity matrix; e denotes the base of the natural logarithm; the notation $x \sim \mathcal{C}N(0,1)$ means that x is the complex Gaussian random variable with mean zero and variance one; e is the Euler-Mascheroni constant; $\mathrm{Ei}(x) = -\int_{-x}^{\infty} \frac{e^{-t}}{t} dt$ denotes the exponential integral function ([27], Eq. (8.211.1)); $E_h(x) = \int_1^{\infty} t^{-h} e^{-xt} dt$ is the exponential integral of order h([27], page xxxv), and $\Gamma(a,x) = \int_x^{\infty} e^{-t} t^{a-1} dt$ is the upper incomplete gamma function ([27], Eq. (8.350.2)).

2 System model

We consider the downlink of a single-cell multi-user mmWave MIMO system where a URA configuration is deployed at the BS as shown in Fig. 1. We assume that the BS is equipped with N_t transmit antennas to simultaneously serve K single-antenna users ($K \ll N_t$). The received signal at the k-th user can be expressed as

$$y_k = \sqrt{\rho g_k}\, \mathbf{h}_k^H \mathbf{w}_k s_k + \sqrt{\rho g_k} \sum_{j=1, j\neq k}^{K} \mathbf{h}_k^H \mathbf{w}_j s_j + n_k, \tag{1}$$

where ρ and g_k ($k = 1, \ldots, K$) denote the average SNR and antenna array gain for user k, $\mathbf{w}_k$, and $\mathbf{w}_j$ is unit-norm precoding vectors of user k and user j, respectively, s_k and s_j denote the transmit zero-mean Gaussian symbols for user k and user j, respectively, $n_k \sim \mathcal{CN}(0,1)$ is the additive noise at the k-th user, and $\mathbf{h}_k \in \mathbb{C}^{N_t \times 1}$ is the mmWave channel vector from the BS to the k-th user. We assume that the inputs must satisfy a transmit power constraint such that $\|\mathbf{w}_j\| = 1$ and $\mathbb{E}\{|s_j|^2\} = 1$ for $j = 1, \ldots, K$ and consider the equal power allocation scheme.

2.1 Channel model

A mmWave channel is composed of a single-path propagation component considering the dominant path and a set of multi-path components because of the highly directional and quasi-optical nature of electromagnetic wave propagation at the mmWave frequency [12]. Hence, $\mathbf{h}_k$ can be accurately modeled as [28–30]

$$\mathbf{h}_k = \beta_k \mathbf{v}\left(\theta_k, \phi_k\right) + \sqrt{\frac{1}{L}} \sum_{l=1}^{L} \beta_{k,l} \mathbf{v}\left(\theta_{k,l}, \phi_{k,l}\right), \tag{2}$$

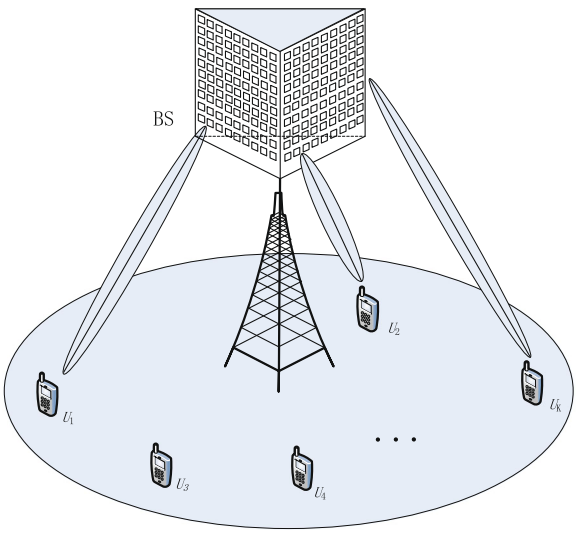

Fig. 1 A schematic diagram of the multi-user MIMO mmWave system, where a BS equipped with a URA serves K users

where β_k and $\beta_{k,l}$ are the complex gain of LoS path and multi-path components for the k-th user, respectively, i.e., β_k ($\beta_{k,l}$) $\sim \mathcal{CN}(0,1)$, L denotes the number of multi-path components, ϕ_k and θ_k are the azimuth and elevation AoD of the k-th user, respectively, and $\mathbf{v}(\theta_k, \phi_k)$ is the steering vector with respects to antenna array configuration.

For a mmWave wireless propagation, the effect of multi-path components are marginal and LoS component is the predominant mode since the path loss of multi-path components is much larger than that of the LoS component. Furthermore, as measurement reports about mmWave channel [28, 31], the magnitude of multi-path components are generally 5 to 10 dB weaker than the LoS predominant component, even there is no multi-path components. Without loss of generality, we neglect the multi-path components and consider only the single-path component; that is, $\beta_{k,l} = 0$ for $l = 1, \ldots, L$. Therefore, the mmWave channel model in (2) can be simplified as

$$\mathbf{h}_k = \beta_k \mathbf{v}_k\left(\theta_k, \phi_k\right). \tag{3}$$

We assume a URA configuration at the BS since it can perform 3D beamforming by employing both the azimuth and elevation dimensions.[1] In the far-field regime, the steering vector for the URA configuration is represented by [32]

$$\mathbf{v}_k\left(\theta_k, \phi_k\right) = \mathbf{v}_{N_x} \otimes \mathbf{v}_{N_y}, \tag{4}$$

where

$$\begin{aligned} \mathbf{v}_{N_x} &= \left[1\ e^{jk_0 d_x \sin\theta_k \cos\phi_k} \cdots e^{j(N_x-1)k_0 d_x \sin\theta_k \cos\phi_k}\right]^T, \\ \mathbf{v}_{N_y} &= \left[1\ e^{jk_0 d_y \sin\theta_k \sin\phi_k} \cdots e^{j(N_y-1)k_0 d_y \sin\theta_k \sin\phi_k}\right]^T \end{aligned} \tag{5}$$

are the steering vectors in the horizontal and vertical directions, respectively. Herein, $k_0 = 2\pi/\lambda$ is the number of waves and λ is the carrier wavelength; N_x and N_y are the numbers of antenna elements placed in the horizontal and vertical dimensions, respectively; d_y and d_x denote the inter-antenna spacing in the horizontal and vertical directions, respectively, which has a linear relationship with λ i.e., $d_x = d_y = \eta\lambda$ with η being an any positive real number. Accordingly, the total number of BS antennas N_t is equal to $N_x \times N_y$. In the following subsection, we present the definition of antenna array gain.

2.2 Antenna array gain

Since we consider a URA configuration at the BS, where an antenna gain changes with the transmit antenna pattern. According to the 3GPP in [33], the horizontal and vertical antenna radiation pattern adopted can be expressed in dB scale as

$$G_H\left(\phi_k\right) = \min\left\{12\left(\frac{\phi_k}{\phi_{3\text{dB}}}\right)^2, A_m\right\} \tag{6}$$

and

$$G_V(\theta_k, \theta_{\text{tilt}}) = \min\left\{12\left(\frac{\theta_k - \theta_{\text{tilt}}}{\theta_{3\text{dB}}}\right)^2, A_m\right\}, \quad (7)$$

where $\phi_{3\text{dB}}$ and $\theta_{3\text{dB}}$ represent the horizontal and vertical half-power beamwidth, respectively, A_m represents the maximum attenuation of the array antennas, and θ_{tilt} denotes the antenna tilting angle which is allowed to be adjusted within the given interval.

