Uplink Power Control in Massive MIMO with Double Scattering Channels

Trinh Van Chien, *Member, IEEE*, Hien Quoc Ngo, *Senior Member, IEEE*, Symeon Chatzinotas, *Senior Member, IEEE*, Björn Ottersten, *Fellow, IEEE*, and Mérouane Debbah, *Fellow, IEEE*

Abstract—Massive multiple-input multiple-output (MIMO) is a key technology for improving the spectral and energy efficiency in 5G-and-beyond wireless networks. For a tractable analysis, most of the previous works on Massive MIMO have been focused on the system performance with complex Gaussian channel impulse responses under rich-scattering environments. In contrast, this paper investigates the uplink ergodic spectral efficiency (SE) of each user under the double scattering channel model. We derive a closed-form expression of the uplink ergodic SE by exploiting the maximum ratio (MR) combining technique based on imperfect channel state information. We further study the asymptotic SE behaviors as a function of the number of antennas at each base station (BS) and the number of scatterers available at each radio channel. We then formulate and solve a total energy optimization problem for the uplink data transmission that aims at simultaneously satisfying the required SEs from all the users with limited data power resource. Notably, our proposed algorithms can cope with the congestion issue appearing when at least one user is served by lower SE than requested. Numerical results illustrate the effectiveness of the closed-form ergodic SE over Monte-Carlo simulations. Besides, the system can still provide the required SEs to many users even under congestion.

Index Terms—Massive MIMO, double scattering channels, total transmit power minimization, congestion issue.

I. INTRODUCTION

Wireless communications has sustained an exponential demand growth in data throughput and reliability over the last decades [2], [3]. The cellular network topology with the assistance of MIMO technology has been evolved over time to indulge the growing demand. However, mobile traffic will increase as foreseen in a short time with 12.3 billion wireless access devices by 2022 [4]. To handle this issue, Massive MIMO, a disruptive technology with commercial deployments started in 2018 [5], not only inherits all the multiplexing gain and spatial diversity of the conventional MIMO but also offers extra degree-of-freedoms as a consequence of equipping base stations (BSs) with many antennas [6]. Massive MIMO, therefore, provides unprecedented spectral and energy efficiency gains of modern wireless networks with only utilizing the contemporary time and frequency resources. Each Massive MIMO BS only exploits a low-cost linear processing technique such as maximum ratio (MR) or zero-forcing (ZF) combining to detect the transmit signals and obtain performance closed to the optimum thanks to the benefits of the use of many more antennas than users [7]. In the uplink transmission, combining vectors for data detection are constructed from channel estimates, and therefore, the overhead is only made practically proportional to the number of users by sending pilot signals in the uplink.

In Massive MIMO, the closed-form expression of the ergodic SE can be obtained in certain scenarios. For rich scattering environments such that propagation channels ideally follow uncorrelated Rayleigh fading, the uplink and downlink SEs were obtained as a function of large-scale fading coefficients when each BS exploits MR or ZF combining as in [8], [9] and references therein. As such, many impacts such as array gains and channel estimation quality are explicitly observed in those ergodic rates, together with the power scaling laws are achieved. However, practical channels usually involve spatial correlation, which is modeled, for example utilizing correlated Rayleigh fading in the isotropic scattering environment where the gathered energy at an antenna array comes from many directions leading to the full ranks of covariance matrices with an overwhelming probability [7], [10], [11]. For rank deficiency occurring in poor scattering conditions, the Kronecker channel model is popularly used to describe the spatial correlations at the transmitter and receiver [12], [13]. The authors in [14] proposed the *double scattering channel* and demonstrated that the channel capacity is also characterized by the structure of scattering in the propagation environment instead of the spatial correlations around the transceiver only.

A few works have studied the effects of low-rank channels in Massive MIMO communications. For the keyhole channels (uncorrelated and rank-deficient), the channel hardening and favorable propagation were investigated in [15] to impress a significant reduction of the ergodic SE compared with that of uncorrelated Rayleigh fading. An extension of this work to

T. V. Chien, S. Chatzinotas, B. Ottersten are with the Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, L-1855 Luxembourg, Luxembourg (email: vanchien.trinh@uni.lu, symeon.chatzinotas@uni.lu, and bjorn.ottersten@uni.lu).

H. Q. Ngo is with the School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast BT7 1NN, United Kingdom (email: hien.ngo@qub.ac.uk).

M. Debbah is with the Technology Innovation Institute, Abu Dhabi, United Arab Emirates. He is also with CentraleSupélec, University Paris-Saclay, 91192 Gif-sur-Yvette, France. (email: merouane.debbah@tii.ae).

communications scenarios where users having multiple antennas has recently reported in [16]. The first work numerically studying the uplink ergodic SE of cellular Massive MIMO systems with the double scattering channels (spatially correlated and rank-deficient) was found in [17]. For theoretical analysis, the authors in [18], [19] computed the asymptotic ergodic SE of a single-cell Massive MIMO system with the different linear precoding techniques when the number of BS antennas, scatterers, and users grow large with the same rate. It is worth emphasizing that these works assumed each user utilizing an orthogonal pilot signal in a single-cell system and the formulations are asymptotically established. To the best of our knowledge, no prior work analyzes the performance of cellular Massive MIMO systems with a finite number of BS antennas, users, and scatterers, where the ergodic SE might have different features than the one at an asymptotic regime.

Many resource allocation tasks in Massive MIMO communications can be implemented on the large-scale fading time scale in place of the small-scale fading one by virtue of the channel hardening [20]. This makes resource allocation feasible to implement in practice. Various optimization problems with different utility functions have been formulated and solved in the Massive MIMO literature [21]–[23]. Notice that the key component of Massive MIMO communications is that it can allow many users to access and share the radio resource at the same time with high quality of service. The max-min fairness optimization is therefore promising to provide uniform service to all the users in the coverage area [24]. However, for large-scale networks with many base stations and users, the fairness level will approach a zero rate [25]. In contrast, one can include separate SE constraints in the optimization problems to simultaneously maintain the quality of service for all the users [26], [27]. However, since the users were randomly distributed, many user locations with poor channel conditions leads the optimization problems to be infeasible. The preliminary work in [28] has indicated that many users are still served by the required SEs if we can detect and relax the constraints of unsatisfied users when solving the problems and analyzing uncorrelated Rayleigh fading only.

By exploiting the double scattering channel model, this paper considers a Massive MIMO system in which a set of orthogonal pilot signals are reused by all the users such that the BSs can estimate channels in the pilot training phase. We then compute the uplink ergodic SE of each user in relation to the channel structure and propagation environment. The ergodic rate is then used to formulate and solve the total energy optimization problem for the uplink data transmission when each BS uses MR combining to detect the desired signals. Our main contributions are summarized as follows:

- A new ergodic SE expression is derived in closed form for a finite number of antennas at each BS while the number of scatterers observed by each user and BS is different from each other. This closed-form SE expression explicitly demonstrates the influence of pilot contamination, channel estimation errors, and limited scatterers. Conforming with the literature, we also analyze the asymptotic closed-form SE expression when the number of antennas and/or scatterers grows large. We analytically

testify the existence of a saturated point in most of the scenarios, but although the system still can offer an unbounded capacity under a certain condition.

- We formulate a total uplink data energy minimization problem subject to the required SE from every user and the power constraints. This problem may have an infeasible domain under the complication of simultaneously serving many users. For user locations and shadow fading realizations, where the our optimization problem is feasible, the global optimum can be obtained in polynomial time owning to its convexity.
- We propose two low computational complexity iterative algorithms that tackle the infeasible optimization problem by relaxing the SE constraints of unsatisfied users. At each iteration, the first algorithm allows users to transmit full data power whenever the required SE constraints are not satisfied. In contrast, the second algorithm gives a procedure to scale down data power assigned to users with the lower SEs than requested.
- Numerical results manifest that the closed-form SE expression overlaps Monte-Carlo simulations in all the system parameter settings. The effectiveness of the proposed data power control algorithms are compared with the interior-point methods. For given user locations and shadow fading realizations that form infeasible problems, the system still can provide satisfactory service to many users after relaxing one or a few the required SE constraints.

This paper is organized as follows: Section II presents the considered Massive MIMO system under the double scattering channels and derives the closed-form expression of the uplink SE for the case each BS utilizing MR combining to decode the transmitted signals. We also compute the asymptotic SE as different factors grow large. Section III formulates the total data energy minimization problem and characterizes its canonical form and feasible domain. The two algorithms to obtain a solution to this problem and handle the congestion issue are proposed in Section IV. Finally, Section V shows extensive numerical results and the main conclusions are drawn in Section VI.

Notation: Upper-case bold face letters are used to denote matrices and lower-case bold face ones for vectors. $\mathbf{I}_M$ is the identity matrix of size $M \times M$. The operation $\mathbb{E}\{\cdot\}$ and $\text{Var}\{\cdot\}$ denotes the expectation and variance of a random variable, respectively. The notation $\|\cdot\|$ is the Euclidean norm of a vector and $\|\cdot\|_2$ is the spectral norm of a matrix. Moreover, $\text{tr}(\cdot)$ is the trace of a matrix. The regular and Hermitian transposes are denoted by $(\cdot)^T$ and $(\cdot)^H$, respectively. Finally, $\mathcal{CN}(\cdot, \cdot)$ denotes the circularly symmetric complex Gaussian distribution.

II. MASSIVE MIMO SYSTEM WITH DOUBLE SCATTERING CHANNELS

We consider an uplink Massive MIMO system comprising L cells, where cell l has one BS equipped with M antennas and serving K single-antenna users. Even though the propagation channels change over time and frequency, we use a quasi-static channel model where the time-frequency plane is divided

into coherence blocks. Each coherence block comprises τ_c symbols such that the channel between an arbitrary user and the BS is static and frequency flat. This paper assumes that instantaneous channels are not known at the BSs. Therefore, in each coherence block, the τ_p symbols are dedicated to the pilot training phase and the remaining $\tau_c - \tau_p$ symbols are used for the uplink data transmission. The channel between user k in cell l and BS l' is modeled by the double scattering channel model [13], [17], which is[1]

$$\mathbf{h}_{lk}^{l'} = \sqrt{\frac{\beta_{lk}^{l'}}{S_{lk}^{l'}}} \left(\mathbf{R}_{lk}^{l'}\right)^{1/2} \mathbf{G}_{lk}^{l'} \left(\widetilde{\mathbf{R}}_{lk}^{l'}\right)^{1/2} \mathbf{g}_{lk}^{l'}, \qquad (1)$$

where $\beta_{lk}^{l'}$ is the large-scale fading coefficient, which models the effects of the pathloss due to long distance and shadow fading due to obstacles. The integer parameter $S_{lk}^{l'}$ is the number of scatters generating the channel between BS l' and user k in cell l. The matrix $\mathbf{R}_{lk}^{l'} \in \mathbb{C}^{M \times M}$ represents the correlation between the BS antennas and its scatterers; $\mathbf{G}_{lk}^{l'} \in \mathbb{C}^{M \times S_{lk}^{l'}}$ includes the small-scale fading coefficients between BS l' and its scattering cluster. The matrix $\widetilde{\mathbf{R}}_{lk}^{l'} \in \mathbb{C}^{S_{lk}^{l'} \times S_{lk}^{l'}}$ stands for the correlation between the transmit and receive scatterers and $\mathbf{g}_{lk}^{l'} \in \mathbb{C}^{S_{lk}^{l'}}$ represents the small-scale fading between the user and its scattering cluster. The elements of both $\mathbf{G}_{lk}^{l'}$ and $\mathbf{g}_{lk}^{l'}$ are independent and identically distributed as $\mathcal{CN}(0,1)$ by constraints on the trace of the covariance matrices.

Remark 1. *The double scattering channel model in* (1) *reflects three important aspects of Massive MIMO channel propagation: the rank deficiency at the transceiver, the spatial fading correlation, and the signal attenuation by controlling multiple factors such as the number of scatterers in the environment, the correlation matrices, and the large-scale fading coefficients. It is more an involved channel model than in previous non-line-of-sight models to describe the sensitivity of the actual channel capacity to both the fading correlation and scattering structure in real propagation environments [14], [17]. This model spans scenarios from uncorrelated Rayleigh to the single-keyhole channels. In practical systems, the covariance matrices can be estimated by averaging over many realizations of instantaneous channels, while the number of scatterers can be obtained by formulating and solving, for example, an ℓ_p−norm optimization problem, which matches the double scattering channel model with measurement data [29].*

The further interesting statistical information of the double scattering channels, which is later useful for computing the uplink ergodic SE expression in a closed form, is presented in the following lemma.

Lemma 1. *Let us consider the two random channel vectors $\mathbf{h}_{l'k'}^l$ and $\mathbf{h}_{l''k''}^l$ generated by the double scattering channel*

model and a deterministic matrix $\mathbf{B} \in \mathbb{C}^{M \times M}$. If $(l', k') \neq (l'', k'')$, it holds that

$$\mathbb{E}\left\{\left|\left(\mathbf{h}_{l'k'}^l\right)^{\mathrm{H}} \mathbf{B} \mathbf{h}_{l''k''}^l\right|^2\right\} = b_{l'k'}^l \mathrm{tr}\left(\mathbf{B} \mathbf{R}_{l''k''}^l \mathbf{B}^{\mathrm{H}} \mathbf{R}_{l'k'}^l\right), \qquad (2)$$

with $d_{l'k'}^l = \mathrm{tr}(\widetilde{\mathbf{R}}_{l'k'}^l)/S_{l'k'}^l$ and $b_{l'k'}^l = \beta_{l'k'}^l d_{l'k'}^l \beta_{l''k''}^l d_{l''k''}^l$. Moreover, for the channel $\mathbf{h}_{l'k'}^l$, it holds that

$$\mathbb{E}\left\{\left|\left(\mathbf{h}_{l'k'}^l\right)^{\mathrm{H}} \mathbf{B} \mathbf{h}_{l'k'}^l\right|^2\right\} = \left(\beta_{l'k'}^l\right)^2 \left(\left(d_{l'k'}^l\right)^2 + \frac{\mathrm{tr}\left(\left(\widetilde{\mathbf{R}}_{l'k'}^l\right)^2\right)}{\left(S_{l'k'}^l\right)^2}\right) \times$$
$$\left(\left|\mathrm{tr}(\mathbf{R}_{l'k'}^l \mathbf{B})\right|^2 + \mathrm{tr}\left(\mathbf{R}_{l'k'}^l \mathbf{B} \mathbf{R}_{l'k'}^l \mathbf{B}^{\mathrm{H}}\right)\right). \quad (3)$$

Proof. The proof is to compute the moments of non-Gaussian random variables and available in Appendix A. □

In (2), the second moment obtained for the inner product of two different channel vectors is a deterministic value, which depends on their covariance matrices and scales up with the number of antennas installed at BS l, say M. Meanwhile, the weighted forth moment in (3) indicates a scaling factor of M^2. This moment is also inversely proportional to the number of scatterers. The moments of channels in Lemma 1 are utilized to compute the closed-form expression on the uplink ergodic rate of an arbitrary user.

A. Uplink Pilot Training

In each coherence block, each BS needs instantaneous channel state information for the uplink data detection. The τ_p symbols are dedicated to the uplink pilot training, which can create τ_p mutually orthogonal pilot signals. User k in cell l uses the deterministic pilot signal $\boldsymbol{\phi}_{lk} \in \mathbb{C}^{\tau_p}$ with $\|\boldsymbol{\phi}_{lk}\|^2 = \tau_p$. This pilot signal is also reused by other users in multiple cells and we can define the pilot reuse set as

$$\mathcal{P}_{lk} = \left\{(l', k') : \boldsymbol{\phi}_{l'k'} = \boldsymbol{\phi}_{lk}, l = 1, \ldots, L, k' = 1, \ldots, K\right\}, \quad (4)$$

which contains the indices of all users sharing the same pilot signal as user k in cell l, including (l, k). Mathematically, it observes that

$$\boldsymbol{\phi}_{lk}^{\mathrm{H}} \boldsymbol{\phi}_{l'k'} = \begin{cases} \tau_p, & \text{if } (l', k') \in \mathcal{P}_{lk}, \\ 0, & \text{if } (l', k') \notin \mathcal{P}_{lk}. \end{cases} \quad (5)$$

At BS l, the received pilot signal $\mathbf{Y}_l^p \in \mathbb{C}^{M \times \tau_p}$ with the superscript p standing for the pilot training phase is formulated as

$$\mathbf{Y}_l^p = \sum_{l'=1}^{L} \sum_{k'=1}^{K} \sqrt{\hat{p}_{l'k'}} \mathbf{h}_{l'k'}^l \boldsymbol{\phi}_{l'k'}^{\mathrm{H}} + \mathbf{N}_l^p, \qquad (6)$$

where $\mathbf{N}_l^p \in \mathbb{C}^{M \times \tau_p}$ is additive noise with the independent and identically random elements distributed as $\mathcal{CN}(0, \sigma^2)$. BS l estimates the channel $\mathbf{h}_{l'k'}^l$ from user k' in cell l' by multiplying $\mathbf{Y}_l^p$ with the pilot sequence $\boldsymbol{\phi}_{l'k'}$ as

$$\mathbf{y}_{l'k'}^{l,p} = \mathbf{Y}_l^p \boldsymbol{\phi}_{l'k'} = \sum_{(l'',k'') \in \mathcal{P}_{l'k'}} \sqrt{\hat{p}_{l''k''}} \tau_p \mathbf{h}_{l''k''}^l + \mathbf{N}_l^p \boldsymbol{\phi}_{l'k'}. \quad (7)$$

The minimum mean square error (MMSE) is not straightforward to apply to (7) because of the non Gaussian distributions.

