

Sum-Rate Analysis of the Generalized Spatial Modulation Uplink in Multi-Cell Multi-User Systems in the Face of Pilot Contamination

Xiaoqing Zhang, Shangbin Wu, Shanshan Yu, *Member, IEEE*, Ju Liu, *Senior Member, IEEE* and Lajos Hanzo, *Fellow, IEEE*

Abstract—The sum-rate of the multi-cell multi-user generalized spatial modulation (GSM) uplink is analyzed in the presence of realistic pilot contamination. A practical channel model associated with arbitrary transmit and receive correlation matrices is assumed for each user and base station (BS). Imperfect channel estimation is assumed with pilot contamination. The performance of a single-cell based minimum mean squared error (SMMSE) combiner is analyzed and compared to the ubiquitous zero forcing (ZF) combiner as well as matched filtering (MF). The transmit antenna (TA) index detection and the classic amplitude/phase modulation (APM) signal detection process are carried out separately for the sake of low complexity. Moreover, an algorithm based on order statistics is proposed for calculating the antenna detection probability conditioned on the actual TA. Finally, an approximation of the sum-rate is derived by exploiting the characteristics of massive MIMOs. Simulation results show that SMMSE has the best performance, followed by ZF and MF. Furthermore, the relationships between the system's achievable rate and three systems parameters, namely the number of BS antennas, the signal-to-noise-ratio (SNR) and the interference factor of pilot contamination, are presented. The performance of different number of activated antennas are also compared.

Index Terms—GSM, massive MIMO, sum-rate.

I. INTRODUCTION

MASSIVE multiple-input multiple-output (MIMO) systems are equipped with a large number of antennas. A large antenna array is capable of exploiting the channel hardening, which facilitates low-complexity signal processing. Furthermore, a high sum-rate, high reliability and high energy efficiency can be attained by increasing the number of transmit antennas (TAs) [1]-[4]. These advantages render massive MIMOs one of the most promising technique for fifth generation (5G) wireless communication [5], [6]. However, the high cost of hardware, such as radio frequency (RF) chains and power

consumption, imposes limitations on the number of TAs to be employed [7].

To reduce hardware cost, spatial modulation (SM) was advocated in [8] as a beneficial modulation technique, which exploits the TA index to transmit implicit information. One part of the information bits are used to select a specific TA, while the other part of the information bits are used to perform classic amplitude/phase modulation (APM). The exploitation of the amplitude, phase and spatial domain to modulate the signal has compelling benefits, such as the potential avoidance of inter-antenna interference and synchronization as well as a low number of RF chains [9]. SM was also combined with orthogonal frequency division multiplexing (OFDM) in [10] and its symbol error ratio (SER) performance was analyzed in a single-cell scenario. Jeganathan *et al.* [11] utilized the maximum likelihood (ML) SM detector and calculated the average bit error rate (BER) in a point-to-point transmission model. The corresponding capacity analysis was performed in [12] through separating the mutual information (MI) into two parts, i.e., the MI between the transmitted symbols and the received signal, the MI between the antenna-constellation symbols and the corresponding received signal, according to a multi-input single-output (MISO) channel. Motivated by [12], the system capacity was calculated for a point-to-point MIMO channel by relying on the pair-wise error probabilities (PEP) to approximate the TA detection error probabilities [13]. Given the subject's relative maturity by 2014, a detailed introduction of various SM-MIMO systems was given in [14]. Furthermore, authors in [15] and [16] applied the SM to a single-cell massive MIMO system with the large-scale antenna benefits taken into consideration. In addition to single-cell or point-to-point MIMOs, He *et al.* [16] further studied the achievable uplink (UL) spectral efficiency (SE) of a multi-cell massive SM-MIMO system and derived asymptotic SE lower bounds relying on linear combining schemes [17]. Performance of massive SM-MIMO in high-speed railway was discussed in [18] and massive SM-MIMO with forward error correction (FEC) was investigated in [19].

The authors of [20], [21] proposed the generalized spatial modulation (GSM) concept by activating multiple antennas in a SM system to transmit a flexible number of bits at a time. Generalized space shift keying (GSSK) presented in [22] can be regarded as a special case of GSM with the TA index alone conveying information. Compared to SM, GSM utilizes more links to increase the bits/channel use at the cost of an

L. Hanzo would like to acknowledge the ERC's financial support of his Advanced Fellow Award QuantCom and the EPSRC, UK. J. Liu and X. Zhang would like to acknowledge the financial support of the National Key R&D Plan (2017YFC0803403) and the China Scholarship Council.

X. Zhang, S. Yu and J. Liu are with the School of Information Science and Engineering, Shandong University, Qingdao 266237, China (Corresponding author: juliu@sdu.edu.cn). X. Zhang is also with the Electronics and Computer Science, University of Southampton, SO17 1BJ, UK (e-mail: xiaoqing.sdu@hotmail.com).

S. Wu is with Samsung R&D Institute UK, Staines-upon-Thames, TW18 4QE.

L. Hanzo is with the Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK, and also with King Abdulaziz University, Jeddah 21589, Saudi Arabia (e-mail: lh@ecs.soton.ac.uk).

increase in detection complexity and in the number of RF chains. Hence, GSM reaches a better balance between the low complexity and the multiplexing gain than SM.

Although there have been extensive studies in the area of SM-MIMO, there is a paucity of literature on massive GSM-MIMO in a multi-cell multi-user system. The achievable sum-rate performance was investigated based on zero forcing (ZF) and matched filtering (MF) detection in a multi-cell massive MIMO system relying on GSM and an approximation of the sum-rate was provided for ZF and MF detection [23]. However, the covariance matrices of all base stations (BSs) and of all users are the same in [23], which is impractical. Additionally, the processes of TA index detection and APM signal detection were not considered separately despite its appealingly low complexity. Moreover, the correlation between variables was omitted in the calculation of conditional probabilities, when performing TA index detection. To fill the above-mentioned research gaps, this paper investigates a massive MIMO multi-cell multi-user system relying on GSM. The contributions of this paper can be summarized as follows:

- 1) A practical channel model having an arbitrary transmit correlation matrix and receive correlation matrix for each user and BS is adopted. Both realistic and imperfect channel information and pilot contamination are considered.
- 2) Single-cell based minimum mean squared error (SMMSE) filter is used for detecting the classic APM signals as well as the index of TAs. To the best of our knowledge, this paper is the first one to propose SMMSE filtering, which performs better than the previously used ZF and MF schemes.
- 3) An algorithm is proposed for calculating the TA detection error probability using the order statistics of independent but non-identically distributed random variables.
- 4) The approximate sum-rate expressions of SMMSE, ZF and MF are derived. The analytical relationships between the sum-rate and the number of antennas, the SNR and the interference factor under two type of covariance matrix models are also presented. The performance of different number of activated antennas are compared.

The paper is organized as follows. The system model is introduced in Section II along with the channel estimation scheme, UL data transmission process and the detection process. The system's sum-rate is analyzed in Section III based on order statistics. In Section IV, the large scale antenna effects are exploited to obtain an approximation of the system's sum-rate calculated in Section III. Finally, our numerical results are presented in Section V.

Notations: $\mathbb{C}^{M \times N}$ denotes a complex matrix with M rows and N columns. $\mathcal{CN}(\mathbf{m}, \mathbf{R})$ represents the circularly symmetric complex Gaussian distribution with a mean of \mathbf{m} and covariance matrix \mathbf{R} . The Kronecker product and the Hadamard (element-wise) products are denoted by \otimes and \bullet separately. Boldface lower and upper case symbols represent vectors and matrices, respectively. The trace, transpose, Hermitian transpose and inverse operators are denoted by tr , T , H , and $^{-1}$ respectively. The vectorization of a matrix is denoted by vec , which means rearranging the elements in a column-major

manner. $\mathbb{E}[\cdot]$ represents the expectation operator. Given two sets A and B , then $A \setminus B$ denotes set A excluding set B . $\lfloor x \rfloor$ denotes the largest integer less or equal to x .

II. SYSTEM MODEL

In this paper, the UL of a multi-cell multi-user massive MIMO system is considered, which is shown in Fig. 1. A cellular system having L cells is assumed and each cell includes one BS and K users. Each BS is equipped with M antennas, while each user is equipped with N antennas. It is assumed that $M \gg N$. In the UL transmission process, GSM is adopted in the user side with each user activating the selected antennas and transmitting his/her UL data. Each BS receives the desired UL signal from the users in the same cell, plus the inter-cell interference caused by the UL data signal transmitted by the users in other cells owing to pilot contamination.

Let us denote the channel matrix between the l th BS and the k th user in the z th cell by $\mathbf{G}_{lzk} \in \mathbb{C}^{M \times N}$

$$\mathbf{G}_{lzk} = \sqrt{\beta_{lzk}} \mathbf{R}_l^{\frac{1}{2}} \mathbf{H}_{lzk} \mathbf{R}_{zk}^{\frac{T}{2}}, \quad (1)$$

where β_{lzk} is the large scale fading between the k th user in the z th cell and the l th BS, $\mathbf{H}_{lzk} \in \mathbb{C}^{M \times N}$ is the small scale fading matrix having independent and identically distributed (i.i.d.) elements following $\mathcal{CN}(0, 1)$, \mathbf{R}_l is the covariance matrix at the l th BS side and \mathbf{R}_{zk} is the covariance matrix at the k th user side in the z th cell. The elements in both covariance matrices are generated by a typical exponential model, presented as

$$[\mathbf{R}]_{m,n} = r^{|n-m|} e^{j(n-m)\theta}, \quad (2)$$

where $r \in [0, 1]$ is the correlation factor between the adjacent antennas and θ is the angle-of-arrival (AOA). The subscript m, n is used to specify the m th row and n th column element of the covariance matrix. **Then, \mathbf{G}_{lzk} obeys the distribution of $\mathcal{CN}(0_{MN}, \beta_{lzk} \mathbf{R}_{lzk})$ with $\mathbf{R}_{lzk} = \mathbf{R}_{zk} \otimes \mathbf{R}_l$.** Furthermore, \mathbf{G}_{lzk} can be written as $\mathbf{G}_{lzk} = [\mathbf{g}_{lzk1}, \dots, \mathbf{g}_{lzkN}]$, where \mathbf{g}_{lzkN} is the channel between the l th BS and the n th antenna of the k th user in the z th cell. The channel matrix between the l th BS and all users in the z th cell is denoted as $\mathbf{G}_{lz} = [\mathbf{G}_{lz1}, \dots, \mathbf{G}_{lzK}] \in \mathbb{C}^{M \times NK}$. A Bessel function based covariance matrix was adopted in [23], which is formulated as $[\mathbf{R}_U]_{m,n} = J_0(2\pi d_t |m - n|)$ for all users, with the subscript U representing users and $[\mathbf{R}_B]_{m,n} = J_0(2\pi d_r |m - n|) \bullet e^{-\alpha|m-n|}$ for all BSs with the subscript B denoting BS, where $J_0(\cdot)$ represents the zeroth-order Bessel function of the first kind, d_t and d_r represent the antenna spacing expressed in terms of the number of wavelengths of the user's transmit antenna arrays (TAAs) and the BS's receiving antenna arrays (RAAs), respectively. Finally, $e^{-\alpha|m-n|}$ in $[\mathbf{R}_B]_{m,n}$ denotes the survival probability of the scatterers between the m th and n th antenna. This spatial correlation model will be used as a comparative basis in this paper.

A. Channel Estimation

The UL pilots are assumed to be fully reused in all the cells, while the pilots transmitted within the same cell are orthogonal. For the channel estimation process, a single antenna is

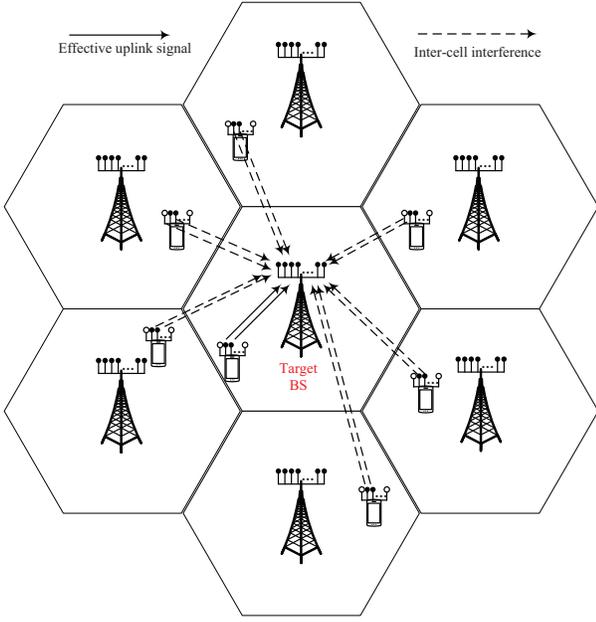


Fig. 1. The uplink data transmission in a multi-cell massive MIMO system under pilot contamination with one user in each cell using the same pilot for example. The solid line denotes the effective uplink signal transmission for the central target BS while the dotted line represents the inter-cell interference from neighboring cells.

activated at a specific channel use. Hence, a total of NK time slots are required for all the users to estimate their channels associated with $NK < T$, where T is the channel's coherence time. The NK time slots can be partitioned into N subframes, where each subframe includes K time slots. Then, the n th TA of each user transmits the pilot sequence of length K during the n th subframe. Thus, the pilot signal \mathbf{Y}_{ln} received at the l th BS in the n th sub-frame is

$$\mathbf{Y}_{ln} = \sum_{z=1}^L \sum_{k=1}^K \sqrt{\rho_{tr}} \mathbf{g}_{lzk} \phi_k + \mathbf{N}_p, \quad (3)$$

where ρ_{tr} is the pilot transmit power and $\phi_k \in \mathbb{C}^{1 \times K}$ is the pilot invoked for the k th user in each cell, which forms $\Phi = [\phi_1^T, \dots, \phi_K^T]$ that satisfies $\Phi^H \Phi = \mathbf{I}$. Furthermore, \mathbf{N}_p is an $(M \times K)$ -element noise matrix with i.i.d. elements following $\mathcal{CN}(0, 1)$.

