

Spectral, Energy, and Economic Efficiency of 5G Multicell Massive MIMO Systems With Generalized Spatial Modulation

Piya Patcharamaneepakorn, *Member, IEEE*, Shangbin Wu, Cheng-Xiang Wang, *Senior Member, IEEE*, el-Hadi M. Aggoune, *Senior Member, IEEE*, Mohammed M. Alwakeel, *Senior Member, IEEE*, Xiaohu Ge, *Senior Member, IEEE*, and Marco Di Renzo, *Senior Member, IEEE*

Abstract—This paper studies generalized spatial modulation (Gen-SM) schemes in multicell multiuser massive multiple-input multiple-output (MIMO) systems as a promising high-throughput and energy-efficient technique for fifth-generation (5G) wireless networks. A detection algorithm for such systems is proposed based on linear processing techniques. By applying the concept of order statistics, a general framework for approximating the achievable sum rates with linear detection is also given. The probability of detecting antenna combinations is analyzed and is used to approximate the sum-rate performance with practical channel conditions, such as antenna correlation, imperfect channel information, and pilot contamination. The fundamental trade-off between spectral efficiency (SE) and energy efficiency (EE) is also investigated. Despite offering less SE, spatial modulation (SM) with a single active antenna per user is shown to be the most energy-efficient transmission mode among the Gen-SM class. Within the operating range of SM (i.e., in the

low-to-moderate SE regime), SM is demonstrated to achieve better EE compared with conventional massive MIMO schemes. The performance in terms of economic efficiency, indicating economic profitability (in monetary unit per second), is also analyzed and is shown to serve as a complementary performance metric, enabling an implicit trade-off between SE and EE.

Index Terms—Economic efficiency (ECE), energy efficiency (EE), generalized spatial modulation (Gen-SM), massive multiple-input multiple-output (MIMO), spectral efficiency (SE).

I. INTRODUCTION

THE fifth-generation (5G) cellular network is expected to offer 1000 times the system capacity; 10 times the spectral efficiency (SE), energy efficiency (EE), and data rate; and 25 times the average cell throughput, compared with the fourth-generation network [1], [2]. This has attracted considerable interest in researching revolutionary wireless technologies to meet these challenging targets [3].

Massive multiple-input multiple-output (MIMO) technology is one of the key potential techniques to dramatically improve system capacity and reliability [4]–[6]. By deploying a large number of antennas at the base stations (BSs), e.g., hundreds of antenna elements compared with tens of single-antenna user terminals, surplus degrees of freedom can be achieved, leading to increased spatial multiplexing and beamforming gains [4], [5]. At an asymptotically large number of BS antennas, channel vectors from the BS to different users were shown to be nearly orthogonal, and the impacts of small-scale fading, thermal noise, and uncorrelated interference diminish [5], [6]. Nevertheless, intercell interference caused by *pilot contamination*, i.e., due to the reuse of pilot sequences in multicell environments, does not vanish and becomes the limiting factor to the system performance [6]. It was also shown in [4]–[6] that simple linear processing methods, such as matched filter (MF) precoding and detection, can achieve near-optimal performance in the regime with a large number of antennas. The performance of massive MIMO with linear processing has been extensively investigated in the literature. In particular, the uplink performance of massive MIMO systems with Rayleigh fading channels was studied in [7] and [8], whereas that with Ricean fading channels was analyzed in [9]. The impact of pilot contamination was also investigated in [10]–[12]. Furthermore, practical channel assumptions, such as finite-dimensional channels and antenna correlations, were considered in [12] and [13]. The investigations on downlink performance were also given in [13]–[15].

Manuscript received July 8, 2015; revised November 19, 2015; accepted January 4, 2016. Date of publication February 8, 2016; date of current version December 14, 2016. This work was supported in part by the EU H2020 ITN 5G Wireless Project (No. 641985); by the EU FP7 QUICK Project (No. PIRSES-GA-2013-612652); by the Sensor Networks and Cellular Systems Research Center, University of Tabuk, Saudi Arabia; by the Ministry of Science and Technology of China through the 863 Project in 5G Wireless Networking (No. 2014AA012101); by the National Natural Science Foundation of China through the Major International Joint Research Project (No. 61210002), the Fundamental Research Funds for the Central Universities (No. 2015XJGH011), the EU FP7-PEOPLE-IRSES WINDOW Project (No. 318992) and the CROWN Project (No. 610524); by the National International Scientific and Technological Cooperation Base of Green Communications and Networks (No. 2015B01008); and by the Hubei International Scientific and Technological Cooperation Base of Green Broadband Wireless Communications. The review of this paper was coordinated by Dr. C. Xing.

P. Patcharamaneepakorn was with the Institute of Sensors, Signals, and Systems, School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, U.K. He is now with True Corporation, Bangkok 10310, Thailand (e-mail: piya_pat@truecorp.co.th).

S. Wu and C.-X. Wang are with the Institute of Sensors, Signals, and Systems, School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, U.K., and also with the Sensor Networks and Cellular Systems (SNCS) Research Center, University of Tabuk, Tabuk 71491, Saudi Arabia (e-mail: sw271@hw.ac.uk; cheng-xiang.wang@hw.ac.uk).

e.-H. M. Aggoune and M. M. Alwakeel are with the Sensor Networks and Cellular Systems (SNCS) Research Center, University of Tabuk, Tabuk 71491, Saudi Arabia (e-mail: haggoune.snscs@ut.edu.sa; alwakeel@ut.edu.sa).

X. Ge is with the Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: xhge@mail.hust.edu.cn).

M. Di Renzo is with the Laboratoire des Signaux et Systèmes, CNRS, CentraleSupélec, Univ Paris Sud, Université Paris-Saclay, 91192 Gif-sur-Yvette, France (e-mail: marco.direnzo@12s.centralesupelec.fr).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TVT.2016.2526628

Spatial modulation (SM) is another promising candidate for 5G radio access technology [1]. In SM, only a single transmit antenna is allowed to be active at each time instant, and user data are conveyed by two information-carrying units, namely, the antenna index and the symbol modulation [16]–[18]. A special form of SM, which is known as space-shift keying (SSK), is referred to the case when the user information is only carried on the antenna indexes. Compared with the conventional MIMO systems, SM and SSK schemes offer several advantages, including mitigation of interchannel interference, relaxation of interantenna synchronization requirements, reduction of receiver complexity, and the use of a single radio-frequency (RF) chain at the transmitter [19], [20]. The concept of SM is extended to allow multiple active transmit antennas, leading to generalized spatial modulation (Gen-SM) schemes [22]–[24]. The performance of SM has been intensively studied in the literature. Most of the existing works, however, focus on the *single-user* (point-to-point) MIMO context. For instance, the bit error rate (BER) performance of SM was analyzed in [16]–[21], and the achievable sum rate of SM was discussed in [25] and [26]. In contrast, limited works have considered SM schemes in the *multiuser* context, most of which only focused on the BER performance, such as in [27]–[29].

Despite extensive research on individual technologies, the combination of massive MIMO and SM has been less investigated in the literature. In [30], the performance of SM and Gen-SM schemes was analyzed in single-user massive MIMO scenarios. A study of SM in multiuser massive MIMO systems was given in [31], whereby SM was used to enhance uplink transmission, with multiple antennas (with a single RF chain) being deployed at each user and a massive-antenna array being considered at the BS. By applying SM in uplink, the saving of user transmit power was shown to be achieved [31]. This potentially leads to significant improvement in EE and enhanced battery life at mobile terminals, which is one of the main targets specified in 5G requirements. Nevertheless, the work in [31] only studied SM in the *single-cell* environment, and the performance was evaluated only in terms of BER. To the best of the authors' knowledge, the evaluation of the achievable sum rate of Gen-SM in *multicell multiuser* massive MIMO systems has yet to be considered in the literature. Motivated by the benefits of SM in massive MIMO systems, this paper extends the idea in [31] and evaluates the potential advantages of Gen-SM in *multicell multiuser* massive MIMO scenarios. Linear receive processing is considered for low computational complexity, and a comprehensive performance evaluation using various performance metrics, including SE, EE, and economic efficiency (ECE), is also given. The contributions of this paper can be summarized as follows.

- 1) A detection algorithm based on linear processing is proposed for *multicell multiuser* massive MIMO systems with Gen-SM. A general framework for evaluating the sum-rate performance, considering practical scenarios, such as antenna correlation, imperfect channel state information (CSI), and pilot contamination, is also presented.
- 2) Following the proposed framework and applying the concept of order statistics, the probability of antenna-

combination detection, based on two common linear processing, namely, MF and zero-forcing (ZF) detection, is analyzed and is adopted in the approximation of the achievable sum rate. The approximation is shown to be in good agreement with simulation results.

- 3) Trade-off between SE and EE is analyzed for the first time for *multicell multiuser* massive MIMO systems with Gen-SM. It is shown that, although offering less SE, SM (with one active antenna per user) tends to be the most efficient mode among the entire class of the Gen-SM scheme. Compared with conventional massive MIMO schemes (deploying spatial multiplexing with a single antenna per user), the feasible SE of SM is limited to a smaller range, i.e., in the low-to-moderate SE values, since SM generally requires a longer training period (equivalently, shorter period for data transmission). Nevertheless, within its feasible SE regime, SM is shown to potentially achieve better EE than conventional massive MIMO. Furthermore, ECE is evaluated as a complementary performance measure to SE and EE, offering an inherent trade-off between SE and EE.

Notation: $\text{Tr}(\cdot)$, $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^*$, and $\text{vec}(\cdot)$ denote the trace, transpose, Hermitian, complex conjugate, and vectorization operations, respectively. $\binom{\cdot}{\cdot}$ denotes the binomial coefficient, and $a \rightarrow b$ denotes that a asymptotically converges to b . $|\cdot|$ represents the complex modulus or the cardinality of a set, and $\|\cdot\|$ denotes the Frobenius norm. The probability of an event is expressed as $P(\cdot)$, and the expectation operation is denoted as $E(\cdot)$. Given two sets S_1 and S_2 with $S_2 \subset S_1$, $S_1 \setminus S_2$ represents a set containing all elements of S_1 that are not in S_2 . The (i, j) th element and the j th column vector of a matrix \mathbf{A} are denoted as $[\mathbf{A}]_{i,j}$ and $[\mathbf{A}]_{*,j}$, respectively. Furthermore, $[\mathbf{A}]_{k,k}$ and $[\mathbf{A}]_{k,k}^{(N \times N)}$ denote the k th diagonal entry and the k th diagonal block (of size $N \times N$) of matrix \mathbf{A} . Additionally, \otimes and \circ denote the Kronecker and Hadamard (element-wise) products, respectively, and $\Re\{\cdot\}$ represents the real part of a complex number.

II. SYSTEM MODEL

Consider an uplink multicell MIMO system with L BSs, where each is equipped with M receive antennas and serves K users, each with N transmit antennas. A user u in cell j is denoted as (j, u) . A similar channel model as given in [12] is considered and can be summarized as follows. Let $\mathbf{G}_{lju} = \sqrt{\beta_{lju}} \mathbf{H}_{lju} \in \mathbb{C}^{M \times N}$ be the channel matrix from a user (j, u) to BS l , with $\sqrt{\beta_{lju}}$ representing large-scale fading and \mathbf{H}_{lju} denoting small-scale fading as given by

$$\mathbf{H}_{lju} = \mathbf{R}_R^{\frac{1}{2}} \mathbf{H}_{lju}^w \mathbf{R}_T^{\frac{T}{2}} \quad (1)$$

where \mathbf{R}_T and \mathbf{R}_R are transmit and receive spatial correlation matrices, respectively, with $\text{Tr}(\mathbf{R}_T) = N$ and $\text{Tr}(\mathbf{R}_R) = M$. Each element of \mathbf{H}_{lju}^w is an independent and identically distributed (i.i.d.) Gaussian random variable (RV) with zero mean and unit variance. It follows that $\mathbf{H}_{lju} \sim \mathcal{CN}(\mathbf{0}_{M \times N}, \mathbf{R}_H)$,

with $\mathbf{R}_H = \mathbf{R}_T \otimes \mathbf{R}_R$. The channel matrix from all users in cell j to BS l can be given by $\mathbf{G}_{lj} = [\mathbf{G}_{lj1}, \dots, \mathbf{G}_{ljK}]$, and the overall channel matrix to BS l is denoted as $\mathbf{G}_l = [\mathbf{G}_{l1}, \dots, \mathbf{G}_{lL}]$. For Gen-SM systems, the information can be conveyed via active antenna indexes, in addition to the conventional signal modulation. For each user, independent multiple data streams are simultaneously transmitted over $1 \leq A \leq N$ active antennas, i.e., the data vector of user (j, u) , which is denoted as $\mathbf{x}_{ju} = [x_{ju1}, \dots, x_{juN}]^T \in \mathbb{C}^{N \times 1}$, contains A nonzero entries. Each nonzero entry is assumed to be an i.i.d. Gaussian RV with zero mean and unit variance. Let $\bar{\mathbf{x}}_{ju} \in \mathbb{C}^{A \times 1}$ be a subvector of \mathbf{x}_{ju} corresponding to nonzero data entries and $\bar{\mathbf{G}}_{lju} \in \mathbb{C}^{M \times A}$ denote a submatrix of \mathbf{G}_{lju} with each column associated to user u 's active antennas. The overall data vector of users in cell j can be given by $\bar{\mathbf{x}}_j = [\bar{\mathbf{x}}_{j1}^H, \dots, \bar{\mathbf{x}}_{jK}^H]^H$. The composite active channel can be written as $\bar{\mathbf{G}}_{lj} = [\bar{\mathbf{G}}_{lj1}, \dots, \bar{\mathbf{G}}_{ljK}] \in \mathbb{C}^{M \times AK}$ and $\bar{\mathbf{G}}_l = [\bar{\mathbf{G}}_{l1}, \dots, \bar{\mathbf{G}}_{lL}] \in \mathbb{C}^{M \times AKL}$. The received signal vector at BS l can be expressed as

$$\mathbf{y}_l = \sqrt{\rho_{ul}} \sum_{j=1}^L \bar{\mathbf{G}}_{lj} \bar{\mathbf{x}}_j + \mathbf{n}_l \quad (2)$$

where $\mathbf{n}_l \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_M)$, and $\rho_{ul} = \rho/A$, with ρ being the transmitted power per user.

The following notations are further used to facilitate subsequent discussions. With $1 \leq A \leq N$ active antennas, the number of possible antenna combinations for each user is given by $C = \binom{N}{A}$. The set of all possible antenna-index combinations is denoted as $\mathcal{A} = \{\mathcal{A}_1, \dots, \mathcal{A}_C\}$, whereby \mathcal{A}_i is a set of active-antenna indexes for antenna combination i . The transmitted and detected antenna combinations for user (l, k) are represented as $A_{lk}, \hat{A}_{lk} \in \mathcal{A}$, respectively.

A. Channel Estimation

Uplink training (pilot) sequences are considered for channel estimation. It is assumed that the same set of orthogonal pilot sequences is used in all L cells. Following [7], the received pilot matrix after removing the transmitted pilot symbols for user k at BS l can be written as

$$\tilde{\mathbf{Y}}_{lk} = \sqrt{\rho_t} \sum_{j=1}^L \mathbf{G}_{lj} + \tilde{\mathbf{N}}_{lk} \quad (3)$$

where $\rho_t = \rho\tau/N$ denotes the transmitted pilot power per user, $NK \leq \tau \leq T$ represents the duration of the pilot sequences with T being the coherence time, and $\tilde{\mathbf{N}}_{lk} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_{MN})$. Note that $\tilde{\mathbf{Y}}_{lk}$ contains the channel matrices of other-cell users using the same pilot sequence as user (l, k) , causing *pilot contamination* to the channel estimation of user (l, k) . It follows from [32] that the minimum mean square error (MMSE) channel estimate can be obtained as

$$\text{vec}(\hat{\mathbf{G}}_{lk}) = \mathbf{A} \left(\frac{1}{\sqrt{\rho_t}} \text{vec}(\tilde{\mathbf{Y}}_{lk}) \right) \quad (4)$$

with

$$\mathbf{A} = \beta_{lk} \mathbf{R}_H \left(\sum_{j=1}^L \beta_{lj} \mathbf{R}_H + \frac{\sigma^2}{\rho_t} \mathbf{I}_{MN} \right)^{-1}. \quad (5)$$

Using the orthogonal property of MMSE channel estimates, it is shown in [32] that the actual channel matrix can be decomposed as $\mathbf{G}_{lk} = \hat{\mathbf{G}}_{lk} + \tilde{\mathbf{G}}_{lk}$, with $\hat{\mathbf{G}}_{lk}$ and $\tilde{\mathbf{G}}_{lk}$ being statistically independent. Furthermore, to ease subsequent discussions, let $\mathbf{g}_{llks} = [\mathbf{G}_{llk}]_{*,s}$, $\hat{\mathbf{g}}_{llks} = [\hat{\mathbf{G}}_{llk}]_{*,s}$, and $\tilde{\mathbf{g}}_{llks} = [\tilde{\mathbf{G}}_{llk}]_{*,s}$ for the rest of this paper.

