electronics

Article

Energy Efficiency Optimization for Massive MIMO Non-Orthogonal Unicast and Multicast Transmission with Statistical CSI

Wenjin Wang *, Yufei Huang, Li You, Jiayuan Xiong, Jiamin Li and Xiqi Gao

National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China
* Correspondence: wangwj@seu.edu.cn; Tel.: +86-025-8379-0506

Received: 16 July 2019; Accepted: 30 July 2019; Published: 1 August 2019

Abstract: We study the energy efficiency (EE) optimization problem in non-orthogonal unicast and multicast transmission for massive multiple-input multiple-output (MIMO) systems with statistical channel state information of all receivers available at the transmitter. Firstly, we formulate the EE maximization problem. We reduce the number of variables to be solved and simplify this large-dimensional-matrix-valued problem into a real-vector-valued problem. Next, we lower the computational complexity significantly by replacing the objective with its deterministic equivalent to avoid the high-complex expectation operation. With guaranteed convergence, we propose an iterative algorithm on beam domain power allocation using the minorize maximize algorithm and Dinkelbach's transform and derive the locally optimal power allocation strategy to achieve the optimal EE. Finally, we illustrate the significant EE performance gain of our EE maximization algorithm compared with the conventional approach through conducting numerical simulations.

Keywords: energy efficiency; non-orthogonal unicast and multicast transmission; statistical channel state information; massive MIMO; beam domain

1. Introduction

As mobile data expands rapidly, it is expected that global wireless data traffic will surpass 100 exabytes per month by 2023 [1]. A considerable proportion of the data traffic, such as massive software updating and sports broadcasting, is of common interest, which stimulates the demand for services that can deliver the same data to a group of user terminals (UTs) efficiently. Since physical layer multicasting can provide efficient point-to-multipoint wireless transmission, it has great potential for future mobile communication systems [2–4].

Recently, non-orthogonal unicast and multicast (NOUM) transmission has been gaining increasing interest [5–7]. At the transmitter, the unicast and multicast signals are precoded and then sent out to the receivers simultaneously, sharing the same time-frequency resources. Compared with the conventional orthogonal unicast and multicast (OUM) transmission, NOUM transmission is more spectrum-efficient, and more suitable for scenarios where both multicast and unicast signals are needed by a UT. Massive multiple-input multiple-output (MIMO) has become one of the core technologies of the fifth generation wireless system for its significant performance in energy efficiency (EE) and spectral efficiency [8,9]. Therefore, there has been considerable research on the combination of multicast transmission and massive MIMO systems [6,10,11]. Please note that mutual coupling is a major concern in massive MIMO because it can weaken the system performance [12–16]. We assume perfect isolation between the antennas without loss of generality.

EE has become a significant design criterion for wireless communication systems [17–19]. The broad-scale antenna arrays equipped at the base station (BS) cause the power consumption to increase in massive MIMO systems, and the energy consumed by wireless communications is responsible for greenhouse gas emissions [20], which motivates the need to design energy-efficient systems [8,21,22]. EE of a massive MIMO system was considered in [8]. However, it ignores the power consumed by the BS circuit, while in [21], research on maximizing the EE and power transfer efficiency for wireless-powered systems was analyzed, taking the circuit power consumption into account. In [22], how the system parameters (number of antennas, transmitted power and number of UTs) affect the EE of a multi-user MIMO system was investigated.

There are also previous works that studied energy-efficient NOUM transmission in massive MIMO systems [23–25]. In [23], energy-efficient NOUM beamforming in multi-cell multi-user MIMO scenario was studied. An optimization beamforming algorithm was proposed in [24] to optimize the EE in the multi-cell multicast system. The extension of the problem was investigated in [25], which takes antenna selection into consideration.

Please note that most of the previous works made the assumption that the UTs' instantaneous channel state information (CSI) is available at the BS. However, in realistic systems, obtaining good estimates of instantaneous CSI is a challenging job [26–28]. Compared with obtaining instantaneous CSI, the acquisition of statistical CSI is easier and more precise. In [11], rate maximization problem for NOUM massive MIMO transmission was considered, and the EE maximization problem for physical-layer multicast transmission was investigated in [29], both assumed that the BS only has access to the UTs' statistical CSI.