Correlating both sides of (3) with ϕ_k^H , we can get the channel estimation vector between the n th TA of the k th user in the l th cell and the l th BS, namely \mathbf{Y}_{lkn} , based on the received observation, given as

$$\mathbf{Y}_{lkn} = \mathbf{Y}_{ln} \phi_k^H = \sum_{z=1}^L \sqrt{\rho_{tr}} \mathbf{g}_{lzk} + \mathbf{n}_{pn}, \quad (4)$$

where $\mathbf{n}_{pn} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ for ϕ_k does not change the distribution of \mathbf{N}_p .

Given $\mathbf{Y}_{lk} = [\mathbf{Y}_{lk1}^T, \mathbf{Y}_{lk2}^T, \dots, \mathbf{Y}_{lkN}^T]^T$, the classic MMSE channel estimation method may be used for processing \mathbf{Y}_{lk} , which is contaminated by the pilots arriving from neighboring

cells. Then, a channel estimation vector $\text{vec}(\hat{\mathbf{G}}_{lzk})$ of size $(MN \times 1)$ is obtained as

$$\begin{aligned} \text{vec}(\hat{\mathbf{G}}_{lzk}) &= \beta_{lzk} \mathbf{R}_{lzk} \mathbf{Q}_{lk}^{-1} \left(\frac{1}{\sqrt{\rho_{tr}}} \mathbf{Y}_{lk} \right) \\ &= \beta_{lzk} \mathbf{R}_{lzk} \mathbf{Q}_{lk}^{-1} \left(\sum_{z=1}^L \text{vec}(\mathbf{G}_{lzk}) + \frac{1}{\sqrt{\rho_{tr}}} \mathbf{n}_p \right) \\ &\sim \mathcal{CN}(0, \Phi_{lzk}), \end{aligned} \quad (5)$$

where we have

$$\mathbf{Q}_{lk} = \sum_{z=1}^L \beta_{lzk} \mathbf{R}_{lzk} + \frac{1}{\rho_{tr}} \mathbf{I}_{MN}, \quad (6)$$

and

$$\Phi_{lzk} = \beta_{lzk}^2 \mathbf{R}_{lzk} \mathbf{Q}_{lk}^{-1} \mathbf{R}_{lzk}. \quad (7)$$

Upon exploiting the property of the MMSE channel estimation method, namely that the estimated channel vector is orthogonal to the channel estimation error vector, we arrive at

$$\text{vec}(\tilde{\mathbf{G}}_{lzk}) \sim \mathcal{CN}(0, \beta_{lzk} \mathbf{R}_{lzk} - \Phi_{lzk}), \quad (8)$$

where $\tilde{\mathbf{G}}_{lzk}$ is the channel estimation error vector which satisfies $\mathbf{G}_{lzk} = \hat{\mathbf{G}}_{lzk} + \tilde{\mathbf{G}}_{lzk}$.

Moreover, let $\Phi_{l_z-z'k}$ denote the covariance matrix of the estimated UL channels between the k th user in different cells and the l th BS, which is expressed as

$$\begin{aligned} \Phi_{l_z-z'k} &= \mathbb{E}[\text{vec}(\hat{\mathbf{G}}_{lzk}) \text{vec}(\hat{\mathbf{G}}_{l_z'k})^H] \\ &= \beta_{lzk} \beta_{l_z'k} \mathbf{R}_{lzk} \mathbf{Q}_{lk}^{-1} \mathbf{R}_{l_z'k}. \end{aligned} \quad (9)$$

B. Uplink Data Transmission

The UL transmission data vector of the k th user in the z th cell is denoted as $\mathbf{x}_{zk} = [x_{zk1}, \dots, x_{zkN}]^T \in \mathbb{C}^{N \times 1}$, consisting of S nonzero entries. Each nonzero entry obeys $\mathcal{CN}(0, 1)$. Upon denoting the actual UL transmitted signal as $\check{\mathbf{x}}_{zk} \in \mathbb{C}^{S \times 1}$ and the UL signal transmitted by all the K users in the z th cell as $\check{\mathbf{x}}_z = [\check{\mathbf{x}}_{z1}^T, \dots, \check{\mathbf{x}}_{zK}^T]^T \in \mathbb{C}^{SK \times 1}$, the corresponding active channel between the l th BS and the users in the z th cell can be written as $\check{\mathbf{G}}_{lz} \in \mathbb{C}^{M \times SK}$. Moreover, there is a total of $C = \lfloor C_N^S \rfloor$ antenna groups to be selected according to the information bits. These antenna combination groups can be listed as $(\Pi_1, \Pi_2, \dots, \Pi_C)$. Every antenna combination group pattern has the same probability to be activated. Let us consider $N = 4$ TAs as an example. Then, if $S = 2$ TAs are activated at a time, the total number of antenna groups is $\lfloor C_4^2 \rfloor = 6$ and the TA combination groups are as follows: $[(1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (3, 4)]$.

In the UL data transmission process, each user transmits his/her UL data to the corresponding BS. Thus, the $(M \times 1)$ -element signal vector \mathbf{y}_l received at the l th BS becomes:

$$\mathbf{y}_l = \sqrt{\rho_{ul}} \sum_{z=1}^L \check{\mathbf{G}}_{lz} \check{\mathbf{x}}_z + \mathbf{n}_{ul}, \quad (10)$$

where ρ_{ul} is the UL data transmit power of each antenna and \mathbf{n}_{ul} is the UL noise vector obeying $\mathcal{CN}(\mathbf{0}, \mathbf{I})$.

C. Antenna Index Detection and Amplitude/Phase Modulated Signal Detection

For the detection process, three linear filters, namely SMMSE, ZF and MF, are adopted for detecting both the indices of TAs as well as the classic APM signals.

1) *SMMSE*: The SMMSE detector jointly minimizes the effects of interferences and noise. The SMMSE detector $\hat{\mathbf{W}}_{lk}^{\text{SMMSE}} \in \mathbb{C}^{M \times N}$ of the k th user in the l th cell can be expressed as

$$\hat{\mathbf{W}}_{lk}^{\text{SMMSE}} = \left(\sum_{k=1}^K \hat{\mathbf{G}}_{llk} \hat{\mathbf{G}}_{llk}^H + \mathbf{Z}_l \right)^{-1} \hat{\mathbf{G}}_{llk}, \quad (11)$$

where

$$\mathbf{Z}_l = \sum_{k=1}^K (\beta_{lk} \mathbf{R}_{llk}^d - \Phi_{llk}^d) + \sum_{\substack{z=1 \\ z \neq l}}^L \sum_{k=1}^K (\beta_{lzk} \mathbf{R}_{lzk}^d) + \frac{1}{\rho_{ul}} \mathbf{I}_M, \quad (12)$$

while \mathbf{R}_{lzk}^d in \mathbf{Z}_l represents the multiplication of \mathbf{R}_l and the sum of diagonal elements in \mathbf{R}_{zk} .

To clarify the role of Φ_{llk}^d in \mathbf{Z}_l , Φ_{llk} can be rewritten in the form of a block matrix as

$$\Phi_{llk} = \begin{bmatrix} \Phi_{llk}^{1,1} & \cdots & \cdots \\ \cdots & \ddots & \cdots \\ \cdots & \cdots & \Phi_{llk}^{N,N} \end{bmatrix}, \quad (13)$$

where $\Phi_{llk}^{i,j}$ is an $(M \times M)$ -element matrix, which is taken from the $[(i-1)M+1 \sim iM]$ th rows, $[(j-1)M+1 \sim jM]$ th columns of Φ_{llk} . Then, Φ_{llk}^d is the sum of the diagonal block matrices, i.e., $\Phi_{llk}^d = \sum_{j=1}^N \sum_{i=1}^N \Phi_{llk}^{i,j}$.

2) *ZF*: The ZF detector aims for nulling the intra-cell interference, which is the pseudo-inverse of the estimated channel matrix, neglecting the interference emanating from the other cells and noise. The zero forcing detector $\hat{\mathbf{W}}_l^{\text{ZF}} \in \mathbb{C}^{M \times NK}$ of the l th cell can be formulated as

$$\hat{\mathbf{W}}_l^{\text{ZF}} = \hat{\mathbf{G}}_{ll} \left(\hat{\mathbf{G}}_{ll}^H \hat{\mathbf{G}}_{ll} \right)^{-1}. \quad (14)$$

Then, the detector of the k th user in the l th cell, namely $\hat{\mathbf{W}}_{lk}^{\text{ZF}}$, spans from the $[(k-1)N+1]$ th column to the kN th column of $\hat{\mathbf{W}}_l^{\text{ZF}}$, which can be expressed as $\hat{\mathbf{W}}_{lk}^{\text{ZF}} = \left[\hat{\mathbf{W}}_l^{\text{ZF}} \right]_{[(k-1)N+1:kN]} = [\hat{\mathbf{w}}_{lk1} \cdots \hat{\mathbf{w}}_{lkN}]$, where each $\hat{\mathbf{w}}_{lkd}$ is an $M \times 1$ vector.

3) *MF*: The MF detector aims for maximizing the signal power received via the estimated channel, which can be expressed as,

$$\hat{\mathbf{W}}_{lk}^{\text{MF}} = \hat{\mathbf{G}}_{llk}. \quad (15)$$

The TA index detection and APM signal detection can be performed either jointly at a high complexity or separately at a reduced complexity. Here we opt for the latter as shown in Fig. 2, where the information bits after SM are transmitted via the channel and then arrive at the receiver side. On the one hand, the received signals pass through an arbitrary linear

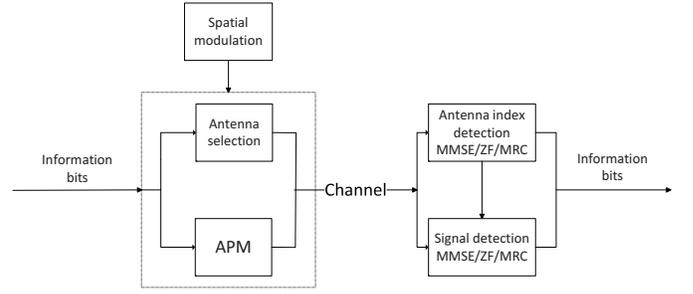


Fig. 2. Diagram of signal transmission and signal detection.

filter for TA index detection; on the other hand, they also pass through an arbitrary linear filter given the already detected TA index for APM signal detection. After that, the information bits can be recovered.

4) *Antenna Index Detection*: For UL data transmission, after the l th BS receives the data signals \mathbf{y}_l from all users, the filtered signal $\hat{\mathbf{y}}_{lk}$ is obtained by correlating \mathbf{y}_l with a linear filter $\hat{\mathbf{W}}_{lk}$ according to:

$$\hat{\mathbf{y}}_{lk} = \hat{\mathbf{W}}_{lk}^H \mathbf{y}_l, \quad (16)$$

where $\hat{\mathbf{y}}_{lk} = [\hat{y}_{lk1}, \dots, \hat{y}_{lkN}]^T$.

Then, a TA combination $\hat{\Pi}_{lk}$, which contains the largest S elements in $\hat{\mathbf{y}}_{lk}$ will be selected as

$$\hat{\Pi}_{lk} = \arg \max_{\Pi_i} \sum_{\forall d \in \Pi_i} |\hat{y}_{lkd}|^2. \quad (17)$$

The antennas in $\hat{\Pi}_{lk}$ will be regarded as the TAs.

5) *Amplitude/Phase Modulated Signal Detection*: For the APM signal detection process, as shown in Fig. 2, either the same or a different linear filter can be applied to the corresponding elements of $\hat{\mathbf{y}}_{lk}$ selected by $\hat{\Pi}_{lk}$, which is denoted as \mathbf{y}_{ls} . Thus, the detected signal $\hat{\mathbf{r}}_{lk}$ can be expressed as

$$\hat{\mathbf{r}}_{lk} = \hat{\mathbf{W}}_{lk}^H \mathbf{y}_{ls}, \quad (18)$$

which will be regarded as the transmitted signals.