B. Linear Filters

Two linear filters, namely, MF and ZF, are considered for the detection of Gen-SM signals in this paper. By constructing the matrix of channel estimates as the receiver matrix, the MF performs maximum ratio combining across multiple receive antennas, aiming to maximize the received signal energy for each detected data stream. This, however, may lead to high interference between multiple data streams. The ZF, in contrast, applies the matrix of channel inversion (the pseudoinverse of the channel estimates) as the receiver matrix to ensure zero interference among different data streams. Consequently, considering the case of imperfect CSI, using the channel estimate $\hat{\mathbf{G}}_{ll} = [\hat{\mathbf{G}}_{ll1}, \dots, \hat{\mathbf{G}}_{llK}]$ with $\hat{\mathbf{G}}_{llk}$ being obtained from (4), the MF and ZF receivers can, respectively, be written as

$$\hat{\mathbf{B}}_l^{\text{MF}} = \hat{\mathbf{G}}_{ll} \quad (6)$$

$$\hat{\mathbf{B}}_l^{\text{ZF}} = \hat{\mathbf{G}}_{ll} \left(\hat{\mathbf{G}}_{ll}^H \hat{\mathbf{G}}_{ll} \right)^{-1}. \quad (7)$$

Linear filters for the case of perfect CSI, which are denoted as \mathbf{B}_l , can also be given in a similar manner, with $\hat{\mathbf{G}}_{ll}$ being replaced by \mathbf{G}_{ll} . In the following discussion, both analytical results and proofs will be given for the case of imperfect CSI. For the case of perfect CSI, only the results will be provided since the proofs can be obtained by applying similar analysis (e.g., as a special case of imperfect CSI) and can therefore be omitted.

II. DETECTION ALGORITHM WITH LINEAR FILTERS AND ACHIEVABLE RATE ANALYSIS

A. Detection Algorithm With Linear Filters

For signal detection in *multicell multiuser* massive MIMO systems with Gen-SM, the optimal maximum-likelihood (ML) detection tends to be impractical due to its high complexity. Motivated by the work in [16] and [24], to reduce complexity, a suboptimal detection algorithm based on linear processing is proposed in this section and can be divided into two stages, i.e., the detection of the active antennas and the detection of signal constellation from the active antennas previously obtained in the first stage. The proposed algorithm can be considered as an extension of the algorithms in [16] and [24] from *single-user*

single-cell SM systems to multicell multiuser Gen-SM scenarios. The proposed procedure can be summarized as follows.

- 1) A linear filter, e.g., either MF or ZF, is applied to the received signal vector. The output vector for a user k in BS l can be written as

$$\hat{\mathbf{y}}_{lk} = \hat{\mathbf{B}}_{lk}^H \mathbf{y}_l \quad (8)$$

with $\mathbf{B}_{lk} \in \mathbb{C}^{M \times N}$ being a linear filter of user k in BS l . Letting $\hat{\mathbf{y}}_{lk} = [\hat{y}_{lk1}, \dots, \hat{y}_{lkN}]$, \hat{A}_{lk} can be obtained by selecting a valid antenna combination i that provides the maximum received power, e.g.,

$$\hat{A}_{lk} = \arg \max_{\mathcal{A}_i} \sum_{\forall d \in \mathcal{A}_i} |\hat{y}_{lkd}|^2. \quad (9)$$

- 2) Once the antenna combination \hat{A}_{lk} is decoded, the corresponding received signal vector of the selected antennas can be expressed as

$$\hat{\mathbf{r}}_{lk} = \text{vec} \left(\{\hat{\mathbf{y}}_{lkd}\}_{d \in \hat{A}_{lk}} \right). \quad (10)$$

It follows that the detected received symbols can be obtained by

$$\hat{\mathbf{x}}_{lk} = Q(\hat{\mathbf{r}}_{lk}) \quad (11)$$

with $Q(\cdot)$ being the constellation quantization (slicing) function [16].

B. Achievable Rate Analysis

The two-stage detection algorithm, as described in Section III-A, can be viewed as a mapping function from the postprocessed received signal vector ($\hat{\mathbf{y}}_{lk} = \hat{\mathbf{B}}_{lk}^H \mathbf{y}_l$) to the detected antenna combination (\hat{A}_{lk}^{IP}) and the corresponding output vector ($\hat{\mathbf{r}}_{lk}^{\text{IP}}$), i.e., $\{\hat{A}_{lk}^{\text{IP}}, \hat{\mathbf{r}}_{lk}^{\text{IP}}\} = f_{\text{mapping}}(\hat{\mathbf{B}}_{lk}^H \mathbf{y}_l)$. The superscript $(\cdot)^{\text{IP}}$ indicates that the linear filter with imperfect CSI knowledge, i.e., $\hat{\mathbf{B}}_{lk}$, is applied to \mathbf{y}_l . In the case of perfect CSI, the superscript $(\cdot)^{\text{P}}$ is adopted.

For imperfect CSI, the average achievable uplink rate of a user (l, k) can be obtained through the amount of mutual information as given in the following:

$$\begin{aligned} & \left(1 - \frac{\tau}{T}\right) E \left\{ I \left(\bar{\mathbf{x}}_{lk}, \bar{\mathbf{G}}_{llk}; \hat{A}_{lk}^{\text{IP}}, \hat{\mathbf{r}}_{lk}^{\text{IP}} \right) \right\} = \left(1 - \frac{\tau}{T}\right) \\ & \times \left(E \left\{ I \left(\bar{\mathbf{x}}_{lk}; \hat{A}_{lk}^{\text{IP}}, \hat{\mathbf{r}}_{lk}^{\text{IP}} | \bar{\mathbf{G}}_{llk} \right) \right\} + E \left\{ I \left(\bar{\mathbf{G}}_{llk}; \hat{A}_{lk}^{\text{IP}}, \hat{\mathbf{r}}_{lk}^{\text{IP}} \right) \right\} \right) \end{aligned} \quad (12)$$

where the coefficient $(1 - (\tau/T))$ reflects the percentage of time being used for data transmission.

Proposition 1: The average achievable rate of a user (l, k) in (12) can be lower bounded by

$$\begin{aligned} & \left(1 - \frac{\tau}{T}\right) E \left\{ I \left(\bar{\mathbf{x}}_{lk}, \bar{\mathbf{G}}_{llk}; \hat{A}_{lk}^{\text{IP}}, \hat{\mathbf{r}}_{lk}^{\text{IP}} \right) \right\} \\ & \geq \left(1 - \frac{\tau}{T}\right) R_{lk}^{\text{IP}} \triangleq \left(1 - \frac{\tau}{T}\right) (R_{lk,1}^{\text{IP}} + R_{lk,2}^{\text{IP}}) \end{aligned} \quad (13)$$

where

$$R_{lk,1}^{\text{IP}} = E \left\{ \frac{1}{C} \sum_{c=1}^C \sum_{i=1}^C \varrho_{lk(i|c)}^{\text{IP}} \left[\sum_{d \in \mathcal{A}_i} \log_2 \left(1 + \frac{\mu_{lkd|c}^{\text{IP}}}{\nu_{lkd|c}^{\text{IP}}} \right) \right] \right\} \quad (14)$$

$$\begin{aligned} R_{lk,2}^{\text{IP}} &= \log_2(C) \\ &- E \left\{ \frac{1}{C} \sum_{c=1}^C \sum_{i=1}^C \varrho_{lk(i|c)}^{\text{IP}} \left[\log_2 \frac{\sum_{c_2=1}^C \varrho_{lk(i|c_2)}^{\text{IP}}}{\varrho_{lk(i|c)}^{\text{IP}}} \right] \right\} \end{aligned} \quad (15)$$

$$\mu_{lkd|c}^{\text{IP}} = \rho_{ul} |\hat{\mathbf{b}}_{lkd}^H \hat{\mathbf{g}}_{llka}|^2 \quad (16)$$

$$\begin{aligned} \nu_{lkd|c}^{\text{IP}} &= \sum_{s \in \mathcal{A}_c \setminus \{a\}} \rho_{ul} |\hat{\mathbf{b}}_{lkd}^H \hat{\mathbf{g}}_{llks}|^2 \\ &+ \sum_{(j,u) \neq (l,k)} \sum_{s \in \mathcal{A}_{ju}} \rho_{ul} |\hat{\mathbf{b}}_{lkd}^H \hat{\mathbf{g}}_{ljus}|^2 \\ &+ \sum_{j=1}^L \sum_{u=1}^K \sum_{s \in \mathcal{A}_{ju}} \rho_{ul} |\hat{\mathbf{b}}_{lkd}^H \tilde{\mathbf{g}}_{ljus}|^2 + \|\hat{\mathbf{b}}_{lkd}\|^2 \sigma^2 \end{aligned} \quad (17)$$

with $\hat{\mathbf{b}}_{lkd} = [\hat{\mathbf{B}}_{lk}]_{*,d}$, $\hat{\mathbf{g}}_{ljus} = [\hat{\mathbf{G}}_{lju}]_{*,d}$ for any user (j, u) , and $\varrho_{lk(i|c)}^{\text{IP}} = P(\hat{A}_{lk}^{\text{IP}} = \mathcal{A}_i | A_{lk} = \mathcal{A}_c)$ denoting the probability of detecting antenna combination i when antenna combination c is actually transmitted, and $d \in \mathcal{A}_i$, $a \in \mathcal{A}_c$ representing a detected antenna and the corresponding transmitted antenna, respectively.

Proof: See Appendix A. \blacksquare

Corollary 1: Following the analysis in Proposition 1, a lower bound on the average achievable rate of a user (l, k) in the case of perfect CSI can be obtained as

$$E \left\{ I \left(\bar{\mathbf{x}}_{lk}, \bar{\mathbf{G}}_{llk}; \hat{A}_{lk}^{\text{P}}, \hat{\mathbf{r}}_{lk}^{\text{P}} \right) \right\} \geq R_{lk}^{\text{P}} \triangleq R_{lk,1}^{\text{P}} + R_{lk,2}^{\text{P}} \quad (18)$$

whereby $R_{lk,1}^{\text{P}}$ and $R_{lk,2}^{\text{P}}$ can be defined in a similar manner as (14) and (15), respectively, with

$$\mu_{lkd|c}^{\text{P}} = \rho_{ul} |\mathbf{b}_{lkd}^H \mathbf{g}_{llka}|^2 \quad (19)$$

$$\begin{aligned} \nu_{lkd|c}^{\text{P}} &= \sum_{s \in \mathcal{A}_c \setminus \{a\}} \rho_{ul} |\mathbf{b}_{lkd}^H \mathbf{g}_{llks}|^2 \\ &+ \sum_{(j,u) \neq (l,k)} \sum_{s \in \mathcal{A}_{ju}} \rho_{ul} |\mathbf{b}_{lkd}^H \mathbf{g}_{ljus}|^2 + \sigma^2 \|\mathbf{b}_{lkd}\|^2 \end{aligned} \quad (20)$$

and $\mathbf{b}_{lkd} = [\mathbf{B}_{lk}]_{*,d}$, $\mathbf{g}_{ljus} = [\mathbf{G}_{lju}]_{*,d}$ for any user (j, u) , and $\varrho_{lk(i|c)}^{\text{P}} = P(\hat{A}_{lk}^{\text{P}} = \mathcal{A}_i | A_{lk} = \mathcal{A}_c)$. Note that the coefficient $(1 - (\tau/T))$ is absent from (18) as the training phase is not considered for the case of perfect CSI.

C. Spatial Multiplexing as a Special Case of Gen-SM

From (13), the achievable rate of Gen-SM schemes consists of two contributing factors, i.e., the information rate due to symbol modulation ($R_{lk,1}^{\text{IP}}$) and the information rate due to antenna modulation ($R_{lk,2}^{\text{IP}}$). It can be interpreted from (14) that $R_{lk,1}^{\text{IP}}$ is the average of symbol modulation information rates over all possible antenna combination events. From (15) and (63), $R_{lk,2}$ can also be viewed as the information rate for transmitting codewords (antenna indexes) across a C -ary

asymmetric channel. For a special case of $A = N$ ($C = 1$), it can be seen that $R_{lk,2}^{\text{IP}}$ becomes 0 and $R_{lk,1}^{\text{IP}}$ reduces its form to the information rate of spatial multiplexing schemes. Hence, spatial multiplexing can be considered as a special case of Gen-SM to which the analysis in Proposition 1 and the subsequent corollaries remain applicable.

IV. ASYMPTOTIC PERFORMANCE APPROXIMATION FOR LARGE ANTENNA SYSTEMS

Considering the case of asymptotically large numbers of BS antennas, the ergodic achievable rate in (13) is approximated in this section. A general framework for approximating the achievable rate is first described and is then applied to practical linear receivers, i.e., MF and ZF detectors. Note that the analysis in this section involves basic knowledge in order statistics, which is summarized in Appendix E for a quick reference. More details are available in [33].

A. General Framework

Here, the average probability of detecting antenna combination i when transmitting antenna combination c will be evaluated for all pair (i, c) and will be used for estimation of SE in (13). Consider the case of imperfect CSI. Assuming that, at the receiver, a transmitted antenna $d \in \mathcal{A}_i$ is detected for user (l, k) with the antenna $a \in \mathcal{A}_c$ being actually transmitted, the corresponding received symbol can be written as

$$\begin{aligned} \hat{y}_{lkd}^{\text{IP}} &= \sqrt{\rho_{ul}} \hat{\mathbf{b}}_{lkd}^H \hat{\mathbf{g}}_{llka} x_{lka} + \sum_{s \in \mathcal{A}_c \setminus \{a\}} \sqrt{\rho_{ul}} \hat{\mathbf{b}}_{lkd}^H \hat{\mathbf{g}}_{llks} x_{lks} \\ &+ \sum_{(j,u) \neq (l,k)} \sum_{s \in \mathcal{A}_{ju}} \sqrt{\rho_{ul}} \hat{\mathbf{b}}_{lkd}^H \hat{\mathbf{g}}_{ljus} x_{jus} \\ &+ \sum_{j=1}^L \sum_{u=1}^K \sum_{s \in \mathcal{A}_{ju}} \sqrt{\rho_{ul}} \hat{\mathbf{b}}_{lkd}^H \tilde{\mathbf{g}}_{ljus} x_{jus} + \hat{\mathbf{b}}_{lkd}^H \mathbf{n}_l. \end{aligned} \quad (21)$$