To our knowledge, the research on EE optimization of NOUM transmission for massive MIMO systems with statistical CSI at the transmitter has not been studied yet. We investigate this problem in our work, and the major contributions we provide in this paper are listed as follows:

- With statistical CSI, we formulate the EE maximization problem for NOUM transmission in the massive MIMO scenario.
- We determine the optimal transmit directions of the multicast and unicast transmission in closed-form, respectively, and then simplify the large-scale complex-matrix-valued precoding design problem into a real-vector-valued power allocation problem in the beam domain.
- We reduce the computational complexity of the EE optimization problem significantly by replacing the objective function with its deterministic equivalent (DE).
- With guaranteed convergence, we propose an algorithm on beam domain power allocation using the minorize maximize (MM) algorithm and Dinkelbach's transform. We deal with the EE optimization problem by iteratively solving a series of related convex optimization problems.

The remainder of the paper is constructed as follows. The channel model is introduced in Section 2. The EE maximization problem is formulated and investigated in Section 3. Numerical simulations are conducted in Section 4. Section 5 summarizes the paper.

Column vectors and matrices are represented by lower and upper case boldface letters, respectively, whereas italic letters stand for scalars, and the following are other notations used in this paper.

- We adopt $\mathbb{R}^{M\times N}$ to represent $M \times N$ real-valued vector space and $\mathbb{C}^{M\times N}$ to denote $M \times N$ complex-valued vector space.
- $\mathbf{I}_M$ represents the identity matrix of size $M \times M$.
- $\mathbf{X} \succeq \mathbf{0}$ indicates that matrix $\mathbf{X}$ is positive semidefinite.
- $\mathsf{E}\{.\}$ represents the expectation operation.
- $\odot$ denotes the Hadamard product.
- Denote $\mathrm{tr}\{.\}$ as the trace operation, $(.)^T$ as the transpose operation, $(.)^*$ as the conjugate operation, $(.)^H$ as the conjugate-transpose operation, and $\det\{.\}$ as the determinant operation.
- $\sim$ stands for "be distributed as", and $\triangleq$ stands for "be defined as".

2. System Model

Consider a single cell massive MIMO system with an M-antenna BS, jointly serves K UTs. Denote by $\mathcal{K} \triangleq \{1, 2, \ldots, K\}$ the UT set, where the kth UT is equipped with N_k antennas. The multicast and unicast services are carried out with the same time-frequency resources. The BS sends a multicast signal that is of common interest to all the UTs in the cell while delivering unique messages to UTs according to each UT's demand during the downlink transmission, as shown in Figure 1.

Figure 1. System model of NOUM.

Assume the downlink signal sent by the BS is denoted by

$$\mathbf{x} = \mathbf{x}^{\mathrm{m}} + \sum_{k \in \mathcal{K}} \mathbf{x}_k^{\mathrm{u}} \in \mathbb{C}^{M \times 1}, \tag{1}$$

where $\mathbf{x}^{\mathrm{m}} \in \mathbb{C}^{M \times 1}$ represents the multicast signal and $\mathbf{x}_k^{\mathrm{u}} \in \mathbb{C}^{M \times 1}$ denotes the unicast signal sent to the kth UT. Assume that $\mathbf{x}^{\mathrm{m}}$ and $\mathbf{x}_k^{\mathrm{u}}$ are mutually uncorrelated, zero-mean, and their covariance matrices are $\mathbf{Q}^{\mathrm{m}}$ and $\mathbf{Q}_k^{\mathrm{u}}$, respectively. Define $\mathrm{tr}\left\{\mathbf{Q}^{\mathrm{m}}\right\}$ as the multicast transmission power and $\mathrm{tr}\left\{\mathbf{Q}_k^{\mathrm{u}}\right\}$ as the unicast transmission power. At the kth UT, the received signal is denoted by

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x} + \mathbf{n}_k \in \mathbb{C}^{N_k \times 1}, \tag{2}$$

where $\mathbf{H}_k$ is the downlink channel matrix of size $N_k \times M$, and $\mathbf{n}_k \sim \mathcal{CN}\left(\mathbf{0}, \sigma^2 \mathbf{I}_{N_k}\right)$ represents the additive circularly symmetric complex-valued Gaussian noise with the variance being σ^2.