III. ACHIEVABLE SUM-RATE ANALYSIS

Our analysis will be carried out by following these steps:

1. At the transmitter side, the classic APM signals $\check{\mathbf{x}}_{lk}$ of the k th user in the l th cell pass through the channel $\check{\mathbf{G}}_{llk}$ between the k th user in the l th cell and the l th BS. It is noted that $\check{\mathbf{x}}_{lk}$ carries one part of information and $\check{\mathbf{G}}_{llk}$ carries the other part of information implied in the TA index.

2. At the receiver side, the detected TA index $\hat{\Pi}_{lk}$ and the detected classic APM signals $\hat{\mathbf{r}}_{lk}$ are obtained via (17) and (18) separately.

3. The recovering of information is a mapping $\left\{ \hat{\Pi}_{lk}, \hat{\mathbf{r}}_{lk} \right\} = f_{\text{mapping}} \left\{ \check{\mathbf{x}}_{lk}, \check{\mathbf{G}}_{llk} \right\}$. Thus, The sum-rate of the k th user in

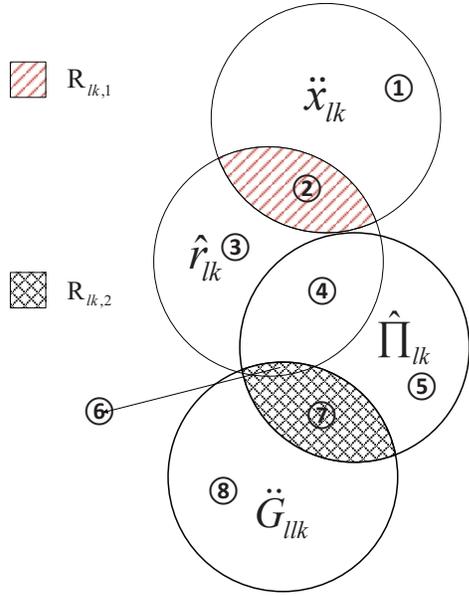


Fig. 3. The relationship of four parameters, i.e., the transmitted signals $\ddot{\mathbf{x}}_{lk}$, transmission channels $\ddot{\mathbf{G}}_{llk}$, detected TA index $\hat{\Pi}_{lk}$ and detected classic signals $\hat{\mathbf{r}}_{lk}$.

the l th cell can be calculated as the mutual information (MI) between $\{\hat{\Pi}_{lk}, \hat{\mathbf{r}}_{lk}\}$ and $\{\ddot{\mathbf{x}}_{lk}, \ddot{\mathbf{G}}_{llk}\}$, which is expressed as

$$\begin{aligned} & \left(1 - \frac{KN}{T}\right) \mathbb{E} \left\{ I \left(\ddot{\mathbf{x}}_{lk}, \ddot{\mathbf{G}}_{llk}; \hat{\mathbf{r}}_{lk}, \hat{\Pi}_{lk} \right) \right\} \\ &= \left(1 - \frac{KN}{T}\right) \left(\mathbb{E} \left\{ I \left(\ddot{\mathbf{x}}_{lk}; \hat{\mathbf{r}}_{lk}, \hat{\Pi}_{lk} | \ddot{\mathbf{G}}_{llk} \right) \right\} \right. \\ & \quad \left. + \mathbb{E} \left\{ I \left(\ddot{\mathbf{G}}_{llk}; \hat{\mathbf{r}}_{lk}, \hat{\Pi}_{lk} \right) \right\} \right), \end{aligned} \quad (19)$$

where $\left(1 - \frac{KN}{T}\right)$ is the effective data transmission time.

To elaborate on the MI calculation more explicitly, the relationship of the four variables in (19) can be characterized using Fig. 3 where the four circles represent four variables respectively and they split the area into eight non-overlapping regions which are labelled in Fig. 3. Since the choice of the TAs and the classic transmitted APM signals is controlled by independent information bits, $\ddot{\mathbf{x}}_{lk}$ has no intersection with $\ddot{\mathbf{G}}_{llk}$. By contrast, $\hat{\mathbf{r}}_{lk}$ has intersection (4) + (6) with $\hat{\Pi}_{lk}$ and the intersection area with $\ddot{\mathbf{G}}_{llk}$ (6) is part of the intersection area (6) + (7) of $\hat{\Pi}_{lk}$ and $\ddot{\mathbf{G}}_{llk}$, since the APM signal detection directly depends on the TA index detection process. Moreover, $\hat{\Pi}_{lk}$ has an intersection (6) + (7) with $\ddot{\mathbf{G}}_{llk}$ excluding $\ddot{\mathbf{x}}_{lk}$, since the TA index detection is only related to the true TA index. The MI between $\{\hat{\Pi}_{lk}, \hat{\mathbf{r}}_{lk}\}$ and $\{\ddot{\mathbf{x}}_{lk}, \ddot{\mathbf{G}}_{llk}\}$ is represented by the intersection area between them, i.e., (2) + (6) + (7), where (2) is the first term in (19), which represents the sum-rate of the classic signal detection process expressed as $R_{lk,1}$ and (6) + (7) is the second term in (19), which represents the sum-rate of the TA index detection process expressed as $R_{lk,2}$.

1) Calculation of $R_{lk,1}$: Let $R_{lk,1} = \mathbb{E} \left\{ I \left(\ddot{\mathbf{x}}_{lk}; \hat{\mathbf{r}}_{lk}, \hat{\Pi}_{lk} | \ddot{\mathbf{G}}_{llk} \right) \right\}$, which represents the MI between

the transmit signals $\ddot{\mathbf{x}}_{lk}$ and the detected TA index $\hat{\Pi}_{lk}$ as well as the detected APM signals $\hat{\mathbf{r}}_{lk}$, when the channel $\ddot{\mathbf{G}}_{llk}$ (true TA index) is known. The area corresponding to $R_{lk,1}$ has been fish-bone-shaded in Fig. 3. $R_{lk,1}$ can be divided into two terms as

$$\begin{aligned} R_{lk,1} &= \mathbb{E} \left\{ I \left(\ddot{\mathbf{x}}_{lk}; \hat{\mathbf{r}}_{lk} | \ddot{\mathbf{G}}_{llk}, \hat{\Pi}_{lk} \right) \right\} \\ & \quad + \mathbb{E} \left\{ I \left(\ddot{\mathbf{x}}_{lk}; \hat{\Pi}_{lk} | \ddot{\mathbf{G}}_{llk} \right) \right\}. \end{aligned} \quad (20)$$

Because the detected TA index $\hat{\Pi}_{lk}$ is directly related to the channel $\ddot{\mathbf{G}}_{llk}$ (true TA index) and it is independent of the transmit APM signals $\ddot{\mathbf{x}}_{lk}$, it can be seen that the second term of (20) obeys $\mathbb{E} \left\{ I \left(\ddot{\mathbf{x}}_{lk}; \hat{\Pi}_{lk} | \ddot{\mathbf{G}}_{llk} \right) \right\} = 0$. It can also be observed that according to Fig. 3, there is no intersection between $\ddot{\mathbf{x}}_{lk}$ and $\hat{\Pi}_{lk}$.

The first term of (20) can be interpreted as the MI between the transmit APM signal $\ddot{\mathbf{x}}_{lk}$ and the detected APM signal $\hat{\mathbf{r}}_{lk}$ given the channel $\ddot{\mathbf{G}}_{llk}$ (true TA index). Furthermore, in view of Fig. 3, both $\hat{\Pi}_{lk}$ and $\ddot{\mathbf{G}}_{llk}$ are uncorrelated with the mutual information (2) between $\ddot{\mathbf{x}}_{lk}$ and $\hat{\mathbf{r}}_{lk}$. Hence, the appearance of $\hat{\Pi}_{lk}$ and $\ddot{\mathbf{G}}_{llk}$ in the first term of (20) has no effect on the MI. The first term can be written as

$$\begin{aligned} \mathbb{E} \left\{ I \left(\ddot{\mathbf{x}}_{lk}; \hat{\mathbf{r}}_{lk} | \ddot{\mathbf{G}}_{llk}, \hat{\Pi}_{lk} \right) \right\} &= \mathbb{E} \left\{ I \left(\ddot{\mathbf{x}}_{lk}; \hat{\mathbf{r}}_{lk} | \ddot{\mathbf{G}}_{llk} \right) \right\} \\ &= \mathbb{E} \left\{ \frac{1}{C} \sum_{c=1}^C \sum_{t \in \Pi_c} \log_2 (1 + \text{SINR}_{lkt}) \right\}. \end{aligned} \quad (21)$$

Hence, we arrive at

$$\begin{aligned} R_{lk,1} &= \mathbb{E} \left\{ \frac{1}{C} \sum_{c=1}^C \sum_{t \in \Pi_c} \log_2 (1 + \text{SINR}_{lkt}) \right\} \\ &\approx \frac{1}{C} \sum_{c=1}^C \sum_{t \in \Pi_c} \log_2 \{ 1 + \mathbb{E} [\text{SINR}_{lkt}] \} \end{aligned} \quad (22)$$

To obtain the SINR_{lkt} when the channel $\ddot{\mathbf{G}}_{llk}$ is perfectly known, the APM signal detected received via the t th antenna of the k th user in the l th cell can be written as

$$\begin{aligned} \hat{y}_{lkt} &= \sqrt{\rho_{ul}} \hat{\mathbf{w}}_{lkt}^H \hat{\mathbf{g}}_{llkt} x_{lkt} + \sum_{s \in \Pi_c \setminus \{t\}} \sqrt{\rho_{ul}} \hat{\mathbf{w}}_{lkt}^H \hat{\mathbf{g}}_{llks} x_{lks} \\ & \quad + \sum_{(z,u) \neq (l,k)} \sum_s \sqrt{\rho_{ul}} \hat{\mathbf{w}}_{lkt}^H \hat{\mathbf{g}}_{lzus} \gamma_{zus} x_{zus} \\ & \quad + \sum_{z=1}^L \sum_{u=1}^K \sum_s \sqrt{\rho_{ul}} \hat{\mathbf{w}}_{lkt}^H \tilde{\mathbf{g}}_{lzus} \gamma_{zus} x_{zus} + \hat{\mathbf{w}}_{lkt}^H \mathbf{n}_l. \end{aligned} \quad (23)$$

However, the real channel is split into an estimated channel part and an estimation error part due to the specific nature of the MMSE filter. In (23), γ_{zus} has two legitimate status, i.e., $\gamma \in \{0, 1\}$, representing the inactive and active scenarios separately. The first term of (23) represents the target signal gleaned from the t th TA of the k th user in the l th cell. The second term represents the inter-antenna interference imposed by the other TAs of the k th user in the l th cell. The interference impinging from all the users except for the k th user in the l th cell is expressed by the third term. Moreover, the interference

caused by the channel estimation error of all users in all cells is formulated by the fourth term.

According to (23), the SINR of y_{lkt} , namely SINR_{lkt} , can be derived as in (24). Then, upon substituting (24) into (22), $R_{l_{k,1}}$ can be obtained.

2) *Calculation of $R_{l_{k,2}}$* : Let $R_{l_{k,2}} = \mathbb{E} \left\{ I(\ddot{\mathbf{G}}_{llk}; \hat{\mathbf{r}}_{lk}, \hat{\Pi}_{lk}) \right\}$, which represents the MI between the channel $\ddot{\mathbf{G}}_{llk}$ (true TA index) and the detected TA index $\hat{\Pi}_{lk}$, detected APM signals $\hat{\mathbf{r}}_{lk}$. The area corresponding to $R_{l_{k,1,2}}$ has also been shown by a specific pattern in Fig. 3. Similar to $R_{l_{k,1}}$, $R_{l_{k,2}}$ can be divided into two terms,

$$R_{l_{k,2}} = \mathbb{E} \left\{ I(\ddot{\mathbf{G}}_{llk}; \hat{\Pi}_{lk}) \right\} + \mathbb{E} \left\{ I(\ddot{\mathbf{G}}_{llk}; \hat{\mathbf{r}}_{lk} | \hat{\Pi}_{lk}) \right\}. \quad (25)$$

For the second term in (25), given the detected TA index $\hat{\Pi}_{lk}$, the detected signal $\hat{\mathbf{r}}_{lk}$ is independent of the channel $\ddot{\mathbf{G}}_{llk}$ (TA index). Hence, we have $\mathbb{E} \left\{ I(\ddot{\mathbf{G}}_{llk}; \hat{\mathbf{r}}_{lk} | \hat{\Pi}_{lk}) \right\} = 0$. It can also be seen from Fig. 3 that \mathbf{G}_{llk} has no intersection area with $\hat{\mathbf{r}}_{lk}$ excluding $\hat{\Pi}_{lk}$. Thus, we have

$$\begin{aligned} R_{l_{k,2}} &= \mathbb{E} \left\{ I(\ddot{\mathbf{G}}_{llk}; \hat{\Pi}_{lk}) \right\} \\ &\approx \sum_{i=1}^C \sum_{c=1}^C \frac{1}{C} P(i|c) \log_2 \frac{P(i|c)}{P(i)} \\ &= \sum_{i=1}^C \sum_{c=1}^C \frac{1}{C} P(i|c) \log_2 \frac{P(i|c)}{\sum_c \frac{1}{C} P(i|c)}, \end{aligned} \quad (26)$$

where $P(i|c)$ represents the conditional probability that TAs $\in \Pi_i$ is detected, when actually TAs $\in \Pi_c$ transmits the signal and $P(i|c)$ is a value averaged over the channel realizations, which results in the approximate equality in Eq. (26).