For large antenna systems, it can be argued that all of the above inner products in the form of $\hat{\mathbf{b}}_{lkd}^H \hat{\mathbf{g}}_{ljus}$ and $\hat{\mathbf{b}}_{lkd}^H \tilde{\mathbf{g}}_{ljus}$ for all j, u, s asymptotically converge to deterministic constants. Since the transmitted symbols \mathbf{x}_{jus} are assumed to be Gaussian distributed, $\hat{y}_{lkd}^{\text{IP}}$ can be approximated as a sum of Gaussian RVs and is therefore Gaussian distributed with zero mean and variance, i.e.,

$$\begin{aligned} \eta_{lkd|c}^{\text{IP}} &\triangleq E \{ \hat{y}_{lkd}^{\text{IP}} \hat{y}_{lkd}^{\text{IP}*} \} \\ &= E \{ \mu_{lkd|c}^{\text{IP}} + \nu_{lkd|c}^{\text{IP}} \} \end{aligned} \quad (22)$$

with $\mu_{lkd|c}^{\text{IP}}$ and $\nu_{lkd|c}^{\text{IP}}$ being given in (16) and (17), respectively. Similar arguments can be made for the case of perfect CSI. Hence, the superscripts $(\cdot)^{\text{IP}}$ are omitted to generalize the discussion and to simplify the notation. It follows that $|\hat{y}_{lkd}|^2$ can be approximated as exponentially distributed RVs with the probability distribution function (pdf), i.e., $f_{|\hat{y}_{lkd}|^2}(z)$, and the cumulative distribution function (cdf), i.e., $F_{|\hat{y}_{lkd}|^2}(z)$, given by

$$f_{|\hat{y}_{lkd}|^2}(z) = \frac{1}{\eta_{lkd|c}} e^{-\frac{z}{\eta_{lkd|c}}} \quad (23)$$

$$F_{|\hat{y}_{lkd}|^2}(z) = 1 - e^{-\frac{z}{\eta_{lkd|c}}}. \quad (24)$$

Let $\tilde{\mathcal{A}}_i$ be a complement of a set \mathcal{A}_i , i.e., containing antenna indexes that are not included in \mathcal{A}_i . The cardinality of $\tilde{\mathcal{A}}_i$ is, therefore, equal to $N - A$. It can be shown from (9) that the conditions for an antenna combination \mathcal{A}_i to be detected can be written as

$$|\hat{y}_{lkd}|^2 > \max_{d \in \tilde{\mathcal{A}}_i} \{ |\hat{y}_{lk\tilde{d}}|^2 \}, \forall d \in \mathcal{A}_i. \quad (25)$$

Following a similar idea as in [16], by treating \hat{y}_{lkd} as being statistically independent for all d , it can be seen that $|\hat{y}_{lk\tilde{d},\max}|^2 \triangleq \max_{d \in \tilde{\mathcal{A}}_i} \{ |\hat{y}_{lk\tilde{d}}|^2 \}$ is the $(N - A)$ th-order statistic, i.e., the largest value of the $N - A$ RVs in $\tilde{\mathcal{A}}_i$. It follows from (80) that the pdf of $|\hat{y}_{lk\tilde{d},\max}|^2$ is given by

$$\begin{aligned} f_{|\hat{y}_{lk\tilde{d},\max}|^2}(z) &= \frac{1}{(N - A - 1)!} + \\ &\left\{ \begin{array}{ccc} F_{|\hat{y}_{lk\tilde{d}_1}|^2}(z) & \cdots & F_{|\hat{y}_{lk\tilde{d}_{N-A}}|^2}(z) \\ \vdots & & \vdots \\ F_{|\hat{y}_{lk\tilde{d}_1}|^2}(z) & \cdots & F_{|\hat{y}_{lk\tilde{d}_{N-A}}|^2}(z) \\ f_{|\hat{y}_{lk\tilde{d}_1}|^2}(z) & \cdots & f_{|\hat{y}_{lk\tilde{d}_{N-A}}|^2}(z) \end{array} \right\} \begin{array}{l} (N - A - 1) \text{ rows} \\ + \\ 1 \text{ row} \end{array} \end{aligned} \quad (26)$$

where $|\mathbf{A}|^+$ denotes the permanent of matrix \mathbf{A} , which is similarly defined as the determinant, except that all signs are positive [33]. It follows that the probability of detecting antenna combination i when antenna combination c is transmitted can be written as

$$\begin{aligned} E \{ \rho_{lk(i|c)} \} &= P \left(|\hat{y}_{lkd}|^2 > |\hat{y}_{lk\tilde{d},\max}|^2 \text{ for all } d \in \mathcal{A}_i \right) \\ &\approx \int_0^\infty f_{|\hat{y}_{lk\tilde{d},\max}|^2}(z) \cdot P \left(|\hat{y}_{lkd}|^2 > z \text{ for all } d \in \mathcal{A}_i \right) dz \\ &\approx \int_0^\infty f_{|\hat{y}_{lk\tilde{d},\max}|^2}(z) \cdot \prod_{d \in \mathcal{A}_i} e^{-\frac{z}{\eta_{lkd|c}}} dz \triangleq P_{lk(i|c)}. \end{aligned} \quad (27)$$

Proposition 2: The probability of detecting antenna combination i conditioned on that antenna combination c is transmitted for user (l, k) can be approximated by

$$P_{lk(i|c)} = \sum_{\substack{\forall \mathcal{A}_n \subseteq \tilde{\mathcal{A}}_i \\ \mathcal{A}_n \neq \emptyset}} \frac{\Upsilon(\mathcal{A}_n \cup \mathcal{A}_i)}{\Upsilon(\mathcal{A}_n)} (-1)^{(|\mathcal{A}_n|+1)} \quad (28)$$

whereby the summation in (28) is evaluated for all possible subsets of $\tilde{\mathcal{A}}_i$, and $\Upsilon(\mathcal{S})$, for $\mathcal{S} \neq \emptyset$, is defined as

$$\Upsilon(\mathcal{S}) = \frac{1}{\sum_{d \in \mathcal{S}} \frac{1}{\eta_{lkd|c}}}. \quad (29)$$

Proof: See Appendix B. ■

Using the approximation of $P_{lk(i|c)}$ in (28) for all pairs (i, c) and following the approximation in [9, Lemma 1], i.e., $E \{ \log_2(1 + (\sum_i X_i / \sum_i Y_i)) \} \approx \log_2(1 + (\sum_i E \{ X_i \} / \sum_i E \{ Y_i \}))$

$\sum_i E\{Y_i\}$), which tends to be relatively accurate for large antenna systems, the achievable rate in (15) can be estimated as

$$R_{lk,1}^{\text{IP}} \approx \frac{1}{C} \sum_{c=1}^C \sum_{i=1}^C P_{lk(i|c)}^{\text{IP}} \left[\sum_{d \in \mathcal{A}_i} \log_2 \left(1 + \frac{E \left\{ \mu_{lkd|c}^{\text{IP}} \right\}}{E \left\{ \nu_{lkd|c}^{\text{IP}} \right\}} \right) \right] \triangleq \tilde{R}_{lk,1}^{\text{IP}}. \quad (30)$$

Furthermore, applying the log sum inequality [34, Th. 2.7.1] to (15), it can be shown that (15) can be lower bounded by

$$R_{lk,2}^{\text{IP}} \geq \log_2(C) - \frac{1}{C} \sum_{c=1}^C \sum_{i=1}^C E \left\{ \varrho_{lk(i|c)}^{\text{IP}} \right\} \times \log_2 \frac{\sum_{c_2=1}^C E \left\{ \varrho_{lk(i|c_2)}^{\text{IP}} \right\}}{E \left\{ \varrho_{lk(i|c)}^{\text{IP}} \right\}} \approx \log_2(C) - \frac{1}{C} \sum_{c=1}^C \sum_{i=1}^C P_{lk(i|c)}^{\text{IP}} \log_2 \frac{\sum_{c_2=1}^C P_{lk(i|c_2)}^{\text{IP}}}{P_{lk(i|c)}^{\text{IP}}} \triangleq \tilde{R}_{lk,2}^{\text{IP}}. \quad (31)$$

Using (30) and (31) as the estimates of (14) and (15), respectively, an approximation of the ergodic achievable rate (13) can be obtained, as can be illustrated for the case of MF and ZF detection in the following section.

B. Ergodic Achievable Rate Approximation

Following the assumption in [12], where the uplink is considered to be interference limited, it can be assumed that the pilot power is sufficiently high, i.e., $(\sigma^2/\rho_t) \rightarrow 0$. It follows, from (4), that $\mathbf{A} \rightarrow (\beta_{llk}/\sum_{j=1}^L \beta_{ljk}) \mathbf{I}_{MN}$ and the MMSE channel estimate in (5) can be approximated as

$$\hat{\mathbf{G}}_{llk} \approx \frac{\beta_{llk}}{\beta_{l(k)}} \left(\sum_{j=1}^L \mathbf{G}_{ljk} + \frac{\tilde{\mathbf{N}}_{lk}}{\sqrt{\rho_t}} \right) \quad (32)$$

with $\beta_{l(k)} = \sum_{j=1}^L \beta_{ljk}$. It is seen that $\hat{\mathbf{G}}_{llk}$ is a zero-mean complex Gaussian matrix. The covariance properties of $\hat{\mathbf{G}}_{llk}$ can be obtained by applying the following statistical properties derived from (1), i.e.,

$$E \left\{ \mathbf{g}_{ljus} \mathbf{g}_{llkd}^H \right\} = \begin{cases} \beta_{llk} t_{sd} \mathbf{R}_H, & (j, u) = (l, k) \\ \mathbf{0}, & (j, u) \neq (l, k) \end{cases} \quad (33)$$

with $t_{sd} = [\mathbf{R}_T]_{s,d}$. Specifically, it can be shown that

$$E \left\{ \hat{\mathbf{g}}_{llkd} \hat{\mathbf{g}}_{llkd}^H \right\} = \frac{\beta_{llk}^2}{\beta_{l(k)}^2} \left(\sum_{j=1}^L \beta_{ljk} t_{dd} \mathbf{R}_H + \frac{\sigma^2}{\rho_t} \mathbf{I}_M \right) \approx \frac{\beta_{llk}^2}{\beta_{l(k)}} t_{dd} \mathbf{R}_H \quad (34)$$

$$E \left\{ \hat{\mathbf{g}}_{ljus} \hat{\mathbf{g}}_{llkd}^H \right\} = \begin{cases} \frac{\beta_{ljk} \beta_{llk}}{\beta_{l(k)}} t_{sd} \mathbf{R}_H, & u = k \text{ and } s \neq d \\ \mathbf{0}, & u \neq k \end{cases} \quad (35)$$

$$E \left\{ \hat{\mathbf{g}}_{llkd} \hat{\mathbf{g}}_{llkd}^H \right\} = E \left\{ \mathbf{g}_{llkd} \mathbf{g}_{llkd}^H \right\} - E \left\{ \hat{\mathbf{g}}_{llkd} \hat{\mathbf{g}}_{llkd}^H \right\} \approx \left(\beta_{llk} - \frac{\beta_{llk}^2}{\beta_{l(k)}} \right) t_{dd} \mathbf{R}_H. \quad (36)$$

1) *MF Detection*: Letting $[\mathbf{R}_T]_{d,d} = t_{dd}$, $[\mathbf{R}_T]_{*,d} = \mathbf{t}_d$, and $\phi_{llk} = \beta_{llk}^2/\beta_{l(k)}$, an approximation of the ergodic achievable rate with MF detection can be given in the following.

Corollary 2: A lower bound on the achievable rate of user (l, k) with MF detection can be approximated by $(1 - (\tau/T)) R_{lk}^{\text{IP, MF}} \approx (1 - (\tau/T)) (\tilde{R}_{lk,1}^{\text{IP, MF}} + \tilde{R}_{lk,2}^{\text{IP, MF}})$, where $\tilde{R}_{lk,1}^{\text{IP, MF}}$ and $\tilde{R}_{lk,2}^{\text{IP, MF}}$ are obtained from (30) and (31), respectively, using $P_{i|c}^{\text{IP}}$ from (28) with

$$E \left\{ \mu_{lkd|c}^{\text{IP}} \right\} \approx \begin{cases} \rho_{ul} \phi_{llk}^2 t_{dd}^2, & d = a, a \in \mathcal{A}_c \\ \rho_{ul} \phi_{llk}^2 \left(\frac{1}{A} \right) \sum_{s \in \mathcal{A}_c} t_{sd}^2, & d \neq a, \forall a \in \mathcal{A}_c \end{cases} \quad (37)$$

$$E \left\{ \nu_{lkd|c}^{\text{IP}} \right\} \approx \rho_{ul} M^2 \phi_{llk}^2 t_{sd}^2 + \rho_{ul} M^2 \phi_{llk}^2 \left(\frac{A}{N} \right) \|\mathbf{t}_d\|^2 \times \sum_{j \neq l} \frac{\beta_{ljk}^2}{\beta_{llk}^2} + \rho_{ul} \phi_{llk} t_{dd} \left(\frac{A}{N} \right) \sum_{j=1}^L \sum_{u=1}^K \sum_{s=1}^N \times \Re \left\{ \text{Tr} \left[\mathbf{Q}_{ljus} \left(\mathbf{R}_H + \frac{\sigma^2}{\rho_t t_{dd} \beta_{l(k)}} \mathbf{I}_M \right) \right] \right\} + \sigma^2 M \phi_{llk} t_{dd} \quad (38)$$

$$t_{sd}^2 = \begin{cases} \sum_{s \in \mathcal{A}_c \setminus \{a\}} t_{sd}^2, & d = a, a \in \mathcal{A}_c \\ \left(\frac{A-1}{A} \right) \sum_{s \in \mathcal{A}_c} t_{sd}^2, & d \neq a \quad \forall a \in \mathcal{A}_c \end{cases} \quad (39)$$

$$\mathbf{Q}_{ljus} \triangleq \begin{cases} \beta_{ljk} t_{ss} \mathbf{R}_H, & u = k \\ \left(\beta_{lju} - \frac{\beta_{lju}^2}{\beta_{l(u)}} \right) t_{ss} \mathbf{R}_H, & u \neq k. \end{cases} \quad (40)$$

Proof: See Appendix C. \blacksquare

Corollary 3: For the perfect-CSI case, an approximation on the achievable rate of user (l, k) with MF detection can be obtained as $R_{lk}^{\text{P, MF}} \approx \tilde{R}_{lk,1}^{\text{P, MF}} + \tilde{R}_{lk,2}^{\text{P, MF}}$, with $\tilde{R}_{lk,1}^{\text{P, MF}}$ and $\tilde{R}_{lk,2}^{\text{P, MF}}$ obtained from (30) and (31), respectively, using $P_{i|c}^{\text{P}}$ from (28) with

$$E \left\{ \mu_{lkd|c}^{\text{P}} \right\} \approx \rho_{ul} \beta_{llk}^2 M^2 t_{ad}^2 \quad (41)$$

$$E \left\{ \nu_{lkd|c}^{\text{P}} \right\} \approx \rho_{ul} \beta_{llk}^2 M^2 \sum_{s \in \mathcal{A}_c \setminus \{a\}} t_{sd}^2 + \rho_{ul} \beta_{llk} t_{dd} \left(\frac{A}{N} \right) \sum_{(j,u) \neq (l,k)} \sum_{s=1}^N \times \Re \left\{ \text{Tr} \left[\mathbf{Q}_{ljus} \mathbf{R}_H \right] \right\} + \sigma^2 \beta_{llk} M t_{dd} \quad (42)$$

and $\mathbf{Q}_{ljus} = \beta_{ljk} t_{ss} \mathbf{R}_H$.

2) *ZF Detection*: Letting $[\mathbf{R}_T^{-T}]_{d,d} = z_{dd}$, $[\mathbf{R}_T^{-T}]_{*,d} = \mathbf{z}_d$, and $\psi = 1$, an approximation of the ergodic achievable rate with ZF detection can be given in the following. Note that, despite being equal to 1, a multiplier ψ is included for future use in improving the approximation, as will be elaborated upon in Section IV-C.

Corollary 4: A lower bound on the achievable rate of user (l, k) with ZF detection can be approximated by $(1 - (\tau/T)) R_{lk}^{\text{IP, ZF}} \approx (1 - (\tau/T)) (\tilde{R}_{lk,1}^{\text{IP, ZF}} + \tilde{R}_{lk,2}^{\text{IP, ZF}})$, where $\tilde{R}_{lk,1}^{\text{IP, ZF}}$ and $\tilde{R}_{lk,2}^{\text{IP, ZF}}$

are obtained from (30) and (31), respectively, using $P_{i|c}^{\text{IP}}$ from (28) with

$$E \left\{ \mu_{lkd|c}^{\text{IP}} \right\} = \begin{cases} \rho_{ul}, & d = a \\ 0, & d \neq a \end{cases} \quad (43)$$

$$E \left\{ \nu_{lkd|c}^{\text{IP}} \right\} \approx \frac{\rho_{ul} z_{dd} \psi^2}{M^2 \phi_{llk}} \left(\frac{A}{N} \right) \sum_{j=1}^L \sum_{u=1}^K \sum_{s=1}^N \times \Re \left\{ \text{Tr} \left[\mathbf{Q}_{ljus} \left(\mathbf{R}_H + \frac{\sigma^2 \|\mathbf{z}_d\|^2}{\rho_t z_{dd} \beta_{l(u)}} \mathbf{I}_M \right) \right] \right\} + \rho_{ul} \left(\frac{A}{N} \right) \sum_{j \neq l} \frac{\beta_{ljk}^2}{\beta_{llk}^2} + \frac{\sigma^2 \psi}{M \phi_{llk}} z_{dd} \quad (44)$$

and $\mathbf{Q}_{ljus} = (\beta_{lju} - (\beta_{lju}^2 / \beta_{l(u)})) t_{ss} \mathbf{R}_H$.