---

[1]This outdoor channel model was initiated for conventional MIMO systems under a far-field region and dedicated sub 6-GHz bands for mobile services. In cellular Massive MIMO communications, the far-field effects are still observed since many antenna components can be practically installed in a small compact array [5].

Nonetheless, the processed received signal $\mathbf{y}_{l'k'}^{l,p} \in \mathbb{C}^M$ has sufficient statistics to obtain a channel estimate of the origin $\mathbf{h}_{l'k'}^l$ by utilizing linear MMSE (LMMSE). We now consider the channel estimates under assumptions of statistical channel knowledge available at each BS.

Lemma 2. *By utilizing the LMMSE estimation, the channel estimate $\hat{\mathbf{h}}_{l'k'}^l \in \mathbb{C}^M$ from user k' in cell l' and BS l is*

$$\hat{\mathbf{h}}_{l'k'}^l = \sqrt{\hat{p}_{l'k'}}\beta_{l'k'}^l d_{l'k'}^l \mathbf{R}_{l'k'}^l \mathbf{\Psi}_{l'k'}^l \mathbf{y}_{l'k'}^{l,p}, \qquad (8)$$

where $\mathbf{\Psi}_{l'k'}^l \in \mathbb{C}^{M \times M}$ is

$$\mathbf{\Psi}_{l'k'}^l = \left(\sum_{(l'',k'') \in \mathcal{P}_{l'k'}} a_{l''k''}^l \mathbf{R}_{l''k''}^l + \sigma^2 \mathbf{I}_M. \right)^{-1}. \qquad (9)$$

with $a_{l''k''}^l = \tau_p \hat{p}_{l''k''}\beta_{l''k''}^l d_{l''k''}^l$. The covariance matrix of the channel estimate $\hat{\mathbf{h}}_{l'k'}^l$ is computed as

$$\mathbb{E}\left\{ \hat{\mathbf{h}}_{l'k'}^l (\hat{\mathbf{h}}_{l'k'}^l)^{\mathrm{H}} \right\} = \hat{p}_{l'k'} (\beta_{l'k'}^l)^2 (d_{l'k'}^l)^2 \tau_p \mathbf{R}_{l'k'}^l \mathbf{\Psi}_{l'k'}^l \mathbf{R}_{l'k'}^l. \qquad (10)$$

Proof. The proof is based on the LMMSE estimation of non-Gaussian random variables [30], but adapted to our framework with the channel vector in (1) and the pilot reuse in (4). The detail proof is available in Appendix B. □

Lemma 2 shows the concrete expression of the channel estimate of each user together with the statistical information, which are used to formulate the combining vectors and computing the closed-form expression on the uplink ergodic SE hereafter. It should be noticed that our channel estimation considers the influence of coherent interference caused by the pilot contamination in multi-cell Massive MIMO scenarios, which is a generalization of the previous result in [18], [19] that assumed the orthogonal pilot signals for all the users in a single cell. Along with the statistical information in Lemma 1, the channel estimates and estimation errors in Lemma 2 are utilized to compute the closed-form uplink SE expression hereafter.

B. Uplink Data Transmission

During the uplink data transmission, user k in cell l sends a data symbol s_{lk} with $\mathbb{E}\{|s_{lk}|^2\} = 1$ and the received data signal $\mathbf{y}_l \in \mathbb{C}^M$ at BS l is a superposition of all the transmitted signals from all the users as

$$\mathbf{y}_l = \sum_{l'=1}^L \sum_{k'=1}^K \sqrt{p_{l'k'}} \mathbf{h}_{l'k'}^l s_{l'k'} + \mathbf{n}_l, \qquad (11)$$

where $p_{l'k'}$ is the transmit power of user k' in cell l' assigned to each data symbol and $\mathbf{n}_l$ is additive noise distributed as $\mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_M)$. By utilizing a combining vector $\mathbf{v}_{lk} \in \mathbb{C}^M$

based on the channel estimates, BS l decodes the desired signal from user k in cell l as

$$\mathbf{v}_{lk}^H \mathbf{y}_l = \sqrt{p_{lk}}\mathbb{E}\{\mathbf{v}_{lk}^H \mathbf{h}_{lk}^l\} s_{lk} + \sqrt{p_{lk}}\left(\mathbf{v}_{lk}^H \mathbf{h}_{lk}^l - \mathbb{E}\{\mathbf{v}_{lk}^H \mathbf{h}_{lk}^l\} \right) s_{lk}$$

$$+ \sum_{k'=1,k'\neq k}^K \mathbf{v}_{lk}^H \sqrt{p_{lk'}} \mathbf{h}_{lk'}^l s_{lk'}$$

$$+ \sum_{l'=1,l'\neq l}^L \sum_{k'=1}^K \mathbf{v}_{lk}^H \sqrt{p_{l'k'}} \mathbf{h}_{l'k'}^l s_{l'k'} + \mathbf{v}_{lk}^H \mathbf{n}_l, \quad (12)$$

where the first term contains the desired signal by virtue of the channel hardening [31]. The second term describes the beamforming uncertainty effects, while the remaining terms are mutual interference and noise. As shown in [20], the uplink ergodic SE is obtained by the use-and-then-forget channel capacity bounding technique as

$$R_{lk} = \left(1 - \frac{\tau_p}{\tau_c} \right) \log_2 \left(1 + \mathrm{SINR}_{lk} \right), \, [\mathrm{b/s/Hz}], \qquad (13)$$

where the effective signal-to-interference-and-noise ratio (SINR) value is computed as in (14). The expectations in (14) are taking over all the sources of randomness and (13) is an achievable rate since it is a lower bound on the channel capacity. Furthermore, this achievable rate can be computed numerically for any combining scheme. The main demerit of (13) is high computational complexity since many instantaneous channels need to be gathered such that several expectations can be numerically estimated.

C. Uplink Spectral Efficiency Analysis

If MR combining is used by each BS, i.e., $(\mathbf{v}_{lk} = \hat{\mathbf{h}}_{lk}^l), \forall l, k$, we obtain the closed-form expression for the uplink SE in (13) as shown by Theorem 1.[2]

Theorem 1. *When BS l uses the MR combing vector to decode the desired signal from user k in cell l, the achievable uplink SE obtained in (13) with the closed-form expression of the SINR value computed as*

$$\mathrm{SINR}_{lk} = \frac{p_{lk} z_{lk}^l \left| \mathrm{tr}\left(\mathbf{R}_{lk}^l \mathbf{\Psi}_{lk}^l \mathbf{R}_{lk}^l \right) \right|^2}{\mathrm{NI}_{lk} + \mathrm{CI}_{lk} + \mathrm{NO}_{lk}}, \qquad (15)$$

where NI_{lk}, CI_{lk}, and NO_{lk} are respectively the non-coherent interference, coherent interference, and noise, which are com-

---

[2]The framework in this paper can be easily extended to the downlink data transmission.

$$\text{SINR}_{lk} = \frac{p_{lk}\left|\mathbb{E}\{\mathbf{v}_{lk}^H\mathbf{h}_{lk}^l\}\right|^2}{\sum_{l'=1}^{L}\sum_{k'=1}^{K} p_{l'k'}\mathbb{E}\left\{\left|\mathbf{v}_{lk}^H\mathbf{h}_{l'k'}^l\right|^2\right\} - p_{lk}\left|\mathbb{E}\{\mathbf{v}_{lk}^H\mathbf{h}_{lk}^l\}\right|^2 + \sigma^2\mathbb{E}\{\|\mathbf{v}_{lk}\|^2\}}. \tag{14}$$

puted in the closed-form expression as

$$\text{NI}_{lk} = \sum_{l'=1}^{L}\sum_{k'=1}^{K} p_{l'k'} m_{l'k'}^l \text{tr}(\mathbf{R}_{lk}^l\boldsymbol{\Psi}_{lk}^l\mathbf{R}_{lk}^l\mathbf{R}_{l'k'}^l), \tag{16}$$

$$\text{CI}_{lk} = \sum_{(l',k')\in\mathcal{P}_{lk}\backslash(l,k)} p_{l'k'} z_{l'k'}^l \left|\text{tr}(\mathbf{R}_{l'k'}^l\boldsymbol{\Psi}_{lk}^l\mathbf{R}_{lk}^l)\right|^2 +$$

$$\sum_{(l',k')\in\mathcal{P}_{lk}} p_{l'k'}\frac{z_{l'k'}^l\text{tr}\left((\widetilde{\mathbf{R}}_{l'k'}^l)^2\right)}{(d_{l'k'}^l S_{l'k'}^l)^2}\left|\text{tr}(\mathbf{R}_{l'k'}^l\boldsymbol{\Psi}_{lk}^l\mathbf{R}_{lk}^l)\right|^2 +$$

$$\sum_{(l',k')\in\mathcal{P}_{lk}} p_{l'k'}z_{l'k'}^l\frac{\text{tr}\left((\widetilde{\mathbf{R}}_{l'k'}^l)^2\right)}{(S_{l'k'}^l)^2}\text{tr}\left(\mathbf{R}_{l'k'}^l\boldsymbol{\Psi}_{lk}^l\mathbf{R}_{lk}^l\mathbf{R}_{l'k'}^l\mathbf{R}_{lk}^l\boldsymbol{\Psi}_{lk}^l\right), \tag{17}$$

$$\text{NO}_{lk} = \sigma^2\hat{p}_{lk}(\beta_{lk}^l)^2(d_{lk}^l)^2\tau_p\text{tr}(\mathbf{R}_{lk}^l\boldsymbol{\Psi}_{lk}^l\mathbf{R}_{lk}^l), \tag{18}$$

with the values $m_{l'k'}^l$ and $z_{l'k'}^l, \forall l', k', l$, defined as

$$m_{l'k'}^l = \beta_{l'k'}^l d_{l'k'}^l \hat{p}_{lk}(\beta_{lk}^l)^2(d_{lk}^l)^2\tau_p, \tag{19}$$

$$z_{l'k'}^l = \hat{p}_{l'k'}(\beta_{l'k'}^l)^2(d_{l'k'}^l)^2\hat{p}_{lk}(\beta_{lk}^l)^2(d_{lk}^l)^2\tau_p^2. \tag{20}$$

Proof. The proof is obtained by computing the expectations of non-Gaussian random variables in (14). The detailed proof is available in Appendix C. □

The SINR expression (15) is explicitly influenced by many factors such as channel covariance matrices, the number of scatters, pilot reuse, channel estimation quality, which are hidden in the general formulation (14). Specifically, the numerator of (15) shows the contribution of both channel estimation quality and covariance matrix of user k in cell l. Moreover, the effectiveness of the array gain is verified since the numerator scales up with the number of antennas thanks to the spatial covariance property in (23). The first part in the denominator of (15) demonstrates the degradation of the received signal quality due to non-coherent interference. The second part presents the contributions of coherent interference caused by reusing the pilot signals among the users that is defined by the pilot reuse set $\mathcal{P}_{lk}$. Unlike previous works with many scatterers [10], this part also shows that a small number of scatterers have significant contributions to increasing non-coherent interference. If the coherent blocks are large enough such that pilot sequences allocated to all users are pairwisely orthogonal, i.e., $\tau_p \geq LK$, the SINR value of user k in cell l is still computed as (15), but the following parameters are

reformulated as

$$\boldsymbol{\Psi}_{lk}^l = \left(a_{lk}^l\mathbf{R}_{lk}^l + \sigma^2\mathbf{I}_M\right)^{-1}, \tag{21}$$

$$\text{CI}_{lk} = p_{lk}\frac{z_{lk}^l\text{tr}\left((\widetilde{\mathbf{R}}_{lk}^l)^2\right)}{(d_{lk}^l S_{lk}^l)^2}\left|\text{tr}(\mathbf{R}_{lk}^l\boldsymbol{\Psi}_{lk}^l\mathbf{R}_{lk}^l)\right|^2 +$$

$$p_{lk}z_{lk}^l\frac{\text{tr}\left((\widetilde{\mathbf{R}}_{lk}^l)^2\right)}{(S_{lk}^l)^2}\text{tr}\left(\mathbf{R}_{lk}^l\boldsymbol{\Psi}_{lk}^l\mathbf{R}_{lk}^l\mathbf{R}_{lk}^l\mathbf{R}_{lk}^l\boldsymbol{\Psi}_{lk}^l\right), \tag{22}$$

which demonstrates the influences of a finite scatterer number to the uplink SE. Finally, the last part in the denominator of (15) presents the additive noise effects.

Remark 2. *We consider the MR combining technique due to its low computational complexity. This linear combining technique allows the execution of SE analysis in the closed form with a finite set of BS antennas, users, and scatterers. In addition, it can be implemented by only using the local channel state information, and therefore, easy to implement in a distributed manner.*

D. Asymptotic Analysis

In order to observe the uplink SE at an asymptotic regime and also compare with previous works, we now investigate the uplink asymptotic SE of each user when $M \to \infty$ and $S_{l'k'}^l \to \infty, \forall l, l', k'$. Aligned with previous works [32], the general preliminary settings on the covariance matrices are given in Assumption 1.

Assumption 1. *For $l, l' = 1, \ldots, L$ and $k' = 1, \ldots, K$, the spatial covariance matrices $\mathbf{R}_{l'k'}^l$ and $\widetilde{\mathbf{R}}_{l'k'}^l$ satisfy*

$$\limsup_M \left\|\mathbf{R}_{l'k'}^l\right\|_2 < \infty, \quad \liminf_M \frac{\text{tr}(\mathbf{R}_{l'k'}^l)}{M} > 0, \tag{23}$$

$$\limsup_{S_{l'k'}^l} \left\|\widetilde{\mathbf{R}}_{l'k'}^l\right\|_2 < \infty, \quad \liminf_{S_{l'k'}^l} \frac{\text{tr}(\widetilde{\mathbf{R}}_{l'k'}^l)}{S_{l'k'}^l} > 0. \tag{24}$$

Assumption 1 is established based on the fact that a double scattering channel has two covariance matrices on the definition. This assumption is extended from the standard form in the asymptotic analysis for Massive MIMO communications with a single covariance matrix [7]. Physically, the gathered signal energy at a BS originates from many spatial directions and is proportional to the number of antennas. We also utilize the spatial orthogonality between two covariance matrices to seek for a convergence point at the asymptotic regime as shown in Definition 1.

Definition 1. *The two covariance matrices $\mathbf{R}_{l'k'}^l$ and $\mathbf{R}_{lk}^l, \forall l', l, k, k'$ are asymptotically spatially orthogonal if*

$$\frac{1}{M}\text{tr}\left(\mathbf{R}_{l'k'}^l\mathbf{R}_{lk}^l\right) \to 0, M \to \infty. \tag{25}$$

As pointed out in previous works [33], [34], the condition (25) indicates the two users having orthogonal correlation eigenspaces. This holds for a network where each BS is equipped with antennas in a uniform linear array and the supports of the multi-path angular distributions of the two users are strictly non-overlapping. The convergence of the uplink SE for each user is stated in Theorem 2.