To obtain $P(i|c)$, the detected signal should be rewritten. In (23), $\ddot{\mathbf{G}}_{llk}$ is already known. While when $\ddot{\mathbf{G}}_{llk}$ is now unknown, the detected APM signal \hat{y}_{lkd} can be rewritten as ,

$$\begin{aligned} \hat{y}_{lkd} &= \sqrt{\rho_{ul}} \hat{\mathbf{w}}_{lkd}^H \hat{\mathbf{g}}_{llkt} x_{lkt} + \sum_{s \in \Pi_c \setminus \{t\}} \sqrt{\rho_{ul}} \hat{\mathbf{w}}_{lkd}^H \hat{\mathbf{g}}_{llks} x_{lks} \\ &+ \sum_{(z,u) \neq (l,k)} \sum_s \sqrt{\rho_{ul}} \hat{\mathbf{w}}_{lkd}^H \hat{\mathbf{g}}_{lzus} \gamma_{zus} x_{zus} \\ &+ \sum_{z=1}^L \sum_{u=1}^K \sum_s \sqrt{\rho_{ul}} \hat{\mathbf{w}}_{lkd}^H \tilde{\mathbf{g}}_{lzus} \gamma_{zus} x_{zus} + \hat{\mathbf{w}}_{lkd}^H \mathbf{n}_l, \end{aligned} \quad (27)$$

when the TA $t \in \Pi_c$ is actually activated, but TA $d \in \Pi_i$ is detected.

The power of the first term can be written as $\zeta_{lkd|t}$, i.e. the useful signal power at the receiver becomes:

$$\zeta_{lkd|t} = \rho_{ul} |\hat{\mathbf{w}}_{lkd}^H \hat{\mathbf{g}}_{llkt}|^2. \quad (28)$$

The power of the interference term $\xi_{lkd|t}$ can be written as,

$$\begin{aligned} \xi_{lkd|t} &= \sum_{s \in \Pi_c \setminus \{t\}} \rho_{ul} |\hat{\mathbf{w}}_{lkd}^H \hat{\mathbf{g}}_{llks}|^2 \\ &+ \sum_{(z,u) \neq (l,k)} \sum_s \rho_{ul} |\hat{\mathbf{w}}_{lkd}^H \hat{\mathbf{g}}_{lzus}|^2 \\ &+ \sum_{z=1}^L \sum_{u=1}^K \sum_s \rho_{ul} |\hat{\mathbf{w}}_{lkd}^H \tilde{\mathbf{g}}_{lzus}|^2 + \|\hat{\mathbf{w}}_{lkd}\|^2. \end{aligned} \quad (29)$$

Next, the power of \hat{y}_{lkd} can be obtained given $\zeta_{lkd|t}$ and $\xi_{lkd|t}$,

$$\begin{aligned} \kappa_{lkd|t} &\triangleq \mathbb{E} \left\{ \hat{y}_{lkd} \hat{y}_{lkd}^* \right\} \\ &= \mathbb{E} \left\{ \zeta_{lkd|t} + \xi_{lkd|t} \right\}. \end{aligned} \quad (30)$$

According to [23], when the number of TAs is large, the term $\hat{\mathbf{w}}_{lkd}^H \hat{\mathbf{g}}_{lzus}$ will converge to a constant. The probability density function (PDF) of $|\hat{y}_{lkd}|^2$ can be regarded as an exponential function given the transmitted signal x is Gaussian distributed, i.e.,

$$f_{|\hat{y}_{lkd}|^2}(a) = \frac{1}{\kappa_{lkd|t}} e^{-\frac{a}{\kappa_{lkd|t}}} \quad (31)$$

and its corresponding cumulative distribution function (CDF) is

$$F_{|\hat{y}_{lkd}|^2}(a) = 1 - e^{-\frac{a}{\kappa_{lkd|t}}}. \quad (32)$$

To obtain $P(i|c)$, let $\tilde{\Pi}_i$ be the complement of the detected TAs set Π_i , consisting of the TA indices that are not detected. Then, the sufficient and necessary condition to detect the TA set Π_i is that the minimum received signal power $|y_{lkd}|^2$ of TA d in Π_i is larger than the maximum received signal power $|y_{lkd}|^2$ of TA \tilde{d} in $\tilde{\Pi}_i$. This condition can be presented as

$$\min_{d \in \Pi_i} |y_{lkd}|^2 > \max_{\tilde{d} \in \tilde{\Pi}_i} |y_{lkd}|^2. \quad (33)$$

Next, the PDFs of $\min_{d \in \Pi_i} |y_{lkd}|^2$ and $\max_{\tilde{d} \in \tilde{\Pi}_i} |y_{lkd}|^2$ have to be calculated through the order statistics of independent but non-identically distributed random variables [26], which is formulated as:

$$\begin{aligned} f_{|\hat{y}_{lkd, \min}|^2}(a) &= \frac{1}{(S-1)!} + \\ &\left| \begin{array}{ccc} f_{|\hat{y}_{lkd_1}|^2}(a) & \cdots & f_{|\hat{y}_{lkd_S}|^2}(a) \\ 1 - F_{|\hat{y}_{lkd_1}|^2}(a) & \cdots & 1 - F_{|\hat{y}_{lkd_S}|^2}(a) \\ \vdots & & \vdots \\ 1 - F_{|\hat{y}_{lkd_1}|^2}(a) & \cdots & 1 - F_{|\hat{y}_{lkd_S}|^2}(a) \end{array} \right| + \\ &= \sum_{d \in \Pi_i} f_{|\hat{y}_{lkd}|^2}(a) \prod_{n \in \Pi_i \setminus \{d\}} [1 - F_{|\hat{y}_{lkn}|^2}(a)] \\ &= \sum_{d \in \Pi_i} \frac{1}{\kappa_{lkd|t}} e^{-\frac{a}{\kappa_{lkd|t}}} \prod_{n \in \Pi_i \setminus \{d\}} e^{-\frac{a}{\kappa_{lkn|t}}} \\ &= \sum_{d \in \Pi_i} \frac{1}{\kappa_{lkd|t}} e^{-\left(\sum_{d \in \Pi_i} \frac{1}{\kappa_{lkd|t}} \right) a}, \end{aligned} \quad (34)$$

where $+\|\|$ is the permanent [26] of the matrix.

$$\text{SINR}_{lkt} = \frac{\rho_{ul} |\hat{\mathbf{w}}_{lkt}^H \hat{\mathbf{g}}_{llkt}|^2}{\sum_{s \in \Pi_c \setminus \{t\}} \rho_{ul} |\hat{\mathbf{w}}_{lkt}^H \hat{\mathbf{g}}_{llks}|^2 + \sum_{(z,u) \neq (l,k)} \sum_s \gamma_{zus} \rho_{ul} |\hat{\mathbf{w}}_{lkt}^H \hat{\mathbf{g}}_{lzus}|^2 + \sum_{z=1}^L \sum_{u=1}^K \sum_s \gamma_{zus} \rho_{ul} |\hat{\mathbf{w}}_{lkt}^H \hat{\mathbf{g}}_{lzus}|^2 + \|\hat{\mathbf{w}}_{lkt}\|^2} \quad (24)$$

A more compact expression of $f_{|\hat{\mathbf{y}}_{lkd,\min}|^2}(a)$ can be obtained by letting $\lambda = \sum_{d \in \Pi_i} \frac{1}{\kappa_{lkd|t}}$ and substituting λ into (34), leading to:

$$f_{|\hat{\mathbf{y}}_{lkd,\min}|^2}(a) = \lambda e^{-\lambda a}. \quad (35)$$

Then, the CDF of $\mathbf{y}_{lkd,\min}$ can be expressed as

$$F_{|\hat{\mathbf{y}}_{lkd,\min}|^2}(a) = 1 - e^{-\lambda a}. \quad (36)$$

Similar to the derivation of $f_{|\hat{\mathbf{y}}_{lkd,\min}|^2}(a)$, the PDF of $\max_{\bar{d} \in \bar{\Pi}_i} |y_{lkd\bar{d}}|^2$ is derived as

$$\begin{aligned} f_{|\hat{\mathbf{y}}_{lkd,\max}|^2}(a) &= \frac{1}{(N-S-1)!} + \\ & \left| \begin{array}{ccc} F_{|\hat{\mathbf{y}}_{lkd\bar{d}_1}|^2}(a) & \cdots & F_{|\hat{\mathbf{y}}_{lkd\bar{d}_{N-S}}|^2}(a) \\ \vdots & & \vdots \\ F_{|\hat{\mathbf{y}}_{lkd\bar{d}_1}|^2}(a) & \cdots & F_{|\hat{\mathbf{y}}_{lkd\bar{d}_{N-S}}|^2}(a) \\ f_{|\hat{\mathbf{y}}_{lkd\bar{d}_1}|^2}(a) & \cdots & f_{|\hat{\mathbf{y}}_{lkd\bar{d}_{N-S}}|^2}(a) \end{array} \right| + \\ &= \sum_{\bar{d} \in \bar{\Pi}_i} f_{|\hat{\mathbf{y}}_{lkd\bar{d}}|^2}(a) \prod_{n \in \bar{\Pi}_i \setminus \{\bar{d}\}} F_{|\hat{\mathbf{y}}_{lkn}|^2}(a) \\ &= \sum_{\bar{d} \in \bar{\Pi}_i} \frac{1}{\kappa_{lkd\bar{d}|t}} e^{-\frac{a}{\kappa_{lkd\bar{d}|t}}} \prod_{n \in \bar{\Pi}_i \setminus \{\bar{d}\}} \left(1 - e^{-\frac{a}{\kappa_{lkn|t}}}\right). \quad (37) \end{aligned}$$

Given $F_{|\hat{\mathbf{y}}_{lkd,\min}|^2}(a)$ and $f_{|\hat{\mathbf{y}}_{lkd,\max}|^2}(a)$, $P(i|c)$ can be obtained as

$$\begin{aligned} P(i|c) &= P(|\hat{\mathbf{y}}_{lkd,\min}|^2 > |\hat{\mathbf{y}}_{lkd,\max}|^2) \\ &= \int_0^\infty (1 - F_{|\hat{\mathbf{y}}_{lkd,\min}|^2}(a)) f_{|\hat{\mathbf{y}}_{lkd,\max}|^2}(a) da \\ &= \int_0^\infty e^{-\lambda a} \left[\sum_{\bar{d} \in \bar{\Pi}_i} \frac{1}{\kappa_{lkd\bar{d}|t}} e^{-\frac{a}{\kappa_{lkd\bar{d}|t}}} \prod_{n \in \bar{\Pi}_i \setminus \{\bar{d}\}} \left(1 - e^{-\frac{a}{\kappa_{lkn|t}}}\right) \right] da \\ &= \sum_{\bar{d} \in \bar{\Pi}_i} \frac{1}{\kappa_{lkd\bar{d}|t}} \int_0^\infty e^{-\lambda a} e^{-\frac{a}{\kappa_{lkd\bar{d}|t}}} \prod_{n \in \bar{\Pi}_i \setminus \{\bar{d}\}} \left(1 - e^{-\frac{a}{\kappa_{lkn|t}}}\right) da \\ &= \sum_{\bar{d} \in \bar{\Pi}_i} \frac{1}{\kappa_{lkd\bar{d}|t}} \left(\sum_{\bar{\Pi}_n \subset \bar{\Pi}_i \setminus \{\bar{d}\}} \frac{1}{\lambda + \frac{1}{\kappa_{lkd\bar{d}|t}} + (-1)^{|\bar{\Pi}_n|} \Psi(\bar{\Pi}_n)} \right), \quad (38) \end{aligned}$$

where

$$\Psi(\Xi) = \sum_{d \in \Xi} \frac{1}{\kappa_{lkd|t}}. \quad (39)$$

It is noted that $\kappa_{lkd|t}$ in (38) is an average value over channel realizations.

Finally, substituting $P(i|c)$ obtained in (38) into (26), $R_{lk,2}$ can be obtained.

To make this paper clearer and facilitate the analysis, we write down the the final specific expression of the achievable sum-rate R_{sim} as following,

$$\begin{aligned} R_{\text{sim}} &= (1 - \frac{KN}{T})(R_{lk,1} + R_{lk,2}) \\ &\approx (1 - \frac{KN}{T}) \left(\frac{1}{C} \sum_{c=1}^C \sum_{t \in \Pi_c} \log_2 \{1 + \mathbb{E}[\text{SINR}_{lkt}]\} \right. \\ & \quad \left. + \sum_{i=1}^C \sum_{c=1}^C \frac{1}{C} P(i|c) \log_2 \frac{P(i|c)}{\sum_c \frac{1}{C} P(i|c)} \right) \quad (40) \end{aligned}$$

where SINR_{lkt} is shown in Eq. (24) and $P(i|c)$ is shown in Eq. (38).