Let us denote $G_{\max}$ (in dB) as the maximum antenna gain at the antenna boresight. Then, after combining the antenna attenuations and the maximum antenna gain, the resultant antenna gain in dB scale for the k-th user with azimuth angle ϕ_k and the elevation angle θ_k can be formulated as

$$G_k = G_{\max} - \min\left\{G_H(\phi_k) + G_V(\theta_k, \theta_{\text{tilt}}), A_m\right\}. \quad (8)$$

To simplify the analysis, we assume $A_m = \infty$. This assumption is valid when $G_H(\phi_k) + G_V(\theta_k, \theta_{tilt}) < A_m$. With these assumption, we can obtain the antenna gain in linear scale for the k-th user as

$$g_k = 10^{\left(\frac{G_{\max}}{10} - 1.2\left(\frac{\phi_k}{\phi_{3\text{dB}}}\right)^2 - 1.2\left(\frac{\theta_k - \theta_{\text{tilt}}}{\theta_{3\text{dB}}}\right)^2\right)}. \quad (9)$$

From (9), we observe that the antenna gain g_k depends on the maximum antenna gain, the antenna tilting angle, the azimuth and elevation AoD of the k-th user, and the horizontal and vertical half-power beamwidth, respectively. In practice, once the system configuration is complete, it means the maximum antenna gain, the antenna tilting angle, and the horizontal and vertical half-power beamwidth is fixed. For the AoD information which includes the azimuth and elevation of the k-th user, changes very slowly and remains constant over a coherence time intervals. Compared with the complex gain of LoS path, the AoD information is treated as statistical CSI and assumed to be known. We show the achievable ergodic SE of system in the following subsection.

2.3 Achievable ergodic SE

We assume that the BS has the perfect CSI and employs linear precoding to process the signal before transmitting it to the K users, the achievable ergodic SE of the k-th user can be written as

$$R_k = \mathbb{E}\left\{\log_2(1 + \text{SINR}_k)\right\}, \quad (10)$$

where

$$\text{SINR}_k = \frac{\rho g_k \left|\mathbf{h}_k^H \mathbf{w}_k\right|^2}{1 + \rho g_k \sum_{k=1, k\neq j}^{K} \left|\mathbf{h}_k^H \mathbf{w}_j\right|^2}. \quad (11)$$

Hence, the total achievable ergodic SE of the system is given by

$$R_{\text{sum}} = \sum_{k=1}^{K} R_k. \quad (12)$$

To maximize the total achievable ergodic SE in (12) is identical to maximize the achievable ergodic SE of the k-th user in (11) since users are independent of each other. It is worth pointing out that the achievable ergodic SE depends on the choice of the transmit beamforming vector. To maximize the achievable ergodic SE in (11), the optimization aims to find the beamforming vectors $\mathbf{w}_i$ ($i = 1, \ldots, K$) such that the $\mathbf{w}_k$ must simultaneously maximize the numerator and minimize the denominator of (11). Among different choices, achievable rate achieving nonlinear precoders are known to involve high computational complexity. We show the following lemma that presents the optimal beamforming vector maximizing R_{sum} when the number of BS antennas grows without bound.

Lemma 1 *The achievable ergodic SE in* (11) *is maximized when the channel response vectors between different users are orthogonal. With CSI at the BS, the optimal beamforming vector for a URA configuration that maximizes the achievable SE is asymptotically* (*i.e.,* $N_t \to \infty$) *is given by*

$$\mathbf{w}_k^{opt} = \frac{\mathbf{h}_k}{\|\mathbf{h}_k\|}. \quad (13)$$

Proof In order to maximize (12), we find that the optimal beamforming vector $\mathbf{w}_k$ must simultaneously maximize the numerator and minimize the denominator of (11). We first consider the term of the numerator. For LoS channel response vectors, we have

$$\left|\mathbf{h}_k^H \mathbf{w}_k\right| \leqslant \lambda_{\max}\left(\mathbf{h}_k \mathbf{w}_k^H\right) = \sqrt{N_t \beta_k}, \quad (14)$$

where we have used the Rayleigh quotient law. We can see that this maximum of the numerator in (11) is achieved when $\mathbf{w}_k$ is an eigenvector of $\left|\mathbf{h}_k \mathbf{w}_k^H\right|$ corresponding to the maximum eigenvalue, i.e. $\mathbf{w}_k = \mathbf{h}_k / \|\mathbf{h}_k\|$. At the same time, to maximize the SINR, we must have the following criterion

$$\left|\mathbf{h}_k^H \mathbf{w}_j\right| \to 0. \quad (15)$$

For a URA configuration, substituting (4) into (15), the inner product term can be calculated as

$$\left|\mathbf{h}_k^H \mathbf{w}_j\right| = \frac{1}{\sqrt{N_t}} \frac{\beta_k}{\sqrt{\beta_j}} \sum_{n=1}^{N_x} \sum_{m=1}^{N_y} e^{j(n-1)k_0 d_x \delta_x + j(m-1)k_0 d_y \delta_y}, \quad (16)$$

where

$$\begin{aligned} \delta_x &= \sin\theta_k \cos\phi_k - \sin\theta_j \cos\phi_j \\ \delta_y &= \sin\theta_k \sin\phi_k - \sin\theta_j \sin\phi_j. \end{aligned} \quad (17)$$

By applying Eular's formula ([27], Eq. (1.222.2)), we obtain the following simplified expression

$$\left|\mathbf{h}_k^H \mathbf{w}_j\right| = \frac{1}{\sqrt{N_t}} \frac{\beta_k}{\sqrt{\beta_j}} \left| \frac{\sin\left(\frac{N_x k_0 d_x}{2}\delta_x\right)}{\sin\left(\frac{k_0 d_x}{2}\delta_x\right)} \frac{\sin\left(\frac{N_y k_0 d_y}{2}\delta_y\right)}{\sin\left(\frac{k_0 d_y}{2}\delta_y\right)} \right|, \tag{18}$$

which converges to zero when the number of BS antennas grows without bound and the AoDs of users are distinct (i.e. $\phi_k \neq \phi_j$ and $\theta_k \neq \theta_j$),[2] which leads to $\delta_x \neq 0$ and $\delta_y \neq 0$. Thus, the above criterion in (15) is satisfied. □

From Lemma 1, it is evident that the optimal beamforming vector $\mathbf{w}_k^{\text{opt}}$ is obtained by adopting MRT precoding. This is because when the number of BS antennas grows without bound, the channel steering vectors between different users become asymptotically orthogonal to each other. This implies that the interference between different users is effectively suppressed. Then, the achievable SE attains its maximum value.

By utilizing (3) and (13), the achievable ergodic SE of the k-th user can be rewritten as

$$R_k = \mathbb{E}\left\{\log_2\left(1+\frac{\rho g_k \left|c(\beta_k \mathbf{v}_k)^H \beta_k \mathbf{v}_k\right|^2}{1+\rho g_k \sum_{j=1,j\neq k}^{K} \left|c(\beta_k \mathbf{v}_k)^H \beta_j \mathbf{v}_j\right|^2}\right)\right\}, \tag{19}$$

where $c = 1/\|\mathbf{h}_k\| = 1/\sqrt{|\beta_k| N_t}$ denotes the normalized factor of the precoding vector while the expectation in (19) is across all channel realizations of the complex gain β_k and steering vector $\mathbf{v}_k$. We assume the azimuth and elevation angles of users are known a priori information because they can be obtained in practice via feedback.[3]

3 Achievable SE analysis

In this section, we derive a new exact closed-form expression of the achievable ergodic SE for an arbitrary number of BS antennas and investigate its behavior in the high-SNR regime. Based on the analytical expression, we also evaluate the maximum achievable ergodic SE by a user scheduling method and the system's minimum achievable ergodic SE in a dense user scenario.

3.1 Achievable SE analysis

In this subsection, we focus on deriving closed-form expressions of the achievable ergodic SE by analyzing the SINR in (19). The following theorem is useful to calculate on the achievable ergodic SE when a URA configuration is employed at the BS.