Theorem 2. *Under Assumption 1, the uplink SE of user k in cell l can be asymptotically observed by the following cases:*

a) *As $M \to \infty$ and a given set of finite scatterers, the achievable rate of user k in cell l converges to*

$$R_{lk} = \left(1 - \frac{\tau_p}{\tau_c}\right) \times$$
$$\log_2\left(1 + \frac{p_{lk} z_{lk}^l \left|\mathrm{tr}\left(\mathbf{R}_{lk}^l \mathbf{\Psi}_{lk}^l \mathbf{R}_{lk}^l\right)\right|^2}{\mathsf{CI}_{lk}}\right), \, [b/s/Hz]. \quad (26)$$

b) *As $M \to \infty$, a limited number of scatterers, and the two covariance matrices $\mathbf{R}_{l'k'}^l$ and $\mathbf{R}_{lk}^l$ are asymptotically orthogonal for all $(l', k') \in \mathcal{P}_{lk} \setminus (l, k)$, the achievable rate of user k in cell l converges to*

$$R_{lk} = \left(1 - \frac{\tau_p}{\tau_c}\right) \log_2\left(1 + \frac{\left(d_{lk}^l S_{lk}^l\right)^2}{\mathrm{tr}\left(\left(\widetilde{\mathbf{R}}_{l'k'}^l\right)^2\right)}\right), \, [b/s/Hz]. \quad (27)$$

c) *As $M \to \infty$ and $S_{l'k'}^l \to \infty, \forall l', k' \in \mathcal{P}_{lk}$, the achievable rate of user k in cell l converges to*

$$R_{lk} = \left(1 - \frac{\tau_p}{\tau_c}\right) \times$$
$$\log_2\left(1 + \frac{p_{lk} z_{lk}^l \left|\mathrm{tr}\left(\mathbf{R}_{lk}^l \mathbf{\Psi}_{lk}^l \mathbf{R}_{lk}^l\right)\right|^2}{\mathsf{NI}_{lk}}\right), \, [b/s/Hz], \quad (28)$$

where $\mathsf{NI}_{lk} = \sum_{(l',k') \in \mathcal{P}_{lk} \setminus (l,k)} p_{l'k'} z_{l'k'}^l \left|\mathrm{tr}\left(\mathbf{R}_{l'k'}^l \mathbf{\Psi}_{lk}^l \mathbf{R}_{lk}^l\right)\right|^2$.

d) *As $M \to \infty$, $S_{l'k'}^l \to \infty, \forall l', k' \in \mathcal{P}_{lk}$, and the two covariance matrices $\mathbf{R}_{l'k'}^l$ and $\mathbf{R}_{lk}^l$ are asymptotically orthogonal for all $(l', k') \in \mathcal{P}_{lk} \setminus (l, k)$, the achievable rate of user k in cell l grows without bound as*

$$R_{lk} \to \infty, \, [b/s/Hz]. \quad (29)$$

Proof. The proof is to compute the asymptotic SE of each user in the network with Assumption 1 and Definition 1 when the number of antennas and/or scatterers increases. The detailed proof is available in Appendix D. $\square$

Theorem 2 reveals that the uplink SE at an asymptotic regime is dependent on both the number of antennas at each BS and scatterers in propagation environments as well. For a limited number of scatterers at each communication link, the uplink SE of user k in cell l is bounded when the number of antennas increases due to the pilot contamination effects. Different from [35], the SE converges to a finite point as shown (27) even when the asymptotically orthogonality among covariance matrices holds because of lacking the scatterers.

For a rich scattering environment, the limitation is mainly from reusing the pilot signals among users causing coherent interference, which is dominant at an asymptotic regime. The fundamental difference of the double scattering channels compared with other spatial fading models as correlated Rayleigh fading or local scattering fading is that the unbounded channel capacity is obtained when the covariance matrices are asymptotically orthogonal as well as both numbers of antennas at each BS and scatterers go asymptotically.

III. UPLINK TOTAL DATA ENERGY CONSUMPTION MINIMIZATION

This section expresses an uplink energy consumption minimization problem by assuming that user k in cell l requests a SE $\xi_{lk} > 0, \forall l, k$, and has a maximum power $P_{\max,lk} > 0$. Investigating this optimization problem, we further manifest the feasibility for user locations, where all the users are served with the requested SE under the limited power budget. In contrast, the infeasibility is manifested for certain user locations, where users may be served with the SE lower than what has been requested.

A. Problem Formulation

The main goal of 5G-and-beyond systems is to provide the high SEs to all users with a minimal power consumption. In this paper, we formulate a total data energy optimization problem for the uplink data transmission as follows

$$\begin{aligned} \underset{\{p_{lk} \geq 0\}}{\text{minimize}} \quad & (\tau_c - \tau_p) \sum_{l=1}^{L} \sum_{k=1}^{K} p_{lk} \\ \text{subject to} \quad & R_{lk} \geq \xi_{lk}, \forall l, k, \\ & p_{lk} \leq P_{\max,lk}, \forall l, k, \end{aligned} \quad (30)$$

where $P_{\max,lk}$ is the maximum power level that user k in cell l can allocate to each data symbol. Problem (30) constrains on the rate requirement and limited power budget of each user. The per-user power constraints implicitly indicate that the total transmit power in the network should be upper bounded. In addition, the objective function of problem (30) ensures the minimal network power consumption. Therefore, our proposed optimization problem is able to reduce the mutual interference on other networks.

Remark 3. *Note that, in (30), we consider the per-user power constraints. It is also interesting to additionally consider a network power constraint so that the mutual interference on other networks can be controlled more effectively. For this case, the feasibility of our optimization problem is a main issue. We may first check if the network power constraint would be active in the selected point, i.e., if the network power constraint is satisfied under the optimized individual constraints. If it is inactive, the solution remains unaffected. If it is active, a heuristic approach would be to reduce the number of users, increase the number of antennas, or relax the per-user SE requirements. This potential extension is left for the future work. In this paper, we assume that the network power constraint is always satisfied and only handling a scenario that the per-user powers are constrained.*

By setting $v_{lk} = 2^{\xi_{lk}\tau_c/(\tau_c-\tau_p)} - 1$ and removing the constant $\tau_c - \tau_p$ in the objective function, problem (30) is converted from the SE constraints into the equivalent SINR constraints as

$$\begin{aligned}
&\underset{\{p_{lk} \geq 0\}}{\text{minimize}} && \sum_{l=1}^{L}\sum_{k=1}^{K} p_{lk} \\
&\text{subject to} && \text{SINR}_{lk} \geq v_{lk}, \forall l, k, \\
&&& p_{lk} \leq P_{\max,lk}, \forall l, k.
\end{aligned} \tag{31}$$

Instead of optimizing the energy consumption as (30), problem (31) minimizes the total transmit powers, which all users consume for the uplink data transmission. Due to the universe of all SINR expressions $\{\text{SINR}_{lk}\}$, problem (31) is in a general form for any combining technique. We now focus on MR combining technique as the corresponding SINRs have been derived in closed-form as obtained in Theorem 1. The concrete optimization problem is reformulated by utilizing the SINR expression (15) into (31) as

$$\begin{aligned}
&\underset{\{p_{lk} \geq 0\}}{\text{minimize}} && \sum_{l=1}^{L}\sum_{k=1}^{K} p_{lk} \\
&\text{subject to} && \frac{p_{lk} z_{lk}^{l} \left| \text{tr}\left(\mathbf{R}_{lk}^{l}\mathbf{\Psi}_{lk}^{l}\mathbf{R}_{lk}^{l}\right)\right|^2}{\text{NI}_{lk} + \text{CI}_{lk} + \text{NO}_{lk}} \geq v_{lk}, \forall l, k, \\
&&& p_{lk} \leq P_{\max,lk}, \forall l, k.
\end{aligned} \tag{32}$$

We stress that problem (32) jointly optimizes the powers to satisfy the requested SINRs from all the users. The required SINR levels $v_{lk}, \forall l, k$, are distinct from each other in practice and the global optimum is only found when all the users are simultaneously served by the required SEs. This problem can be either feasible or infeasible for a given set of user locations and shadow fading realizations as presented hereafter.[3]

B. Feasible and Infeasible Problems

When problem (32) has a non-empty feasible set meaning that the network is able to simultaneously provide the required SEs to all the users conditioned on the power constraints. We can find the global optimal solution to problem (32). Indeed, the objective function is a linear combination of all the power variables $\{p_{lk}\}, \forall l, k$. In addition, the power budget constraint functions are affine while the SINR constraints, $\forall l, k$, are reformulated as

$$v_{lk}\text{NI}_{lk} + v_{lk}\text{CI}_{lk} + v_{lk}\text{NO}_{lk} \leq p_{lk}z_{lk}^{l}\left|\text{tr}\left(\mathbf{R}_{lk}^{l}\mathbf{\Psi}_{lk}^{l}\mathbf{R}_{lk}^{l}\right)\right|^2, \tag{33}$$

which are also affine functions. Consequently, (32) is a linear program on standard form [36]. We hence enable to solve (32) to the global optimality in polynomial time, for instance, utilizing a general interior-point optimization toolbox as CVX [37]. Problem (32) includes the KL optimization variables and the $2KL$ constraints and as such it has the computational complexity of the order $O\left(N_i 2K^3 L^3\right)$, where N_i is the number of

---

[3]The congestion issue may appear in the other optimization problems as the spectral or energy efficiency maximization subject to the SE requirements and/nor the limited power budget constraints. The key argument of our framework is to point out that many users might still be served with their SE requirements in Massive MIMO communications if there is a strategic policy to deal with a few unsatisfied users.

Newton iterations needed to obtain a predetermined precision, typically in the order of tens [36, Chapter 11]. It should be noticed that all the KL users will spend non-zero data powers at the global optimum when problem (32) is feasible owning to the non-zero SE requirements.

For a specific realization of user locations and the power budgets, there may be a situation that all the users cannot be simultaneously served by the SE requirements. We emphasize that only one unfortunate user served with a lower SE suffices to create an empty feasible domain for the total transmit power optimization problem. Alternatively, problem (32) lacks a feasible solution [36, Section 4.1]. The unsatisfied SE is caused by high mutual interference in cellular networks and/or extreme locations as the cell edge leading to some users having a weak channel. Moreover, a user may require a too high SE for which the system cannot provide this service even spending maximum data power. Fortunately, a feasible solution of the data powers might still exist for most of the users with the required SEs, while only one or a few users are *unsatisfied*. Consequently, it may be sufficient to remove or reduce the required SEs of those unsatisfied users to convert an infeasible problem to a feasible one. However, it is not trivial to identify which users are unsatisfied to completely remove during solving problem (32). As one of the main contributions, this paper develops the power allocation strategies to handle such infeasible instances by allowing the corresponding SINR constraints to be violated.

IV. CONGESTION SOLUTION BASED ON ALTERNATING OPTIMIZATION

This section proposes the two algorithms attaining a fixed-point solution to problem (32) with either empty or non-empty feasible set. When the feasible set is empty, the SINR constraints of users, which potentially make the congestion issue are relaxed: The first approach is spending the maximum power on unsatisfied users. In contrast, the second approach is reducing the data power of those unsatisfied users. We now introduce important notations which will be widely utilized in this paper to construct the proposed algorithms as shown in Definition 2.

Definition 2. *Let us denote* $\mathbf{z}$ *and* $\mathbf{z}'$ *the real vectors of size* $KL \times 1$, *for which the* n-*th elements are* z_n *and* z_n', *respectively. The notation* $\mathbf{z} \geq \mathbf{z}'$ *indicates element-wise inequality* $z_n \geq z_n', \forall n = 1, \ldots, KL$. *Meanwhile, the notation* $\mathbf{z} \leq \mathbf{z}'$ *indicates* $z_n \leq z_n', \forall n = 1, \ldots, KL$.

A. Spending Maximum Transmit Power on Unsatisfied Users

For the glorification of simplification in comprehension, problem (32) with a non-empty feasible domain is first considered. We stack all the data powers into a vector $\mathbf{p} = [p_{11}, \ldots, p_{LK}]^T \in \mathbb{R}_+^{LK}$, then the SINR constraint of user k in cell l is reformulated as

$$p_{lk} \geq I_{lk}(\mathbf{p}), \tag{34}$$

where $I_{lk}(\mathbf{p})$ is so-called a standard interference function, which is given by

$$I_{lk}(\mathbf{p}) = \frac{\nu_{lk}\mathsf{NI}_{lk}(\mathbf{p}) + \nu_{lk}\mathsf{CI}_{lk}(\mathbf{p}) + \nu_{lk}\mathsf{NO}_{lk}}{z_{lk}^l \left| \mathrm{tr}\left(\mathbf{R}_{lk}^l \boldsymbol{\Psi}_{lk}^l \mathbf{R}_{lk}^l\right)\right|^2}. \quad (35)$$

In (35), the detailed expressions of $\mathsf{NI}_{lk}(\mathbf{p})$ and $\mathsf{CI}_{lk}(\mathbf{p})$ have been already expressed in (16) and (17), but we here emphasize them as the functions of data power variables stacked in $\mathbf{p}$. We now introduce the definition of a standard interference function for which an low complexity algorithm to obtain a fixed point solution is proposed.

Definition 3 (Standard interference function). *A function $I(\mathbf{z})$ is a standard interference function for all $\mathbf{z} \geq \mathbf{0}$, if the following properties hold: a) Positivity $I(\mathbf{z}) > 0, \forall \mathbf{z} \geq 0$. b) Monotonicity $I(\mathbf{z}) \geq I(\mathbf{z}')$ if $\mathbf{z} \geq \mathbf{z}'$. c) Scalability: $\alpha I(\mathbf{z}) > I(\alpha \mathbf{z}), \forall \alpha > 1$, for all scalar $\alpha > 1$.*

The positivity property is because of the inherent mutual interference and thermal noise in the system, which implies a non-zero value. This means that the transmit data powers are always larger than zero when users request non-zero SEs. The monotonicity property ensures that we can scale up or down (35) by adjusting the data powers. Finally, the scalability property suggests a method to uniformly scale down the data power coefficient of user k in cell l at each iteration by utilizing a positive constant α. We now construct a policy to update the data power of user k in cell l for the given initial values $p_{lk}(0), \forall l, k$, as in Theorem 3.

Theorem 3. *By assuming that the feasible domain is non-empty and $0 \leq I_{lk}(\mathbf{p}) \leq P_{\max,lk}^2$ always holds for all $\mathbf{p}$ in the feasible domain. For the initial values of data powers $p_{lk}(0) = P_{\max,lk}, \forall l, k$, there exist data powers for which each interference function $I_{lk}(\mathbf{p})$ is non-increasing along iterations and converges to a fixed point. Particularly, the data power of user k in cell l, denoted by $p_{lk}(n)$, can be updated at iteration n as*

$$p_{lk}(n) = I_{lk}(\mathbf{p}(n-1)), \forall l, k. \quad (36)$$

Proof. The proof is to testify every function $I_{lk}(\mathbf{p})$ defined in (35) being standard interference, and hence the updated power policy in (36) ensures that this iterative approach will converge to a fixed point. The detailed proof is available in Appendix E. $\square$

Every user in the network has its own standard interference function satisfying the three fundamental properties in Definition 3 and utilizing it to update the data power as in (36). The analysis in Theorem 3 is based on the assumption that problem (32) has the global optimum for which all users are served with their required SEs. The power constraints in (32) ($p_{lk} \leq P_{\max,lk}, \forall l, k$) are tackled by the fact if $I_{lk}(n-1) > P_{\max,lk}$, then the congestion issue appears and leads to an obvious selection $p_{lk}(n) = P_{\max,lk}$. We therefore define the constrained standard interference function used at iteration $n-1$ as

$$\hat{I}_{lk}(\mathbf{p}(n-1)) = \min\left(I_{lk}(\mathbf{p}(n-1)), P_{\max,lk}\right). \quad (37)$$

Algorithm 1 Data power allocation to problem (32) by spending maximum transmit power on unsatisfied users

Input: Define maximum powers $P_{\max,lk}, \forall l, k$; Select initial values $p_{lk}(0) = P_{\max,lk}, \forall l, k$; Compute the total power consumption $P_{\mathrm{tot}}(0) = \sum_{l=1}^L \sum_{k=1}^K p_{lk}(0)$; Set initial value $n = 1$ and tolerance ϵ.

1. User k in cell l computes the standard interference function $I_{lk}(\mathbf{p}(n-1))$ using (35).
2. If $I_{lk}(\mathbf{p}(n-1)) > P_{\max,lk}$, update $p_{lk}(n) = P_{\max,lk}$. Otherwise, update $p_{lk}(n) = I_{lk}(\mathbf{p}(n-1))$.
3. Repeat Steps $1, 2$ with other users, then compute the ratio $\gamma(n) = |P_{\mathrm{tot}}(n) - P_{\mathrm{tot}}(n-1)|/P_{\mathrm{tot}}(n-1)$.
4. If $\gamma_l(n) \leq \epsilon \rightarrow$ Set $p_{lk}^* = p_{lk}(n), \forall l, k$, and Stop. Otherwise, set $n = n + 1$ and go to Step 1.

Output: A fixed point $p_{lk}^*, \forall l, k$.