The physical interpretation of each term in SINR_{lkt} has been introduced right below Eq. (23). The detector $\hat{\mathbf{w}}_{lkd}$ is expected to increase the SINR_{lkt} in Eq. (22), hence increasing $R_{lk,1}$ of Eq. (22). For example, the MF detector is capable of effectively increasing the numerator of SINR_{lkt} of Eq. (24) in Eq. (22) while ZF performs well in terms of reducing the interference, i.e. the denominator of SINR_{lkt} .

As for $R_{lk,2}$, $\kappa_{lkd|t}$ in $P(i|c)$ of Eq. (38), the power received from the N antennas plays a key role, because the power is expected to have a relatively high value for the specific TAs that actually transmit data, thereby increasing the probability of correctly detecting the activated TAs, which is also reminiscent of amplifying the power of useful signal and suppressing the interference by a ZF detector.

IV. APPROXIMATION

When the number of TAs tends to infinity, the approximation becomes accurate. The approximation of the exact achievable sum-rate R_{sim} is actually based on the approximation of the following five terms, $|\hat{\mathbf{w}}_{lkt}^H \hat{\mathbf{g}}_{llkt}|^2$, $|\hat{\mathbf{w}}_{lkt}^H \hat{\mathbf{g}}_{llks}|^2$, $|\hat{\mathbf{w}}_{lkt}^H \hat{\mathbf{g}}_{lzus}|^2$, $|\hat{\mathbf{w}}_{lkt}^H \tilde{\mathbf{g}}_{lzus}|^2$, $\|\hat{\mathbf{w}}_{lkt}\|^2$ both in the SINR_{lkt} expression of Eq. (24) and in $\kappa_{lkd|t}$ of Eq. (30). Upon obtaining these five approximate terms associated with the SMMSE, ZF and MF detectors, the approximate SINR_{lkt} and approximate $\kappa_{lkd|t}$ expressed as $\text{SINR}_{lkt}^{\text{appr}}$ and $\kappa_{lkd|t}^{\text{appr}}$ can be readily acquired. Then, the specific expression of the approximate sum-rate R^{appr} is given by

$$\begin{aligned} R^{\text{appr}} &= (1 - \frac{KN}{T})(R_{lk,1}^{\text{appr}} + R_{lk,2}^{\text{appr}}) \\ &= (1 - \frac{KN}{T}) \left(\frac{1}{C} \sum_{c=1}^C \sum_{t \in \Pi_c} \log_2 \{1 + \text{SINR}_{lkt}^{\text{appr}}\} \right. \\ & \quad \left. + \sum_{i=1}^C \sum_{c=1}^C \frac{1}{C} P(i|c)^{\text{appr}} \log_2 \frac{P(i|c)^{\text{appr}}}{\sum_c \frac{1}{C} P(i|c)^{\text{appr}}} \right) \quad (41) \end{aligned}$$

where $P(i|c)^{\text{appr}}$ is obtained upon replacing $\kappa_{lkd|t}$ in the $P(i|c)$ expression of Eq. (38) with $\kappa_{lkd|t}^{\text{appr}}$.

To derive the approximate five terms above mentioned, let us first define Φ_{lzu} and Φ_{lz-lk} as a block matrix in the following form,

$$\Phi_{lzu}/\Phi_{lz-lk} = \begin{bmatrix} \Delta_{11} & \Delta_{12} & \cdots & \Delta_{1N} \\ \Delta_{21} & \Delta_{22} & \cdots & \Delta_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ \Delta_{N1} & \Delta_{N2} & \cdots & \Delta_{NN} \end{bmatrix}, \quad (42)$$

where each block matrix Δ_{ij} is an $(M \times M)$ -element matrix and represents the block matrix in the i th row and j th column of Φ_{lzu} or Φ_{lz-lk} , expressed as $\Phi_{lzu(i,j)}$ or $\Phi_{lz-lk(i,j)}$.

The following is the approximate process of the five terms respectively for SMMSE, ZF and MF.

A. SMMSE approximation

For the SMMSE detector, the product of the estimated channel and its Hermitian is approximated by its correlation matrix. Then $\hat{\mathbf{W}}_{lk}^{\text{SMMSE}}$ can be approximated as

$$\begin{aligned} \hat{\mathbf{W}}_{lk}^{\text{SMMSE}} &= \left(\sum_{k=1}^K \hat{\mathbf{G}}_{lk} \hat{\mathbf{G}}_{lk}^H + \mathbf{Z}_l \right)^{-1} \hat{\mathbf{G}}_{lk} \\ &\approx \left(\sum_{z=1}^L \sum_{k=1}^K \beta_{lzk} \mathbf{R}_{lzk}^d + \frac{1}{\rho_{ul}} \mathbf{I}_M \right)^{-1} \hat{\mathbf{G}}_{lk}, \end{aligned} \quad (43)$$

where

$$\mathbf{Z}_l = \sum_{k=1}^K (\beta_{lk} \mathbf{R}_{lk}^d - \Phi_{llk}^d) + \sum_{\substack{z=1 \\ z \neq l}}^L \sum_{k=1}^K (\beta_{lzk} \mathbf{R}_{lzk}^d) + \frac{1}{\rho_{ul}} \mathbf{I}_M \quad (44)$$

The approximation result of (43) is obtained by using $\sum_{k=1}^K (-\Phi_{llk}^d)$ in \mathbf{Z}_l to approximate $\sum_{k=1}^K \hat{\mathbf{G}}_{lk} \hat{\mathbf{G}}_{lk}^H$.

Let us use the shorthand of $\mathbf{X} = \left(\sum_{z=1}^L \sum_{k=1}^K \beta_{lzk} \mathbf{R}_{lzk}^d + \frac{1}{\rho_{ul}} \mathbf{I}_M \right)$. Then substituting it into (43), we can get $\hat{\mathbf{W}}_{lk}^{\text{SMMSE}} = \mathbf{X} \hat{\mathbf{G}}_{lk}$. Hence, the detector for the t th TA of the k th user in the l th cell can be written as

$$\hat{\mathbf{w}}_{lkt}^H = \hat{\mathbf{g}}_{lkt}^H \mathbf{X}. \quad (45)$$

To obtain the SINR $_{lkt}$ for calculating $R_{lk,1}$ and $\kappa_{lkd|t}$ for calculating $R_{lk,2}$, the following five terms in Eq. (24) are approximated which can also be applied to Eq. (28) and Eq. (29).

The first term is the useful information signal in the numerator of Eq. (24) which has a mean power of

$$\begin{aligned} \mathbb{E}[|\hat{\mathbf{w}}_{lkt}^H \hat{\mathbf{g}}_{llkt}|^2] &= \mathbb{E}\left[|\hat{\mathbf{g}}_{llkt}^H \mathbf{X} \hat{\mathbf{g}}_{llkt}|^2\right] \\ &= \left\{ \text{tr} [\mathbf{X} \Phi_{llk(t,t)}] \right\}^2, \end{aligned} \quad (46)$$

for $\mathbb{E}\left[|\hat{\mathbf{g}}_{llkt}^H \hat{\mathbf{g}}_{llkt}|\right] = \Phi_{llk(t,t)}$.

The second term is the mean power of the inter-antenna interference signal in the denominator of Eq. (24) imposed by the same user

$$\begin{aligned} \mathbb{E}[|\hat{\mathbf{w}}_{lkt}^H \hat{\mathbf{g}}_{llks}|^2] &= \mathbb{E}\left[|\hat{\mathbf{g}}_{llkt}^H \mathbf{X} \hat{\mathbf{g}}_{llks}|^2\right] \\ &= \left\{ \text{tr} [\mathbf{X} \Phi_{llk(s,t)}] \right\}^2, \end{aligned} \quad (47)$$

for $\mathbb{E}\left[|\hat{\mathbf{g}}_{llks}^H \hat{\mathbf{g}}_{llkt}|\right] = \Phi_{llk(s,t)}$.

The third term is the mean power of the interference signal in the denominator of Eq. (24) imposed by the other users in all cells. Since $\hat{\mathbf{g}}_{llkt}$ is independent of $\hat{\mathbf{g}}_{lzus}$ when $u \neq k$, we can get

$$\begin{aligned} \mathbb{E}[|\hat{\mathbf{w}}_{lkt}^H \hat{\mathbf{g}}_{lzus}|^2] &= \mathbb{E}\left[|\hat{\mathbf{g}}_{llkt}^H \mathbf{X} \hat{\mathbf{g}}_{lzus}|^2\right] \\ &= \begin{cases} \text{tr} \left\{ \mathbb{E} \left[|\mathbf{X} \hat{\mathbf{g}}_{lzus} \hat{\mathbf{g}}_{lzus}^H \mathbf{X} \hat{\mathbf{g}}_{llkt} \hat{\mathbf{g}}_{llkt}^H| \right] \right\} & \text{if } u \neq k \\ \text{tr} \left\{ \mathbb{E} \left[\mathbf{X} \hat{\mathbf{g}}_{lzus} \hat{\mathbf{g}}_{llkt}^H \right] \right\}^2 & \text{if } u = k \end{cases} \\ &= \begin{cases} \text{tr} \left\{ |\mathbf{X} \Phi_{lzu(s,s)} \mathbf{X} \Phi_{llk(t,t)}| \right\} & \text{if } u \neq k \\ \text{tr} \left\{ \mathbf{X} \Phi_{lz-lk(s,t)} \right\}^2 & \text{if } u = k. \end{cases} \end{aligned} \quad (48)$$

The fourth term is the mean power of the interference signal in the denominator of Eq. (24) due to the channel estimation error of all cells. Since $\tilde{\mathbf{g}}_{ljus}$ is independent of $\tilde{\mathbf{g}}_{llkt}$, we arrive at:

$$\begin{aligned} \mathbb{E}[|\hat{\mathbf{w}}_{lkt}^H \tilde{\mathbf{g}}_{lzus}|^2] &= \mathbb{E}\left[|\hat{\mathbf{g}}_{llkt}^H \mathbf{X} \tilde{\mathbf{g}}_{lzus}|^2\right] \\ &= \text{tr} \left[\mathbf{X} \tilde{\mathbf{g}}_{lzus} \tilde{\mathbf{g}}_{lzus}^H \mathbf{X} \hat{\mathbf{g}}_{llkt} \hat{\mathbf{g}}_{llkt}^H \right] \\ &= \text{tr} \left[\mathbf{X} (\mathbf{R}_{lzu(s,s)} - \Phi_{lzu(s,s)}) \mathbf{X} \Phi_{llk(t,t)} \right]. \end{aligned} \quad (49)$$

The fifth term, i.e. the last term in the denominator of (24) and the last term in (29), can be calculated as

$$\begin{aligned} \mathbb{E}\left[|\hat{\mathbf{w}}_{lkt}|^2\right] &= \mathbb{E}\left[|\hat{\mathbf{g}}_{llkt}^H \mathbf{X} \hat{\mathbf{g}}_{llkt}|^2\right] \\ &= \text{tr} \left[\mathbf{X} \mathbf{X} \Phi_{llk(t,t)} \right]. \end{aligned} \quad (50)$$

Substituting (46), (47), (48), (49), (50) into (24) and (30), the approximated $R_{lk,1}$ and $R_{lk,2}$ using the SMMSE can be obtained.

B. ZF approximation

The zero forcing detector $\hat{\mathbf{W}}_{lk}^{\text{ZF}}$ of the k th user in the l th cell can be formulated as

$$\hat{\mathbf{W}}_{lk}^{\text{ZF}} = \left[\hat{\mathbf{G}}_{ll} \left(\hat{\mathbf{G}}_{ll}^H \hat{\mathbf{G}}_{ll} \right)^{-1} \right]_{[(k-1)MN+1:kMN]} \quad (51)$$

For the mean signal power received from the l th cell, $\mathbb{E}[|\hat{\mathbf{w}}_{lkt}^H \hat{\mathbf{g}}_{llus}|^2]$ equals 1 only when the signal is transmitted from the t th antenna of the k th user in the l th cell, otherwise it equals 0, i.e.,

$$\mathbb{E}[|\hat{\mathbf{w}}_{lkt}^H \hat{\mathbf{g}}_{llus}|^2] = \begin{cases} 1 & \text{if } (u, s) = (k, t) \\ 0 & \text{otherwise.} \end{cases} \quad (52)$$

The relationship between $\hat{\mathbf{G}}_{lzk}$ and $\hat{\mathbf{G}}_{llk}$ can be obtained according to (5), which is given by:

$$\begin{aligned} \text{vec}(\hat{\mathbf{G}}_{lzk}) &= \frac{\beta_{lzk}}{\beta_{llk}} \mathbf{R}_{lzk} \mathbf{R}_{llk}^{-1} \text{vec}(\hat{\mathbf{G}}_{llk}) \\ &= \frac{\beta_{lzk}}{\beta_{llk}} [(\mathbf{R}_{lzk} \mathbf{R}_{llk}^{-1}) \otimes \mathbf{I}_M] \text{vec}(\hat{\mathbf{G}}_{llk}). \end{aligned} \quad (53)$$

Then, the averaged interference power $\mathbb{E}[|\hat{\mathbf{w}}_{lkt}^H \hat{\mathbf{g}}_{lzus}|^2]$ impinging from other cells can be calculated as

$$\begin{aligned} \mathbb{E}[|\hat{\mathbf{w}}_{lkt}^H \hat{\mathbf{g}}_{lzus}|^2] &= \begin{cases} \left[\frac{\beta_{lzk}}{\beta_{lik}} (\mathbf{R}_{zk} \mathbf{R}_{lk}^{-1})_{(s,t)} \right]^2 & \text{if } u = k \\ \text{tr} \left(\mathbb{E} [\hat{\mathbf{g}}_{lzus} \hat{\mathbf{g}}_{lzus}^H \hat{\mathbf{w}}_{lkt} \hat{\mathbf{w}}_{lkt}^H] \right) & \text{if } u \neq k \end{cases} \\ &= \begin{cases} \left[\frac{\beta_{lzk}}{\beta_{lik}} (\mathbf{R}_{zk} \mathbf{R}_{lk}^{-1})_{(s,t)} \right]^2 & \text{if } u = k \\ \text{tr} \left(\Phi_{lzu(s,s)} \hat{\mathbf{w}}_{lkt} \hat{\mathbf{w}}_{lkt}^H \right) & \text{if } u \neq k. \end{cases} \end{aligned} \quad (54)$$

Exploiting the characteristic of ZF that $(\hat{\mathbf{W}}_{lk}^{\text{ZF}})^H \hat{\mathbf{G}}_{ll} = \mathbf{I}$, both (52) and the $u = k$ case of (54) can be obtained.