Theorem 1 *An exact analytical expression for the achievable ergodic SE with a URA configuration at the BS and MRT precoding is given by*

$$R_k = \log_2(e)\left(\sum_{j=1}^{K} \frac{a_j^{K-2} e^{1/\rho g_k a_j}}{\prod_{k\neq j} a_j - a_k} \mathrm{Ei}\left(\frac{1}{\rho g_k a_j}\right) - \sum_{j=1,j\neq k}^{K} \frac{a_j^{K-3} e^{1/\rho g_k a_j}}{\prod_{k\neq j} a_j - a_k} \mathrm{Ei}\left(\frac{1}{\rho g_k a_j}\right)\right), \tag{20}$$

where $\{a_j\}_{j=1}^{K}$ *denote the crosstalk coefficients, i.e.,*

$$a_j = \frac{\left|\mathbf{v}_k^H \mathbf{v}_j\right|^2}{\|\mathbf{v}_j\|^2}, \tag{21}$$

where the steering vector $\mathbf{v}_k$ *is defined in* (4)*. Note that when the steering vector* $\mathbf{v}_k$ *equals to the itself* $\mathbf{v}_k$*, we have* $a_k = N_t$*.*

Proof We first factorize (19) as

$$R_k = \underbrace{\left\{\log_2\left(1+\rho g_k \underbrace{\sum_{j=1}^{K} \left|c(\beta_k \mathbf{v}_k)^H \beta_j \mathbf{v}_j\right|^2}_{X}\right)\right\}}_{I_1} - \underbrace{\left\{\log_2\left(1+\rho g_k \underbrace{\sum_{j=1,j\neq k}^{K} \left|c(\beta_k \mathbf{v}_k)^H \beta_j \mathbf{v}_j\right|^2}_{Y}\right)\right\}}_{I_2}. \tag{22}$$

Let

$$X_j = \left|c(\beta_k \mathbf{v}_k)^H \beta_j \mathbf{v}_j\right|^2 \tag{23}$$

for $j = 1, \ldots, K$. Then, we evaluate X and Y in (22) to obtain

$$X = \sum_{j=1}^{K} X_j = \sum_{j=1}^{K} a_j \frac{\left|\beta_k^H \beta_j\right|^2}{\left|\beta_j\right|^2} \tag{24}$$

and

$$Y = \sum_{j=1,j\neq k}^{K} X_j = \sum_{j=1,j\neq k}^{K} a_j \frac{\left|\beta_k^H \beta_j\right|^2}{\left|\beta_j\right|^2}. \tag{25}$$

Since the users' AoDs change very slowly and is approximately constant over a coherence time interval, it is obvious that the squared inner products between different steering vectors $a_1, a_2, \ldots, a_K$ are known a priori.

Therefore, $X_1, X_2, \ldots, X_K$ are i.i.d. exponential distributed random variables with mean $1/N_t$ and variance $1/N_t^2$. In addition, with the help of the results in ([34], Page 552), the pdfs of $X = \sum_{j=1}^{K} X_j$ and $Y = \sum_{j=1, j\neq k}^{K} X_j$ can be expressed as

$$f_X(x) = \sum_{j=1}^{K} \frac{a_j^{K-2}}{\prod_{k\neq j} a_j - a_k} e^{-\frac{x}{a_j}} \tag{26}$$

and

$$f_Y(y) = \sum_{j=1, j\neq k}^{K} \frac{a_j^{K-3}}{\prod_{k\neq j} a_j - a_k} e^{-\frac{y}{a_j}}, \tag{27}$$

respectively. By utilizing the pdfs in (26) and (27), we can evaluate I_1 and I_2 according to

$$I_1 = \int_0^{\infty} \log_2\left(1 + \rho g_k X\right) f_X(x)\, dx, \tag{28}$$

and

$$I_2 = \int_0^{\infty} \log_2\left(1 + \rho g_k Y\right) f_Y(y)\, dy. \tag{29}$$

By substituting (26) and (27) into (28) and (29), respectively, and applying the integral identity ([35], Eq. (13))

$$\int_0^{\infty} \ln(1 + a\lambda)\lambda^{q-1} e^{-b\lambda} d\lambda = (q-1)!\, e^{b/a} b^{-q} \sum_{h=1}^{q} E_h\left(\frac{b}{a}\right). \tag{30}$$

We can obtain

$$I_1 = \log_2(e) \sum_{j=1}^{K} \frac{a_j^{K-2} e^{\frac{1}{\rho g_k a_j}}}{\prod_{k\neq j} a_j - a_k} \mathrm{Ei}\left(\frac{1}{\rho g_k a_j}\right), \tag{31}$$

and

$$I_2 = \log_2(e) \sum_{j=1, j\neq k}^{K} \frac{a_j^{K-3} e^{\frac{1}{\rho g_k a_j}}}{\prod_{k\neq j} a_j - a_k} \mathrm{Ei}\left(\frac{1}{\rho g_k a_j}\right). \tag{32}$$

Subsequently, by substituting (31) and (32) into (22), followed by some basic algebraic manipulations, we arrive at the desired result. □

In Theorem 1, we observe that the closed-form expression for the achievable ergodic SE is a function of the average SNR the number of users and crosstalk coefficients, which depends on the number of BS antennas, inter-antenna spacing and the azimuth and elevation AoDs. This makes it very different to understand the impact when the number of BS antennas is increased. To gain further insights, we have the following result.

Theorem 2 *In the high-SNR regime (i.e., as $\rho \to \infty$), the achievable ergodic SE can be written as*

$$\lim_{\rho\to\infty} R_k = \log_2(e) \left(\sum_{j=1}^{K} \frac{a_j^{K-1}\left(\ln(a_j) - e\right)}{\prod_{k\neq j} a_j - a_k} - \sum_{j=1, j\neq k}^{K} \frac{a_j^{K-2}\left(\ln(a_j) - e\right)}{\prod_{k\neq j} a_j - a_k} \right), \tag{33}$$

where a_j was defined in (21).

Proof When SNR increases without bound, (19) can be rewritten as

$$\lim_{\rho\to\infty} R_k = \mathbb{E}\left\{ \log_2 \left(\frac{\sum_{j=1}^{K} \left| c(\beta_k \mathbf{v}_k)^H \beta_j \mathbf{v}_j \right|^2}{\sum_{j=1, j\neq k}^{K} \left| c(\beta_k \mathbf{v}_k)^H \beta_j \mathbf{v}_j \right|^2} \right) \right\} \tag{34}$$

Then, we can factorize it into as

$$\lim_{\rho\to\infty} R_k = \underbrace{\left\{ \log_2 \left(\underbrace{\sum_{j=1}^{K} \left| c(\beta_k \mathbf{v}_k)^H \beta_j \mathbf{v}_j \right|^2}_{X} \right) \right\}}_{I_3} - \underbrace{\left\{ \log_2 \left(\underbrace{\sum_{j=1, j\neq k}^{K} \left| c(\beta_k \mathbf{v}_k)^H \beta_j \mathbf{v}_j \right|^2}_{Y} \right) \right\}}_{I_4}, \tag{35}$$

where the pdfs of the random variable X and Y are given in (26) and (27), respectively. We now evaluate I_3 and I_4 according to

$$I_3 = \int_0^{\infty} \log_2(x) f_X(x)\, dx \tag{36}$$

and

$$I_4 = \int_0^{\infty} \log_2(y) f_Y(y)\, dy. \tag{37}$$

By substituting the pdfs of the random variable X and Y into (37) and applying the integral identity ([27], Eq. (4.331.2))

$$\int_0^{\infty} \ln(\lambda) e^{-b\lambda} d\lambda = -\frac{1}{b}(e + \ln b). \tag{38}$$