For a cellular Massive MIMO system with the power budget constraints and the initial data power vector $\mathbf{p}(0)$ with the entries $p_{lk}(0) = P_{\max,lk}, \forall l, k$, iteration n updates the data power of user k in cell l as

$$p_{lk}(n) = \hat{I}_{lk}(\mathbf{p}(n-1)). \quad (38)$$

Combining (37) and (38), we observe that if $\hat{I}_{lk}(\mathbf{p}(n-1)) = P_{\max,lk}$, the update $p_{lk}(n) = P_{\max,lk}$ maintains the non-increasing objective function of problem (32). Otherwise, it holds that $\hat{I}_{lk}(\mathbf{p}(n-1)) = I_{lk}(\mathbf{p}(n-1))$, and hence user k in cell l consumes less power than the maximum. This procedure will be applied to all the KL users, which results in an alternating approach is summarized in Algorithm 1. Since the convergence of the update $p_{lk}(n) = P_{\max,lk}$ is trivial, the proposed algorithm converges to a fixed point follows a similar methodology as [38, Theorem 7]. By assuming that the channel statistic information is computed in advance and available in the network, we can compute the total number of operations that dominate the computational complexity of this algorithm as $O\left(N_m L^2 K^2 + 3N_m \left|\mathcal{P}_{lk}\right| LK\right)$, where N_m is the number of iterations needed to reach the fixed point in polynomial time. Notice that, in Algorithm 1, when users cannot be served by the required SEs, one still lets them utilize the maximum power. This policy aims at maximizing the SE of a particular user, however producing more mutual interference to the other users.

B. Softly Removing Unsatisfied Users

Instead of allowing potential unsatisfied users to spend full data power, one can reduce their power with the goal to degrade mutual interference to the others. This policy might ameliorate the number of satisfied users in the entire network. The idea is in detail that: At first, every user improves the transmission quality by spending more power to each data symbol. This target can be achieved by, for example, simply constructing the standard inference functions as in the previous subsection. If at the limited power budget, the required SE cannot be achieved, unsatisfied users will reduce data power. We then mathematically suggest an update of the data powers along iterations as in Theorem 4.

Theorem 4. *From the initial values $p_{lk}(0) = P_{\max,lk}, \forall l, k$, if the data power of user k in cell l is updated at iteration n as*

$$
\begin{aligned}
p_{lk}(n) &= f_{lk}(\mathbf{p}(n-1)) \\
&= \begin{cases}
I_{lk}(\mathbf{p}(n-1)), & \text{if } I_{lk}(\mathbf{p}_l(n-1)) \leq P_{\max,lk}, \\
\frac{P^2_{\max,lk}}{I_{lk}(\mathbf{p}(n-1))}, & \text{if } I_{lk}(\mathbf{p}(n-1)) > P_{\max,lk},
\end{cases}
\end{aligned}
\quad (39)
$$

then the iterative approach converges to a fixed point.

Proof. The proof is first to confirm that the updated power policy in (39) follows a so-called two-sided function and the convergence is then established. The detailed proof is available in Appendix F. □

This theorem provides a procedure to minimize the total transmit power in the network and coping with the congestion issue based on the standard interference function defined for each user as in (35). If $I_{lk}(\mathbf{p}(n-1))$ is less than the maximum power $P_{\max,lk}$ then the data power of user k in cell l is updated based on (36), same as what has done in Algorithm 1. The main distinction is to prevent any unsatisfied user from transmitting full power whenever the congestion issue happens, i.e. $I_{lk}(\mathbf{p}(n-1)) > P_{\max,lk}$. In particular, the data power of a unsatisfied user scales down with the total mutual interference and noise level, which contains in $I_{lk}(\mathbf{p}(n-1))$. By doing this, the mutual interference from this unsatisfied user to the others should be reduced, and hence there is chance for the remaining users to get their required SEs. The proposed optimization approach is summarized in Algorithm 2. The per iteration complexity is $O\left(L^2 K^2 + 3|\mathcal{P}|LK\right)$, thus the computational complexity of Algorithm 2 is in the order of $O\left(N_s L^2 K^2 + 3N_s|\mathcal{P}|LK\right)$, where N_s is the number of iterations needed for this algorithm converges. Furthermore, Theorem 4 analytically proves the convergence to a fixed point, whose property is stated in Remark 4.

Remark 4. *The proposed algorithms enable to work in both feasible and infeasible domain such that a fixed point to problem (32) can be obtained. For realizations of user locations that result in feasible domains, the fixed point obtained by those algorithms is unique, which is the global optimum. The main difference between the two algorithms is at the policy to assign data powers whenever the congestion issue appears. While Algorithm 1 allocates the maximum data power to users when their SINR constraints are not satisfied, Algorithm 2 reduces the data power. As a consequence, for an infeasible domain to problem (32), the fixed point obtained by each algorithm may be different from each other.*

We notice that it is straightforward to extend the proposed algorithms to the total downlink energy consumption optimization problem with the per-user power constraints. The extension is not trivial if one considers the per-BS total limited power budgets and a primal-dual decomposition approach might be utilized to allocate the downlink power coefficients based on the standard interference functions.

V. NUMERICAL RESULTS

We consider a Massive MIMO system with $L = 4$ square cells in a 1 km^2 area, each serving $K = 5$ users. All the users

---

Algorithm 2 Data power allocation to problem (32) by softly removing unsatisfied users

Input: Define maximum powers $P_{\max,lk}, \forall l, k$; Select initial values $p_{lk}(0) = P_{\max,lk}, \forall l, k$; Compute the total power consumption $P_{\text{tot}}(0) = \sum_{l=1}^{L} \sum_{k=1}^{K} p_{lk}(0)$; Set initial value $n = 1$ and tolerance ϵ.

1. User k in cell l computes the standard interference function $I_{lk}(\mathbf{p}(n-1))$ using (35).
2. If $I_{lk}(\mathbf{p}(n)) < P_{\max,lk}$, update $p_{lk}(n) = I_{lk}(\mathbf{p}(n-1))$. Otherwise, update $p_{lk}(n) = P^2_{\max,lk}/I_{lk}(\mathbf{p}(n-1))$.
3. Repeat Steps $1, 2$ with other users, then compute the ratio $\gamma(n) = |P_{\text{tot}}(n) - P_{\text{tot}}(n-1)|/P_{\text{tot}}(n-1)$.
4. If $\gamma_l(n) \leq \epsilon \to$ Set $p^*_{lk} = p_{lk}(n), \forall l, k$, and Stop. Otherwise, set $n = n + 1$ and go to Step 1.

Output: A fixed point $p^*_{lk}, \forall l, k$.

---

are uniformly distributed within its cell with the distance to the BS no less than 35 m. Each coherence book has $\tau_c = 200$ symbols and there are $\tau_p = 5$ orthogonal pilot signals with the power $\hat{p}_{lk} = P_{\max,lk} = 200$ mW, $\forall l, k$. Without the loss of generality, the users with same index in all cells sharing a orthogonal pilot signal. The system bandwidth is 20 MHz and the noise variance is -96 dBm with the noise figure 5 dB. The large-scale fading coefficient [dB] of user k in cell l and BS l' is modeled based on the 3GPP LTE specifications [39] as

$$
\beta^{l'}_{lk} = -128.1 - 37.6 \log_{10}\left(d^{l'}_{lk}/1\text{km}\right) + z^{l'}_{lk}, \quad (40)
$$

where $d^{l'}_{lk} > 35$ m is the distance between user k in cell l and BS l'; $z^{l'}_{lk}$ is the shadow fading coefficient, which follows a Gaussian distribution with zero mean and standard deviation 7 dB. The covariance matrices are computed by using [17, (13) and (16)]. In the proposed algorithms (Algorithms 1 and 2), we set $\epsilon = 0.001$, except Fig. 5 which visualizes the convergence property. For feasible systems, the global optimum obtained by utilizing interior point methods from previous works like [40], [41] are included for comparison.[4]

Figure 1 shows the cumulative distribution function (CDF) of SE per user [b/s/Hz] to verify the correctness of the closed-form expression of the uplink SE for each user obtained in Theorem 1. There are 21 scatterers per communication link and all users spend full power for the data transmission. Particularly, the closed-form expression result matches very well Monte-Carlo simulation result for all the considered number of BS antennas. This figure also illustrates the SE per user getting better when each BS is equipped with more antennas. Each user can be served by a data rate increasing from 1.3 [b/s/Hz] to 1.8 [b/s/Hz] on average if the number of BS antennas increases from 50 to 150, which is a 38.5% data rate improvement. From this amount of antennas added, the median SE gets significantly better with a 60% data rate

---

[4]In [40], [41], user locations and shadow fading realizations resulting in a feasible domain have been considered for conveniences to utilize the interior-point methods. If only one user is not satisfied with its SE requirement, it is sufficient to create an infeasible set. Consequently, the problem lacks a feasible solution.
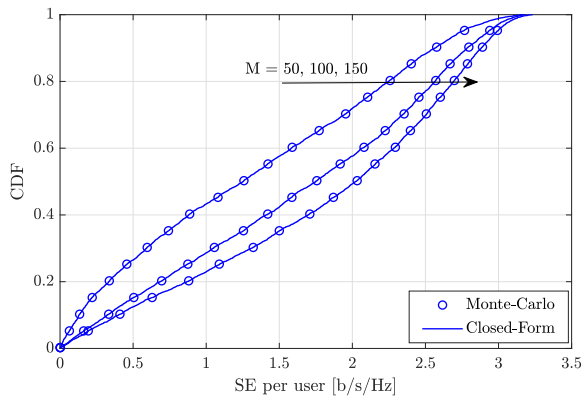
Fig. 1. The CDF of the uplink SE per user [b/s/Hz] with Monte-Carlo simulation and closed-form expression with $S_{lk}^{l'} = 21, \forall l, l', k$.

Fig. 2. The CDF of the uplink SE per user [b/s/Hz] with Monte-Carlo simulation and closed-form expression with $M = 100$.

improvement as a consequence of the SE per user increasing from 1.25 [b/s/Hz] to 2 [b/s/Hz].

Figure 2 plots the CDF of SE per user [b/s/Hz] with a different number of scatterers. Each BS is equipped with 100 antennas. All the Monte-Carlo simulations producing the same SE as the closed-form expression verifies the correctness of Theorem 1 when the number of scatterers varies. Clearly, the SE per user gets better for rich scattering environments. On average, a notable gain of 1.25× in SE is obtained if each channel has 21 scatterers instead of 11 scatterers. However, the SE has a small gai, e.g., with only 6.6% if the propagation environment has 31 scatterers. Therefore, Fig. 2 unveils a slow growth of the SE as a function of the scatterer number. At 95%-likely, the three considered scenarios provide the same SE with 0.16 [b/s/Hz] without data power control. Consequently, it seems that poor scattering environments affect the worst SE slightly.

Figure 3 shows the CDF of SE per user [b/s/Hz] for a system with either MR or ZF combining technique with a small number of scatterers per each propagation channel. The transmit power per symbol is 50 mW and the large-scale fading coefficients are computed similar to (40) but with the penetration loss of 20 dB. ZF generally provides better performance than MR since it cancels out mutual interference more effectively [17]. On average, a system with MR combining is still the baseline that offers less than that of utilizing ZF combining. Nonetheless, Fig. 3 demonstrates the sensitivity of ZF when the propagation environment lacks scatterers in many user locations and shadow fading realizations which result in low-rank channels. Consequently, MR outperforms ZF about 45.5% at the median SE.

Figure 4 presents the CDF of SE per user by utilizing the different spatial correlation channel models. There are 21 scatterers for each propagation link with the double scattering channel model. The exponential correlation model is defined as in [10] with the correlation magnitude 0.9, while the local scattering channel model is defined in [20] with 6 scattering clusters, the angular standard deviation 5°, and the antenna spacing of the half wavelength. By assuming that the scattering clusters are in the half-space in front of the BSs, the local scattering channel model offers the highest SE per user with up to 2.1 [b/s/Hz] on average. The exponential correlation model
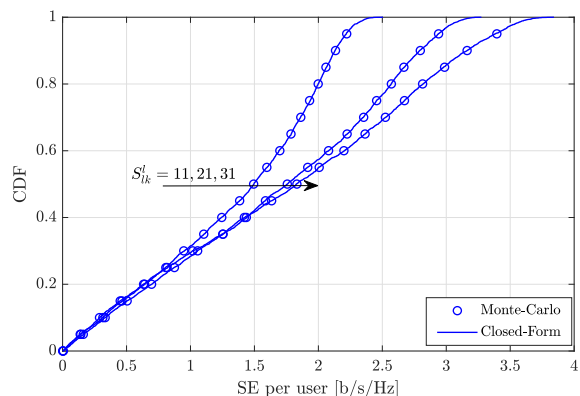
provides the SE of about 1.8 [b/s/Hz] per user. Meanwhile, the double scattering model yields to the lowest SE with only 1.6 [b/s/Hz] due to taking both the local scattering property and rank deficiency into account.

Figure 5 illustrates the convergence of Algorithms 1 and 2 by utilizing two different required SEs. They converge fast to a fixed point after a few tens of iterations. If each user requests a SE 1 [b/s/Hz], the proposed algorithms need less than 10 iterations to reach convergence, which is the same fixed point. This fixed point is the global optimum since the optimization problem is always feasible for the user locations and shadow fading realizations have been generated. When the required SEs expand to 2 [b/s/Hz], the proposed algorithms require around 40 iterations to approach the optimum. The convergence rate is therefore slower when the SE requirements enlarge. This SE setting also manifests the benefits of Algorithm 2, which yields 20% less the total transmit power than Algorithm 1. On the other hand, the fixed point obtained by each algorithm is different from each other.

We show the CDF of the data power consumption [mW] consumed by each user in Fig. 6 for feasible systems with the two different required SEs. Matched well with the claim in Remark 4 for feasible systems, the proposed algorithms provide a unique fixed point that is the global optimum as what has obtained by the interior-point methods. Additionally, data power escalates when users require higher SEs. With the required SE 1.5 [b/s/Hz], each user only spends 5.2 mW for each data symbol on average. However, it drastically grows to 11.4 mW (corresponding to 2.2× more power) with the required SE 1.75 [b/s/Hz]. Both the considered SE settings illustrate significant reductions of transmit power compared to the scenario dedicating full power to the data symbols. Particularly, all the users consume 38.5× and 17.5× less power than the full power transmission with the two considered SEs, respectively.

Figure 7 displays the CDF of the data power consumption [mW] per user for infeasible systems. It is the main interest of this paper when working with multiple access in Massive MIMO communications since there is no global optimum to obtain or compare against. All the users consume non-zero powers at the fixed points identified Algorithms 1 and 2. The trend that more data power is needed when the users
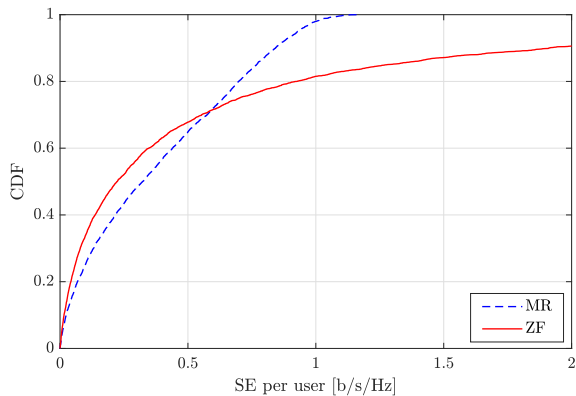
Fig. 3. The CDF of SE per user with the different linear combining techniques, $M = 100$, and $S_{lk}^{l'} = 3, \forall l, l', k$.

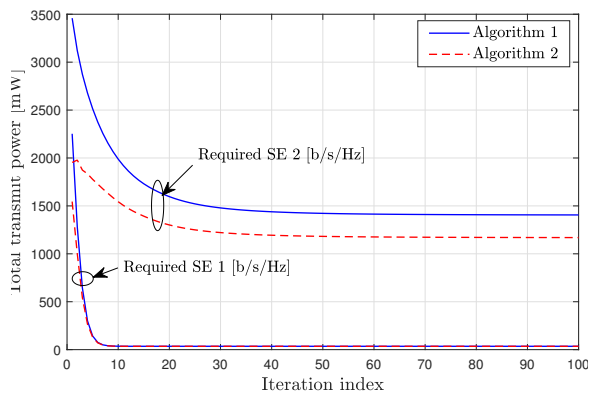Fig. 4. The CDF of SE per user with the different spatially correlated models, $M = 100$.

Fig. 5. The convergence of Algorithms 1 and 2 with the different required SEs at the users, $M = 100$, and $S_{lk}^{l'} = 21, \forall l, l', k$.