Since $\hat{\mathbf{w}}_{lkt}$ is independent of $\hat{\mathbf{g}}_{lzus}$ and $\tilde{\mathbf{g}}_{lzus}$, the $u \neq k$ case in (54) and the average interference power imposed due to the non-zero channel estimation error, namely $\mathbb{E}[|\hat{\mathbf{w}}_{lkt}^H \tilde{\mathbf{g}}_{lzus}|^2]$, can be obtained as

$$\begin{aligned} \mathbb{E}[|\hat{\mathbf{w}}_{lkt}^H \tilde{\mathbf{g}}_{lzus}|^2] &= \mathbb{E}[\hat{\mathbf{w}}_{lkt}^H \tilde{\mathbf{g}}_{lzus} \tilde{\mathbf{g}}_{lzus}^H \hat{\mathbf{w}}_{lkt}] \\ &= \text{tr} \left(\mathbb{E} [\tilde{\mathbf{g}}_{lzus} \tilde{\mathbf{g}}_{lzus}^H \hat{\mathbf{w}}_{lkt} \hat{\mathbf{w}}_{lkt}^H] \right) \\ &= \text{tr} \left((\mathbf{R}_{lzu(s,s)} - \Phi_{lzu(s,s)}) \hat{\mathbf{w}}_{lkt} \hat{\mathbf{w}}_{lkt}^H \right). \end{aligned} \quad (55)$$

The square of the Frobenius norm item, i.e. the last term in the denominator of (24) and the last term in (29), can be approximated as

$$\begin{aligned} \mathbb{E} [\|\hat{\mathbf{w}}_{lkt}\|^2] &= \mathbb{E} \left[\left(\hat{\mathbf{G}}_{llk}^H \hat{\mathbf{G}}_{llk} \right)^{-1} \right]_{[t,t]} \\ &\approx (\Phi_{llk}^{\text{tr}})^{-1}_{[t,t]}, \end{aligned} \quad (56)$$

where Φ_{llk}^{tr} is the matrix obtained by calculating the trace of each block in (42), i.e.,

$$\Phi_{llk}^{\text{tr}} = \begin{bmatrix} \text{tr}(\Delta_{11}) & \text{tr}(\Delta_{12}) & \cdots & \text{tr}(\Delta_{1N}) \\ \text{tr}(\Delta_{21}) & \text{tr}(\Delta_{22}) & \cdots & \text{tr}(\Delta_{2N}) \\ \vdots & \vdots & \vdots & \vdots \\ \text{tr}(\Delta_{N1}) & \text{tr}(\Delta_{N2}) & \cdots & \text{tr}(\Delta_{NN}) \end{bmatrix}. \quad (57)$$

Let us now exploit that the item $\hat{\mathbf{w}}_{lkt} \hat{\mathbf{w}}_{lkt}^H$ in (54) and (55) can be approximated as

$$\mathbb{E} [\hat{\mathbf{w}}_{lkt} \hat{\mathbf{w}}_{lkt}^H] \approx \mathbb{E} \left\{ [\hat{\mathbf{G}}_{llk} (\Phi_{llk}^{\text{tr}})^{-1}]_{[t,t]} [(\Phi_{llk}^{\text{tr}})^{-1} \hat{\mathbf{G}}_{llk}^H]_{[t,t]} \right\}. \quad (58)$$

Assuming that ϕ_{lkt} is the t th column of $(\Phi_{llk}^{\text{tr}})^{-1}$, $\mathbb{E} [\hat{\mathbf{w}}_{lkt} \hat{\mathbf{w}}_{lkt}^H]$ can be approximated by

$$\begin{aligned} \mathbb{E} [\hat{\mathbf{w}}_{lkt} \hat{\mathbf{w}}_{lkt}^H] &\approx \mathbb{E} \left[\hat{\mathbf{G}}_{llk} \phi_{lkt} \phi_{lkt}^H \hat{\mathbf{G}}_{llk}^H \right] \\ &= \text{Sum}_{M \times M} \{ [(\phi_{lkt} \phi_{lkt}^H) \otimes \mathbf{I}_M] \bullet \Phi_{llk} \}, \end{aligned} \quad (59)$$

where $\text{Sum}_{M \times M}(\cdot)$ represents dividing the matrix into M -by- M matrices, which are then summed up. In detail, let $B = [(\phi_{lkt} \phi_{lkt}^H) \otimes \mathbf{I}_M] \bullet \Phi_{llk}$ be an $MN \times MN$ matrix which can be written as the following block matrix,

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \cdots & \mathbf{B}_{1N} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \cdots & \mathbf{B}_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{B}_{N1} & \mathbf{B}_{N2} & \cdots & \mathbf{B}_{NN} \end{bmatrix}, \quad (60)$$

where each matrix \mathbf{B}_{ij} is an $M \times M$ matrix taken from the $((i-1)M+1 \sim iM)$ th rows, $((j-1)M+1 \sim jM)$ th columns of \mathbf{B} . Then, $\text{Sum}_{M \times M}(\mathbf{B})$ can be expressed as the sum of all the matrix blocks, i.e., $\text{Sum}_{M \times M}(\mathbf{B}) = \sum_j \sum_i \mathbf{B}_{ij}$.

Substituting (52), (54), (55) and (56) into (24) and (30) separately, the approximated $R_{lk,1}$ and $R_{lk,2}$ using ZF can be obtained.

C. MF approximation

For the MF detector, the approximation method is similar to that of the MMSE method, but the difference is that in contrast to Eq. (43), no \mathbf{X} matrix is invoked in the MF detector. Hence, the MF detector's $\hat{\mathbf{W}}_{lk}^{\text{MF}}$ is expressed as

$$\hat{\mathbf{W}}_{lk}^{\text{MF}} = \hat{\mathbf{G}}_{llk}. \quad (61)$$

The mean power of the useful signal term in the numerator of Eq. (24) can be calculated using the covariance matrix $\Phi_{llk(t,t)}$ to approximate the expectation of the estimated channel correlation $\mathbb{E}[\hat{\mathbf{g}}_{llkt} \hat{\mathbf{g}}_{llkt}^H]$, we have:

$$\mathbb{E}[|\hat{\mathbf{w}}_{lkt}^H \hat{\mathbf{g}}_{llkt}|^2] = \mathbb{E}[\text{tr}^2(\hat{\mathbf{g}}_{llkt} \hat{\mathbf{g}}_{llkt}^H)] = \text{tr}^2(\Phi_{llk(t,t)}). \quad (62)$$

The mean inter-antenna interference signal power term in the denominator of Eq. (24) is computed via using the covariance matrix $\Phi_{llk(s,t)}$ between different antennas of the same user to approximate the expectation of the estimated channel correlation, i.e.,

$$\mathbb{E}[|\hat{\mathbf{w}}_{lkt}^H \hat{\mathbf{g}}_{llks}|^2] = \mathbb{E}[\text{tr}^2(\hat{\mathbf{g}}_{llks} \hat{\mathbf{g}}_{llks}^H)] = \text{tr}^2(\Phi_{llk(s,t)}). \quad (63)$$

For the mean power of the interference arriving from other cells, if it arrives from a user relying on the same pilot, the covariance matrix $\Phi_{lz-lk(s,t)}$ between different estimated channels can be used to approximate the expectation of the different cells' estimated channel correlation. Otherwise, since $\hat{\mathbf{g}}_{lzus}$ is independent of $\hat{\mathbf{g}}_{llkt}$, the $u \neq k$ case of (64) can be obtained:

$$\begin{aligned} \mathbb{E}[|\hat{\mathbf{w}}_{lkt}^H \hat{\mathbf{g}}_{lzus}|^2] &= \begin{cases} \text{tr} \left(\mathbb{E} [\hat{\mathbf{g}}_{lzus} \hat{\mathbf{g}}_{lzus}^H \hat{\mathbf{w}}_{lkt} \hat{\mathbf{w}}_{lkt}^H] \right) & \text{if } u \neq k \\ \left\{ \mathbb{E}[\text{tr}(\hat{\mathbf{g}}_{lzus} \hat{\mathbf{g}}_{lzus}^H)] \right\}^2 & \text{if } u = k \end{cases} \\ &= \begin{cases} \text{tr} \left(\Phi_{lzu(s,s)} \Phi_{llk(t,t)} \right) & \text{if } u \neq k \\ \left\{ \text{tr} \left[\Phi_{lz-lk(s,t)} \right] \right\}^2 & \text{if } u = k. \end{cases} \end{aligned} \quad (64)$$

For the mean interference power imposed by the non-zero channel estimation error term, since $\tilde{\mathbf{g}}_{lzus}$ is independent of $\hat{\mathbf{g}}_{llkt}$, we can express the mean power $\mathbb{E}[|\hat{\mathbf{w}}_{lkt}^H \tilde{\mathbf{g}}_{lzus}|^2]$ of the interference caused by the non-zero channel estimation error as follows:

$$\begin{aligned} \mathbb{E}[|\hat{\mathbf{w}}_{lkt}^H \tilde{\mathbf{g}}_{lzus}|^2] &= \mathbb{E}[\hat{\mathbf{g}}_{llkt}^H \tilde{\mathbf{g}}_{lzus} \tilde{\mathbf{g}}_{lzus}^H \hat{\mathbf{g}}_{llkt}] \\ &= \text{tr} \left(\mathbb{E} [\tilde{\mathbf{g}}_{lzus} \tilde{\mathbf{g}}_{lzus}^H \hat{\mathbf{g}}_{llkt} \hat{\mathbf{g}}_{llkt}^H] \right) \\ &= \text{tr} \left((\mathbf{R}_{lzu(s,s)} - \Phi_{lzu(s,s)}) \Phi_{llk(t,t)} \right). \end{aligned} \quad (65)$$

As for the fifth item, i.e. the last term in the denominator of (24) and the last term in (29), it can be approximated as

$$\begin{aligned} \mathbb{E} [\|\hat{\mathbf{w}}_{lkt}\|^2] &= \mathbb{E} [\hat{\mathbf{g}}_{llkt}^H \hat{\mathbf{g}}_{llkt}] \\ &= \text{tr} [\Phi_{llk(t,t)}]. \end{aligned} \quad (66)$$

Upon substituting (62), (63), (64) and (66) into (24) and (30) separately, the approximate $R_{lk,1}$ and $R_{lk,2}$ expression of using the MF can be obtained.

V. NUMERICAL RESULTS

In this section, we consider a seven-cell example to study the sum-rate in the face of the inter-cell interference caused by pilot contamination and in the presence of inter-antenna interference. In particular, only a single user is considered in each cell because pilots in each cell are orthogonal. Furthermore, each user is equipped with $N = 4$ TAs. The channel coherence time is assumed to $T = 16$. The large scale fading β_{lk} for its own cell will be normalized to 1. Numerical results section is divided into two parts: in the first part, we only consider the sum-rate when two antennas are activated at any instant. In the second part, different number of activated antennas are compared.

A. sum-rate when two antennas are activated

Two types of channel covariance matrix are considered in this paper. For the correlation factor of the exponential model based covariance matrix at the user's side and at the BS's side are both set to $r = 0.2$. For the Bessel function based covariance matrix, the antenna spacing at the user's side and at the BS's side are both set to $d_t = d_r = 0.1$ and the survival factor α is set to be 0.3. Moreover, the pilot transmit power and the UL data transmit power are both assumed to be 10dB.