We obtain

$$I_3 = \log_2(e) \sum_{j=1}^{K} \frac{a_j^{K-2}\left(\ln(a_j) - e\right)}{\prod_{k\neq j} a_j - a_k} \tag{39}$$

and

$$I_4 = \log_2(e) \sum_{j=1, j\neq k}^{K} \frac{a_j^{K-3}\left(\ln\left(a_j\right) - e\right)}{\prod_{k\neq j} a_j - a_k}. \tag{40}$$

Finally, substituting (39) and (40) into (35) concludes the proof. □

Based on Theorem 2, we observe that in high-SNR regime, the achievable ergodic SE is a function of the number of BS antennas, the number of users and the crosstalk, which converges to a constant. The reason is both the signal power and the interference power increase as the SNR. This implies that the performance of the system with MRT precoding can severely deteriorate in a high SNR regime. More importantly, from Theorems 1 and 2, we observe the involvement of the crosstalk coefficients. According to the definition of a_j in (21), we see that the crosstalk coefficient mainly depends on the particular antenna array configuration at the BS. Once the BS deployment is completed, the crosstalk coefficient a_j can be easily acquired. In addition, other antenna array configuration such as UCA or ULA can be easily applied to Theorems 1 and 2 by plugging the corresponding crosstalk coefficients. The corresponding crosstalk coefficients a_j for these antenna arrangements can be found in [32, 36].[4]

In order to proceed, substituting (4) into (21), the crosstalk coefficients a_j for URA configuration can be expressed as

$$a_j = \frac{1}{N_t}\left|\frac{\sin\left(\frac{N_x k_0 d_x}{2}\delta_x\right)}{\sin\left(\frac{k_0 d_x}{2}\delta_x\right)}\frac{\sin\left(\frac{N_y k_0 d_y}{2}\delta_y\right)}{\sin\left(\frac{k_0 d_y}{2}\delta_y\right)}\right|^2, \tag{41}$$

where δ_x and δ_y are defined in (17). The crosstalk coefficients a_j can be further simplified as

$$a_j = \frac{1}{N_t}\left|\frac{\left(1-\cos\left(N_x k_0 d_x \delta_x\right)\right)}{\left(1-\cos\left(k_0 d_x \delta_x\right)\right)}\frac{\left(1-\cos\left(N_y k_0 d_y \delta_y\right)\right)}{\left(1-\cos\left(k_0 d_y \delta_y\right)\right)}\right|. \tag{42}$$

According to (42), we see that the effect of the crosstalk coefficient on the achievable ergodic SE is difficult to derive for the general case. Alternatively, we first observe that the normalized crosstalk coefficient depends on the number of BS antennas in the horizontal and vertical planes, inter-antenna spacing in the corresponding horizontal and vertical planes, and users' AoD. In the following, we present the effect of users' AoD and inter-antenna spacing on the normalized crosstalk coefficient.

Figure 2 shows the normalized crosstalk versus the azimuth angle of the j-th user and different number of BS antennas. In the simulations, the users' elevation angles are fixed to be $\phi_k = \phi_j = \pi/4$, the azimuth angle of the k-th user is fixed to be $\theta_k = \pi/6$, the inter-antenna spacing in horizontal direction is set to $d_x = d_y = \lambda/2$, and the number of BS antennas is $36, 64, 100, 400$, and 1000, respectively. The fluctuations of the curves are due to cos function in (42). The crosstalk coefficient decreases when the number of BS antennas increase or when angle difference between two steering vectors increases. In particular, for the case with very large antenna arrays, a_j decreases rapidly. In this case, nearby users can be easily distinguished by the azimuth angle.

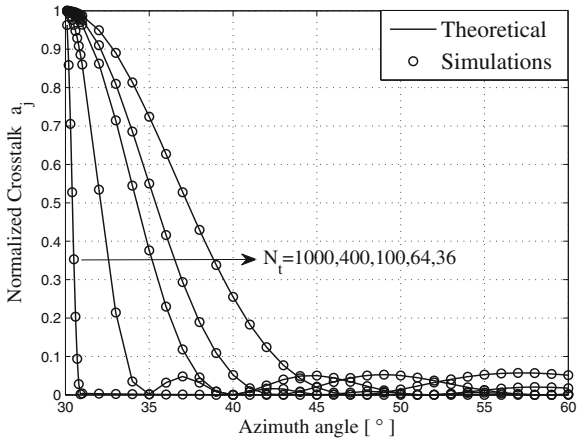

Fig. 2 The normalized crosstalk coefficient versus the azimuth angle of the j-th user for different N_t

Figure 3 shows the normalized crosstalk versus inter-antenna spacing d_y in the vertical direction. In the simulations, we choose fixed angles corresponding to user positions $\theta_k = \theta_j = \pi/4$, $\phi_k = \phi_j = \pi/4$, and the inter-antenna spacing in horizontal direction is set to $d_x = \lambda/2$. The number of antennas in both horizontal and the vertical directions is eight; thus, the total number of antennas

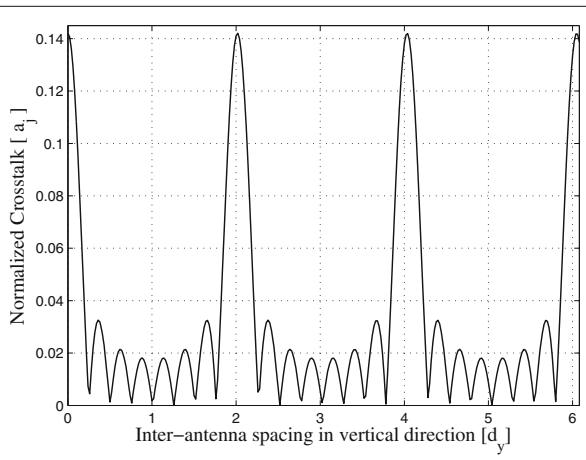

Fig. 3 The normalized crosstalk versus inter-antenna spacing in the vertical direction

is $8 \times 8 = 64$. We see that a_j is a periodic function, due to the properties of the cos function. To minimize the value of a_j, the vertical inter-antenna spacing d_y is adjusted according to the BS size given fixed d_x and N_t because the users' AoDs are not distinguishable in some special scenarios such as a dense region or hot spot.

In order to improve the achievable ergodic SE, we now propose a user scheduling scheme by exploiting azimuth and elevation AoDs of users. The work in [36] has demonstrated that perfect orthogonality between different user channel vectors does exist for ULA and URA configurations under LoS propagation conditions. Herein, we only consider the URA configuration in the paper since ULA configuration is a special of the URA configuration.

3.2 Maximum and minimum achievable SE

Although Theorems 1 and 2 provide a result on the achievable ergodic SE, reducing the normalized crosstalk to zero is difficult in a practical system with finite antenna elements. To improve the achievable ergodic SE under MRT precoding, we propose a user scheduling scheme by exploiting users' AoD information. User scheduling is well recognized as a promising technique to improve the system performance in communication systems. If performed effectively, then user scheduling can be an attractive solution to mitigate interference [37, 38]. In mmWave MIMO systems, the steering vector is predominant and the users' AoD changes slowly [10, 11]. Therefore, our user scheduling scheme shall use the information of the steering vectors to determine a set of users with low crosstalk coefficients. The user scheduling scheme can be outlined as follows:

- The scheduler acquires AoD information of the users and according to the acquired AoD information, the users satisfying the orthogonal criterion are selected as the scheduled group. The specific process is that the AoD information is firstly estimated at mobile stations by using the estimated algorithm (such as multiple signal classification (MUSIC), estimation of signal parameters via rotational invariance techniques (ESPRIT), and subspace algorithms [10, 39]) and then fed back to BS via feedback of downlink channel. In closed-loop frequency division duplex (FDD) MIMO system, downlink AoD information is usually fed back to BS in forms of codebook or channel quality indicator (CQI)[40, 41].
- According to (42), if the steering vectors satisfy the orthogonal criterion for given an arbitrary finite number of BS antennas, then

$$\cos(N_x k_0 d_x \delta_x) = 1 \text{ or/and } \cos(N_y k_0 d_y \delta_y) = 1. \tag{43}$$

As a result, we have the following conditions for δ_x and δ_y, respectively

$$\frac{N_x d_x \delta_x}{\lambda} = n_x \text{ or/and } \frac{N_y d_y \delta_y}{\lambda} = n_y, \tag{44}$$

where n_x and n_y are any positive numbers. From (43), we obtain the conditions for the orthogonal criterion as follows:

$$N_x \eta \left(\sin\theta_k \cos\phi_k - \sin\theta_j \cos\phi_j\right) = n_x, \tag{45}$$

or/and

$$N_y \eta \left(\sin\theta_k \sin\phi_k - \sin\theta_j \sin\phi_j\right) = n_y. \tag{46}$$

where η is defined in (4).