Fig. 6. The CDF of the power consumption per user [mW] for feasible systems with the different required SEs at the users, $M = 100$, and $S_{lk}^{l'} = 21, \forall l, k, l'$.
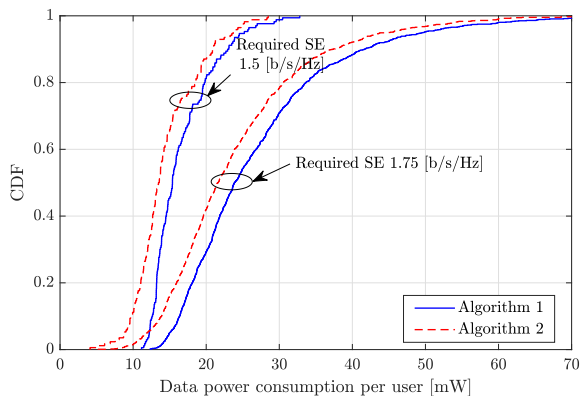

Fig. 7. The CDF of the power consumption per user [mW] for infeasible systems with the different required SEs at the users, $M = 100$, and $S_{lk}^{l'} = 21, \forall l, k, l'$.

Fig. 8. The satisfied SE probability versus the different required SE per user for a system with $M = 100$ and $S_{lk}^{l'} = 21, \forall l, k, l'$.

require higher SEs has still remained. In more detail, the data power obtained by Algorithm 1 grows 1.6× from 16.6 mW to 27.0 mW when the required SE increases from 1.5 [b/s/Hz] to 1.75 [b/s/Hz]. The data power increases 1.7× from 14.5 mW to 24.1 mW if Algorithm 2 is exploited. Moreover, the data power consumption per user obtained by Algorithm 1 is 12.3% and 15.1% higher than by Algorithm 2.

Figure 8 plots the satisfied SE probability defined as the fraction of random user locations and shadow fading realizations in which the users can be served by the required SEs. If each user requires an SE 1.5 [b/s/Hz], all the benchmarks provide an overwhelming satisfied SE probability. For instance,
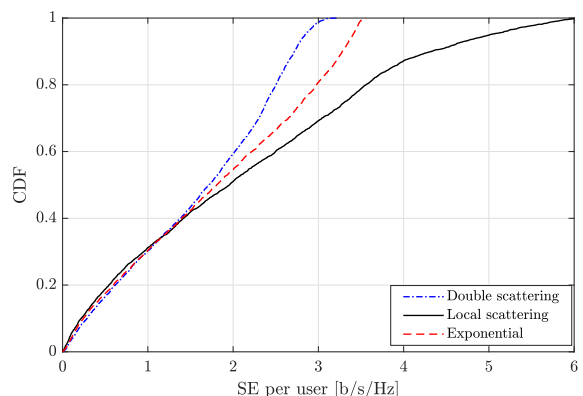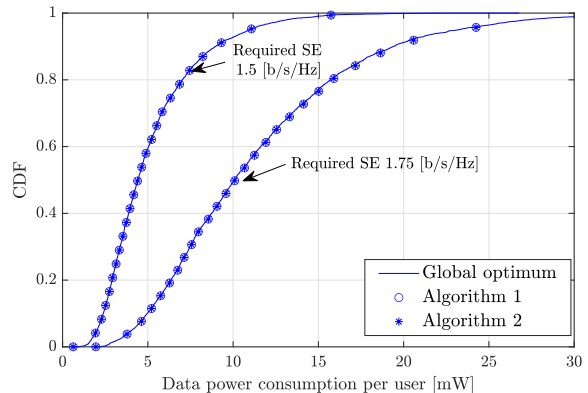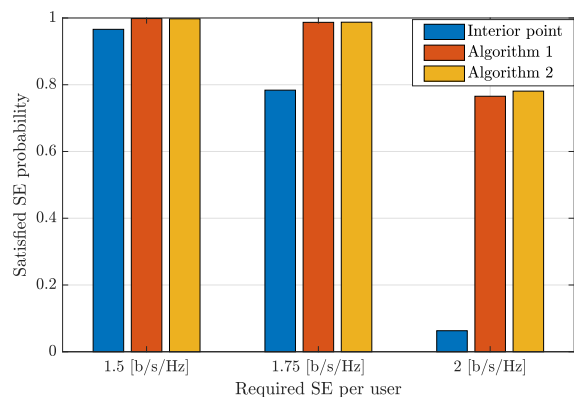
the interior-point methods offer 96.7% user locations and shadow fading realizations with the required SEs. Meanwhile, the proposed algorithms offer a satisfied SE probability 99.8%. However, the interior-point methods will perform worse with higher SE requirements since only one user is sufficient to create an empty feasible set as aforementioned in Section III-B, especially only 6.3% users satisfied the required SE 2 [b/s/Hz]. In contrast, the proposed algorithms still offer a satisfied SE probability of more than 75%. Furthermore, Algorithm 2 slightly performs better than Algorithm 1 in those required SE settings.

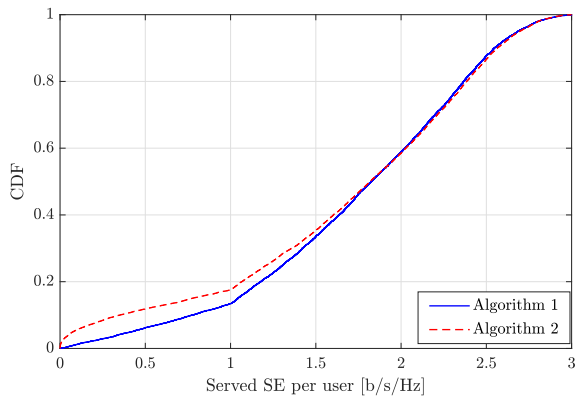Figure 9 provides the served SE per user [b/s/Hz] when

Fig. 9. The CDF of served SE per user [b/s/Hz] with $M = 100$, $S_{lk}^{l'} = 21$, $\forall l, k, l'$, and the required SEs uniformly varying in the range $[1, 3]$ [b/s/Hz].

Fig. 10. The CDF of data power consumption [mW] with $M = 100$, $S_{lk}^{l'} = 21$, $\forall l, k, l'$, and the required SEs uniformly varying in the range $[1, 3]$ [b/s/Hz].
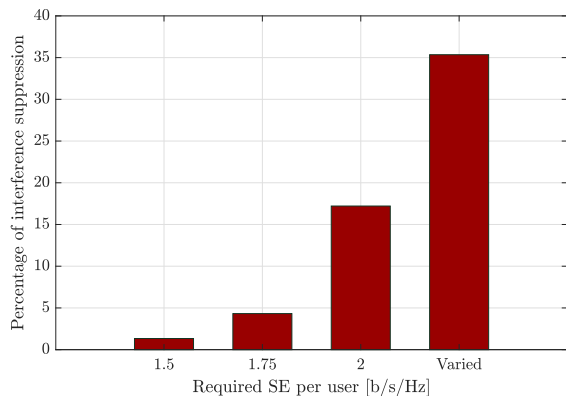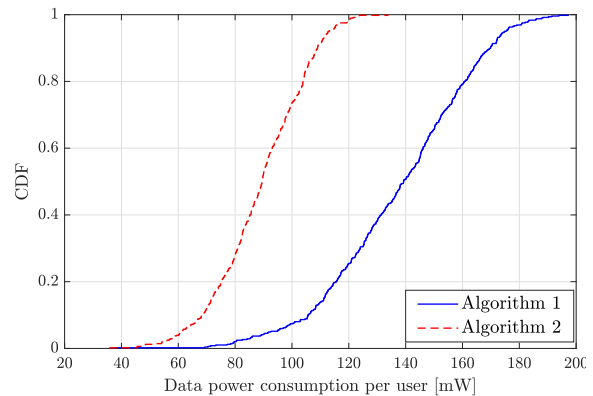

Fig. 11. The interference suppression obtained by Algorithm 2 compared to Algorithm 1 as a function of the required SE per user with $M = 100$ and $S_{lk}^{l'} = 21$, $\forall l, k, l'$.

the users have different required SEs, which are uniformly distributed in the range $[1, 3]$ [b/s/Hz] over many user locations and shadowing fading realizations. The interior-point methods are not included since the optimization problem always has an empty feasible domain in this complicated scenario. Interestingly, Algorithm 1 performs pretty better than Algorithm 2 since the former gives 86.5% users satisfied their SEs, while the latter is only 82.5%. However, Fig. 10 indicates that Algorithm 2 produces a fixed point that has much lower power consumption than Algorithm 1. The saving power is about 54.7% on average thanks to the data reduction policy in (39) whenever the congestion issue appears.

Figure 11 shows the percentage of interference suppression obtained by Algorithm 2 in a comparison to Algorithm 1 by utilizing the different required SEs per user. Softly removing unsatisfied users generates less mutual interference than spending the maximum transmit power on those users, especially when the SE requirements are high. For instance, mutual interference from Algorithm 2 is only 1.3% less than that of Algorithm 1 if the required SE per user is 1.5 [b/s/Hz]. However, the mutual interference suppression gains up to 17.2% with the SE requirement 2 [b/s/Hz]. In particular, Algorithm 2 suppresses mutual interference significantly when each user has its own SE requirement varied in the range from 1 [b/s/Hz] to 3 [b/s/Hz] with the mutual interference suppression of about

35.4%. We therefore conclude the effectiveness of the second algorithm compared with the first one.

VI. CONCLUSION

This paper has analyzed the system performance of Massive MIMO systems with an arbitrary number of BS antennas, users, and scatterers by utilizing the double scattering channel model, rather than the asymptotic regime as in previous works. The closed-form expression of the uplink SE per user was first computed, then the asymptotic performance was obtained. We further formulated and solved a total transmit power minimization problem with the required SE constraints and limited power budget. We proposed two algorithms to handle effectively the congestion issue that often happens since multiple users are simultaneously connecting to the network and sharing the same time and frequency resources. The solutions to those algorithms are quite similar to each other if the required SEs can be almost satisfied with the given power budget. In contrast, Algorithm 2 outperforms Algorithm 1 in phenomena where the SE requirements are vastly different and many users cannot be served with the required SEs.

APPENDIX

A. Proof of Lemma 1

For a given matrix $\mathbf{B}$, we first compute the statistical information of the two channels $\mathbf{h}_{l'k'}^{l}$ and $\mathbf{h}_{l''k''}^{l}$ when $(l', k') \neq (l'', k'')$ by averaging over the different realizations of small-fading coefficients as

$$
\mathbb{E}\left\{\left|(\mathbf{h}_{l'k'}^{l})^{\mathrm{H}}\mathbf{B}\mathbf{h}_{l''k''}^{l}\right|^2\right\} = \\
\mathrm{tr}\left(\mathbf{B}\mathbb{E}\left\{\mathbf{h}_{l''k''}^{l}(\mathbf{h}_{l''k''}^{l})^{\mathrm{H}}\right\}\mathbf{B}^{\mathrm{H}}\mathbb{E}\left\{\mathbf{h}_{l'k'}^{l}(\mathbf{h}_{l'k'}^{l})^{\mathrm{H}}\right\}\right). \quad (41)
$$

The first expectation in the right-hand side of (41) is computed by plugging the definition of the double-scattering channel

model in (1) as

$$
\begin{aligned}
\mathbb{E}\left\{\mathbf{h}_{l''k''}^{l}(\mathbf{h}_{l''k''}^{l})^{\mathrm{H}}\right\} &= \frac{\beta_{l''k''}^{l}}{S_{l''k''}^{l}}\mathbb{E}\left\{(\mathbf{R}_{l''k''}^{l})^{1/2}\mathbf{G}_{l''k''}^{l}(\widetilde{\mathbf{R}}_{l''k''}^{l})^{1/2}\times\right. \\
&\quad \left.\mathbb{E}\left\{\mathbf{g}_{l''k''}^{l}(\mathbf{g}_{l''k''}^{l})^{\mathrm{H}}\right\}(\widetilde{\mathbf{R}}_{l''k''}^{l})^{1/2}(\mathbf{G}_{l''k''}^{l})^{\mathrm{H}}(\mathbf{R}_{l''k''}^{l})^{1/2}\right\} \\
&= \frac{\beta_{l''k''}^{l}}{S_{l''k''}^{l}}(\mathbf{R}_{l''k''}^{l})^{1/2}\mathbb{E}\left\{\mathbf{G}_{l''k''}^{l}\widetilde{\mathbf{R}}_{l''k''}^{l}(\mathbf{G}_{l''k''}^{l})^{\mathrm{H}}\right\}(\mathbf{R}_{l''k''}^{l})^{1/2} \\
&= \beta_{l''k''}^{l}d_{l''k''}^{l}\mathbf{R}_{l''k''}^{l},
\end{aligned}
\tag{42}
$$

where the last equality of (42) is obtained by utilizing [32, Lemma 8] to compute the covariance matrix of the circularly symmetric complex Gaussian matrix $\mathbf{G}_{l''k''}^{l}$ for a given deterministic matrix $\mathbf{R}_{l''k''}^{l}$. Following a similar manner, the second expectation in the right-hand side of (41) is computed in closed form as

$$
\mathbb{E}\left\{\mathbf{h}_{l'k'}^{l}(\mathbf{h}_{l'k'}^{l})^{\mathrm{H}}\right\} = \beta_{l'k'}^{l}d_{l'k'}^{l}\mathbf{R}_{l'k'}^{l}.
\tag{43}
$$

Plugging (42) and (42) into (41), we obtain the result as shown in (2). For a given deterministic matrix $\mathbf{B}$, the statistical information of the channel $\mathbf{h}_{l'k'}^{l}$ is computed as

$$
\begin{aligned}
\mathbb{E}\left\{\left|(\mathbf{h}_{l'k'}^{l})^{\mathrm{H}}\mathbf{B}\mathbf{h}_{l'k'}^{l}\right|^{2}\right\} &= \frac{(\beta_{l'k'}^{l})^{2}}{(S_{l'k'}^{l})^{2}}\mathbb{E}\left\{\left|(\mathbf{g}_{l'k'}^{l})^{\mathrm{H}}(\widetilde{\mathbf{R}}_{l'k'}^{l})^{1/2}(\mathbf{G}_{l'k'}^{l})^{\mathrm{H}}\right.\right. \\
&\quad \left.\left.(\mathbf{R}_{l'k'}^{l})^{1/2}\mathbf{B}(\mathbf{R}_{l'k'}^{l})^{1/2}\mathbf{G}_{l'k'}^{l}(\widetilde{\mathbf{R}}_{l'k'}^{l})^{1/2}\mathbf{g}_{l'k'}^{l}\right|^{2}\right\} \\
&= \frac{(\beta_{l'k'}^{l})^{2}}{(S_{l'k'}^{l})^{2}}\mathbb{E}\left\{\left\|(\widetilde{\mathbf{R}}_{l'k'}^{l})^{1/2}\mathbf{g}_{l'k'}^{l}\right\|^{4}\left|\frac{(\mathbf{g}_{l'k'}^{l})^{\mathrm{H}}(\widetilde{\mathbf{R}}_{l'k'}^{l})^{1/2}(\mathbf{G}_{l'k'}^{l})^{\mathrm{H}}}{\left\|(\widetilde{\mathbf{R}}_{l'k'}^{l})^{1/2}\mathbf{g}_{l'k'}^{l}\right\|}\right.\right. \\
&\quad \left.\left.\times(\mathbf{R}_{l'k'}^{l})^{1/2}\mathbf{B}(\mathbf{R}_{l'k'}^{l})^{1/2}\frac{\mathbf{G}_{l'k'}^{l}(\widetilde{\mathbf{R}}_{l'k'}^{l})^{1/2}\mathbf{g}_{l'k'}^{l}}{\left\|(\widetilde{\mathbf{R}}_{l'k'}^{l})^{1/2}\mathbf{g}_{l'k'}^{l}\right\|}\right|^{2}\right\},
\end{aligned}
\tag{44}
$$

where the last equality of (44) is obtained by utilizing the normalization term $\left\|(\widetilde{\mathbf{R}}_{l'k'}^{l})^{1/2}\mathbf{g}_{l'k'}^{l}\right\|$. Let us introduce the new optimization variable $\mathbf{z}_{l'k'}^{l}$, which is defined as

$$
\mathbf{z}_{l'k'}^{l} = \frac{\mathbf{G}_{l'k'}^{l}(\widetilde{\mathbf{R}}_{l'k'}^{l})^{1/2}\mathbf{g}_{l'k'}^{l}}{\left\|(\widetilde{\mathbf{R}}_{l'k'}^{l})^{1/2}\mathbf{g}_{l'k'}^{l}\right\|},
\tag{45}
$$

then it is straightforward to prove that $\mathbf{z}_{l'k'}^{l} \sim \mathcal{CN}(\mathbf{0},\mathbf{I}_{M})$, and is independent of $\mathbf{g}_{l'k'}^{l}$. Thus, (44) is equivalent to the following expression

$$
\begin{aligned}
\mathbb{E}\left\{\left|(\mathbf{h}_{l'k'}^{l})^{\mathrm{H}}\mathbf{B}\mathbf{h}_{l'k'}^{l}\right|^{2}\right\} &= \frac{(\beta_{l'k'}^{l})^{2}}{(S_{l'k'}^{l})^{2}}\mathbb{E}\left\{\left\|(\widetilde{\mathbf{R}}_{l'k'}^{l})^{1/2}\mathbf{g}_{l'k'}^{l}\right\|^{4}\right\}\times \\
&\quad \mathbb{E}\left\{\left|(\mathbf{z}_{l'k'}^{l})^{\mathrm{H}}(\mathbf{R}_{l'k'}^{l})^{1/2}\mathbf{B}(\mathbf{R}_{l'k'}^{l})^{1/2}\mathbf{z}_{l'k'}^{l}\right|^{2}\right\} \\
&= \frac{(\beta_{l'k'}^{l})^{2}}{(S_{l'k'}^{l})^{2}}\left(\left|\mathrm{tr}(\widetilde{\mathbf{R}}_{l'k'}^{l})\right|^{2} + \mathrm{tr}\left((\widetilde{\mathbf{R}}_{l'k'}^{l})^{2}\right)\right)\times \\
&\quad \left(\left|\mathrm{tr}(\mathbf{R}_{l'k'}^{l}\mathbf{B})\right|^{2} + \mathrm{tr}(\mathbf{R}_{l'k'}^{l}\mathbf{B}\mathbf{R}_{l'k'}^{l}\mathbf{B}^{H})\right),
\end{aligned}
\tag{46}
$$

where the last equality in (46) is obtained by utilizing [32, Lemma 9] to compute the forth moment of zero-mean complex

Gaussian variables, and then the result is obtained as in (3) after doing some algebra.