Proof: See Appendix D. ■

Corollary 5: For the perfect-CSI case, an approximation on the achievable rate of user (l, k) with ZF detection can be obtained as $R_{lk}^{P,ZF} \approx \tilde{R}_{lk,1}^{P,ZF} + \tilde{R}_{lk,2}^{P,ZF}$, with $\tilde{R}_{lk,1}^{P,ZF}$ and $\tilde{R}_{lk,2}^{P,ZF}$ obtained from (30) and (31), respectively, using $P_{i|c}^P$ from (28) with

$$E \left\{ \mu_{lkd|c}^P \right\} = \begin{cases} \rho_{ul}, & d = a \\ 0, & d \neq a \end{cases} \quad (45)$$

$$E \left\{ \nu_{lkd|c}^P \right\} \approx \frac{\rho_{ul} z_{dd} \psi^2}{M^2 \beta_{llk}} \left(\frac{A}{N} \right) \sum_{j \neq l} \sum_{u=1}^K \sum_{s=1}^N \Re \left\{ \text{Tr} [\mathbf{Q}_{ljus} \mathbf{R}_H] \right\} + \frac{\sigma^2 \psi}{M \beta_{llk}} z_{dd} \quad (46)$$

and $\mathbf{Q}_{ljus} = \beta_{lju} t_{ss} \mathbf{R}_H$.

Note that, for uncorrelated Rayleigh channels with perfect CSI, it can be shown from (28) that $P_{i|c}^P \rightarrow 1$ for $i = c$ and $P_{i|c}^P \rightarrow 0$ for $i \neq c$ as $M \rightarrow \infty$. It follows that $R_{lk}^P \rightarrow \log_2(C)$, which is the theoretical upper bound [26], suggesting that the upper bound of the achievable rate can be achieved with linear (MF, ZF) detection at an asymptotically large M value.

C. Discussion on Accuracy and Improvement of the Approximation

1) *MF Detection:* The analysis in Section IV-A applies the concept of order statistics by treating \hat{y}_{lkd} , for all values of d , in (21) as being statistically independent. For MF detection, however, the received symbols $\{\hat{y}_{lkd}\}$ are seen to be correlated since, for any d , \hat{y}_{lkd} contains a linear combination of the transmitted symbols x_{lks} , for $\forall s \in \mathcal{A}_c$. Consequently, (26) appears to underestimate the probability of correctly detecting antenna combinations since $\max_{\tilde{d} \in \tilde{\mathcal{A}}_i} \{|\hat{y}_{lkd}|^2\}$ given in (25) tends to have smaller variation due to the correlation of $\{\hat{y}_{lkd}\}$. Simulation results suggest that the approximation of ergodic achievable rates in Section IV-B1, which is labeled as *approximation 1*, remains relatively accurate when $A \geq (N/2)$, as shown in Fig. 1 with the case of $A = 2$ and $N = 4$. For $A < (N/2)$, the approximation can be improved by using the following heuristic approach. To reflect smaller variation on $\max_{\tilde{d} \in \tilde{\mathcal{A}}_i} \{|\hat{y}_{lkd}|^2\}$, a scaling factor can be incorporated into the variance of $|\hat{y}_{lkd}|^2$ for $\tilde{d} \in \tilde{\mathcal{A}}_i$. Experimental observation

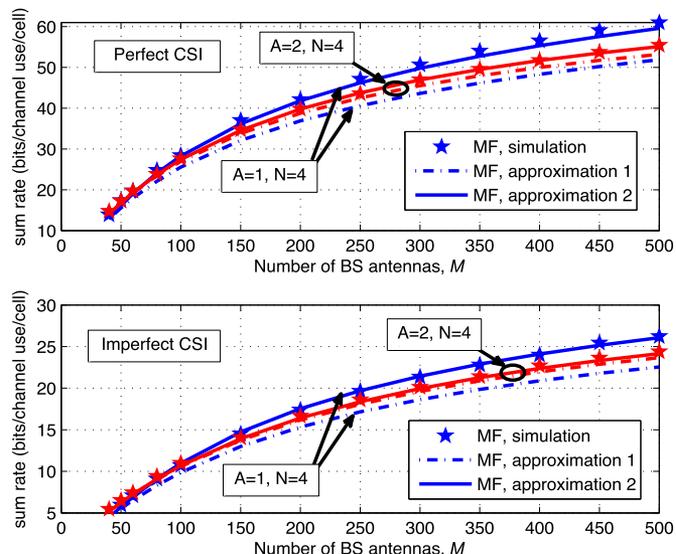


Fig. 1. Analytical approximations of sum-rate performance with MF detection.

suggests that the modified variance, which is denoted as $\hat{\eta}_{lkd|c}$, may be expressed as

$$\hat{\eta}_{lkd|c} = \begin{cases} \eta_{lkd|c}, & d \in \mathcal{A}_i \\ \min \left\{ 1, 1 - \frac{1}{A} \left(1 - \frac{A}{N} \right)^\theta \log_2 \left(\frac{M}{NK} \right) \right\} \eta_{lkd|c}, & d \in \tilde{\mathcal{A}}_i \end{cases} \quad (47)$$

with θ observed to be 2 and 3 for the case of perfect and imperfect CSI, respectively. With $\hat{\eta}_{lkd|c}$, the antenna detection probability given by (28) is denoted as $\hat{P}_{lk(i|c)}$. By modifying $\eta_{lkd|c}$ to $\hat{\eta}_{lkd|c}$, $\sum_{i=1}^C \hat{P}_{lk(i|c)}$, for any given c , may no longer be equal to 1. To ensure that the summation of the probabilities becomes unity, the following adjustment is also imposed:

$$\hat{P}_{lk(i|c)} = \begin{cases} \hat{P}_{lk(c|c)}, & i = c \\ \left(\frac{C}{\sum_{\substack{n=1 \\ n \neq c}}^C \hat{P}_{lk(n|c)}} \right)^{-1} \hat{P}_{lk(i|c)}, & i \neq c. \end{cases} \quad (48)$$

Note that, despite (47) and (48) being obtained from heuristic observation, they tend to result in relatively accurate approximation as supported by extensive simulation for $2 \leq N \leq 8$ and can be demonstrated with the label *approximation 2* in Fig. 1.

2) *ZF Detection:* In contrast to MF detection, the received symbols \hat{y}_{lkd} tend to be uncorrelated with ZF detection. This can be seen by noticing that, with ZF detection, the first and second terms on the right-hand side (RHS) of (21) become zero when d does not match with one of the transmitted antennas in set \mathcal{A}_c , i.e., $\hat{\mathbf{b}}_{lkd}^H \hat{\mathbf{g}}_{llks} = 0$ for $d \neq s$, $s \in \mathcal{A}_c$. As a result, \hat{y}_{lkd} contains the transmitted symbol x_{lks} with $s = d$ only, leading to \hat{y}_{lkd} being uncorrelated for different values of d . Hence, the analysis in Section IV-A is generally valid for ZF detection.

It can be seen, however, that the approximation in Section IV-B2 is mainly based on the case of large antenna systems. In particular, the inverse of the matrix $(\hat{\mathbf{G}}_{ll}^H \hat{\mathbf{G}}_{ll})^{-1}$ in (74) is considered to be a block-diagonal matrix since the off-diagonal blocks converge to zero, i.e., $\hat{\mathbf{G}}_{llk}^H \hat{\mathbf{G}}_{llu} \rightarrow 0$ for $k \neq u$ as $M \rightarrow \infty$. Simulation results suggest that this results in a close approximation for large M values but tends to be

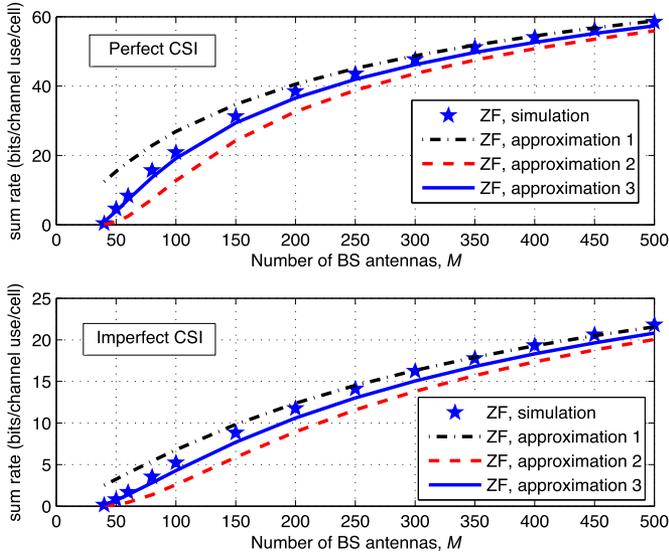


Fig. 2. Analytical approximations of sum-rate performance with ZF detection.

overestimated for smaller M values, as illustrated with the label *approximation 1* in Fig. 2. This may be improved by revisiting the approximation of the matrix inversion as follows. Using the identities of block matrix inversion, i.e., $(\hat{\mathbf{G}}_{ll}^H \hat{\mathbf{G}}_{ll})^{-1} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} W & X \\ Y & Z \end{bmatrix}$ with $W = (A - BD^{-1}C)^{-1}$, and the knowledge that A and D converge to block-diagonal matrices, the diagonal blocks of $(\hat{\mathbf{G}}_{ll}^H \hat{\mathbf{G}}_{ll})^{-1}$ can be approximated as

$$\begin{aligned} & \left[(\hat{\mathbf{G}}_{ll}^H \hat{\mathbf{G}}_{ll})^{-1} \right]_{k,k}^{(N \times N)} \\ & \approx \left(\hat{\mathbf{G}}_{llk}^H \left[\mathbf{I}_M - \sum_{u \neq k} \hat{\mathbf{G}}_{llu} (\hat{\mathbf{G}}_{llu}^H \hat{\mathbf{G}}_{llu})^{-1} \hat{\mathbf{G}}_{llu}^H \right] \hat{\mathbf{G}}_{llk} \right)^{-1} \\ & = \left(\hat{\mathbf{G}}_{llk}^H \left[\mathbf{I}_M - \sum_{u \neq k} \hat{\mathbf{U}}_{llu} \hat{\mathbf{U}}_{llu}^H \right] \hat{\mathbf{G}}_{llk} \right)^{-1} \end{aligned} \quad (49)$$

with $\hat{\mathbf{U}}_{llu} \in \mathbb{C}^{M \times N}$ being a matrix containing N eigenvectors of $\hat{\mathbf{G}}_{llu}^H \hat{\mathbf{G}}_{llu}$. Due to $\hat{\mathbf{G}}_{llu}$ being independent for all u , $\hat{\mathbf{U}}_{llu}$ are statistically independent. Using $E\{\hat{\mathbf{U}}_{llu} \hat{\mathbf{U}}_{llu}^H\} = (N/M)\mathbf{I}_M$ [35] as an estimate of $\hat{\mathbf{U}}_{llu} \hat{\mathbf{U}}_{llu}^H$, (49) can be rewritten as

$$\begin{aligned} \left[(\hat{\mathbf{G}}_{ll}^H \hat{\mathbf{G}}_{ll})^{-1} \right]_{k,k}^{(N \times N)} & \approx \left(\hat{\mathbf{G}}_{llk}^H \left(1 - \frac{(K-1)N}{M} \right) \hat{\mathbf{G}}_{llk} \right)^{-1} \\ & \approx \psi_2 (\phi_{llk} M)^{-1} \mathbf{R}_T^{-T} \end{aligned} \quad (50)$$

with $\psi_2 = (1 - ((K-1)N)/M)^{-1}$. Notice that (50) resembles (75) with ψ being replaced by ψ_2 . Thus, a similar approximation as in Section IV-B2 can be obtained with the modified scaling factor ψ_2 . It is observed by simulation that, in contrast to ψ , using ψ_2 leads to an underestimation for small values of M , suggesting that an optimized value would lie between ψ and ψ_2 . Motivated by this observation, a geometric mean $\psi_3 = \sqrt{\psi \psi_2} = (1 - ((K-1)N)/M)^{-(1/2)}$ can be applied and is shown to result in a close approximation, as shown in Fig. 2, with the labels *approximation 2* and *approximation 3* corresponding to the use of ψ_2 and ψ_3 , respectively.

D. Discussion on System Performance and Impact of Spatial Correlation

The general framework given in Section IV-A provides an analytical approach to evaluating the performance of Gen-SM systems, as demonstrated in Section IV-B and C. More insights into system performance can also be obtained based on these analytical results. As an illustration, the impact of spatial correlation to the ergodic achievable rate will be discussed in this section. From (30) and (31) in Section IV-A, the achievable rate is seen to depend on three main factors, i.e., the average signal power $E\{\mu_{lkd|c}^{\text{IP}}\}$, the average interference-plus-noise power $E\{\nu_{lkd|c}^{\text{IP}}\}$, and the antenna-combination detection probability $P_{lk(i|c)}^{\text{IP}}$. It can be shown that $E\{\mu_{lkd|c}^{\text{IP}}\}$ and $P_{lk(i|c)}^{\text{IP}}$ become worse with higher spatial correlation. In particular, for an MF receiver, the terms t_{sd} and $\|\mathbf{t}_d\|$ increase with an increasing spatial correlation at the BS, resulting in a higher value of $E\{\nu_{lkd|c}^{\text{IP}}\}$ in (38). It follows from (22), (28), and (29) that the probability of detecting the antenna combination erroneously $P_{lk(i|c)}^{\text{IP}}$ worsens. Consequently, the ergodic achievable rate is seen to degrade with an increasing spatial correlation. A similar conclusion can be drawn for the case of ZF by observing the increasing values of z_{dd} and $\|\mathbf{z}_d\|$ in (44) when spatial correlation increases.

It should also be highlighted that the analysis in this paper does not impose restrictions on the structures of \mathbf{R}_T and \mathbf{R}_R . The analysis can, therefore, be applied to spatial correlation matrices with any specific structures, e.g., to the case of practical channel models with finite numbers of angles of arrival (AoAs) at the BS in [12] and [36]–[38].

V. SPECTRAL, ENERGY, AND ECONOMIC EFFICIENCY TRADE-OFF

In typical communication systems, SE can generally be improved with increasing power consumption, resulting in decreasing EE. Thus, there is fundamentally a trade-off between SE and EE as investigated in various literature studies, e.g., [7] and [39]–[43]. Here, a trade-off between SE and EE for massive MIMO systems with Gen-SM will be examined. The discussion will be restricted to the case of imperfect CSI. An analogy can be drawn for the case of perfect CSI.

A. SE Versus EE

SE can be defined as the sum rate per cell (in bits per channel use per cell), as given by

$$R^{\text{IP}} = \left(\frac{1}{L} \right) \sum_{l=1}^L \sum_{k=1}^K \left(1 - \frac{\tau}{T} \right) R_{lk}^{\text{IP}}. \quad (51)$$

EE (in bits per Joule) is defined as a ratio of the SE (in bits per channel use per cell) to the transmit power consumption (in Joule per channel use per cell), as can be mathematically expressed by

$$\varepsilon^{\text{IP}} = \frac{1}{p_{\text{ms}}} R^{\text{IP}} \quad (52)$$

whereby p_{ms} denotes the total power expended by mobile stations for data transfer.