- The BS transmits data streams to selected users with equal power allocation.

Remark 1 *The conditions in* (45) *and* (46) *guarantee that the users in the selected set $\mathcal{S}$ are mutually orthogonal. Combined with the user scheduling, the MRT precoding is able to obtain a similar performance to the ZF precoding for the reason that the selected user interferences have been completely canceled. Finally, note that our proposed scheduling method focuses on maximizing the achievable ergodic SE; however, in doing so, the fairness among the users is not guaranteed.*

We now analyze the asymptotic performance of the proposed scheduling method.

Proposition 1 *When a set of selected users are mutually orthogonal based on the user scheduling scheme, the inner product of the steering vectors between different links tends to zero. Therefore, the maximum achievable ergodic SE of the system can be expressed as*

$$R_k^{max} = \log_2(e)\, e^{\frac{1}{\rho g_k N_t}} E_1\left(\frac{1}{\rho g_k N_t}\right). \tag{47}$$

Proof The set of selected users are orthogonal to each other. Hence, the achievable ergodic SE of the proposed scheduling scheme is given by

$$R_k^{\max} = \mathbb{E}\left\{\log_2\left(1 + \rho g_k \underbrace{\left|\mathbf{h}_k^H \mathbf{w}_k\right|^2}_{Z}\right)\right\}. \tag{48}$$

We note that the probability density function (pdf) of the random variable Z is given by

$$f_Z(z) = \frac{1}{N_t} e^{-\frac{z}{N_t}}. \tag{49}$$

With (49) in hand, we can evaluate $R_k^{\max}$ as

$$R_k^{\max} = \int_0^\infty \log_2\left(1 + \rho g_k z\right) f_Z(z)\, dz. \tag{50}$$

Substituting the pdf of Z into (50) and with the help of ([27], Eq. (4.337.2)), we obtain

$$R_k^{\max} = -\log_2(e)\, e^{\frac{1}{\rho g_k N_t}} \mathrm{Ei}\left(-\frac{1}{\rho g_k N_t}\right). \tag{51}$$

By applying the identity $E_1(x) = -\mathrm{Ei}(-x)$, we complete the proof. □

Remark 2 *From Proposition* 1*, we observe that the achievable ergodic SE of MRT precoding with the user scheduling method is identical to that of the ZF precoding* [42]*. This is because the MRT transmission scheme with the scheduling criterion in Proposition* 1 *facilitates inter-user interference cancelation. In addition, compared to the ZF precoding, the MRT precoding enjoys a much lower computational complexity and does not involve matrix inverse calculations, whereas the user scheduling scheme described above emphasizes the importance of selecting users for multi-user mmWave MIMO systems.*

Corollary 1 *For the special case of $N_t \to \infty$ or $\rho \to \infty$, the achievable SE reduces to a function with the number of BS antennas and the SNR as follows:*

$$\lim_{N_t\to\infty, or \rho\to\infty} R_k^{max} \to \log_2(e)\left(-\ln\left(\frac{1}{\rho g_k N_t}\right) + O(1)\right). \tag{52}$$

Proof As the number of antennas or the SNR grows without bound, i.e., $N_t \to \infty$, or $\rho \to \infty$, we have

$$\frac{1}{\rho N_t} \to 0. \tag{53}$$

By utilizing the following identity for the exponential integral function given in ([43], Eq. (06.34.06.0007.01), when $z \to 0$, we get

$$E_1(z) \propto -\ln(z) + O(1). \tag{54}$$

By making the number of BS antennas and the SNR grow without bound, and performing by some basic algebraic manipulations, we prove the result. □

From Corollary 1, we see that in the high SNR regime, MRT precoding with orthogonal user scheduling not only reduces the consumption of transmit power but also ensures a high achievable ergodic SE. More importantly, the MRT scheme with user scheduling only needs a small number of channel feedback bits to perform near-ideal interference cancelation because the steering vectors become deterministic. These observations clearly reveal the effectiveness of user scheduling.

We now focus on a dense user deployment. In a dense user scenario, e.g., conference hall, railway station, air-plane, or subway entrances, many devices could be active within close proximity [44]. When users are co-located, the users' AoDs shall become nearly identical. This extreme case introduces very high inter-user interference. In the following proposition, we analyze this specific case and evaluate the minimum achievable ergodic SE of the system.

Proposition 2 *When the users are co-located, the inner product of the steering vectors achieves its maximum value, i.e., $a_j = \left|\mathbf{v}_j^H \mathbf{v}_j\right|^2 / \| \mathbf{v}_j \|^2 \approx N_t$, then the minimum achievable ergodic SE is given by*

$$R_k^{min} = \log_2(e)\, e^{\frac{1}{\rho g_k N_t}} E_K\left(\frac{1}{\rho g_k N_t}\right). \tag{55}$$

Proof Let us begin by rewriting the achievable ergodic SE in (19) as

$$R_k^{\min} = \mathbb{E}\left\{\log_2\left(1 + \frac{\overbrace{\left|c(\beta_k \mathbf{v}_k)^H \beta_k \mathbf{v}_k\right|^2}^{Z}}{\frac{1}{\rho g_k} + \underbrace{\sum_{j=1, j\neq k}^{K} \left|c(\beta_k \mathbf{v}_k)^H \beta_j \mathbf{v}_j\right|^2}_{V}}\right)\right\}. \tag{56}$$

For the sake of simplicity, we define

$$S = \frac{Z}{1/\rho g_k + V}, \tag{57}$$

where the pdf of Z is given by

$$f_Z(z) = \frac{1}{N_t} e^{-\frac{z}{N_t}}. \tag{58}$$

We have $a_j = \left|\mathbf{v}_j^H \mathbf{v}_j\right|^2 / \| \mathbf{v}_j \|^2 \approx N_t$, for $j = 1, \ldots, K$ when all users are co-located. Also $X_j = \left|c(\beta_k \mathbf{v}_k)^H \beta_j \mathbf{v}_j\right|^2$ for $j = 1, \ldots, K$ and $j \neq k$ are i.i.d. exponential random variables. Let $V = \sum_{j=1, j\neq k}^{K} X_j$ and the pdf of V can be expressed as

$$f_V(v) = \frac{v^{K-2} e^{-\frac{v}{N_t}}}{N_t^{K-1}(K-2)!}. \tag{59}$$

Conditioning on V, the pdf of S can be written as

$$f_S(s) = \int_0^\infty f_{S|V}(s|v) f_V(v)\, dv. \tag{60}$$

Substituting the pdf of V in (59) into (60), we have

$$f_S(s) = \int_0^\infty \left(\frac{1}{\rho g_k N_t} + v\right) e^{-\left(\frac{1}{\rho g_k N_t} + v\right)s} \frac{s^{K-2} e^{-\frac{v}{N_t}}}{N_t^{K-1}(K-2)!} dv. \tag{61}$$