B. Proof of Lemma 2

Following the similar approach as [32, Lemma 3], we can compute the correlation matrix of two channel vectors $\mathbf{h}_{lk}^{l'}$ and $\mathbf{h}_{l''k''}^{l'}$ by averaging over the different realizations of small-scale fading coefficients as

$$
\mathbb{E}\left\{\mathbf{h}_{lk}^{l'}(\mathbf{h}_{l''k''}^{l'})^{\mathrm{H}}\right\} = \begin{cases} \beta_{lk}^{l'}d_{lk}^{l'}\mathbf{R}_{lk}^{l'}, & \text{if } (l,k) = (l'',k''), \\ \mathbf{0}, & \text{if } (l,k) \neq (l'',k''). \end{cases}
\tag{47}
$$

The LMMSE estimate $\hat{\mathbf{h}}_{l'k'}^{l}$ is obtained by, first, computing the cross-covariance matrix between the two random variables $\mathbf{h}_{l'k'}^{l}$ and $\mathbf{y}_{l'k'}^{l,p}$ as

$$
\mathbb{E}\left\{\mathbf{h}_{l'k'}^{l}(\mathbf{y}_{l'k'}^{l,p})^{\mathrm{H}}\right\} = \sqrt{\hat{p}_{l'k'}}\tau_{p}\beta_{l'k'}^{l}d_{l'k'}^{l}\mathbf{R}_{l'k'}^{l}.
\tag{48}
$$

In fact, (48) is obtained by utilizing the formulation of $\mathbf{y}_{l'k'}^{l,p}$ in (7) and the channel correlation property in (47). The covariance matrix of the signal $\mathbf{y}_{l'k'}^{l,p}$ is computed as

$$
\mathbb{E}\left\{\mathbf{y}_{l'k'}^{l,p}(\mathbf{y}_{l'k'}^{l,p})^{\mathrm{H}}\right\} = (\mathbf{\Psi}_{l'k'}^{l})^{-1}\tau_{p}^{-1}.
\tag{49}
$$

By utilizing (48) and (49) into the Bayesian Gauss-Markov theorem [30, Theorem 12.1], i.e.,

$$
\hat{\mathbf{h}}_{l'k'}^{l} = \mathbb{E}\left\{\mathbf{h}_{l'k'}^{l}(\mathbf{y}_{l'k'}^{l,p})^{\mathrm{H}}\right\}\left(\mathbb{E}\left\{\mathbf{y}_{l'k'}^{l,p}(\mathbf{y}_{l'k'}^{l,p})^{\mathrm{H}}\right\}\right)^{-1}\mathbf{y}_{l'k'}^{l,p},
\tag{50}
$$

and doing some algebra, we obtain the expression of the channel estimate $\hat{\mathbf{h}}_{l'k'}^{l}$ as shown in the lemma.

C. Proof of Theorem 1

We compute the expectation in the numerator of (14) with noting that $\mathbf{v}_{lk} = \hat{\mathbf{h}}_{lk}^{l}$ as

$$
\mathbb{E}\left\{\mathbf{v}_{lk}^{\mathrm{H}}\mathbf{h}_{lk}^{l}\right\} = \mathbb{E}\left\{\|\mathbf{v}_{lk}\|^{2}\right\} = \sqrt{z_{lk}^{l}}\mathrm{tr}\left(\mathbf{R}_{lk}^{l}\mathbf{\Psi}_{lk}^{l}\mathbf{R}_{lk}^{l}\right),
\tag{51}
$$

where the last equality in (51) is obtained by using the covariance property in (10). The first part of the denominator of (14) is decomposed into the coherent and non-coherent interference based on the pilot reuse pattern as

$$
\begin{aligned}
\sum_{l'=1}^{L}\sum_{k'=1}^{K}p_{l'k'}\mathbb{E}\left\{\left|\mathbf{v}_{lk}^{\mathrm{H}}\mathbf{h}_{l'k'}^{l}\right|^{2}\right\} &= \sum_{(l',k')\notin\mathcal{P}_{lk}}p_{l'k'}\mathbb{E}\left\{\left|\mathbf{v}_{lk}^{\mathrm{H}}\mathbf{h}_{l'k'}^{l}\right|^{2}\right\} \\
&\quad + \sum_{(l',k')\in\mathcal{P}_{lk}}p_{l'k'}\mathbb{E}\left\{\left|\mathbf{v}_{lk}^{\mathrm{H}}\mathbf{h}_{l'k'}^{l}\right|^{2}\right\}.
\end{aligned}
\tag{52}
$$

The first expectation in the right-hand side of (52) is non-coherent interference and computed in closed form by the independence of two random variables $\mathbf{v}_{lk}$ and $\mathbf{h}_{l'k'}^{l}$ as

$$
\begin{aligned}
&\sum_{(l',k')\notin\mathcal{P}_{lk}}p_{l'k'}\mathbb{E}\left\{\left|\mathbf{v}_{lk}^{\mathrm{H}}\mathbf{h}_{l'k'}^{l}\right|^{2}\right\} \\
&= \sum_{(l',k')\notin\mathcal{P}_{lk}}p_{l'k'}\mathrm{tr}\left(\mathbb{E}\left\{\mathbf{h}_{l'k'}^{l}(\mathbf{h}_{l'k'}^{l})^{\mathrm{H}}\right\}\mathbb{E}\left\{\mathbf{v}_{lk}\mathbf{v}_{lk}^{\mathrm{H}}\right\}\right) \\
&= \sum_{(l',k')\notin\mathcal{P}_{lk}}p_{l'k'}m_{l'k'}^{l}\mathrm{tr}\left(\mathbf{R}_{lk}^{l}\mathbf{\Psi}_{lk}^{l}\mathbf{R}_{lk}^{l}\mathbf{R}_{l'k'}^{l}\right).
\end{aligned}
\tag{53}
$$

The second expectation in the right-hand side of (52) is coherent interference and computed by utilizing the channel estimate in (8) to construct the combining vector as

$$
\sum_{(l',k')\in\mathcal{P}_{lk}} p_{l'k'}\mathbb{E}\left\{\left|\mathbf{v}_{lk}^{\mathrm{H}}\mathbf{h}_{l'k'}^{l}\right|^2\right\}
$$

$$
= \sum_{(l',k')\in\mathcal{P}_{lk}} p_{l'k'}\mathbb{E}\left\{\left|(\mathbf{y}_{lk}^{l,p})^{\mathrm{H}}(\mathbf{B}_{lk}^{l})^{\mathrm{H}}\mathbf{h}_{l'k'}^{l}\right|^2\right\}
$$

$$
= \sum_{(l',k')\in\mathcal{P}_{lk}} p_{l'k'}\mathbb{E}\left\{\left|\sum_{(l'',k'')\in\mathcal{P}_{lk}}\sqrt{\hat{p}_{l''k''}}\tau_p(\mathbf{h}_{l''k''}^{l})^{\mathrm{H}}(\mathbf{B}_{lk}^{l})^{\mathrm{H}}\mathbf{h}_{l'k'}^{l}\right.\right.
$$
$$
\left.\left.+\boldsymbol{\phi}_{lk}^{\mathrm{H}}(\mathbf{N}_{l}^{p})^{\mathrm{H}}(\mathbf{B}_{lk}^{l})^{\mathrm{H}}\mathbf{h}_{l'k'}^{l}\right|^2\right\}
$$

$$
= \sum_{(l',k')\in\mathcal{P}_{lk}}\sum_{(l'',k'')\in\mathcal{P}_{lk}\setminus(l',k')} p_{l'k'}\hat{p}_{l''k''}\tau_p^2
$$
$$
\times\mathbb{E}\left\{\left|(\mathbf{h}_{l''k''}^{l})^{\mathrm{H}}(\mathbf{B}_{lk}^{l})^{\mathrm{H}}\mathbf{h}_{l'k'}^{l}\right|^2\right\}
$$
$$
+\sum_{(l',k')\in\mathcal{P}_{lk}} p_{l'k'}\hat{p}_{l'k'}\tau_p^2\mathbb{E}\left\{\left|(\mathbf{h}_{l'k'}^{l})^{\mathrm{H}}(\mathbf{B}_{lk}^{l})^{\mathrm{H}}\mathbf{h}_{l'k'}^{l}\right|^2\right\}
$$
$$
+\sum_{(l',k')\in\mathcal{P}_{lk}} p_{l'k'}\mathbb{E}\left\{\left|\boldsymbol{\phi}_{lk}^{\mathrm{H}}(\mathbf{N}_{l}^{p})^{\mathrm{H}}(\mathbf{B}_{lk}^{l})^{\mathrm{H}}\mathbf{h}_{l'k'}^{l}\right|^2\right\},
$$
$$(54)$$

where $\mathbf{B}_{lk}^{l}=\sqrt{\hat{p}_{lk}}\beta_{lk}^{l}d_{lk}^{l}\mathbf{R}_{lk}^{l}\boldsymbol{\Psi}_{lk}^{l}$ and the last equality in (54) is decomposed based on the correlation among the channels, and the uncorrelation between the channels and noise. In the last equation of (54), the first expectation is computed by using the independence of two random variables $\mathbf{h}_{l''k''}^{l}$ and $\mathbf{h}_{l'k'}^{l}$ as

$$
\mathbb{E}\left\{\left|(\mathbf{h}_{l''k''}^{l})^{\mathrm{H}}(\mathbf{B}_{lk}^{l})^{\mathrm{H}}\mathbf{h}_{l'k'}^{l}\right|^2\right\}
$$
$$
= \beta_{l''k''}^{l}d_{l''k''}^{l}\beta_{l'k'}^{l}d_{l'k'}^{l}\mathrm{tr}\left((\mathbf{B}_{lk}^{l})^{\mathrm{H}}\mathbf{R}_{l'k'}^{l}\mathbf{B}_{lk}^{l}\mathbf{R}_{l''k''}^{l}\right) \quad (55)
$$
$$
= \beta_{l''k''}^{l}d_{l''k''}^{l}\frac{m_{l'k'}^{l}}{\tau_p}\mathrm{tr}\left(\boldsymbol{\Psi}_{lk}^{l}\mathbf{R}_{lk}^{l}\mathbf{R}_{l'k'}^{l}\mathbf{R}_{lk}^{l}\boldsymbol{\Psi}_{lk}^{l}\mathbf{R}_{l''k''}^{l}\right).
$$

In order to obtain the result in (55), we have borrowed (2) in Corollary 1. The second expectation of (54) is computed by exploiting (3) as

$$
\mathbb{E}\left\{\left|(\mathbf{h}_{l'k'}^{l})^{\mathrm{H}}(\mathbf{B}_{lk}^{l})^{\mathrm{H}}\mathbf{h}_{l'k'}^{l}\right|^2\right\} = (\beta_{l'k'}^{l})^2\left((d_{l'k'}^{l})^2+\frac{\mathrm{tr}\left((\widetilde{\mathbf{R}}_{l'k'}^{l})^2\right)}{(S_{l'k'}^{l})^2}\right)
$$
$$
\times\left(\left|\mathrm{tr}\left(\mathbf{R}_{l'k'}^{l}(\mathbf{B}_{lk}^{l})^{\mathrm{H}}\right)\right|^2+\mathrm{tr}\left(\mathbf{R}_{l'k'}^{l}(\mathbf{B}_{lk}^{l})^{\mathrm{H}}\mathbf{R}_{l'k'}^{l}\mathbf{B}_{lk}^{l}\right)\right)
$$
$$
= (\beta_{l'k'}^{l})^2\hat{p}_{lk}(\beta_{lk}^{l})^2(d_{lk}^{l})^2\left((d_{l'k'}^{l})^2+\frac{\mathrm{tr}\left((\widetilde{\mathbf{R}}_{l'k'}^{l})^2\right)}{(S_{l'k'}^{l})^2}\right)\times
$$
$$
\left(\left|\mathrm{tr}\left(\mathbf{R}_{l'k'}^{l}\boldsymbol{\Psi}_{lk}^{l}\mathbf{R}_{lk}^{l}\right)\right|^2+\mathrm{tr}\left(\mathbf{R}_{l'k'}^{l}\boldsymbol{\Psi}_{lk}^{l}\mathbf{R}_{lk}^{l}\mathbf{R}_{l'k'}^{l}\mathbf{R}_{lk}^{l}\boldsymbol{\Psi}_{lk}^{l}\right)\right).
$$
$$(56)$$

Thanks to the independence between the channel and noise, the last expectation of (54) is computed as

$$
\mathbb{E}\left\{\left|\boldsymbol{\phi}_{lk}^{\mathrm{H}}(\mathbf{N}_{l}^{p})^{\mathrm{H}}(\mathbf{B}_{lk}^{l})^{\mathrm{H}}\mathbf{h}_{l'k'}^{l}\right|^2\right\}
$$
$$
= \mathrm{tr}\left((\mathbf{B}_{lk}^{l})^{\mathrm{H}}\mathbb{E}\left\{\mathbf{h}_{l'k'}^{l}(\mathbf{h}_{l'k'}^{l})^{\mathrm{H}}\right\}\mathbf{B}_{lk}^{l}\mathbb{E}\left\{\mathbf{N}_{l}^{p}\boldsymbol{\phi}_{lk}\boldsymbol{\phi}_{lk}^{\mathrm{H}}(\mathbf{N}_{l}^{p})^{\mathrm{H}}\right\}\right) \quad (57)
$$
$$
= \sigma^2 m_{l'k'}^{l}\mathrm{tr}\left(\boldsymbol{\Psi}_{lk}^{l}\mathbf{R}_{lk}^{l}\mathbf{R}_{l'k'}^{l}\mathbf{R}_{lk}^{l}\boldsymbol{\Psi}_{lk}^{l}\right).
$$

Plugging (55)-(57) into (54) and doing some algebra, the coherent interference term (54) is obtained in closed form as

$$
\sum_{(l',k')\in\mathcal{P}_{lk}} p_{l'k'}\mathbb{E}\left\{\left|\mathbf{v}_{lk}^{H}\mathbf{h}_{l'k'}^{l}\right|^2\right\} =
$$
$$
\sum_{(l',k')\in\mathcal{P}_{lk}} p_{l'k'}m_{l'k'}^{l}\mathrm{tr}\left(\mathbf{R}_{lk}^{l}\boldsymbol{\Psi}_{lk}^{l}\mathbf{R}_{lk}^{l}\mathbf{R}_{l'k'}^{l}\right)+\sum_{(l',k')\in\mathcal{P}_{lk}} p_{l'k'}
$$
$$
\times\frac{z_{l'k'}^{l}}{(d_{l'k'}^{l})^2}\left((d_{l'k'}^{l})^2+\frac{\mathrm{tr}\left((\widetilde{\mathbf{R}}_{l'k'}^{l})^2\right)}{(S_{l'k'}^{l})^2}\right)\left|\mathrm{tr}\left(\mathbf{R}_{l'k'}^{l}\boldsymbol{\Psi}_{lk}^{l}\mathbf{R}_{lk}^{l}\right)\right|^2+
$$
$$
\sum_{(l',k')\in\mathcal{P}_{lk}} p_{l'k'}z_{l'k'}^{l}\frac{\mathrm{tr}\left((\widetilde{\mathbf{R}}_{l'k'}^{l})^2\right)}{(S_{l'k'}^{l})^2}\mathrm{tr}\left(\mathbf{R}_{l'k'}^{l}\boldsymbol{\Psi}_{lk}^{l}\mathbf{R}_{lk}^{l}\mathbf{R}_{l'k'}^{l}\mathbf{R}_{lk}^{l}\boldsymbol{\Psi}_{lk}^{l}\right).
$$
$$(58)$$