1) *Power Consumption Model*: Motivated by the work in [44], the sources of power consumption can be divided into three main groups, namely, radiated power, circuit power, and static power. For a required radiated power ρ_u , the power consumption of a practical power amplifier can be given by ρ_u/η_{PA} , with η_{PA} being the efficiency of the power amplifier. Circuit power concerns the power consumed by RF circuits and baseband processing. The power consumption due to RF circuits linearly increases with the number of RF chains as given by $AP_{\text{RF}}^{\text{ms}}$ per user. Baseband processing involves channel coding and modulation, also leading to power consumption as a linear function of the number of data streams as expressed by AP_{cod} . Static power considers fixed power consumption, i.e., P_0^{ms} , incurred by mobile stations for other purposes than for data transmission, such as control signaling and load-independent baseband processing, which does not scale with the number of RF chains and data streams. Using the given power consumption model, the total transmit power for K users in a cell can be expressed as

$$p_{\text{ms}} = K \left(\frac{\rho_u}{\eta_{\text{PA}}} + P_0^{\text{ms}} + AP_1 \right) \quad (53)$$

with $P_1 = P_{\text{RF}}^{\text{ms}} + P_{\text{cod}}$.

2) *SE and EE Trade-off*: Following the ideas in [7] and [40], the trade-off curves between SE and EE can be obtained by seeking a set of network parameters, such as transmit power (ρ_u) and number of users (K), that maximize the EE for a given SE. This leads to the following optimization problem:

$$\arg \max_{K, \rho_u} \varepsilon^{\text{IP}}, \quad \text{s.t. } R^{\text{IP}} = \text{constant}. \quad (54)$$

It will be shown in Section VI-C that a trade-off curve between SE and EE contains the *Pareto-optimal frontier* [45], representing the *efficient* (so-called *Pareto optimal*) operating points, in the sense that no points in the feasible SE and EE region can achieve better values of both SE and EE simultaneously. To gain a benefit in one particular aspect, the performance in the other aspect needs to be traded off. This motivates a search for a complementary performance metric that can provide a balanced trade-off between SE and EE along the Pareto-optimal frontier.

B. ECE

Inspired by the work in [40] and [46], ECE, measuring the profitability of the system (in monetary unit per second), is considered here as a complement to SE and EE. ECE implicitly takes into account SE and EE as both metrics that reflect revenues and costs of the system; therefore, it potentially offers a good trade-off between SE and EE. Despite having been considered in [40] and [46], the evaluation of ECE in these works appears to be limited to a single traffic class. In this paper, the definition of ECE is generalized to incorporate a multitude of traffic classes and tariff deals, as can be given by

$$\varepsilon^{\text{IP}} = \sum_{v=1}^V \kappa_r(v)U(v, R^{\text{IP}}) - (C_0 + \kappa_c p_{\text{bs}}) \quad (55)$$

where V denotes the number of traffic classes, $U(v, R^{\text{IP}})$ represents an effective chargeable data throughput of traffic class v (in bits per second), and $\kappa_r(v)$ is the revenue per bit of traffic class v . Furthermore, p_{bs} denotes the power consumption expended by BSs in the data transmission, κ_c is the energy cost per Joule (Watt-second), and C_0 involves other costs (in monetary unit per second) in addition to the energy cost.

To illustrate the concept of ECE in (55), three traffic classes ($V = 3$), representing voice traffic (class 1), data traffic with limited volume (class 2), and data traffic with unlimited volume (class 3), are considered in this paper. The revenue generated by each traffic class largely depends on marketing strategies and deals. Typically, voice service is charged per minute usage. Thus, the amount of chargeable traffic tends to be a linear function of the voice traffic volume. In contrast, based on the observation in [46], the attainable revenue for data service with limited quota seems to grow logarithmically with the data volume as users tend to consider acquiring higher traffic quota only with a slight increase in the premium. Moreover, by intuition, the generated revenue for data service with unlimited quota seems to be independent of the amount of data usage. Consequently, the effective chargeable data throughput of traffic classes 1, 2, and 3 can be modeled as linear, logarithmic, and constant functions, respectively, as can be expressed as

$$U(v, R^{\text{IP}}) = \begin{cases} R_1, & v = 1 \\ R_2^{\text{ref}} \cdot \log_2 \left(1 + \frac{R_v}{R_2^{\text{ref}}} \right), & v = 2 \\ R_3^{\text{ref}} \cdot u(R_3), & v = 3 \end{cases} \quad (56)$$

whereby $u(\cdot)$ is the Heaviside step function, and R_2^{ref} and R_3^{ref} represent referenced data rates for traffic classes 2 and 3, respectively. In (56), R_v denotes the amount of the cell throughput contributed to each traffic class, as can be modeled by

$$R_v = \begin{cases} \min \left\{ R_1^{\text{lim}}, \frac{\alpha_1}{T_{cu}} R^{\text{IP}} \right\}, & v = 1 \\ \frac{\alpha_2}{T_{cu}} R^{\text{IP}}, & v = 2 \\ \frac{1}{T_{cu}} R^{\text{IP}} - \min \left\{ r_1, \frac{\alpha_1}{T_{cu}} R^{\text{IP}} \right\} - \frac{\alpha_2}{T_{cu}} R^{\text{IP}}, & v = 3 \end{cases} \quad (57)$$

with α_1 and α_2 being the estimated traffic volume (in percent) for traffic classes 1 and 2, respectively, and T_{cu} denoting the time duration per channel use (c.u.). Since recent traffic trends [47] show that voice traffic tends to reach a saturate level, the upper limit of voice traffic R_1^{lim} is also included in (57).

The evaluation of energy cost in (55) considers the power consumption at the BS to reflect network operators' costs. This power is required for the detection of transmitted data from the users. Using the model in Section V-A1, it can be seen that radiated power can be omitted since only signal detection is considered. The power consumption of receiver chains can be written as $MP_{\text{RF}}^{\text{bs}}$. Baseband processing involves channel decoding and demodulation, incurring the power of KAP_{dec} . Additional processing is also required at the BS for uplink channel estimation and receive processing. Following the work in [44], the additional power consumption can be given by $MNK((1 + T - \tau)/ZT) + P_{\text{RX}}$, with P_{RX} being $2NKM/ZT$ and $((3N^2K^2M + 2NKM)/ZT) + (2N^3K^3/3ZT)$ for MF and ZF detection, respectively, and Z denoting the computational

TABLE I
POWER CONSUMPTION AND COST PARAMETERS

Power Consumption Parameters	Value	Cost Parameters	Value
$P_{\text{RF}}^{\text{ms}}$	0.3 W	$\kappa_r(1)$	1.43×10^{-5} pence/bit
$P_{\text{cod}}^{\text{ms}}$	0.5 W	$\kappa_r(2)$	7.75×10^{-7} pence/bit
P_0^{ms}	0.3 W	$\kappa_r(3)$	2.93×10^{-8} pence/bit
$P_{\text{RF}}^{\text{bs}}$	1 W	κ_c	4.22×10^{-6} pence/J
$P_{\text{dec}}^{\text{bs}}$	4 W	C_0	4.15×10^{-4} pence/s
P_0^{dec}	4 W	R_1^{lim}	50 Kbps
η_{PA}	0.3	R_2^{ref}	10 Kbps
T_{cu}	$\frac{1}{10^7}$ s/c.u.	R_3^{ref}	100 Kbps
σ^2	10^{-20} J/c.u.	α_1	1 %
Z	10^9 operations/J	α_2	80 %

efficiency in operations per Joule (J). Moreover, the static power consumption at the BS can be written as P_0^{bs} . Based on the given assumption, the power consumption can be given by

$$p_{\text{bs}} = P_0^{\text{bs}} + AKP_{\text{dec}}^{\text{bs}} + MP_{\text{RF}}^{\text{bs}} + MNK \left(\frac{1 + T - \tau}{ZT} \right) + P_{\text{RX}}. \quad (58)$$

By evaluating the ECE in (55) along the SE–EE trade-off curves, a sensible operating point achieving a good trade-off between SE and EE can be obtained, as will be illustrated in Section VI-C.

VI. SIMULATION RESULTS

Numerical examples are given in this section. Simulation parameters are described in Section VI-A, and results are presented in Section VI-B and C.

A. Simulation Parameters

1) *Network Parameters*: A multicell system with $L = 7$ cells, each with $K = 10$ users, are considered. Each user is assumed to be equipped with $N = 4$ antennas, unless stated otherwise. Following the assumptions in [7], [12], and [13], path gains of the serving links are set to 1, and those of interfering links are set to β , i.e., $\beta_{llk} = 1$ and $\beta_{lju} = \beta = 0.1 \forall j \neq l$. Furthermore, following the work in [6], the coherence time is set to $T = 98$ c.u.. It is also assumed that $\tau = NK$, and the signal-to-noise ratio (SNR) is defined as ρ_u/σ^2 .

2) *Antenna Correlation Parameters*: For a large antenna array, different antenna elements in the array may experience different sets of scatterers. The evolution of scatterers on the array axis can be modeled based on the birth–death process [48]. Here, the spatial correlation matrices of the received antennas can be given by $\mathbf{R}_R = \mathbf{E}_R \circ \mathbf{T}_R$, where $[\mathbf{E}_R]_{m,n} = e^{-\alpha|m-n|}$ represents the survival probability of the scatterers between the m th and the n th antenna and $[\mathbf{T}_R]_{m,n} = J_0(2\pi d_r|m-n|)$ is the correlation between the m th and the n th antenna based on the conventional Kronecker model, with $J_0(\cdot)$ denoting the zeroth-order Bessel function of the first kind and d_r being the antenna spacing of the receive-antenna arrays in wavelengths. For transmit antenna arrays, a few number of antenna elements (i.e., no evolution of scatterers) are considered. Thus, $\mathbf{R}_T = \mathbf{T}_T$ with $[\mathbf{T}_T]_{m,n} = J_0(2\pi d_t|m-n|)$. Note that R_R and R_T hold under the assumption of an isotropic scattering environment, i.e., the AoAs and angles of departure are uniformly distributed. Here, $\alpha = 0.3$, $d_r = d_t = 0.3$ are assumed.

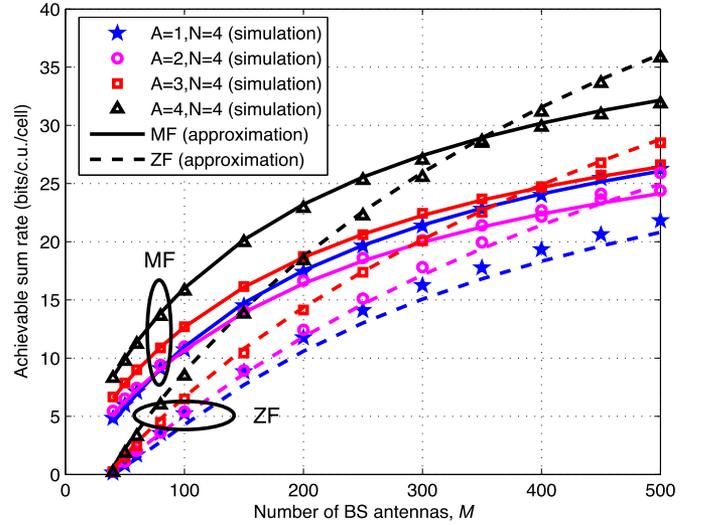


Fig. 3. Ergodic achievable sum rate versus number of BS antennas for imperfect CSI with $K = 10$ and $\text{SNR} = 10$ dB.

3) *Power Consumption and Cost Parameters*: Following the model in Section V, power consumption parameters are given in Table I. For a practical evaluation of ρ_u , users are assumed to be located at distance $(2/3)r$, with a cell radius of $r = 1000$ m and the path-loss exponent being 3.7.

Motivated by the work in [47], [49], and [50], the cost parameters are summarized in Table I. Note that these values indicate the current pricing and are referred to as *cost set-up 1*. The economic trends, however, suggest that the revenue per bit tends to decrease, whereas the network cost appears to be increasing. To evaluate future economic figures and the impact on ECE, an extreme set of parameters, which is referred to as *cost set-up 2*, is assumed by dividing the revenue per bit ($\kappa_r(1)$, $\kappa_r(2)$, and $\kappa_r(3)$) and multiplying the cost parameters (κ_c and C_0) by a factor of 10.

B. Ergodic Achievable Sum Rate

The ergodic sum rates in the case of imperfect CSI for different values of M and SNR are given in Figs. 3 and 4, respectively. The theoretical analysis with the improved approximation as presented in Section IV-C is used and is shown to agree well with the simulation results. For a small to moderate M value, MF detection tends to outperform ZF detection since, unlike ZF detection, the modulated symbols can be weakly detected with MF detection, even when the antenna combination

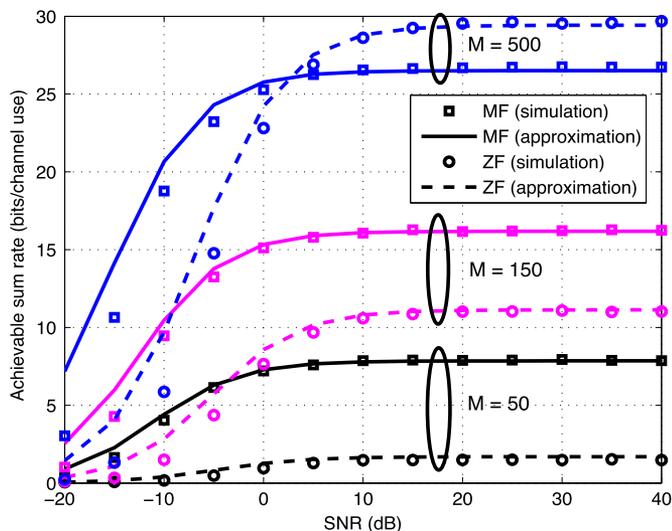


Fig. 4. Ergodic achievable sum rate versus SNR for imperfect CSI with $K = 10$, $N = 4$, and $A = 3$.

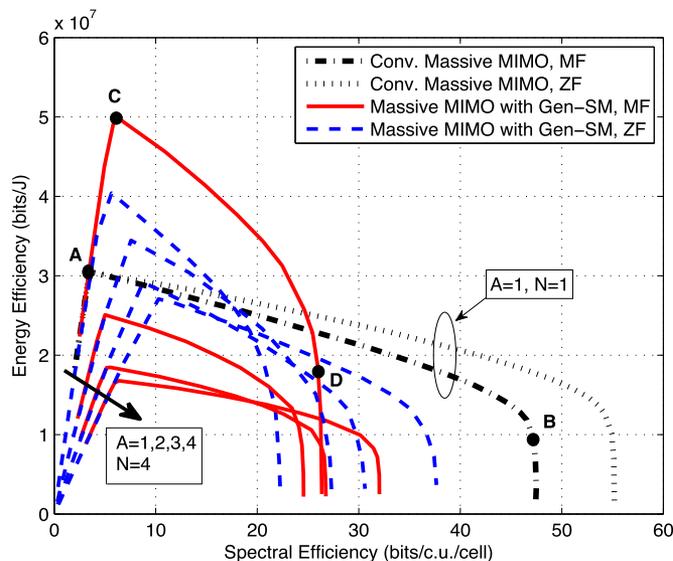


Fig. 6. EE and SE trade-off for imperfect CSI with $M = 500$.

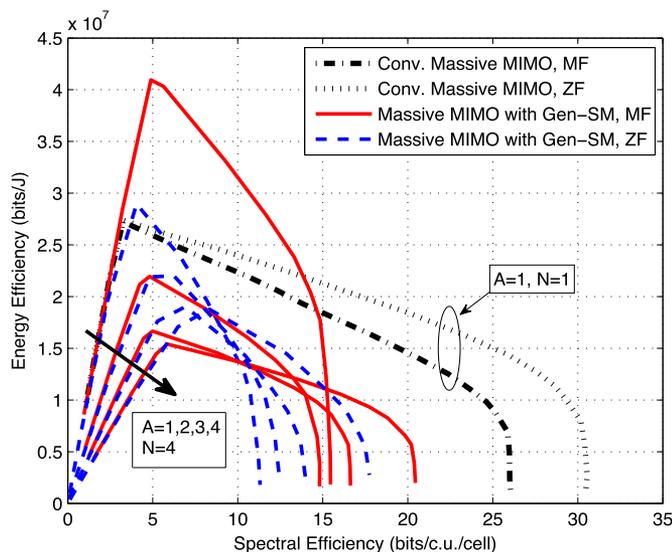


Fig. 5. EE and SE trade-off for imperfect CSI with $M = 150$.

is incorrectly decoded, as can be seen by comparing (37) with (43). In contrast, for a large M value with sufficiently high probability of correctly detecting the antenna combination, ZF detection becomes superior as it can suppress interstream interference more effectively compared with MF detection. Notice that, for ZF detection, the sum rate increases with the number of active antennas, i.e., A . For MF detection, on the other hand, since interstream interference cannot be mitigated at the receiving side, deploying a smaller A value may lead to a higher sum rate, as can be illustrated by comparing the cases of $A = 1$ with $A = 2$ for MF detection in Fig. 3.