With the help of the integration identity ([27], Eq. (3.310)) and ([27], Eq. (3.381.4)), we obtain

$$f_S(s) = \frac{e^{-\frac{s}{\rho g_k N_t}}}{(1+s)^K}\left(\frac{1}{\rho g_k N_t}(1+s)+K-1\right). \quad (62)$$

By calculating the integral in (62), we can easily show that the cdf of S is given by

$$F_S(s) = 1 - \frac{e^{-\frac{s}{\rho g_k N_t}}}{(1+s)^{(K-1)}}. \quad (63)$$

The achievable ergodic SE in (56) can be rewritten as ([45], Eq. (4))

$$R_k^{\min} = \log_2(e)\int_0^\infty \frac{1-F_S(s)}{1+s}ds. \quad (64)$$

Substituting the cdf of S in (63) into (64), and with the aid of ([27], Eq. (3.382.4))

$$\int_0^\infty \frac{e^{-bx}}{(1+x)^N}dx = e^b b^{(N-1)}\Gamma(-N+1,b). \quad (65)$$

We obtain

$$R_k^{\min} = \log_2(e)\, e^{\frac{1}{\rho g_k N_t}}\left(\frac{1}{\rho g_k N_t}\right)^{K-1}\Gamma\left(-(K-1),\frac{1}{\rho g_k N_t}\right). \quad (66)$$

By applying the identity $E_h(x) = x^{h-1}\Gamma(1-h,x)$, we complete the proof. □

From Proposition 2, we draw an interesting conclusion that the $R_k^{\min}$ is function of the SNR, the number of BS antennas and users. $R_k^{\min}$ decreases as the number of users increases. The reason is that $E_K(\cdot)$ is a monotonically decreasing function of K, and contributes toward increasing the inter-user interference. Therefore, increasing number of users in a dense user scenario cannot benefit the achievable SE. This observation is different from the case when the number of BS antennas is increased. Clearly, if we do not perform user scheduling, then the achievable SE tends to zero. The following corollary presents the impact of the SNR and the number of BS antennas on the downlink achievable SE.

Corollary 2 *For the special case of $N_t \to \infty$ or $\rho \to \infty$, the minimum achievable SE is reduced to*

$$\lim_{N_t\to\infty, or\ \rho\to\infty} R_k^{min} = \frac{\log_2(e)}{K-1}. \quad (67)$$

Proof The proof starts by recalling the properties of $E_h(\cdot)$ from ([43], Eq. (06.34.03.0002.01)), when $h > 1$, we have

$$E_h(0) = \frac{1}{h-1}. \quad (68)$$

For the special case of $N_t \to \infty$ or $\rho \to \infty$, which leads to $1/(\rho g_k N_t) \to 0$. Then, applying the above identity, we have

$$E_K\left(\frac{1}{\rho g_k N_t}\right) = \frac{1}{K-1}. \quad (69)$$

Substituting (69) into (55) yields the desired result. □

Corollary 2 showcases that the fixed the number of users, $R_k^{\min}$ converges to constant as the number of BS antennas grows without bound. This because the co-located users brings high inter-user interference and degrades the performance. Therefore, improving the number of BS antennas and high SNR regime in this scenario cannot contribute to the achievable SE.

4 Numerical results

In this section, we provide simulation results to validate the derived analytical expressions. All simulation results were obtained by averaging over 100,000 independent channel realizations. In our simulations, we adopt the micro cell environment in [33] with slight modifications. The maximum antenna gain at the antenna boresight $G_{\max}$ is set to 20 dB, the horizontal and vertical half-power beamwidth $\phi_{3\text{dB}}$ and $\theta_{3\text{dB}}$ are set to 65° and 65°, respectively, and the antenna tilting angle θ_{tilt} is set to 90°. Furthermore, we assume that the users are randomly distributed in a circular cell. The azimuth and elevation angles of user are randomly drawn from $[0, 2\pi]$ and $[-\pi/2, \pi/2]$ interval, respectively [46]. We select eight users from user grouping, the azimuth and elevation angles of whose are listed in Table 1, and channel vectors are generated according to (3) in Section 3.

In Fig. 4, the simulated total achievable SE for MRT and ZF precoding (MRT with scheduling method), as well as, the derived analytical expressions of Theorem 1 and Proposition 1 are plotted against the SNR. In the simulations, the number of BS antennas is set to $8 \times 8 = 64$ and $10 \times 10 = 100$, and the inter-antenna spacing along the horizontal and the vertical direction are set to $d_x = d_y = \lambda/2$. We see that the theoretical results show a good match with the Monte-Carlo simulations across the entire SNR regime, which validates the analytical results.

Table 1 The azimuth and elevation angle of the users ($K = 8$)

	U_1	U_2	U_3	U_4	U_5	U_6	U_7	U_8
ϕ	0.9681	2.1770	3.0080	1.7151	1.0428	0.7703	0.1790	1.2570
θ	−1.5008	0.6122	−0.4638	0.2146	0.5428	−0.7163	0.4769	1.0956

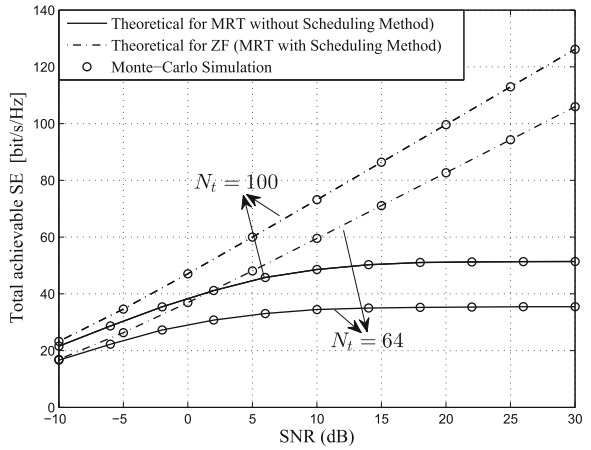

Fig. 4 The total achievable SE versus SNR for $N_t = 64, 100$ and $K = 8$. Results are showed for MRT without scheduling method and ZF precoding (MRT with scheduling method)

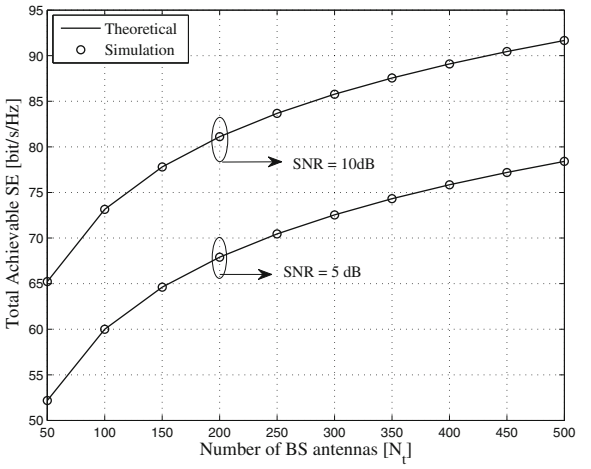

Fig. 5 The total achievable ergodic SE versus the number of BS antennas for SNR = 5 and 10 dB

We also see that for a fixed number of BS antennas and inter-antenna spacing, the total achievable SE for ZF precoding increases with the SNR for the reason that the ZF precoding scheme is able to completely cancel out inter-user interference, whereas the total achievable SE for MRT precoding without scheduling method converges to a saturated value in the high SNR regime because the inter-user interference becomes dominant as the SNR increases. Compared with ZF precoding, the total achievable SE for MRT precoding without scheduling method is almost identical at SNR = -10 dB, but suffers severe loss (about 200%) at SNR = 30 dB. For comparison, we depict the total achievable SE for different number of antennas $N_t = 64$ and $N_t = 100$, respectively. We observe that deploying more antennas at the BS always provides additional achievable SE because the large antenna arrays facilitate the asymptotic orthogonal condition between different steering vectors. This observation implies that equipping at the BS with large-scale antenna arrays becomes critical for achieving a good total achievable SE.