Fig. 4 shows the multi-cell SM's sum-rate versus the number of TAs at each BS with the inter-cell large scale fading interference factor $\beta_{ljk} = 0.25$ under the exponential covariance matrix and the Bessel covariance matrix separately. The number of TAs ranges from 10 to 359. It is observed in Fig. 4 that for both type of covariance matrix model, SMMSE shows the best performance for an arbitrary number of TAs. MF is better than ZF for a lower number of TAs. When M is gradually increasing, ZF surpasses MF and approaches SMMSE. Since the growth rate of the sum of inter-antenna interference and of the inter-cell interference along with M is higher than that of the useful signal, the system becomes interference-limited upon increasing the number of antennas. The MF is designed to amplify the desired signal, while ZF is designed to suppress the interference. Hence, ZF gradually starts to exhibit an improved performance as M grows. SMMSE combines the advantages of ZF and MF, which makes it our proposed option. Furthermore, for the exponential covariance matrix, the approximated values of both SMMSE and ZF are close to the simulation based values. While for the Bessel covariance matrix, the approximated ZF and MF cases are both close to their simulated values, while the approximated value of SMMSE deviates from its corresponding simulated value. This indicates that the accuracy of the approximation is sensitive to the covariance matrix model assumed.

Fig. 5 shows the sum-rate versus the SNR with the interference factor of $\beta_{ljk} = 0.25$ and correlation factor of $r = 0.2$ under the exponential covariance matrix for different

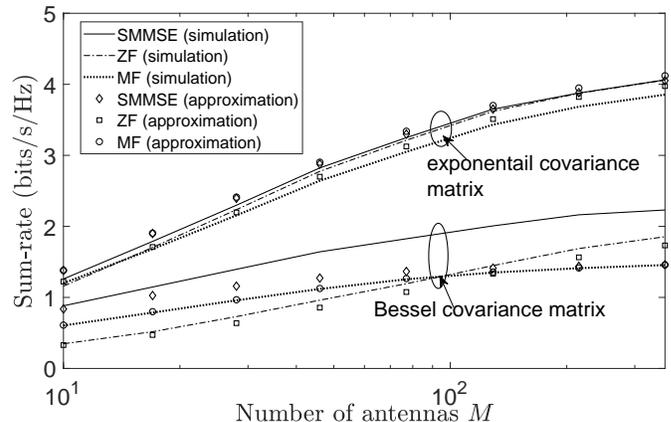


Fig. 4. Sum-rate w.r.t. the number of antennas at each BS under the exponential covariance matrix and the Bessel covariance matrix.

values of M . It is shown that the sum-rate increases rapidly when the SNR increases from -20 dB to 10 dB, but beyond 10 dB it becomes flat even when the SNR continues to grow which indicates that it is meaningful to increase the SNR in a certain range. The scenarios associated with $M = 50$ and $M = 100$ are also compared. They show similar trends w.r.t. the SNR and the sum-rate of $M = 100$ is always higher than that of $M = 50$, the SMMSE is better than ZF and ZF is better than MF for both $M = 100$ and $M = 50$ while ZF is closer to SMMSE which is consistent with Fig. 4. In contrast, Fig. 6 shows the sum-rate versus the SNR for the interference factor of $\beta_{ljk} = 0.25$ and antenna spacing factor of $d_r = d_t = 0.3$ under the Bessel covariance matrix which shows similar characteristics with Fig. 5.

Fig. 7 shows the sum-rate versus the interference factor with the correlation factor of $r = 0.2$ and SNR = 10dB under the exponential covariance matrix for different values of M . The sum-rate of $M = 50$ and $M = 100$ gradually decreases as the interference factor increases, because the pilot contamination becomes more severe. Furthermore, the gap between the curves of $M = 50$ and $M = 100$ becomes smaller as the interference factor increases. This means that the increased number of TAs can not be exploited, when the inter-cell interference becomes more significant. In contrast, Fig. 8 shows the sum-rate versus the interference factor for antenna spacing factor of $d_r = d_t = 0.3$ and SNR = 10dB under the Bessel covariance matrix for $M = 50$ and $M = 100$ respectively.

B. sum-rate when any number of antennas are activated

In this part, we compare the sum-rate w.r.t. the number of antennas M at each BS when different number of antennas are activated at each user side under the two types of channel covariance matrix model. For exponential covariance matrix model, we choose the decoder with admirable approximation performance, i.e., SMMSE and ZF, to display the effect of activating different number of antennas which is given in Fig. 9 and Fig. 10, respectively. In contrast, ZF and MF are chosen to elaborate the effect of different number of activating

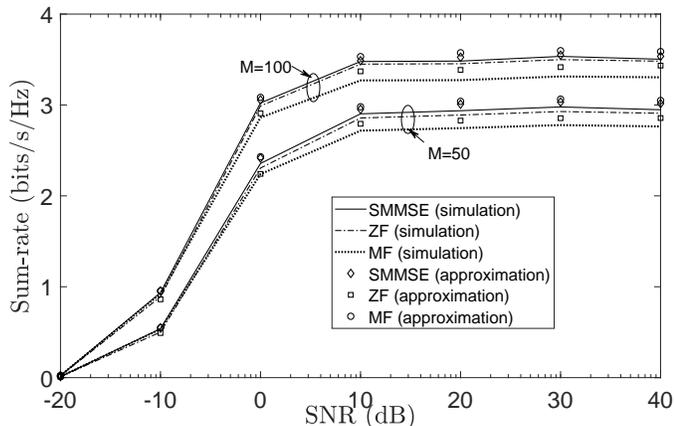


Fig. 5. Sum-rate w.r.t. SNR for $M = 50$ and $M = 100$ under the exponential covariance matrix.

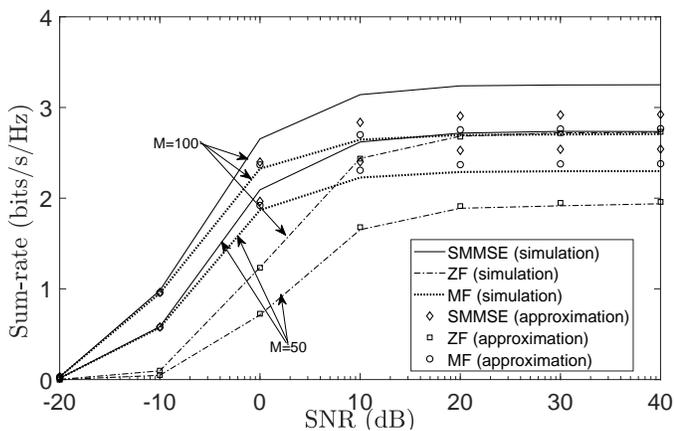


Fig. 6. Sum-rate w.r.t. SNR for $M = 50$ and $M = 100$ under the Bessel covariance matrix.

antennas which is shown in Fig. 12 and 13, respectively. Results from them all show that increasing the number of activated antennas always leads to a higher sum-rate yet with a smaller growth rate since the distance between the two lines is getting smaller. Moreover, more activating antennas also means more power consumption. The energy efficiency is an important factor to be considered for next generation wireless communication systems and the energy efficiency of a system is defined as the sum-rate (the spectral efficiency (bits/s/Hz)) divided by the transmit power (the UL data transmit power of each antenna multiplied by the number of activated antennas (J/s)) as following,

$$\eta = \frac{R_{\text{sim}}}{S \cdot \rho_{ul}} \text{ (bits/J/Hz)} \quad (67)$$

The relationship between the energy efficiency and the spectral efficiency with $M = 100$ is given in Fig. 11 and Fig. 14 respectively for exponential covariance matrix and Bessel covariance matrix. For comparison of the different decoders, SMMSE gives the best energy efficiency across the entire spectral efficiency range while ZF gives the worst performance but will exceed MF as the number of activated antennas increases and the spectral efficiency increases. For comparison of different number of activated antennas, it is observed

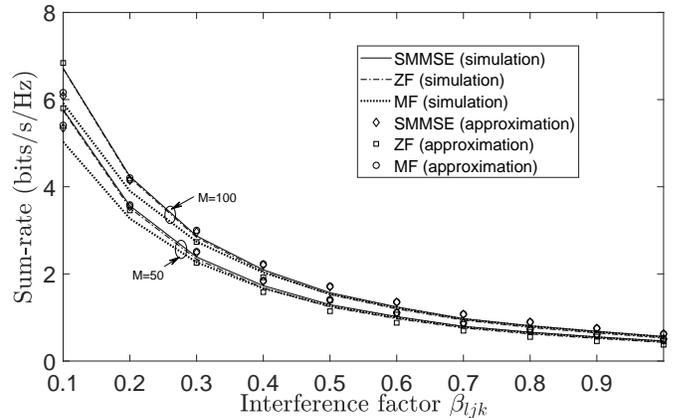


Fig. 7. Sum-rate w.r.t. interference factor for $M = 50$ and $M = 100$ under the exponential covariance matrix.

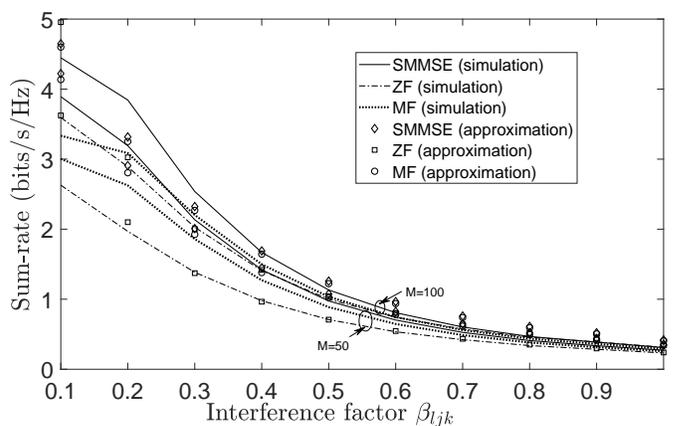


Fig. 8. Sum-rate w.r.t. interference factor for $M = 50$ and $M = 100$ under the Bessel covariance matrix.

that no matter for which decoder, decreasing the number of activated antennas always achieves higher energy efficiency. Hence, the number of activated antennas need to be selected considering the system sum-rate requirement as well as the power requirement.

VI. CONCLUSIONS

In this paper, we have studied the sum-rate of GSM under a practical channel model in a massive MIMO multi-user multi-cell system. At the GSM detection stage, we separated the TA index detection process and APM signal detection process into two parts and concluded that the SMMSE filter is the best choice for both of them. Furthermore, a sum-rate approximation was derived when the number of TAs tends to infinity. Simulations have demonstrated that SMMSE has the best performance, while MF has better performance than ZF only for a low number of TAs. Moreover, the accuracy of the sum-rate approximation for different linear filters was found to be sensitive to the channel correlation model. The analytical results of SMMSE and ZF exhibited better alignment with the simulations under the exponential covariance model, while the derivations of MF and ZF were better aligned with simulations under the Bessel covariance model. The relationships between

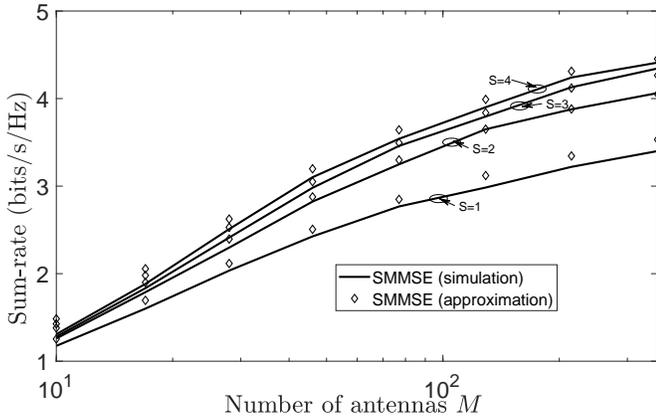


Fig. 9. Sum-rate w.r.t. the number of antennas **at each BS** under the exponential covariance matrix with SMMSE decoder.

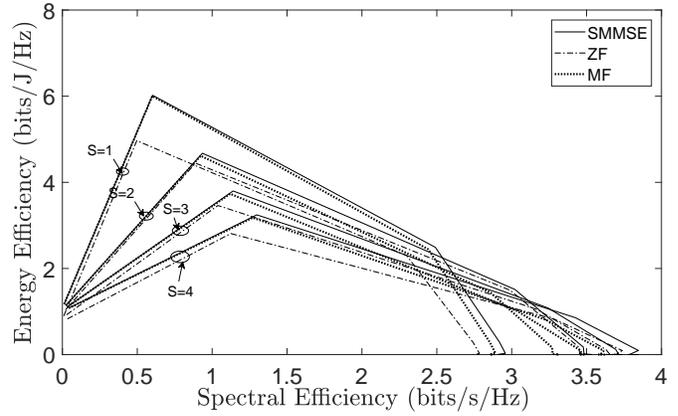


Fig. 11. Energy efficiency w.r.t. spectral efficiency under the exponential covariance matrix with $M = 100$.