C. Trade-off Analysis

A trade-off between EE and SE is evaluated using (54), as shown in Figs. 5 and 6 for $M = 150$ and $M = 500$, respectively. For comparison, results for conventional massive MIMO, i.e.,

$A = 1$ and $N = 1$, are also given. Despite contributing to a higher sum rate, increasing the number of active antennas for Gen-SM schemes is shown to reduce the EE due to the increase in circuit power consumption of RF chains and baseband processing. This suggests that the conventional SM with $A = 1$ tends to be the most energy-efficient mode among the entire class of Gen-SM.

Considering the case of $A = 1$ (SM mode), MF detection achieves better EE than ZF detection for a small value of M , as shown in Fig. 5, due to higher SE being achieved with the same amount of power consumption. The performance gap between MF and ZF becomes smaller as M increases, similar to the observation in Section VI-B, and can be demonstrated in Fig. 6. Due to the significant advantage over ZF detection, MF detection will be focused in the subsequent discussion. Comparing with conventional massive MIMO, SM offers benefits in terms of EE for a certain range (low to moderate values) of SE. Since SM requires longer time for pilot transmission ($\tau = NK$ in contrast to $\tau = K$ for conventional massive MIMO), the duration of data transmission becomes shorter, resulting in a feasible range of SE being limited compared with the conventional massive MIMO scheme. This, in turn, suggests that to combine the benefits of SM (operating with high EE at low-to-moderate SE) and conventional massive MIMO (operating at high SE exceeding the feasible range of SM), a dynamic scheme capable of flexibly switching between SM and conventional massive MIMO should be considered.

A detailed analysis of the trade-off between EE and SE can be demonstrated in Fig. 6. For conventional massive MIMO with MF detection, the *Pareto-optimal frontier* can be indicated as that part on the trade-off curve between points A and B. The points on the curve before point A are seen to be nonoptimal as they are always worse than point A in terms of both EE and SE. The points on the curve after point B are also not optimal for a similar reason. Analogously, the Pareto-optimal frontier for SM with MF detection is indicated with points C and D. As described in Section V-A, the optimal frontiers

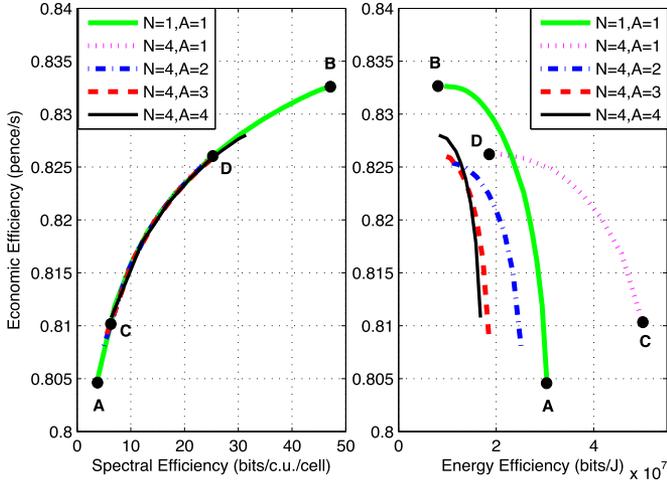


Fig. 7. ECE, evaluated along the Pareto-optimal frontier, for imperfect CSI with $M = 500$ and cost set-up 1.

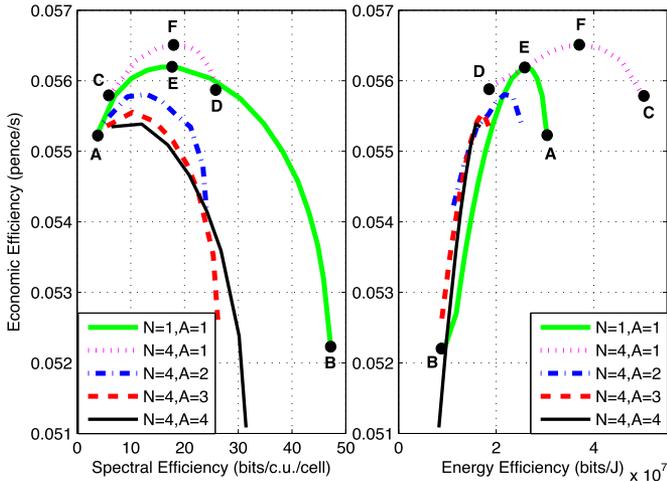


Fig. 8. ECE, evaluated along the Pareto-optimal frontier, for imperfect CSI with $M = 500$ and cost set-up 2.

represent efficient operating points, such that an extra advantage of one performance metric can only be achieved by trading it off with the benefits of the other metric. A sensible operating point on the optimal frontiers can be obtained by evaluating a complementary metric, such as ECE, along the trade-off curve, as shown in Figs. 7 and 8 for the two sets of cost parameters. For cost set-up 1 with the generated revenue dominating the energy cost, the strategy for maximizing ECE is to operate at the point with the highest SE, i.e., points B and D in Fig. 7 for conventional massive MIMO and SM schemes, respectively. In contrast, for cost set-up 2 with decreasing revenue and increasing energy cost, the operating point maximizing ECE tends to move toward the high-EE regime, as shown by points E and F in Fig. 8. In this case, SM (point F) is shown to achieve better ECE compared with the conventional massive MIMO (point E) as a result of significant advantages of SM in terms of EE. This suggests that SM is potentially an attractive candidate, offering a balanced trade-off between EE and SE, for future wireless communications.

VII. CONCLUSION

This paper has investigated the achievable sum-rate performance based on linear detection in multicell multiuser massive MIMO systems with Gen-SM, considering practical channel conditions, including antenna correlation, imperfect channel estimation, and pilot contamination. A general framework for the approximation of sum-rate performance with linear detection has been proposed and evaluated for MF and ZF detection. The proposed methodology has been shown to achieve a close approximation in comparison with simulation results. The fundamental trade-off between SE and EE has also been investigated. Among the entire class of Gen-SM, SM with a single active antenna per user has been demonstrated to be the most energy-efficient strategy as a result of reduced interchannel interference and the use of single RF chain per user. Compared with conventional massive MIMO, the feasible SE range of SM has been shown to be limited to lower values due to longer training duration being required for channel estimation, leaving shorter time slots for useful data transmission. However, within its feasible SE regime, SM tends to achieve better EE compared with conventional massive MIMO schemes. The Pareto-optimal frontiers, representing efficient operating points on SE and EE trade-off curves, have also been analyzed. To achieve a balanced SE–EE operating condition, ECE has been evaluated along the Pareto-optimal frontier and has been demonstrated to potentially provide a sensible trade-off between SE and EE.

VIII. FUTURE WORK

This paper has provided good insights into Gen-SM systems to a certain degree and their trade-off. The analysis in this paper can further be improved and extended in several aspects, some of which are outlined as follows.

- 1) The analysis is mainly based on a homogeneous network, i.e., all the cells and users are equipped with M and N antennas, respectively, and all channels have the same small-scale fading statistics, i.e., complex Gaussian matrix with mean $\mathbf{0}_{M \times N}$ and covariance matrix \mathbf{R}_H . As a result, the MMSE channel estimates can be written as (32), and the corresponding channel statistics can be expressed in a simplified form, as given in (34)–(36). In contrast, for heterogeneous networks with nonidentical small-scale fading statistics, the MMSE channel estimates and their statistical properties are given in a more complicated form, hindering the analysis in Section IV. The analysis of ergodic achievable sum rate for Gen-SM systems in heterogeneous networks remains a challenging task, which will be addressed in future work. Other research challenges and directions of heterogeneous networks are also highlighted in [51].
- 2) It has been assumed that the spatial correlation matrices \mathbf{R}_T and \mathbf{R}_R are known and can be used for MMSE channel estimation. However, the acquisition of these channel correlation matrices in practice would be hampered by the large dimension of massive antenna arrays, which will be studied in future research. The impact of spatial correlation matrices with special structures, e.g., as specified in [36]–[38], will also be considered.

- 3) Only two specific linear receivers, namely, MF and ZF, have been discussed in this paper. Receivers that are more robust to channel uncertainty and interference, such as the MMSE receiver, should be considered. Performance comparison with optimal ML detection will also be discussed in future work.
- 4) With the assumption of orthogonal uplink pilots, Gen-SM generally requires a longer training period, limiting the system performance, compared with conventional MIMO schemes. More advanced pilot designs, e.g., the work in [38], can be incorporated to further improve the performance of the proposed system.

APPENDIX A

The achievable rate, i.e., the RHS of (12), consists of two terms. The first term can be rewritten as

$$I(\bar{\mathbf{x}}_{lk}; \hat{A}_{lk}^{\text{IP}}, \hat{\mathbf{r}}_{lk}^{\text{IP}} | \bar{\mathbf{G}}_{lk}) = I(\bar{\mathbf{x}}_{lk}; \hat{\mathbf{r}}_{lk}^{\text{IP}} | \bar{\mathbf{G}}_{lk}, \hat{A}_{lk}^{\text{IP}}) + I(\bar{\mathbf{x}}_{lk}; \hat{A}_{lk}^{\text{IP}} | \bar{\mathbf{G}}_{lk}). \quad (59)$$

Since it is assumed that the antenna index is chosen independently from the modulated symbols, no information of $\bar{\mathbf{x}}_{lk}$ can be obtained from the detected antenna index \hat{A}_{lk}^{IP} . Thus, $I(\bar{\mathbf{x}}_{lk}; \hat{A}_{lk}^{\text{IP}} | \bar{\mathbf{G}}_{lk}) = 0$. Furthermore, $I(\bar{\mathbf{x}}_{lk}; \hat{\mathbf{r}}_{lk}^{\text{IP}} | \bar{\mathbf{G}}_{lk}, \hat{A}_{lk}^{\text{IP}})$ can be viewed as the mutual information of a *known channel* MIMO system. In particular, conditioned on $\bar{\mathbf{G}}_{lk} = \bar{\mathbf{G}}_{lk,c}$ (or equivalently, $A_{lk} = \mathcal{A}_c$) and $A_{lk}^{\text{IP}} = \mathcal{A}_i$, the received signal symbol of a detected antenna $d \in \mathcal{A}_i$ can be written as (21). Following the work in [7] and [12] and using the knowledge that $p(A_{lk} = \mathcal{A}_c) = 1/C$, $I(\bar{\mathbf{x}}_{lk}; \hat{\mathbf{r}}_{lk}^{\text{IP}} | \bar{\mathbf{G}}_{lk}, \hat{A}_{lk}^{\text{IP}})$ can be lower bounded by

$$I(\bar{\mathbf{x}}_{lk}; \hat{\mathbf{r}}_{lk}^{\text{IP}} | \bar{\mathbf{G}}_{lk} = \bar{\mathbf{G}}_{lk,c}, \hat{A}_{lk}^{\text{IP}} = \mathcal{A}_i) \geq \sum_{d \in \mathcal{A}_i} \log_2 \left(1 + \frac{\mu_{lkd|c}^{\text{IP}}}{\nu_{lkd|c}^{\text{IP}}} \right) \quad (60)$$

whereby $\mu_{lkd|c}^{\text{IP}}$ and $\nu_{lkd|c}^{\text{IP}}$ are defined in (16) and (17), respectively. It follows that

$$I(\bar{\mathbf{x}}_{lk}; \hat{\mathbf{r}}_{lk}^{\text{IP}} | \bar{\mathbf{G}}_{lk}, \hat{A}_{lk}^{\text{IP}}) \geq \frac{1}{C} \sum_{c=1}^C \sum_{i=1}^C \varrho_{lk(i|c)}^{\text{IP}} \times \left[\sum_{d \in \mathcal{A}_i} \log_2 \left(1 + \frac{\mu_{lkd|c}^{\text{IP}}}{\nu_{lkd|c}^{\text{IP}}} \right) \right]. \quad (61)$$

The second term on the RHS of (12) can also be given by

$$I(\bar{\mathbf{G}}_{lk}; \hat{A}_{lk}^{\text{IP}}, \hat{\mathbf{r}}_{lk}^{\text{IP}}) = I(\bar{\mathbf{G}}_{lk}; \hat{\mathbf{r}}_{lk}^{\text{IP}} | \hat{A}_{lk}^{\text{IP}}) + I(\bar{\mathbf{G}}_{lk}; \hat{A}_{lk}^{\text{IP}}). \quad (62)$$

Conditioned on \hat{A}_{lk}^{IP} , i.e., the antenna combination has been decoded, no further information on the spatial symbol $\bar{\mathbf{G}}_k$ can be obtained from $\hat{\mathbf{r}}_{lk}^{\text{IP}}$. Thus, $I(\bar{\mathbf{G}}_{lk}; \hat{\mathbf{r}}_{lk}^{\text{IP}} | \hat{A}_{lk}^{\text{IP}}) = 0$. Furthermore, $I(\bar{\mathbf{G}}_{lk}; \hat{A}_{lk}^{\text{IP}})$ can be rewritten as

$$I(\bar{\mathbf{G}}_{lk}; \hat{A}_{lk}^{\text{IP}}) = H(\bar{\mathbf{G}}_{lk}) - H(\bar{\mathbf{G}}_{lk} | \hat{A}_{lk}^{\text{IP}}) \quad (63)$$

with $H(\bar{\mathbf{G}}_{lk}) = \log_2(C)$ and $H(\bar{\mathbf{G}}_{lk} | \hat{A}_{lk}^{\text{IP}}) = (1/C) \sum_{c=1}^C \sum_{i=1}^C \varrho_{lk(i|c)}^{\text{IP}} \log_2((\sum_{c_2=1}^C \varrho_{lk(i|c_2)}^{\text{IP}}) / \varrho_{lk(i|c)}^{\text{IP}})$. It follows that (13) can be obtained by substituting (61) and (63) into (12).

APPENDIX B

From (26), it can be shown that

$$\begin{aligned} f_{|\hat{\mathbf{y}}_{lk\bar{d},\max}|^2}(z) &= \sum_{\forall d \in \bar{\mathcal{A}}_i} f_{|\hat{y}_{lkd}|^2}(z) \prod_{\forall n \in \bar{\mathcal{A}}_i \setminus \{d\}} F_{|\hat{y}_{lkn}|^2}(z) \\ &= \sum_{\forall d \in \bar{\mathcal{A}}_i} \frac{1}{\eta_{lkd|c}} e^{-\frac{z}{\eta_{lkd|c}}} \prod_{\forall n \in \bar{\mathcal{A}}_i \setminus \{d\}} \left(1 - e^{-\frac{z}{\eta_{lkn|c}}} \right). \end{aligned} \quad (64)$$

By expanding (64) and using the knowledge that $\prod_{\forall d \in \mathcal{S}} e^{-(z/\eta_{lkd|c})} = e^{-(z/(\Upsilon(\mathcal{S})))}$ with $\Upsilon(\mathcal{S})$ being defined in (29), (64) can be rewritten as

$$f_{|\hat{\mathbf{y}}_{lk\bar{d},\max}|^2}(z) = \sum_{\substack{\forall \mathcal{A}_n \subseteq \bar{\mathcal{A}}_i \\ \mathcal{A}_n \neq \emptyset}} \frac{(-1)^{(|\mathcal{A}_n|+1)}}{\Upsilon(\mathcal{A}_n)} e^{-\frac{z}{\Upsilon(\mathcal{A}_n)}}. \quad (65)$$

Substituting (65) into (27) results in

$$\begin{aligned} E\{\varrho_{lk(i|c)}\} &= \int_0^\infty \sum_{\substack{\forall \mathcal{A}_n \subseteq \bar{\mathcal{A}}_i \\ \mathcal{A}_n \neq \emptyset}} \frac{(-1)^{(|\mathcal{A}_n|+1)}}{\Upsilon(\mathcal{A}_n)} e^{-\frac{z}{\Upsilon(\mathcal{A}_n)}} \cdot \prod_{d \in \mathcal{A}_i} e^{-\frac{z}{\eta_{lkd|c}}} dz \\ &= \sum_{\substack{\forall \mathcal{A}_n \subseteq \bar{\mathcal{A}}_i \\ \mathcal{A}_n \neq \emptyset}} \frac{(-1)^{(|\mathcal{A}_n|+1)}}{\Upsilon(\mathcal{A}_n)} \int_0^\infty e^{-\frac{z}{\Upsilon(\mathcal{A}_n \cup \mathcal{A}_i)}} dz \end{aligned} \quad (66)$$

and (27) can be obtained from (66) by substituting $\int_0^\infty e^{-(z/(\Upsilon(\mathcal{A}_n \cup \mathcal{A}_i)))} dz = \Upsilon(\mathcal{A}_n \cup \mathcal{A}_i)$.