Combining (52), (53), and (58), the first part of the denominator of (14) is computed in closed form as

$$
\sum_{l'=1}^{L}\sum_{k'=1}^{K} p_{l'k'}\mathbb{E}\left\{\left|\mathbf{v}_{lk}^{H}\mathbf{h}_{l'k'}^{l}\right|^2\right\} = \sum_{l'=1}^{L}\sum_{k'=1}^{K} p_{l'k'}m_{l'k'}^{l}\times
$$
$$
\mathrm{tr}\left(\mathbf{R}_{lk}^{l}\boldsymbol{\Psi}_{lk}^{l}\mathbf{R}_{lk}^{l}\mathbf{R}_{l'k'}^{l}\right)+\sum_{(l',k')\in\mathcal{P}_{lk}} p_{l'k'}\frac{z_{l'k'}^{l}}{(d_{l'k'}^{l})^2}\times
$$
$$
\left((d_{l'k'}^{l})^2+\frac{\mathrm{tr}\left((\widetilde{\mathbf{R}}_{l'k'}^{l})^2\right)}{(S_{l'k'}^{l})^2}\right)\left|\mathrm{tr}\left(\mathbf{R}_{l'k'}^{l}\boldsymbol{\Psi}_{lk}^{l}\mathbf{R}_{lk}^{l}\right)\right|^2+\sum_{(l',k')\in\mathcal{P}_{lk}} p_{l'k'}
$$
$$
\times z_{l'k'}^{l}\frac{\mathrm{tr}\left((\widetilde{\mathbf{R}}_{l'k'}^{l})^2\right)}{(S_{l'k'}^{l})^2}\mathrm{tr}\left(\mathbf{R}_{l'k'}^{l}\boldsymbol{\Psi}_{lk}^{l}\mathbf{R}_{lk}^{l}\mathbf{R}_{l'k'}^{l}\mathbf{R}_{lk}^{l}\boldsymbol{\Psi}_{lk}^{l}\right).
$$
$$(59)$$

Utilizing (51) and (59) into (14) together with doing some algebra, we obtain the closed-form SINR expression as in the theorem.

D. Proof of Theorem 2

We begin with dividing the numerator and denominator of the SINR expression (15) by $M\mathrm{tr}\left(\mathbf{R}_{lk}^{l}\boldsymbol{\Psi}_{lk}^{l}\mathbf{R}_{lk}^{l}\right)$. The numerator of (15) is $p_{lk}z_{lk}^{l}\mathrm{tr}\left(\mathbf{R}_{lk}^{l}\boldsymbol{\Psi}_{lk}^{l}\mathbf{R}_{lk}^{l}\right)/M$. Meanwhile, the first part in the denominator of (15) becomes

$$
\frac{\mathrm{NI}_{lk}}{M\mathrm{tr}\left(\mathbf{R}_{lk}^{l}\boldsymbol{\Psi}_{lk}^{l}\mathbf{R}_{lk}^{l}\right)} = \frac{\sum_{l'=1}^{L}\sum_{k'=1}^{K} p_{l'k'}m_{l'k'}^{l}\mathrm{tr}\left(\mathbf{R}_{lk}^{l}\boldsymbol{\Psi}_{lk}^{l}\mathbf{R}_{lk}^{l}\mathbf{R}_{l'k'}^{l}\right)}{M\mathrm{tr}\left(\mathbf{R}_{lk}^{l}\boldsymbol{\Psi}_{lk}^{l}\mathbf{R}_{lk}^{l}\right)}
$$
$$
\overset{(a)}{\leq}\frac{1}{M}\sum_{l'=1}^{L}\sum_{k'=1}^{K} p_{l'k'}m_{l'k'}^{l}\|\mathbf{R}_{l'k'}^{l}\|_2
$$
$$
\leq\frac{LK}{M}\max_{(l',k')} p_{l'k'}m_{l'k'}^{l}\|\mathbf{R}_{l'k'}^{l}\|_2,
$$
$$(60)$$

where (a) is obtained by the upper bound of the trace matrix expression [33, Lemma B.7]. By applying Assumption 1 to the last result (60), we observe that this part converges to zero as either $M \to \infty$ or $S_{l'k'}^l \to \infty$. It is also straightforward to prove that the last part in the denominator of the SINR expression (15) converges to zero as either $M \to \infty$ or $S_{l'k'}^l \to \infty$, i.e.,

$$\frac{\mathsf{NI}_{lk}}{M\mathrm{tr}(\mathbf{R}_{lk}^l \mathbf{\Psi}_{lk}^l \mathbf{R}_{lk}^l)} \to 0. \tag{61}$$

Combining (60) and (61), the denominator of (15) is formulated as CI_{lk}, and therefore the asymptotic SINR expression as $M \to \infty$ for a given finite set of the scatterers and covariance matrices as shown in (26).

When $\mathbf{R}_{lk}^l$ is asymptotically orthogonal with all the other covariance matrices of the users sharing the same pilot signal as user k in cell l, the second part in the denominator of (15) converges to as

$$\frac{\mathsf{CI}_{lk}}{M\mathrm{tr}(\mathbf{R}_{lk}^l \mathbf{\Psi}_{lk}^l \mathbf{R}_{lk}^l)} \to \frac{p_{lk} z_{lk}^l \mathrm{tr}\left((\widetilde{\mathbf{R}}_{lk}^l)^2\right) \mathrm{tr}(\mathbf{R}_{lk}^l \mathbf{\Psi}_{lk}^l \mathbf{R}_{lk}^l)}{M(d_{lk}^l S_{lk}^l)^2}$$
$$+ \frac{p_{lk} z_{lk}^l \mathrm{tr}\left((\widetilde{\mathbf{R}}_{lk}^l)^2\right) \mathrm{tr}\left((\mathbf{R}_{lk}^l \mathbf{\Psi}_{lk}^l \mathbf{R}_{lk}^l)^2\right)}{M(d_{lk}^l S_{lk}^l)^2 \mathrm{tr}(\mathbf{R}_{lk}^l \mathbf{\Psi}_{lk}^l \mathbf{R}_{lk}^l)}$$
$$\overset{(a)}{\leq} \frac{p_{lk} z_{lk}^l \mathrm{tr}\left((\widetilde{\mathbf{R}}_{lk}^l)^2\right) \mathrm{tr}(\mathbf{R}_{lk}^l \mathbf{\Psi}_{lk}^l \mathbf{R}_{lk}^l)}{M(d_{lk}^l S_{lk}^l)^2}$$
$$+ \frac{p_{lk} z_{lk}^l \mathrm{tr}\left((\widetilde{\mathbf{R}}_{lk}^l)^2\right) \|\mathbf{R}_{lk}^l \mathbf{\Psi}_{lk}^l \mathbf{R}_{lk}^l\|_2}{M(d_{lk}^l S_{lk}^l)^2}$$
$$\overset{(b)}{\longrightarrow} \frac{p_{lk} z_{lk}^l \mathrm{tr}\left((\widetilde{\mathbf{R}}_{lk}^l)^2\right) \mathrm{tr}(\mathbf{R}_{lk}^l \mathbf{\Psi}_{lk}^l \mathbf{R}_{lk}^l)}{M(d_{lk}^l S_{lk}^l)^2}, \tag{62}$$

where (a) is obtained by [33, Lemma B.7] and (b) is because of our assumptions on the covariance matrices. Consequently, the asymptotic uplink SE of user k in cell l is obtained as in (27).

As both the number of antennas at each BS and scatterers go without bound while the covariance matrices are non-orthogonal, the first and last parts in the denominator of (15) go to zeros, while the second part converges to as

$$\frac{\mathsf{CI}_{lk}}{M\mathrm{tr}(\mathbf{R}_{lk}^l \mathbf{\Psi}_{lk}^l \mathbf{R}_{lk}^l)} \to$$
$$\frac{\sum_{(l',k') \in \mathcal{P}_{lk} \setminus (l,k)} p_{l'k'} z_{l'k'}^l \left|\mathrm{tr}(\mathbf{R}_{l'k'}^l \mathbf{\Psi}_{lk}^l \mathbf{R}_{lk}^l)\right|^2}{M\mathrm{tr}(\mathbf{R}_{lk}^l \mathbf{\Psi}_{lk}^l \mathbf{R}_{lk}^l)}, \tag{63}$$

and hence we obtain the asymptotic SE expression as shown in (28). For the last case in (29) is obtained since the denominator of (15) goes to zeros, while the numerator goes to a constant.

E. Proof of Theorem 3

We first prove that every $I_{lk}(\mathbf{p})$ is a standard interference function as given in Definition 3. Indeed, the positivity prop-

erty is true since it holds for all $\mathbf{p} \geq \mathbf{0}$ that

$$I_{lk}(\mathbf{p}) \geq I_{lk}(\mathbf{0}) \overset{(a)}{=} \frac{\nu_{lk} \mathsf{NO}_{lk}}{z_{lk}^l \left|\mathrm{tr}\left(\mathbf{R}_{lk}^l \mathbf{\Psi}_{lk}^l \mathbf{R}_{lk}^l\right)\right|^2}$$
$$\overset{(b)}{=} \frac{\sigma^2}{\hat{p}_{lk}(\beta_{lk}^l)^2 (d_{lk}^l)^2 \tau_p \mathrm{tr}\left(\mathbf{R}_{lk}^l \mathbf{\Psi}_{lk}^l \mathbf{R}_{lk}^l\right)} > 0, \tag{64}$$

where (a) is obtained since NO_{lk} is independent of the data powers and (b) is obtained after doing some algebra. Let us denote the two vectors $\mathbf{p}$ and $\mathbf{p}'$ having $p_{lk} \geq p'_{lk}, \forall l, k$, then we obtain

$$I_{lk}(\mathbf{p}) - I_{lk}(\mathbf{p}') =$$
$$\frac{\nu_{lk}\left(\mathsf{NI}_{lk}(\mathbf{p}) - \mathsf{NI}_{lk}(\mathbf{p}')\right) + \nu_{lk}\left(\mathsf{CI}_{lk}(\mathbf{p}) - \mathsf{CI}_{lk}(\mathbf{p}')\right)}{z_{lk}^l \left|\mathrm{tr}\left(\mathbf{R}_{lk}^l \mathbf{\Psi}_{lk}^l \mathbf{R}_{lk}^l\right)\right|^2} \geq 0, \tag{65}$$

which means $I_{lk}(\mathbf{p}) \geq I_{lk}(\mathbf{p}')$ and confirms the monotonicity. For the scalability, we observe that

$$\alpha I_{lk}(\mathbf{p}) = \frac{\alpha \nu_{lk} \mathsf{NI}_{lk}(\mathbf{p}) + \alpha \nu_{lk} \mathsf{CI}_{lk}(\mathbf{p}) + \alpha \nu_{lk} \mathsf{NO}_{lk}}{z_{lk}^l \left|\mathrm{tr}\left(\mathbf{R}_{lk}^l \mathbf{\Psi}_{lk}^l \mathbf{R}_{lk}^l\right)\right|^2}$$
$$\overset{(a)}{=} \frac{\nu_{lk} \mathsf{NI}_{lk}(\alpha\mathbf{p}) + \nu_{lk} \mathsf{CI}_{lk}(\alpha\mathbf{p}) + \alpha \nu_{lk} \mathsf{NO}_{lk}}{z_{lk}^l \left|\mathrm{tr}\left(\mathbf{R}_{lk}^l \mathbf{\Psi}_{lk}^l \mathbf{R}_{lk}^l\right)\right|^2}$$
$$\overset{(b)}{\geq} \frac{\nu_{lk} \mathsf{NI}_{lk}(\alpha\mathbf{p}) + \nu_{lk} \mathsf{CI}_{lk}(\alpha\mathbf{p}) + \nu_{lk} \mathsf{NO}_{lk}}{z_{lk}^l \left|\mathrm{tr}\left(\mathbf{R}_{lk}^l \mathbf{\Psi}_{lk}^l \mathbf{R}_{lk}^l\right)\right|^2} \tag{66}$$
$$= I_{lk}(\alpha\mathbf{p}),$$

which confirms that $I_{lk}(\mathbf{p})$ satisfies the monotonicity property. Since every $I_{lk}(\mathbf{p})$ is a interference function, the update procedure in (36) guarantees: First, beginning with the initial data power values $p_{lk}(0) = P_{\max,lk}, \forall l, k$, all the updated power coefficients at iteration n are in the feasible domain. Indeed, we can prove this statement by mathematical induction following similar steps as [28, Lemma 3]. Second, the update in (36) ensures a reduction of the objective function along iterations.

F. Proof of Theorem 4

Before getting in the proof, we recall the so-called two-sided function [42]. Specifically, a function $f(\mathbf{z})$ is a two-sided scalable if for $\forall \alpha > 1$ and $\frac{1}{\alpha}\mathbf{z} \leq \hat{\mathbf{z}} \leq \alpha\mathbf{z}$, implies the following two-sided inequality

$$\frac{1}{\alpha} f(\mathbf{z}) < f(\hat{\mathbf{z}}) < \alpha f(\mathbf{z}). \tag{67}$$

We stress that the authors in [43] gave a toy example of a two-sided scalable function to update the data transmit power for a communication system under perfect channel state information. Unlike the previous works, all the functions $f_{lk}(\mathbf{p}(n-1))$ involve the complicated expressions of many effects from channel estimation, pilot contamination, non-coherent interference, and noise.

We now prove that $f_{lk}(\mathbf{p}(n-1))$ is a two-sided scalable function. If $I_{lk}(\mathbf{p}(n-1)) \leq P_{\max,lk}$, then it is sufficient

to prove that $I_{lk}\left(\mathbf{p}(n-1)\right)$ is a two-side scalable function. Indeed, we have shown in Theorem 1 that $I_{lk}\left(\mathbf{p}(n-1)\right)$ is a standard interference function. Therefore, for $\frac{1}{\alpha}\mathbf{p}_l(n-1) \preceq \hat{\mathbf{p}}(n-1) \preceq \alpha\mathbf{p}_l(n-1)$, we have:

$$I_{lk}\left(\mathbf{p}(n-1)\right) \overset{(a)}{<} I_{lk}\left(\alpha\hat{\mathbf{p}}(n-1)\right) \overset{(b)}{<} \alpha I_{lk}\left(\hat{\mathbf{p}}(n-1)\right), \quad (68)$$

where (a) is obtained by applying the monotonicity property for $\mathbf{p}(n-1) \preceq \alpha\hat{\mathbf{p}}(n-1)$; (b) is obtained by using the scalability property for $\alpha\hat{\mathbf{p}}(n-1)$. As a consequence of (68),

$$\frac{1}{\alpha}I_{lk}\left(\mathbf{p}(n-1)\right) < I_{lk}\left(\hat{\mathbf{p}}(n-1)\right). \quad (69)$$

Similarly, by applying the monotonicity and scalability properties for $\hat{\mathbf{p}}(n-1) \preceq \alpha\mathbf{p}(n-1)$, the following inequalities are obtained as

$$I_{lk}\left(\hat{\mathbf{p}}(n-1)\right) < I_{lk}\left(\alpha\mathbf{p}(n-1)\right) < \alpha I_{lk}\left(\mathbf{p}(n-1)\right), \quad (70)$$

which results in

$$I_{lk}\left(\hat{\mathbf{p}}(n-1)\right) < \alpha I_{lk}\left(\mathbf{p}(n-1)\right). \quad (71)$$

Combining (69) and (71), we attain the two-sided scalable property of $I_{lk}\left(\hat{\mathbf{p}}(n-1)\right)$ as

$$\frac{1}{\alpha}I_{lk}\left(\mathbf{p}(n-1)\right) < I_{lk}\left(\hat{\mathbf{p}}(n-1)\right) < \alpha I_{lk}\left(\mathbf{p}(n-1)\right). \quad (72)$$

We now prove that $P^2_{\max,lk}/I_{lk}\left(\hat{\mathbf{p}}(n-1)\right)$ is also a two-side scalable function. In fact, this is straightforward since $I_{lk}\left(\hat{\mathbf{p}}(n-1)\right)$ satisfies the positivity, an inversion of (72) is

$$\frac{1}{\alpha}\frac{1}{I_{lk}\left(\mathbf{p}(n-1)\right)} < \frac{1}{I_{lk}\left(\hat{\mathbf{p}}(n-1)\right)} < \alpha\frac{1}{I_{lk}\left(\mathbf{p}(n-1)\right)}. \quad (73)$$

Multiplying (73) by $P^2_{\max,lk}$, we obtain the following inequalities

$$\frac{1}{\alpha}\frac{P^2_{\max,lk}}{I_{lk}\left(\mathbf{p}(n-1)\right)} < \frac{P^2_{\max,lk}}{I_{lk}\left(\hat{\mathbf{p}}(n-1)\right)} < \alpha\frac{P^2_{\max,lk}}{I_{lk}\left(\mathbf{p}(n-1)\right)}, \quad (74)$$

which completes the proof that confirms $f_{lk}(\mathbf{p}(n-1))$ being a two-side scalable function. From the initial values $p_{lk}(0) = P_{\max,lk}, \forall l, k$, the update in (39) ensures that the iterative algorithm will converge to a fixed point.