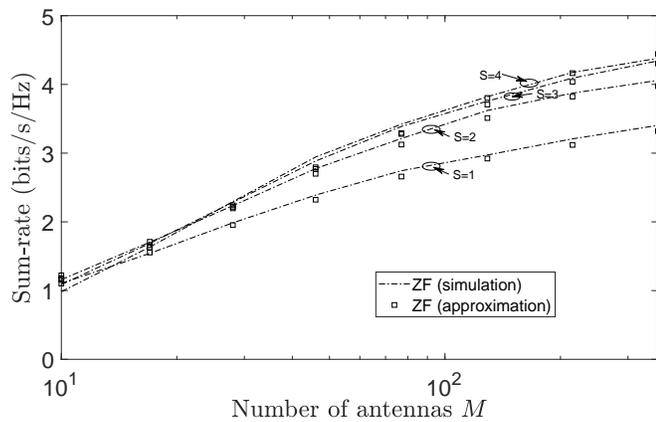


Fig. 10. Sum-rate w.r.t. the number of antennas **at each BS** under the exponential covariance matrix with ZF decoder.

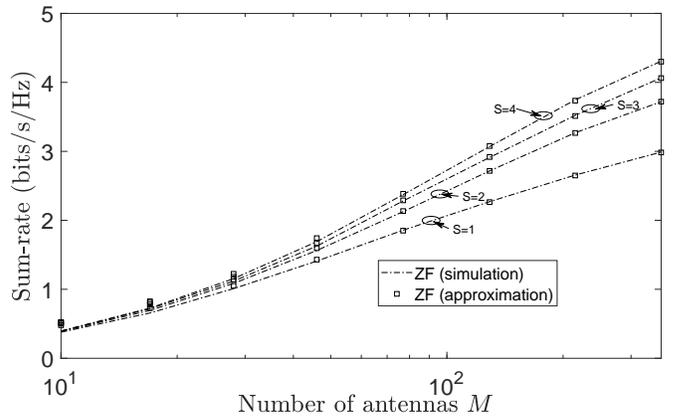


Fig. 12. Sum-rate w.r.t. the number of antennas **at each BS** under the Bessel covariance matrix with ZF decoder.

sum-rate, SNR and interference factor were also discussed. Furthermore, the sum-rate w.r.t. the number of BS antennas and the energy efficiency w.r.t. the spectral efficiency for different number of activated antennas are presented which provide a reference for selecting antennas. For future work, quantitative analysis of the sensitivity of the derivation to the spatial correlation model will be studied.

REFERENCES

- [1] F. Rusek, D. Persson, B.K.Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
- [2] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.
- [3] H. Q. Ngo, E. G. Larsson and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO Systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.
- [4] J. Hoydis, S. ten Brink and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
- [5] J. G. Andrews, S. Buzzi, W. Choi, S.V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, June 2014.
- [6] Z. Zhang, X. Wang, K. Long, A. V. Vasilakos and L. Hanzo, "Large-scale MIMO-based wireless backhaul in 5G networks," *IEEE Wireless Commun.*, vol. 22, no. 5, pp. 58–66, Oct. 2015.

- [7] P. Yang, Y. Xiao, Y. L. Guan, K.V.S. Hari, A. Chockalingam, S. Sugiura, H. Haas, M. Di Renzo, C. Masouros, Z. Liu, L. Xiao, S. Li and L. Hanzo, "Single-carrier SM-MIMO: A promising design for broadband large-scale antenna systems," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1687–1716, July 2016.
- [8] R. Mesleh, H. Haas, C. W. Ahn and S. Yun, "Spatial modulation - A new low complexity spectral efficiency enhancing technique," in *Proc. CHINACOM*, Beijing, China, Oct. 25–27, 2006, pp. 1–5.
- [9] P. Yang, M. Di Renzo, Y. Xiao, S. Li and L. Hanzo, "Design guidelines for spatial modulation," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 6–26, Mar. 2015.
- [10] R. Y. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn and S. Yun, "Spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2228–2241, July 2008.
- [11] J. Jeganathan, A. Ghrayeb and L. Szczecinski, "Spatial modulation: optimal detection and performance analysis," *IEEE Commun. Lett.* vol. 12, no. 8, pp. 545–547, Aug. 2008.
- [12] Y. Yang and B. Jiao, "Information-guided channel-hopping for high data rate wireless communication," *IEEE Commun. Lett.*, vol. 12, no. 4, pp. 225–227, April 2008.
- [13] S. Wu, P. Patcharamaneepakorn, C. X. Wang, E. H. Aggoune, M. M. Alwakeel and Y. He, "A novel method for ergodic sum rate analysis of spatial modulation systems with maximum likelihood receiver," in *Proc. IWCMC'15*, Dubrovnik, Croatia, Aug. 2015, pp. 32–36.
- [14] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura and L. Hanzo, "Spatial modulation for generalized MIMO: Challenges, opportunities, and implementation," *Proc. IEEE*, vol. 102, no. 1, pp. 56–103, Jan. 2014.
- [15] Y. Sun, J. Wang, L. He, and J. Song, "Uplink spectral efficiency analysis and optimization for massive SC-SM MIMO with frequency domain detection," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 3937–3949, May 2018.

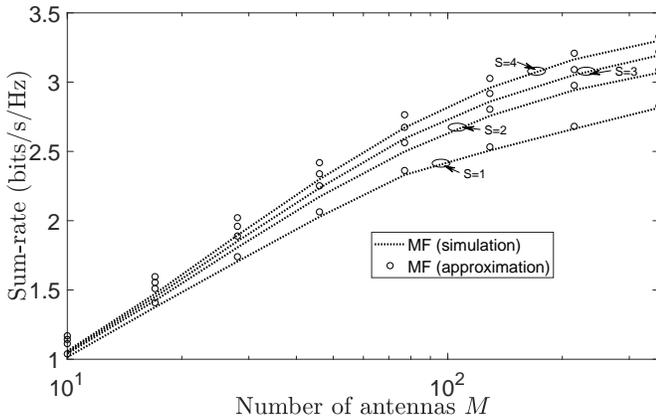


Fig. 13. Sum-rate w.r.t. the number of antennas at each BS under the Bessel covariance matrix with MF decoder.

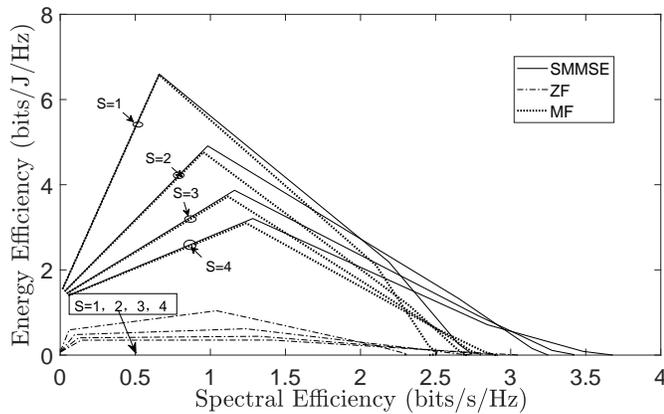


Fig. 14. Energy efficiency w.r.t. spectral efficiency under the Bessel covariance matrix with $M = 100$.

[16] L. He, J. Wang, J. Song and L. Hanzo, "Bandwidth efficiency maximization for single-cell massive spatial modulation MIMO: An adaptive power allocation perspective," *IEEE Access*, vol. 5, pp. 1482–1495, Feb. 2017.

[17] L. He, J. Wang, J. Song and L. Hanzo, "On the multi-user multi-cell massive spatial modulation uplink: How many antennas for each user?," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1437–1451, Mar. 2017.

[18] Y. Cui and X. Fang, "Performance analysis of massive spatial modulation MIMO in high-speed railway," *IEEE Trans. Veh. Technol.*, vol. 65, no. 11, pp. 8925–8932, Nov. 2016.

[19] L. He, J. Wang, and J. Song, "Multiuser detection for FEC-coded massive spatial modulation MIMO: an iterative interference rejection approach," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 9567–9571, Oct. 2017.

[20] J. Wang, S. Jia and J. Song, "Generalised spatial modulation system with multiple active transmit antennas and low complexity detection scheme," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1605–1615, Apr. 2012.

[21] J. Fu, C. Hou, W. Xiang, L. Yan and Y. Hou, "Generalised spatial modulation with multiple active transmit antennas," in *Proc. IEEE Globecom Workshops*, Miami, FL, USA, Dec. 2010, pp. 839–844.

[22] J. Jeganathan, A. Ghayeb and L. Szczecinski, "Generalized space shift keying modulation for MIMO channels," *IEEE International Symposium Personal, Indoor Mobile Radio Commun. (PIMRC'08)*, Cannes, France, Sep. 2008, pp. 1–5.

[23] P. Patcharamaneepakorn, S. Wu, C. X. Wang, H. Aggoune, M. Alwakeel, X. Ge, and M. D. Renzo, "Spectral, energy, and economic efficiency of 5G multicell massive MIMO systems with generalized spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 9715–9731, Dec. 2016.

[24] E. Björnson, J. Hoydis and L. Sanguinetti, "Pilot contamination is not a fundamental asymptotic limitation in massive MIMO," in *Proc. IEEE International Conference on Communications (ICC)*, Paris, May. 2017,

pp. 1–6.

[25] T. M. Cover and J. A. Thomas. Elements of information theory. John Wiley & Sons, 2012.

[26] R. J. Vaughan and W. N. Venables, "Permanent expressions for order statistic densities," *J. R. Statist. Soc.*, vol. 34, no. 2, pp. 308–310, 1972.



Xiaoqing Zhang received the B. Sc. degree in communication engineering from Shandong University, Jinan, China, in 2014. She is currently pursuing the Ph. D degree with the School of Information Science and Engineering, Shandong University, Qingdao, China. From October 2017 to April 2019, funded by the China Scholarship Council, she was a visiting Ph. D student supervised by Prof. L. Hanzo in the School of Electronics and Computer Science, University of Southampton, UK. Her research interests include massive MIMO, precoding, pilot contamination, spatial modulation and generalized spatial modulation.



Shangbin Wu received his B.S. degree in communication engineering from South China Normal University, Guangzhou, China, in 2009, M.Sc. degree in wireless communications with distinction from University of Southampton, Southampton, UK, in 2010, and Ph.D. degree in electrical engineering from the Heriot-Watt University, Edinburgh, UK in 2015. From 2010 to 2011, he worked as a LTE R&D engineer responsible for LTE standardization and system level simulation in New Postcom Equipment Ltd., Guangzhou, China. From October 2011 to August 2012, he was with Nokia Siemens Network, where he worked as a LTE algorithm specialist, mainly focusing on LTE radio resource management algorithm design and system level simulations. He has been with Samsung R&D Institute UK as a 5G researcher since November 2015.



Shanshan Yu received the B.Sc. and M.S. degree from Ocean University of China in 2002 and 2006, respectively. He has been a lecturer of Qufu Normal University from 2006 to 2014. He is currently pursuing the Ph.D. degree with the School of Information Science and Engineering at Shandong University. His research interests include device-to-device communications, physical-layer security and cooperative communications.



Ju Liu received the B.S. and M.S. degrees both in electronic engineering from Shandong University, Jinan, China, in 1986 and 1989, respectively; and the Ph.D. degree in signal processing from Southeast University, Nanjing, China, in 2000. Since July 1989, he has been with the Department of Electronic Engineering, Shandong University, and now he is a Full Professor.

From July, 2002 to December, 2003, he was a Visiting Professor in the Department of Signal Theory and Communication, Polytechnic University of Catalonia, and Telecommunications Technological Centre of Catalonia, Barcelona, Spain. From November 2005 to January 2006, funded by DAAD, he was a Visiting Researcher in the Department of Communication Engineering, University of Bremen, and the Department of Communication Systems, University of Duisburg-Essen, in Germany. From June to December 2009, from July to August 2012, and in January 2019, funded by CSC, he was a Senior Research Fellow in the Department of Electrical Engineering, University of Washington, Seattle, WA, USA, in the Department of Electrical and Computer Engineering, University of California, San Diego, CA, USA, and in the School of Electronics and Computer Science, University of Southampton, UK, respectively.



Lajos Hanzo (<http://www-mobile.ecs.soton.ac.uk>) FREng, F'04, FIET, Fellow of EURASIP, received his 5-year degree in electronics in 1976 and his doctorate in 1983 from the Technical University of Budapest. In 2009 he was awarded an honorary doctorate by the Technical University of Budapest and in 2015 by the University of Edinburgh. In 2016 he was admitted to the Hungarian Academy of Science. During his 40-year career in telecommunications he has held various research and academic posts in Hungary, Germany and the UK. Since 1986 he has

been with the School of Electronics and Computer Science, University of Southampton, UK, where he holds the chair in telecommunications. He has successfully supervised 119 PhD students, co-authored 18 John Wiley/IEEE Press books on mobile radio communications totalling in excess of 10 000 pages, published 1800+ research contributions at IEEE Xplore, acted both as TPC and General Chair of IEEE conferences, presented keynote lectures and has been awarded a number of distinctions. Currently he is directing a 60-strong academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council (EPSRC) UK, the European Research Council's Advanced Fellow Grant and the Royal Society's Wolfson Research Merit Award. He is an enthusiastic supporter of industrial and academic liaison and he offers a range of industrial courses. He is also a Governor of the IEEE ComSoc and VTS. He is a former Editor-in-Chief of the IEEE Press and a former Chaired Professor also at Tsinghua University, Beijing. For further information on research in progress and associated publications please refer to <http://www-mobile.ecs.soton.ac.uk>