APPENDIX C

It can be seen from (6) that, for MF detection, $\hat{\mathbf{b}}_{lkd}^{\text{MF}} = \hat{\mathbf{g}}_{lkd}$. For large antenna systems, by the law of large numbers, $(1/M) \hat{\mathbf{g}}_{lkd}^H \hat{\mathbf{g}}_{llks} \rightarrow (1/M) E\{\hat{\mathbf{g}}_{lkd}^H \hat{\mathbf{g}}_{llks}\}$. Since $E\{\hat{\mathbf{g}}_{lkd}^H \hat{\mathbf{g}}_{llks}\} = E\{\text{Tr}(\hat{\mathbf{g}}_{llks} \hat{\mathbf{g}}_{lkd}^H)\}$, it is obtained from (34) and (35) that $(1/M) \hat{\mathbf{g}}_{lkd}^H \hat{\mathbf{g}}_{llks} \rightarrow (1/M) (\beta_{llk}^2 / \beta_{l(k)}) t_{sd} \text{Tr}(\mathbf{R}_H) = (\beta_{llk}^2 / \beta_{l(k)}) t_{sd}$. Using the asymptotic value as an approximation, it follows that

$$E\left\{ \left| \hat{\mathbf{g}}_{lkd}^H \hat{\mathbf{g}}_{llks} \right|^2 \right\} \approx M^2 \phi_{lk}^2 t_{sd}^2. \quad (67)$$

Consequently, (37) is derived from (67). Note that, with MF detection, useful information can still be received although the detected antenna d is erroneously decoded since $\hat{\mathbf{g}}_{lkd}^H \hat{\mathbf{g}}_{llka} \neq 0$ for $d \neq a$. By considering that the actual transmit antenna a lies in set \mathcal{A}_c with equal probability, the expected value of the useful received power can be obtained by averaging over $a \in \mathcal{A}_c$, resulting in the second condition in (37).

Using the knowledge that, for MF detection, $\hat{\mathbf{g}}_{llkd}$ is statistically independent of $\hat{\mathbf{g}}_{ljus}$, $\tilde{\mathbf{g}}_{ljus}$, and $\mathbf{g}_{ljus} = \hat{\mathbf{g}}_{ljus} + \tilde{\mathbf{g}}_{ljus}$ for all $u \neq k$, (17) can be rewritten as

$$\begin{aligned} \nu_{lkd|c}^{\text{IP}} &= \sum_{s \in \mathcal{A}_c \setminus \{a\}} \rho_{ul} |\hat{\mathbf{g}}_{llkd}^H \hat{\mathbf{g}}_{llks}|^2 + \sum_{j \neq l} \sum_{s \in \mathcal{A}_{ju}} \rho_{ul} |\hat{\mathbf{g}}_{llkd}^H \hat{\mathbf{g}}_{ljks}|^2 \\ &+ \sum_{j=1}^L \sum_{u \neq k} \sum_{s \in \mathcal{A}_{ju}} \rho_{ul} |\hat{\mathbf{g}}_{llkd}^H \mathbf{g}_{ljus}|^2 \\ &+ \sum_{j=1}^L \sum_{s \in \mathcal{A}_{ju}} \rho_{ul} |\hat{\mathbf{g}}_{llkd}^H \tilde{\mathbf{g}}_{ljks}|^2 + \|\hat{\mathbf{g}}_{llkd}\|^2 \sigma^2. \end{aligned} \quad (68)$$

Using (67) with $\hat{\mathbf{g}}_{ljks} = (\beta_{ljk}/\beta_{llk})\hat{\mathbf{g}}_{llks}$, the expected values of the first and second terms on the RHS of (68) can be estimated as

$$\begin{aligned} &E \left\{ \rho_{ul} \sum_{s \in \mathcal{A}_c \setminus \{a\}} |\hat{\mathbf{g}}_{llkd}^H \hat{\mathbf{g}}_{llks}|^2 \right\} \\ &\approx \begin{cases} \rho_{ul} M^2 \phi_{llk}^2 \sum_{s \in \mathcal{A}_c \setminus \{a\}} t_{sd}^2; & d = a, \quad a \in \mathcal{A}_c \\ \rho_{ul} M^2 \phi_{llk}^2 \left(\frac{A-1}{A}\right) \sum_{a \in \mathcal{A}_c} t_{ad}^2; & d \neq a \quad \forall a \in \mathcal{A}_c \end{cases} \quad (69) \\ &E \left\{ \sum_{j \neq l} \sum_{s \in \mathcal{A}_{jk}} \rho_{ul} |\hat{\mathbf{g}}_{llkd}^H \hat{\mathbf{g}}_{ljks}|^2 \right\} \\ &= E \left\{ \sum_{j \neq l} \sum_{s \in \mathcal{A}_{jk}} \rho_{ul} \frac{\beta_{ljk}^2}{\beta_{llk}^2} |\hat{\mathbf{g}}_{llkd}^H \hat{\mathbf{g}}_{llks}|^2 \right\} \\ &\approx \rho_{ul} M^2 \phi_{llk}^2 \left(\frac{A}{N}\right) \|\mathbf{t}_d\|^2 \sum_{j \neq l} \frac{\beta_{ljk}^2}{\beta_{llk}^2}. \end{aligned} \quad (70)$$

Moreover, the summation of the third and fourth terms of (68) can be expressed as $\sum_{j=1}^L \sum_{u=1}^K \sum_{s \in \mathcal{A}_{ju}} \rho_{ul} |\hat{\mathbf{g}}_{llkd}^H \mathbf{q}_{ljus}|^2$, whereby $\mathbf{q} \triangleq \begin{cases} \mathbf{g}_{ljus}, & u \neq k \\ \tilde{\mathbf{g}}_{ljks}, & u = k \end{cases}$, and the corresponding covariance matrix, i.e., \mathbf{Q}_{ljus} , is given in (40). Since $\hat{\mathbf{g}}_{llkd}$ and \mathbf{q}_{ljus} are statistically independent, by using (34) with some mathematical manipulation, it can be shown that

$$\begin{aligned} E \left\{ \rho_{ul} |\hat{\mathbf{g}}_{llkd}^H \mathbf{q}_{ljus}|^2 \right\} &= \rho_{ul} E \left\{ \hat{\mathbf{g}}_{llkd}^H \mathbf{Q}_{ljus} \hat{\mathbf{g}}_{llkd} \right\} \\ &\approx \rho_{ul} \phi_{llk} t_{dd} \Re \left\{ \text{Tr} \left[\mathbf{Q}_{ljus} \left(\mathbf{R}_H + \frac{\sigma^2}{\rho t_{dd} \beta_{l(k)}} \mathbf{I}_M \right) \right] \right\}. \end{aligned} \quad (71)$$

In addition, similar to (67), the last term of (68) can be given by

$$\sigma^2 E \left\{ \|\hat{\mathbf{g}}_{llkd}\|^2 \right\} \approx \sigma^2 M \phi_{llk} t_{dd}. \quad (72)$$

It can be seen that (38) is obtained from (69)–(72).

APPENDIX D

It can be readily seen that (43) is derived from (16) using the property of ZF receive filter given in (7), i.e., $\hat{\mathbf{b}}_{llkd}^H \hat{\mathbf{g}}_{llka}$ is equal to 1 for $d = a$ and to 0 otherwise. Using similar arguments with $\hat{\mathbf{g}}_{ljks} = (\beta_{ljk}/\beta_{llk})\hat{\mathbf{g}}_{llks}$, it can be shown that the first term on

the RHS of (17) is equal to 0 and that the expected value of the second term can be rewritten as

$$\begin{aligned} &E \left\{ \sum_{(j,u) \neq (l,k)} \sum_{s \in \mathcal{A}_{ju}} \rho_{ul} |\hat{\mathbf{b}}_{llkd}^H \hat{\mathbf{g}}_{ljus}|^2 \right\} \\ &= \sum_{j \neq l} \rho_{ul} \left(\frac{\beta_{ljk}^2}{\beta_{llk}^2} \right) \sum_{s \in \mathcal{A}_{jk}} E \left\{ |\hat{\mathbf{b}}_{llkd}^H \hat{\mathbf{g}}_{llks}|^2 \right\} \\ &= \sum_{j \neq l} \rho_{ul} \left(\frac{\beta_{ljk}^2}{\beta_{llk}^2} \right) E \left\{ |\hat{\mathbf{b}}_{llkd}^H \hat{\mathbf{g}}_{llkd}|^2 \right\} p(d \in \mathcal{A}_{jk}) \\ &= \rho_{ul} \left(\frac{A}{N} \right) \sum_{j \neq l} \frac{\beta_{ljk}^2}{\beta_{llk}^2} \end{aligned} \quad (73)$$

whereby $p(d \in \mathcal{A}_{jk})$, denoting the probability that antenna d of user (j, k) is active, can be calculated as the ratio of the number of antenna combinations containing d to the total number of all possible antenna combinations, i.e., $p(d \in \mathcal{A}_{jk}) = \binom{N-1}{A-1} / \binom{N}{A} = A/N$.

Furthermore, for large antenna systems, it can be shown that $\hat{\mathbf{G}}_{ll}^H \hat{\mathbf{G}}_{ll}$ converges to a block-diagonal matrix, i.e., the off-diagonal blocks converge to zero ($\hat{\mathbf{G}}_{llk}^H \hat{\mathbf{G}}_{llu} \rightarrow 0$ for $k \neq u$ as $M \rightarrow \infty$). Consequently, the k th diagonal block of $(\hat{\mathbf{G}}_{ll}^H \hat{\mathbf{G}}_{ll})^{-1}$ can be approximated as

$$\left[(\hat{\mathbf{G}}_{ll}^H \hat{\mathbf{G}}_{ll})^{-1} \right]_{k,k}^{(N \times N)} \approx (\hat{\mathbf{G}}_{llk}^H \hat{\mathbf{G}}_{llk})^{-1} \quad (74)$$

$$\approx \psi (\phi_{llk} M)^{-1} \mathbf{R}_T^{-T} \quad (75)$$

where (75) is derived from (74) and (34), (35), and $\psi = 1$. It follows from (7) that the ZF detector for user (l, k) can be given by

$$\hat{\mathbf{b}}_{llk}^{\text{ZF}} \approx \psi (\phi_{llk} M)^{-1} \hat{\mathbf{G}}_{llk} \mathbf{R}_T^{-T}. \quad (76)$$

Letting $\mathbf{R}_T^{-T} = \mathbf{Z}_T$, $[\mathbf{Z}_T]_{*,d} = \mathbf{z}_d$, and $[\mathbf{Z}_T]_{s,d} = \mathbf{z}_{sd}$, the ZF receive filter corresponding to antenna d of user (l, k) can be written as

$$\begin{aligned} \hat{\mathbf{b}}_{llkd} &\approx \psi (\phi_{llk} M)^{-1} \hat{\mathbf{G}}_{llk} \mathbf{z}_d \\ &= \psi (\phi_{llk} M)^{-1} \frac{\sqrt{\beta_{llk}}}{\beta_{l(k)}} \\ &\quad \times \left[\sum_{j=1}^L \sum_{s=1}^N z_{sd} \sqrt{\beta_{ljk}} \mathbf{h}_{ljks} + \sum_{s=1}^N z_{sd} \tilde{\mathbf{n}}_{lks} \right]. \end{aligned} \quad (77)$$

Using (77) and (33) with some mathematical manipulation, the expected value of the third term on the RHS of (17) can be evaluated as

$$\begin{aligned} E \left\{ \rho_{ul} |\hat{\mathbf{b}}_{llkd}^H \tilde{\mathbf{g}}_{ljus}|^2 \right\} &= \rho_{ul} E \left\{ \hat{\mathbf{b}}_{llkd}^H \mathbf{Q}_{ljus} \hat{\mathbf{b}}_{llkd} \right\} \\ &\approx \frac{\rho_{ul} z_{dd} \psi^2}{M^2 \phi_{llk}} \Re \left\{ \text{Tr} \left[\mathbf{Q}_{ljus} \left(\mathbf{R}_H + \frac{\sigma^2 \|\mathbf{z}_d\|^2}{\rho t_{dd} \beta_{l(k)}} \mathbf{I}_M \right) \right] \right\} \end{aligned} \quad (78)$$

with $\mathbf{Q}_{ljus} = E \{ \tilde{\mathbf{g}}_{ljus} \tilde{\mathbf{g}}_{ljus}^H \} = (\beta_{lju} - (\beta_{lju}^2/\beta_{l(u)})) t_{ss} \mathbf{R}_H$. Furthermore, using (75), the expected value of the fourth term on the RHS of (17) can be given by

$$\sigma^2 E \left\{ \|\hat{\mathbf{b}}_{llkd}\|^2 \right\} \approx \sigma^2 [\psi (\phi_{llk} M)^{-1} \mathbf{R}_T^{-T}]_{d,d} = \frac{\sigma^2 \psi}{M \phi_{llk}} z_{dd}. \quad (79)$$

It follows that (44) is obtained from (73), (78), and (79).

APPENDIX E

BRIEF INTRODUCTION TO ORDER STATISTICS

Suppose X_1, X_2, \dots, X_n are independently (but not necessarily identically) distributed RVs, with X_j having the pdf and the cdf denoted as $f_j(x)$ and $F_j(x)$, respectively. By taking a large number of sample sets, each set containing n samples of each RV that are sorted in nondecreasing order, the q th-order statistic, which is denoted as $X_{(q)}$, represents an RV containing the collection of the q th smallest sample in each sorted set. In particular, the order statistics can be represented as RVs $X_{(1)}, X_{(2)}, \dots, X_{(n)}$, with $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$. It has been analyzed in [33] that the pdf of $X_{(q)}$, which is denoted as $f_{(q)}(x)$ for $q \in \{1, \dots, n\}$, can be given by

$$f_{(q)}(x) = \frac{1}{(q-1)!(n-q)!} \begin{vmatrix} F_1(x) & \cdots & F_n(x) \\ \vdots & & \vdots \\ F_1(x) & \cdots & F_n(x) \\ f_1(x) & \cdots & f_n(x) \\ 1 - F_1(x) & \cdots & 1 - F_n(x) \\ \vdots & & \vdots \\ 1 - F_1(x) & \cdots & 1 - F_n(x) \end{vmatrix}^+ \quad (80)$$

$\left. \begin{array}{l} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} \right\} \begin{array}{l} (q-1) \text{ rows} \\ 1 \text{ row} \\ (n-q) \text{ rows} \end{array}$

where $^+|\mathbf{A}|^+$ denotes the permanent of matrix \mathbf{A} , which is similarly defined as the determinant, except that all signs are positive [33].

REFERENCES

[1] C.-X. Wang *et al.*, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 122–130, Feb. 2014.

[2] Z. Ma, Z. Zhang, Z. Ding, P.-Z. Fan, and H. Li, "Key techniques for 5G wireless communications: Network architecture, physical layer, and MAC layer perspectives," *Sci. China Inf. Sci.*, vol. 58, no. 4, pp. 1–20, Apr. 2015.