Figure 5 depicts the analytical results for the total achievable SE, as well as Monte-Carlo simulation results with different number of BS antennas from 50 to 500. As we can see that the theoretical results remain very tight with the numerical simulation results varies with the number of BS antennas, which validates the theoretical results in Proposition 1. In addition, we also find that with fixed the number of user and the SNR, the total achievable SE grows without bound with the number of BS antennas, which is in accordance with the result in Corollary 1. This observation implies that massive antenna arrays contribute substantially toward improving the total achievable SE, which is especially appealing for mmWave MIMO systems. For comparison, we show the total achievable SE at the different SNR regimes, i.e., 5 and 10 dB. As expected, the achievable SE for the high SNR regime is better than that for the low SNR regime, which is also aligned with the result in Corollary 1.

Finally, Fig. 6 shows the simulation and the theoretical results for the maximum achievable ergodic SE with user scheduling, the minimum achievable ergodic SE with a dense user scenario, and the random user's achievable SE. We observe that in the low SNR regime (e.g., $\rho = -10$ dB), the random user's achievable ergodic SE is almost equal to the case with user scheduling, which slightly increases with the increase in SNR. However, in the high SNR regime, the maximum achievable ergodic SE with the user scheduling method is much greater than that for the random user selection. Moreover, the maximum achievable ergodic SE shows a linear increase, which is in agreement with Corollary 1. The minimum achievable ergodic

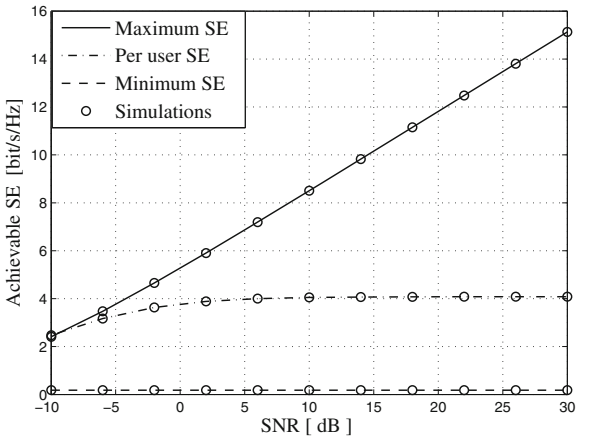

Fig. 6 The total achievable ergodic SE versus SNR. Results are shown for the maximum and minimum rate and the achievable rate of U_1

SE is almost constant in the high SNR regime due to severe interference among users, which is in good agreement with the theoretical analysis in Corollary 2. This is because as the SNR grows, inter-user interference will also increase.

5 Conclusions

This paper investigated the achievable ergodic SE of the downlink of single-cell multi-user mmWave systems, where the BS is equipped with a large number of transmit antennas and service multiple single-antenna users. An exact analytical expression for the achievable ergodic SE was derived. Results showed that the total achievable SE converges to a saturation value in the high SNR regime and increase remarkably in the large antenna regime. For finite number of antennas at the BS, we designed a user scheduling scheme based on users' AoD information, and then derived the corresponding maximum achievable ergodic SE. Under this scheduling scheme, the total achievable SE with MRT precoding can be substantially improved. Furthermore, we presented the minimum achievable ergodic SE of a system based on a dense user case.

Endnotes

[1] Note that although a uniform circular array configuration can also achieve 3D beamforming, channel steering vectors between different users do not achieve orthogonality with a large number of antennas [36].

[2] Note that for randomly distributed users in the circular-shaped cell, the condition $\phi_k \neq \phi_j$ holds with probability one.

[3] The AoDs of the azimuth and elevation angles of the users can be obtained by the BS, and can be treated as constants over a long period [10].

[4] Given a UCA configuration, crosstalk coefficient a_j^c can be expressed by

$$a_j^c = J_0^2\left(k_0 r \delta_c\right), \tag{70}$$

where $\delta_c = \sqrt{\sin^2\theta_k + \sin^2\theta_j - 2\sin\theta_k \sin\theta_j \cos\left(\phi_k - \phi_j\right)}$, and r denotes the radius of the circular transmit array. Given a ULA configuration, crosstalk coefficient a_j^l can be expressed as

$$a_j^l = \frac{1}{N_t}\left|\frac{\sin\left(\frac{N_x k_0 d_x}{2}\delta\right)}{\sin\left(\frac{k_0 d_x}{2}\delta\right)}\right|^2, \tag{71}$$

where $\delta = \cos\phi_k - \cos\phi_j$.

Acknowledgements
We would like to thank the anonymous reviewers for their insightful comments on the paper, as these comments led us to an improvement of the work.

Funding
This work was supported by the National Natural Science Foundation of China under Grant 61531011, the International Science and Technology Cooperation Program of China under Grant 2014DFT10300, and the Guangzhou university project under Grant 27000503123.

Authors' contributions
SJ conceived and designed the idea. WT performed the experiments and analyzed the data. CW and TJ gave valuable suggestions on the structuring of the paper and assisted in the revising and proofreading. All authors read and agreed the manuscript.

Competing interests
The authors declare that they have no competing interests.

Author details
[1] School of Computer Science and Educational Software, Guangzhou University, Guangzhou 510006, People's Republic of China. [2] National Communications Research Laboratory, Southeast University, Nanjing 210096, People's Republic of China. [3] Institute of Communications Engineering, National Sun Yat-sen University, Kaohsiun, 804 Taiwan, People's Republic of China. [4] Wuhan National Laboratory for Optoelectronics, School of Electronics Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, People's Republic of China.

Publisher's Note

Received: 10 March 2017 Accepted: 18 October 2017
Published online: 06 November 2017

References

1. F Boccardi, RW Heath Jr., A Lozano, TL Marzetta, P Popovski, Five disruptive technology directions for 5G. IEEE Commun. Mag. **52**(2), 74–80 (2014)
2. W Roh, J Seol, J Park, B Lee, J Lee, Y Kim, J Cho, K Cheun, Millimeter-Wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results. IEEE Commun. Mag. **52**(2), 106–113 (2014)
3. S Yong, C Chong, An overview of multigigabit wireless through millimeter wave technology: potentials and technical challenges.EURASIP. J. Wireless Commun. Net. **7**(1), 50–60 (2006)
4. Z Pi, F Khan, An introduction to millimeter-wave mobile broadband systems. IEEE Commun. Mag. **49**(6), 101–107 (2011)
5. TL Marzetta, Noncooperative cellular wireless with unlimited numbers of base station antennas. IEEE Trans. Wireless Commun. **9**(11), 3590–3600 (2010)
6. F Rusek, D Persson, BK Lau, EG Larsson, TL Marzetta, O Edfors, F Tufvesson, Scaling up MIMO: opportunities and challenges with very large arrays. IEEE Signal Process. Mag. **30**(1), 40–60 (2013)
7. H Ngo, E Larsson, T Marzetta, The multicell multiuser mimo uplink with very large antenna arrays and a finite-dimensional channel. IEEE Trans. Commun. **61**(6), 2350–2361 (2013)
8. Q Zhang, S Jin, KK Wong, HB Zhu, Power scaling of uplink massive MIMO systems with arbitrary-rank channel means. IEEE J. Sel. Topics Signal Process. **57**(3), 841–849 (2014)
9. Y Nam, B Ng, K Sayana, Y Li, J Zhang, Y Kim, J Lee, Full-dimension MIMO (FD-MIMO) for next generation cellular technology. IEEE Commun. Mag. **21**(2), 172–179 (2013)
10. L Liu, Y Li, J Zhang, in *Proc. IEEE Int. Conf. Signal Proc. Advances in, Wireless Commun. (SPAWC)*. DoA estimation and achievable rate analysis for 3D millimeter wave massive MIMO systems, (Toronto, 2014), pp. 6–11
11. S Rangan, TS Rappaport, E Erkip, Millimeter wave cellular wireless networks: potentials and challenges. Proc. IEEE. **102**(3), 366–385 (2014)