References

[1] T. V. Chien, H. Q. Ngo, S. Chatzinotas, B. Ottersten, and M. Debbah, "Massive MIMO under double scattering channels: Power minimization and congestion controls," in *Proc. IEEE ICC*, 2021.

[2] F. Tariq, M. R. Khandaker, K.-K. Wong, M. A. Imran, M. Bennis, and M. Debbah, "A speculative study on 6G," *IEEE Wireless Commun. Mag.*, vol. 27, no. 4, pp. 118–125, 2020.

[3] T. H. Nguyen, W.-S. Jung, L. T. Tu, T. Van Chien, D. Yoo, and S. Ro, "Performance analysis and optimization of the coverage probability in dual hop LoRa networks with different fading channels," *IEEE Access*, vol. 8, pp. 107 087–107 102, 2020.

[4] C. V. N. Index, "Cisco visual networking index: Global mobile data traffic forecast update, 2017-2022 white paper," *Cisco: San Jose, CA, USA*, 2019.

[5] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis, and T. L. Marzetta, "Massive MIMO is a reality—what is next?: Five promising research directions for antenna arrays," *Digital Signal Processing*, vol. 94, pp. 3–20, 2019.

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

[7] 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, 2013.

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

[9] T. V. Chien, E. Björnson, and E. G. Larsson, "Joint power allocation and user association optimization for Massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6384 – 6399, 2016.

[10] T. Van Chien, C. Mollén, and E. Björnson, "Large-scale-fading decoding in cellular Massive MIMO systems with spatially correlated channels," *IEEE Trans. Commun.*, vol. 67, no. 4, pp. 2746 – 2762, 2019.

[11] A. Kammoun, A. Müller, E. Björnson, and M. Debbah, "Linear precoding based on polynomial expansion: Large-scale multi-cell MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 8, no. 5, pp. 861–875, 2014.

[12] Q.-U.-A. Nadeem, A. Kammoun, M. Debbah, and M.-S. Alouini, "A generalized spatial correlation model for 3D MIMO channels based on the fourier coefficients of power spectrums," *IEEE Trans. Signal Process.*, vol. 63, no. 14, pp. 3671–3686, 2015.

[13] K. Yu and B. Ottersten, "Models for MIMO propagation channels: A review," *Wireless communications and mobile computing*, vol. 2, no. 7, pp. 653–666, 2002.

[14] D. Gesbert, H. Bölcskei, D. Gore, and A. Paulraj, "Outdoor MIMO wireless channels: Models and performance prediction," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 1926–1934, 2002.

[15] H. Q. Ngo and E. G. Larsson, "No downlink pilots are needed in TDD massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2921–2935, 2017.

[16] J. A. C. Sutton, H. Q. Ngo, and M. Matthaiou, "Hardening the channels by precoder design in Massive MIMO with multiple-antenna users," *IEEE Trans. Veh. Technol.*, 2020.

[17] T. Van Chien, E. Björnson, and E. G. Larsson, "Multi-cell Massive MIMO performance with double scattering channels," in *Proc. IEEE CAMAD*. IEEE, 2016, pp. 231–236.

[18] A. Kammoun, M. Debbah, and M.-S. Alouini, "Asymptotic analysis of RZF over double scattering channels with MMSE estimation," *IEEE Trans. Wireless Commun.*, vol. 18, no. 5, pp. 2509–2526, 2019.

[19] J. Ye, Q. Nadeem, A. Kammoun, and M. Alouini, "Asymptotic analysis of MRT over double scattering channels with MMSE estimation," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 7851–7863, 2020.

[20] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends® in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017. [Online]. Available: http://dx.doi.org/10.1561/2000000093

[21] T. V. Chien, E. Björnson, and E. G. Larsson, "Joint pilot sequence design and power control for max-min fairness in uplink Massive MIMO," in *Proc. IEEE ICC*, 2017.

[22] S. Jin, M. Li, Y. Huang, Y. Du, and X. Gao, "Pilot scheduling schemes for multi-cell massive multiple-input-multiple-output transmission," *IET Communications*, vol. 9, no. 5, pp. 689–700, 2015.

[23] T. H. Nguyen, T. K. Nguyen, H. D. Han, and V. D. Nguyen, "Optimal power control and load balancing for uplink cell-free multi-user Massive MIMO," *IEEE Access*, vol. 6, pp. 14 462–14 473, 2018.

[24] R. Sun, M. Hong, and Z. Q. Luo, "Joint downlink base station association and power control for max-min fairness: Computation and complexity," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 6, pp. 1040–1054, 2015.

[25] A. Ghazanfari, H. V. Cheng, E. Björnson, and E. G. Larsson, "A fair and scalable power control scheme in multi-cell Massive MIMO," in *Proc. ICASSP*. IEEE, 2019, pp. 4499–4503.

[26] T. Van Chien, E. Björnson, and E. G. Larsson, "Optimal design of energy-efficient Cell-free Massive MIMO: Joint power allocation and load balancing," in *Proc. IEEE ICASSP*, 2019.

[27] H. Q. Ngo, L.-N. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, "On the total energy efficiency of cell-free Massive MIMO," *IEEE Transactions on Green Communications and Networking*, vol. 2, no. 1, pp. 25–39, 2018.

[28] T. Van Chien, E. Björnson, and H. Q. Ngo, "Uplink power control in cellular Massive MIMO systems: Coping with the congestion issue," in *Proc. Of ICC Workshops*. IEEE, 2020, pp. 1–6.

[29] M. Patzold, *Mobile fading channels*. John Wiley & Sons, Inc., 2003.

[30] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.

[31] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.

[32] T. V. Chien and H. Q. Ngo, "Massive MIMO channels," in *Antennas and Propagation for 5G and Beyond*. IET Publisher, 2020, pp. 1–33.

[33] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends in Signal Processing*, vol. 11, no. 3-4, pp. 154 – 655, 2017.

[34] 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, 2013.

[35] Ö. Özdogan, E. Björnson, and E. G. Larsson, "Massive MIMO with spatially correlated Rician fading channels," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3234–3250, 2019.

[36] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[37] CVX Research Inc., "CVX: Matlab software for disciplined convex programming, academic users," http://cvxr.com/cvx, 2015.

[38] R. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1341–1347, 1995.

[39] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation (Release 9)*. 3GPP TS 36.213, Sep. 2010.

[40] K. Senel, E. Björnson, and E. G. Larsson, "Joint transmit and circuit power minimization in Massive MIMO with downlink SINR constraints: When to turn on Massive MIMO?" *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, pp. 1834 – 1846, 2019.

[41] T. Van Chien, E. Björnson, and E. G. Larsson, "Joint power allocation and load balancing optimization for energy-efficient Cell-free Massive MIMO networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6798 – 6812, 2020.

[42] C. W. Sung and K.-K. Leung, "A generalized framework for distributed power control in wireless networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2625–2635, 2005.

[43] M. Rasti and A. R. Sharafat, "Distributed uplink power control with soft removal for wireless networks," *IEEE Trans. Commun.*, vol. 59, no. 3, pp. 833–843, 2010.

Hien Quoc Ngo received the B.S. degree in electrical engineering from the Ho Chi Minh City University of Technology, Vietnam, in 2007, the M.S. degree in electronics and radio engineering from Kyung Hee University, South Korea, in 2010, and the Ph.D. degree in communication systems from Linköping University (LiU), Sweden, in 2015. In 2014, he visited the Nokia Bell Labs, Murray Hill, New Jersey, USA. From January 2016 to April 2017, Hien Quoc Ngo was a VR researcher at the Department of Electrical Engineering (ISY), LiU. He was also a Visiting Research Fellow at the School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, UK, funded by the Swedish Research Council.

Hien Quoc Ngo is currently a Reader (Associate Professor) at Queen's University Belfast, UK. His main research interests include massive (large-scale) MIMO systems, cell-free massive MIMO, physical layer security, and cooperative communications. He has co-authored many research papers in wireless communications and co-authored the Cambridge University Press textbook *Fundamentals of Massive MIMO* (2016).

Dr. Hien Quoc Ngo received the IEEE ComSoc Stephen O. Rice Prize in Communications Theory in 2015, the IEEE ComSoc Leonard G. Abraham Prize in 2017, and the Best PhD Award from EURASIP in 2018. He also received the IEEE Sweden VT-COM-IT Joint Chapter Best Student Journal Paper Award in 2015. He was an *IEEE Communications Letters* exemplary reviewer for 2014, an *IEEE Transactions on Communications* exemplary reviewer for 2015, and an *IEEE Wireless Communications Letters* exemplary reviewer for 2016. He was awarded the UKRI Future Leaders Fellowship in 2019. Dr. Hien Quoc Ngo currently serves as an Editor for the IEEE Transactions on Wireless Communications, IEEE Wireless Communications Letters, Digital Signal Processing, Elsevier Physical Communication (PHYCOM), and IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences. He was a Guest Editor of IET Communications, special issue on "Recent Advances on 5G Communications" and a Guest Editor of IEEE Access, special issue on "Modelling, Analysis, and Design of 5G Ultra-Dense Networks", in 2017. He has been a member of Technical Program Committees for several IEEE conferences such as ICC, GLOBECOM, WCNC, and VTC.

Trinh Van Chien (S'16-M'20) received the B.S. degree in Electronics and Telecommunications from Hanoi University of Science and Technology (HUST), Vietnam, in 2012. He then received the M.S. degree in Electrical and Computer Enginneering from Sungkyunkwan University (SKKU), Korea, in 2014 and the Ph.D. degree in Communication Systems from Linköping University (LiU), Sweden, in 2020. He is now a research associate at University of Luxembourg. His interest lies in convex optimization problems and machine learning applications for wireless communications and image & video processing. He was an IEEE wireless communications letters exemplary reviewer for 2016 and 2017. He also received the award of scientific excellence in the first year of the 5Gwireless project funded by European Union Horizon's 2020.

Symeon Chatzinotas is currently Full Professor / Chief Scientist I and Head of the SIGCOM Research Group at SnT, University of Luxembourg. He is co-ordinating the research activities on communications and networking, acting as a PI for more than 20 projects and main representative for 3GPP, ETSI, DVB.

In the past, he has been a Visiting Professor at the University of Parma, Italy, lecturing on "5G Wireless Networks". He was involved in numerous R&D projects for NCSR Demokritos, CERTH Hellas and CCSR, University of Surrey.

He was the co-recipient of the 2014 IEEE Distinguished Contributions to Satellite Communications Award and Best Paper Awards at EURASIP JWCN, CROWNCOM, ICSSC. He has (co-)authored more than 450 technical papers in refereed international journals, conferences and scientific books.

He is currently in the editorial board of the IEEE Transactions on Communications, IEEE Open Journal of Vehicular Technology and the International Journal of Satellite Communications and Networking.

Björn Ottersten (S'87–M'89–SM'99–F'04) received the M.S. degree in electrical engineering and applied physics from Linköping University, Linköping, Sweden, in 1986, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 1990. He has held research positions with the Department of Electrical Engineering, Linköping University, the Information Systems Laboratory, Stanford University, the Katholieke Universiteit Leuven, Leuven, Belgium, and the University of Luxembourg, Luxembourg. From 1996 to 1997, he was the Director of Research with ArrayComm, Inc., a start-up in San Jose, CA, USA, based on his patented technology. In 1991, he was appointed Professor of signal processing with the Royal Institute of Technology (KTH), Stockholm, Sweden. Dr. Ottersten has been Head of the Department for Signals, Sensors, and Systems, KTH, and Dean of the School of Electrical Engineering, KTH. He is currently the Director for the Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg. He is a recipient of the IEEE Signal Processing Society Technical Achievement Award, the EURASIP Group Technical Achievement Award, and the European Research Council advanced research grant twice. He has co-authored journal papers that received the IEEE Signal Processing Society Best Paper Award in 1993, 2001, 2006, 2013, and 2019, and 8 IEEE conference papers best paper awards. He has been a board member of IEEE Signal Processing Society, the Swedish Research Council and currently serves of the boards of EURASIP and the Swedish Foundation for Strategic Research. Dr. Ottersten has served as Editor in Chief of EURASIP Signal Processing, and acted on the editorial boards of IEEE Transactions on Signal Processing, IEEE Signal Processing Magazine, IEEE Open Journal for Signal Processing, EURASIP Journal of Advances in Signal Processing and Foundations and Trends in Signal Processing. He is a fellow of EURASIP.

Mérouane Debbah received the M.Sc. and Ph.D. degrees from the Ecole Normale Supérieure Paris-Saclay, France. He was with Motorola Labs, Saclay, France, from 1999 to 2002, and also with the Vienna Research Center for Telecommunications, Vienna, Austria, until 2003. From 2003 to 2007, he was an Assistant Professor with the Mobile Communications Department, Institut Eurecom, Sophia Antipolis, France. In 2007, he was appointed Full Professor at CentraleSupelec, Gif-sur-Yvette, France. From 2007 to 2014, he was the Director of the Alcatel-Lucent Chair on Flexible Radio. From 2014 to 2021, he was Vice-President of the Huawei France Research Center. He was jointly the director of the Mathematical and Algorithmic Sciences Lab as well as the director of the Lagrange Mathematical and Computing Research Center. Since 2021, he is Chief Research Officer at the Technology Innovation Institute in Abu Dhabi. He leads jointly the AI and Telecommunication centers. He has managed 8 EU projects and more than 24 national and international projects. His research interests lie in fundamental mathematics, algorithms, statistics, information, and communication sciences research. He is an IEEE Fellow, a WWRF Fellow, a Eurasip Fellow, an Institut Louis Bachelier Fellow and a Membre émérite SEE. He was a recipient of the ERC Grant MORE (Advanced Mathematical Tools for Complex Network Engineering) from 2012 to 2017. He was a recipient of the Mario Boella Award in 2005, the IEEE Glavieux Prize Award in 2011, the Qualcomm Innovation Prize Award in 2012, the 2019 IEEE Radio Communications Committee Technical Recognition Award and the 2020 SEE Blondel Medal. He received more than 20 best paper awards, among which the 2007 IEEE GLOBECOM Best Paper Award, the Wi-Opt 2009 Best Paper Award, the 2010 Newcom++ Best Paper Award, the WUN CogCom Best Paper 2012 and 2013 Award, the 2014 WCNC Best Paper Award, the 2015 ICC Best Paper Award, the 2015 IEEE Communications Society Leonard G. Abraham Prize, the 2015 IEEE Communications Society Fred W. Ellersick Prize, the 2016 IEEE Communications Society Best Tutorial Paper Award, the 2016 European Wireless Best Paper Award, the 2017 Eurasip Best Paper Award, the 2018 IEEE Marconi Prize Paper Award, the 2019 IEEE Communications Society Young Author Best Paper Award, the 2021 Eurasip Best Paper Award, the 2021 IEEE Marconi Prize Paper Award as well as the Valuetools 2007, Valuetools 2008, CrownCom 2009, Valuetools 2012, SAM 2014, and 2017 IEEE Sweden VT-COM-IT Joint Chapter best student paper awards. He is an Associate Editor-in-Chief of the journal Random Matrix: Theory and Applications. He was an Associate Area Editor and Senior Area Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2011 to 2013 and from 2013 to 2014, respectively. From 2021 to 2022, he serves as an IEEE Signal Processing Society Distinguished Industry Speaker