[3] X. Ge, H. Cheng, M. Guizani, and T. Han, "5G wireless backhaul networks: Challenges and research advances," *IEEE Netw.*, vol. 28, no. 6, pp. 6–11, Nov. 2014.

[4] L. Lu, G. Li, A. Swindlehurst, 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.

[5] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[6] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[7] H. Q. Ngo, E. Larsson, and T. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[8] H. Q. Ngo, M. Matthaiou, T. Duong, and E. Larsson, "Uplink performance analysis of multicell MU-SIMO systems With ZF receivers," *IEEE Trans. Veh. Technol.*, vol. 62, no. 9, pp. 4471–4483, Nov. 2013.

[9] Q. Zhang, S. Jin, K.-K. Wong, H. Zhu, and M. Matthaiou, "Power scaling of uplink massive MIMO systems with arbitrary-rank channel means," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 966–981, Oct. 2014.

[10] N. Krishnan, R. Yates, and N. Mandayam, "Uplink linear receivers for multi-cell multiuser MIMO with pilot contamination: Large system analysis," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4360–4373, Aug. 2014.

[11] H. Q. Ngo, T. Marzetta, and E. Larsson, "Analysis of the pilot contamination effect in very large multicell multiuser MIMO systems for physical channel models," in *Proc. IEEE ICASSP*, Prague, Czech Republic, May 2011, pp. 3464–3467.

[12] H. Ngo, E. Larsson, and T. Marzetta, "The multicell multiuser MIMO uplink with very large antenna arrays and a finite-dimensional channel," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2350–2361, Jun. 2013.

[13] 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.

[14] H. Yang and T. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 172–179, Feb. 2013.

[15] J. Jose, A. Ashikhmin, T. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.

[16] R. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2228–2241, Jul. 2008.

[17] 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.

[18] R. Mesleh and S. Ikki, "On the effect of Gaussian imperfect channel estimations on the performance of space modulation techniques," in *Proc. IEEE VTC—Spring*, Yokohama, Japan, May 2012, pp. 1–5.

[19] J. Jeganathan, A. Ghrayeb, L. Szczecinski, and A. Ceron, "Space shift keying modulation for MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3692–3703, Jul. 2009.

[20] M. Di Renzo and H. Haas, "Space Shift Keying (SSK) modulation with partial channel state information: Optimal detector and performance analysis over fading channels," *IEEE Trans. Commun.*, vol. 58, no. 11, pp. 3196–3210, Nov. 2010.

[21] Y. Fu *et al.*, "BER performance of spatial modulation systems under 3D V2V MIMO channel models," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5725–5730, Jul. 2016.

[22] A. Younis, N. Serafimovski, R. Mesleh, and H. Haas, "Generalised spatial modulation," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2010, pp. 1498–1502.

[23] 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.

[24] Y. Xiao *et al.*, "Low-complexity signal detection for generalized spatial modulation," *IEEE Commun. Lett.*, vol. 18, no. 3, pp. 403–406, Mar. 2014.

[25] 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, Apr. 2008.

[26] R. Rajashekar, K. Hari, and L. Hanzo, "Reduced-complexity ML detection and capacity-optimized training for spatial modulation systems," *IEEE Trans. Commun.*, vol. 62, no. 1, pp. 112–125, Jan. 2014.

[27] N. Serafimovski, S. Sinanovic, A. Younis, M. Di Renzo, and H. Haas, "2-user multiple access spatial modulation," in *Proc. IEEE GC Wkshps*, Houston, TX, USA, Dec. 2011, pp. 343–347.

[28] N. Serafimovski, S. Sinanovic, M. D. Renzo, and H. Haas, "Multiple access spatial modulation," *EURASIP J. Wireless Commun. Netw.*, vol. 2012, p. 299, Sep. 2012.

[29] M. Di Renzo and H. Haas, "Bit error probability of space-shift keying MIMO over multiple-access independent fading channels," *IEEE Trans. Veh. Technol.*, vol. 60, no. 8, pp. 3694–3711, Oct. 2011.

[30] A. Younis, R. Mesleh, M. D. Renzo, and H. Haas, "Generalised spatial modulation for large-scale MIMO," in *Proc. IEEE EUSIPCO*, Lisbon, Portugal, Sep. 2014, pp. 346–350.

[31] T. Narasimhan, P. Raviteja, and A. Chockalingam, "Large-scale multiuser SM-MIMO versus massive MIMO," in *Proc. ITA*, San Diego, CA, USA, Feb. 2014, pp. 1–9.

[32] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

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

[34] T. M. Cover and J. A. Thomas, *John Elements of Information Theory*. New York, NY, USA: Wiley, 1991.

[35] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications*. Breda, The Netherlands: Now, 2004.

[36] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

[37] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing—The large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441–6463, Oct. 2013.

- [38] L. You, X. Gao, X. Gen Xia, N. Ma, and Y. Peng, "Pilot reuse for massive MIMO transmission over spatially correlated Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3352–3366, Jun. 2015.
- [39] X. Hong, Y. Jie, C.-X. Wang, J. Shi, and X. Ge, "Energy–spectral efficiency trade-off in virtual MIMO cellular systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 10, pp. 2128–2140, Oct. 2013.
- [40] I. Ku, C.-X. Wang, and J. Thompson, "Spectral, energy and economic efficiency of relay-aided cellular networks," *IET Commun.*, vol. 7, no. 14, pp. 1476–1486, Sep. 2013.
- [41] X. Ge *et al.*, "Spatial spectrum and energy efficiency of random cellular networks," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 1019–1030, Mar. 2015.
- [42] L. Xiang, X. Ge, C.-X. Wang, F. Y. Li, and F. Reichert, "Energy efficiency evaluation of cellular networks based on spatial distributions of traffic load and power consumption," *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, pp. 961–973, Mar. 2013.
- [43] I. Ku, C.-X. Wang, and J. S. Thompson, "Spectral–energy efficiency trade-off in relay-aided cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 4970–4982, Oct. 2013.
- [44] E. Bjornson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Designing multi-user MIMO for energy efficiency: When is massive MIMO the answer?" in *Proc. IEEE WCNC*, Istanbul, Turkey, Apr. 2014, pp. 242–247.
- [45] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [46] J. Akhtman and L. Hanzo, "Power versus bandwidth-efficiency in wireless communications: The economic perspective," in *Proc. IEEE VTC—Fall*, Sep. 2009, pp. 1–5.
- [47] "Communications market report," OFCOM, London, U.K., Aug. 2013.
- [48] S. Wu, C.-X. Wang, E.-H. Aggoune, M. Alwakeel, and Y. He, "A non-stationary 3-D wideband twin-cluster model for 5G massive MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1207–1218, Jun. 2014.
- [49] "Quarterly energy prices," Dept. Energy Climate Change (DECC), London, U.K., Sep. 2014.
- [50] B. Lee and S.-L. Kim, "Characterizing energy and deployment efficiency relations in cellular systems," in *Proc. ICSPCS*, Gold Coast, Australia, Dec. 2012, pp. 1–5.
- [51] G. Wang, Q. Liu, R. He, F. Gao, and C. Tellambura, "Acquisition of channel state information in heterogeneous cloud radio access networks: Challenges and research directions," *IEEE Wireless Commun. Mag.*, vol. 22, no. 3, pp. 100–107, Oct. 2015.



Piya Patcharamaneepakorn (M'14) received the B.Eng. degree in electrical engineering from Chulalongkorn University, Bangkok, Thailand, in 1999; the M.Eng. degree in telecommunications from the Asian Institute of Technology, Pathumthani, Thailand, in 2001; and the Ph.D. degree in electrical and electronics engineering from the University of Bristol, Bristol, U.K., in 2014.

From 2002 to 2009, he was an RF Optimization Engineer with several companies, including TA-Orange Thailand, T-Mobile Netherlands, and Nokia Siemens Network Argentina. Between 2014 and 2015, he was a Postdoctoral Research Associate with Heriot-Watt University, Edinburgh, U.K. He is currently an engineering expert with TRUE Corporation, Bangkok. His research interests include fifth-generation communication systems, multiuser multiple-input–multiple-output (MIMO) precoding designs, and massive multiple-input–multiple-output systems.



Shangbin Wu received the B.S. degree in communication engineering from South China Normal University, Guangzhou, China, in 2009; the M.Sc. degree in wireless communications with distinction from University of Southampton, Southampton, U.K., in 2010; and the Ph.D. degree in electrical engineering from Heriot-Watt University, Edinburgh, U.K., in 2015.

From 2010 to 2011, he was an LTE R&D engineer responsible for LTE standardization and system-level simulation in New Postcom Equipment Ltd., Guangzhou. From October 2011 to August 2012, he was with Nokia Siemens Network, where he was an LTE algorithm specialist, mainly focusing on LTE radio resource management algorithm design and system level simulations. Since November 2015, Shangbin has been a 5G research engineer with the Samsung R&D Institute, U.K. His research interests include 5G technologies, wireless channel modeling and simulation, and massive multiple-input multiple-output systems.



Cheng-Xiang Wang (S'01–M'0–SM'08) received the B.Sc. and M.Eng. degrees in communication and information systems from Shandong University, Jinan, China, in 1997 and 2000, respectively, and the Ph.D. degree in wireless communications from Aalborg University, Aalborg, Denmark, in 2004.

He has been with Heriot-Watt University, Edinburgh, U.K., since 2005 and was promoted to Professor in 2011. He was a Research Fellow with the University of Agder, Grimstad, Norway, during 2001–2005; a Visiting Researcher with Siemens AG-Mobile Phones, Munich, Germany, in 2004; and a Research Assistant with the Technical University of Hamburg-Hamburg, Germany, during 2000–2001. He has edited one book and published one book chapter and over 240 papers in refereed journals and conference proceedings. His current research interests include wireless channel modeling and simulation, green communications, cognitive radio networks, vehicular communication networks, large multiple-input multiple-output (MIMO), cooperative MIMO, and fifth-generation wireless communications.

Dr. Wang served or is currently serving as an Editor for nine international journals, including the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY (2011–), the IEEE TRANSACTIONS ON COMMUNICATIONS (2015–), and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (2007–2009). He was the leading Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS Special Issue on Vehicular Communications and Networks. He served or is serving as a Technical Program Committee (TPC) Member, a TPC Chair, and a General Chair for over 80 international conferences. He received the Best Paper Award at the 2010 IEEE Global Telecommunications Conference, the 2011 IEEE International Conference on Communication Technology, the 2012 International Conference on Intelligent Transportation Systems Telecommunications, the 2013 IEEE Vehicular Technology Conference (Fall), and the 2015 International Wireless Communications and Mobile Computing Conference. He is a Fellow of the Institution of Engineering and Technology, a Fellow of the Higher Education Academy, and a member of the Engineering and Physical Sciences Research Council Peer Review College.



el-Hadi M. Aggoune (M'83–SM'93) received the M.S. and Ph.D. degrees in electrical engineering from the University of Washington (UW), Seattle, WA, USA.

His work is referred to in many patents, including patents assigned to ABB, Switzerland and EPRI, USA. Currently, he is a Professor and Director of the Sensor Networks and Cellular Systems (SNCS) Research Center, University of Tabuk, Tabuk, Saudi Arabia. His research interests include wireless sensor networks, power system operation and control, simulation and visualization, and neurocomputing and applications. The highest academic ranks he has achieved include Endowed Chair Professor and Vice President and Provost. He has authored papers in IEEE and other journals and conferences.

Dr. Aggoune has served at a number of universities in the USA and abroad. He managed an unmanned aerial vehicle research and development lab that won the Boeing Supplier Excellence Award. He was a winner of the IEEE Professor of the Year Award, UW Branch. He is listed as Inventor in a major patent assigned to the Boeing Company. He is serving on many technical committees and is a Professional Engineer.



Mohammed M. Alwakeel (SM'14) was born in Tabuk, Saudi Arabia. He received the B.S. and M.S. degrees from King Saud University, Riyadh, Saudi Arabia, in 1992 and 1998 respectively, and the Ph.D. degree in electrical engineering from Florida Atlantic University, Boca Raton, FL, USA, in 2005.

He served as Communications Network Manager with the Saudi National Information Center in Riyadh. He also served as faculty member at King Abdulaziz University and then as an Associate Professor and Dean of the Computers and Information Technology College with the University of Tabuk. Currently, he is the Vice Rector for Development and Quality with the University of Tabuk. His current research interests include teletraffic analysis, mobile satellite communications, and sensor networks and cellular systems.



Xiaohu Ge (M'09–SM'11) received the Ph.D. degree in communication and information engineering from Huazhong University of Science and Technology (HUST), Wuhan, China, in 2003.

He is currently a Full Professor with the School of Electronic Information and Communications, HUST. He is an Adjunct Professor with the Faculty of Engineering and Information Technology, University of Technology Sydney, Ultimo, Australia. Since November 2005, he has been with HUST. Prior to that, he was a Researcher with Ajou University, Suwon,

Korea, and the Politecnico di Torino, Turin, Italy, from January 2004 to October 2005. From June to August 2010, he was a Visiting Researcher with Heriot-Watt University, Edinburgh, U.K. He has published more than 100 papers in refereed journals and conference proceedings and has been granted about 15 patents in China. His research interests include mobile communications, traffic modeling in wireless networks, green communications, and interference modeling in wireless communications.

Dr. Ge is a Senior Member of the China Institute of Communications and a member of the National Natural Science Foundation of China and the Chinese Ministry of Science and Technology Peer Review College. He has been actively involved in organizing more than ten international conferences since 2005. He served as the General Chair for the 2015 IEEE International Conference on Green Computing and Communications. He serves as an Associate Editor for the IEEE ACCESS, the *Wireless Communications and Mobile Computing Journal* (Wiley), and the *International Journal of Communication Systems* (Wiley). Moreover, he served as the Guest Editor for the IEEE COMMUNICATIONS MAGAZINE Special Issue on 5G Wireless Communication Systems. He received the Best Paper Award at the 2010 IEEE Global Telecommunications Conference. He is leading several projects funded by the National Natural Science Foundation of China, the China Ministry of Science and Technology, and industry. He is taking part in several international joint projects, such as the EU FP7-PEOPLE-IRSES: project acronym WINDOW (Grant 318992) and project acronym CROWN (Grant 610524).



Marco Di Renzo (S'05–AM'07–M'09–SM'14) received the Laurea (*cum laude*) and Ph.D. degrees in electrical engineering from the University of L'Aquila, L'Aquila, Italy, in 2003 and 2007, respectively, and the Habilitation à Diriger des Recherches from the Université Paris-Sud XI, Orsay, France, in 2013.

He has held various research and academic positions in Italy; in the United States with the Virginia Polytechnic Institute and State University, Blacksburg, VA; in Spain with the Centre Tecnològic de Telecomunicacions de Catalunya, Barcelona; and in the United Kingdom with The University of Edinburgh. Since 2010, he has been a Centre National de la Recherche Scientifique (CNRS) Associate Professor (“Chargé de Recherche Titulaire CNRS”) with the Laboratory of Signals and Systems, Paris-Saclay University-CNRS, Centrale Supélec, Université Paris Sud, France. He is a Distinguished Visiting Fellow with the Royal Academy of Engineering, U.K. He is a cofounder of the university spinoff company WEST Aquila s.r.l. Italy.

Dr. Di Renzo serves as an Editor for the IEEE COMMUNICATIONS LETTERS and the IEEE TRANSACTIONS ON COMMUNICATIONS, where he is the Editor for Heterogeneous Networks Modeling and Analysis of the IEEE Communications Society. He has received several awards, including the Best Paper Award at the 2012 and 2014 IEEE International Workshop on Computer-Aided Modeling, Analysis, and Design of Communication Links and Networks; the 2013 IEEE Vehicular Technology Conference (Fall); the 2014 IEEE International Conference on Advanced Technologies for Communications; and the 2015 IEEE International Conference on Computing, Management, and Telecommunications; the 2013 Network of Excellence NEWCOM Best Paper Award; the 2013 IEEE Communications Society Best Young Researcher Award for Europe, Middle East, and Africa (EMEA Region); the 2015 IEEE Jack Neubauer Memorial Best System Paper Award; and the 2015–2018 CNRS Award for Excellence in Research and in Advising Doctoral Students. He is a member of the European Association for Communications and Networking.