12. J Brady, N Behdad, A Sayeed, Beamspace MIMO for millimeter-wave communications: system architecture, modeling, analysis, and measurements. IEEE Trans. Ant. and Prop. **61**(7), 3814–3827 (2013)
13. MR Akdeniz, Y Liu, MK Samimi, S Sun, S Rangan, TS Rappaport, Millimeter wave channel modeling and cellular capacity evaluation. IEEE J. Sel. Areas Commun. **32**(6), 1164–1179 (2014)
14. J Seo, Y Sung, G Lee, D Kim, in *Proc. IEEE Int. Conf. Signal Proc. Advances in Wireless Commun. (SPAWC)*. Pilot beam sequence design for channel estimation in millimeter-wave MIMO systems: A POMDP framework, (Stockholm, 2015), pp. 236–240
15. A Alkhateeb, O Ayach, G Leus, RW Heath Jr., Channel estimation and hybrid precoding for millimeter wave cellular systems. IEEE J. Sel. Areas Commun. Signal Process. **8**(5), 831–846 (2014)
16. N Ravindran, N Jindal, HC Huang, in *Proc. IEEE Global, Telecommun. Conf. (GLOBECOM)*. Beamforming with finite rate feedback for LOS MIMO downlink channels, (Washington, 2007), pp. 4200–4204
17. X Zhang, A Molisch, S Kung, Variable-phase-shift-based RF baseband codesign for MIMO antenna selection. IEEE Trans. Signal Process. **53**(11), 4091–4103 (2005)
18. X Li, S Jin, HA Suraweera, J Hou, X Gao, Statistical 3-D beamforming for large-scale MIMO downlink systems over Rician fading channels. IEEE Trans. Commun. **64**(5), 1529–1543 (2016)
19. A Alkhateeb, J Mo, NG Prelcic, RW Heath Jr., MIMO precoding and combining solutions for millimeter-wave systems. IEEE Commun. Mag. **52**(12), 122–131 (2013)
20. W Tan, M Matthaiou, S Jin, X Li, Spectral efficiency of DFT-based processing hybrid architectures in massive MIMO. IEEE Wireless Commun. Lett. **6**(5), 586–589 (2017)
21. L Fan, R Zhao, F Gong, N Yang, GK Karagiannidis, Secure multiple amplify-and-forward relaying over correlated fading channels. IEEE Trans. Commun. **65**(7), 2811–2820 (2017)
22. L Fan, X Lei, N Yang, TQ Duong, GK Karagiannidis, Secure multiple amplify-and-forward relaying with cochannel interference. IEEE J. Sel. Topics Signal Process. **10**(8), 1494–1505 (2016)
23. L Fan, S Zhang, TQ Duong, GK Karagiannidis, Secure switch-and-stay combining (SSSC) for cognitive relay networks. IEEE Trans. Commun. **64**(1), 70–82 (2016)
24. X Li, T Jiang, S Cui, J An, Q Zhang, Cooperative communications based on rateless network coding in distributed MIMO systems. IEEE Wireless Commun. **17**(3), 60–67 (2010)
25. T Bai, RW Heath Jr., Coverage and rate analysis for millimeter wave cellular networks. IEEE Trans. Wireless Commun. **14**(2), 1100–1114 (2015)
26. L Liang, Y Dai, W Xu, X Dong, in *Proc. IEEE Int. Conf. Commun. in China (ICCC)*. How to approach zero-forcing under RF chain limitations in large mmWave multiuser systems? (Xi'an, 2014), pp. 518–522
27. IS Gradshteyn, IM Ryzhik, *Table of integrals, series, and products*, 7th edition. (Academic, 2007)
28. AM Sayeed, N Behdad, in *Proc. IEEE Int. Conf. Commun. (ICC)*. Beamspace MIMO for high-dimensional multiuser communication at millimeter wave frequencies, (Budapest, 2013), pp. 3679–3684
29. G Lee, Y Sung, M Kountouris, On the performance of random beamforming in sparse millimeter wave channels. IEEE J. Sel. Topics Signal Process. **10**(3), 1–16 (2016)
30. G Lee, Y Sung, J Seo, Randomly-directional beamforming in millimeter-wave multi-user MISO downlink. IEEE Trans. Wireless Commun. **15**(2), 1086–1100 (2016)
31. TS Rappaport, E Ben-Dor, JN Murdock, Y Qiao, *38 GHz and 60 GHz angle-dependent propagation for cellular & peer-to-peer wireless communications*, (Ottawa, 2012), pp. 679–684
32. W Tan, S Jin, Wang, Y Huang, in *IEEE Wireless Commun. and Networking Conf. (WCNC)*. Achievable sum-rate analysis for massive MIMO systems with different array configurations, (New Orlean, 2015), pp. 316–321
33. 3GPP TR 36.873 V12.1.0, Study on 3D channel model for LTE," (2015)
34. NL Johnson, S Kotz, N Balakrishnan, *Continuous Univariate Distributions Vol. 1*, 2nd Ed. (John Wiley Sons, New York, 1994)
35. G Alfano, A Lozano, AM Tulino, S Verdú, in *Proc. Int. Symp. Information Theory and, Its Applications (ISITA)*. Mutual information and eigenvalue distribution of MIMO Ricean channels, (Parma, 2004), pp. 10–13
36. JH Chen, in *Proc. IEEE Global, Commun. Conf. (GLOBECOM)*. When does asymptotic orthogonality exist for very large arrays? (Atlanta, 2013), pp. 4146–4150
37. T Yoo, A Goldsmith, On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming. IEEE J. Sel. Areas Commun. **24**(5), 528–541 (2006)
38. J Nam, A Adhikary, J Ahn, G Caire, Joint spatial division and multiplexing: opportunistic beamforming, user grouping and simplified downlink scheduling. IEEE J. Sel. Topics Signal Process. **8**, 876–890 (2014)
39. AL Swindlehurst, T Kailath, Azimuth/elevation direction finding using regular array geometries. IEEE Trans. Aerospace Electronic Syst. **29**(1), 145–156 (1993)
40. P Viswanath, DNC Tse, R Laroia, Opportunistic beamforming using dumb antennas. IEEE Trans. Inf. Theory. **48**(6), 1277–1294 (2002)
41. DJ Love, RW Heath Jr., Limited feedback unitary precoding for spatiall multiplexing systems. IEEE Trans. Inf. Theory. **51**(8), 2967–2976 (2005)
42. L Liang, W Xu, X Dong, Low-complexity hybrid precoding in massive multiuser MIMO systems. IEEE Commun. Lett. **3**(6), 653–656 (2014)
43. From Math World-A Wolfram Web Resource. [Online] Available: http://functions.wolfram.com/PDF/ExpIntegralE.pdf
44. A Adhikary, EA Safadi, MK Samimi, R Wang, G Caire, TS Rappaport, AF Molisch, Joint spatial division and multiplexing for mm-wave channels. IEEE J. Sel. Areas Commun. **32**(6), 1239–1248 (2014)
45. HA Suraweera, PJ Smith, M Shafi, Capacity limits and performance analysis of cognitive radio with imperfect channel knowledge. IEEE Trans. Veh. Technol. **59**(4), 1811–1822 (2010)
46. M Shafi, M Zhang, PJ Smith, AL Moustakas, AF Molisch, in *Proc. IEEE Int. Conf. Commun. (ICC)*. The impact of elevation angle on MIMO capacity, (Istanbul, 2006), pp. 4155–4160