We adopt Weichselberger's channel model [30,31] in our work because the correlation properties between the transmit and receive ends of Weichselberger's channel model are jointly considered rather than separately characterized in the Kronecker model. Then, we can write the downlink channel matrix in (2) as

$$\mathbf{H}_k = \mathbf{U}_k \mathbf{G}_k \mathbf{V}_k^H \in \mathbb{C}^{N_k \times M}, \tag{3}$$

where $\mathbf{U}_k \in \mathbb{C}^{N_k \times N_k}$ and $\mathbf{V}_k \in \mathbb{C}^{M \times M}$ are deterministic unitary matrices. $\mathbf{G}_k \in \mathbb{C}^{N_k \times M}$ represents the downlink channel matrix in the beam domain [26,27,32], and the elements of $\mathbf{G}_k$ are independently distributed random variables with zero-mean. Denote $\mathbf{\Omega}_k$ as the beam domain channel power matrix

$$\mathbf{\Omega}_k = \mathsf{E}\left\{\mathbf{G}_k \odot \mathbf{G}_k^*\right\} \in \mathbb{R}^{N_k \times M}, \tag{4}$$

where the average power of $[\mathbf{G}_k]_{i,j}$ is represented by $[\mathbf{\Omega}_k]_{i,j}$. As $\mathbf{\Omega}_k$ has the property of remaining approximately constant while the frequency changes widely, the statistical CSI can be obtained accurately and efficiently [32].

The vast number of antenna arrays employed at the BS brings about new channel properties for massive MIMO systems. For example, as the BS antenna number M tends to infinity, the eigenvector matrices of the transmit correlation matrices between the BS and all UTs tend to be the same and are only affected by the BS array topology [32,33]. Denote the corresponding deterministic unitary matrix as $\mathbf{V}$, and then in the massive MIMO scenario, the downlink channel matrix becomes

$$\mathbf{H}_k \overset{M\to\infty}{=} \mathbf{U}_k\mathbf{G}_k\mathbf{V}^H. \tag{5}$$

Please note that many of the previous works on massive MIMO adopted the channel model mentioned in (5) such as [26,29,34], and it can achieve quite accurate performance [34].

3. NOUM Transmission in Massive MIMO

3.1. Problem Formulation

Consider a NOUM massive MIMO system. We assume that there is only one multicast group without loss of generality. Consider the case when the kth $(\forall k)$ UT knows its own instantaneous CSI with proper pilot design [33], while the BS only has access to the statistic CSI of all UTs.

Rewrite the received signal at the kth UT by inserting (1) into (2) as follows

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}^{\mathrm{m}} + \mathbf{H}_k\mathbf{x}_k^{\mathrm{u}} + \sum_{k'\neq k} \mathbf{H}_k\mathbf{x}_{k'}^{\mathrm{u}} + \mathbf{n}_k. \tag{6}$$

Each UT will decode the common multicast signal and its desired unicast signal in order by applying successive interference cancellation (SIC) method.

During the process of multicast decoding, the kth UT regards the term $\mathbf{H}_k\mathbf{x}^{\mathrm{m}}$ in (6) as the desired message while treating the others as interference. For the covariance matrix of the interference and noise, we have

$$\mathbf{K}_k^{\mathrm{m}} = \underbrace{\sigma^2\mathbf{I}_{N_k}}_{\text{noise}} + \underbrace{\sum_{k'\in\mathcal{K}} \mathsf{E}\left\{\mathbf{H}_k\mathbf{Q}_{k'}^{\mathrm{u}}\mathbf{H}_k^H\right\}}_{\text{interference}} \in \mathbb{C}^{N_k\times N_k}. \tag{7}$$

Since UT k has the knowledge of its own instantaneous CSI and the covariance matrix $\mathbf{K}_k^{\mathrm{m}}$, during the multicast transmission, we denote by R_k^{m} the ergodic rate of the kth UT

$$R_k^{\mathrm{m}} = \mathsf{E}\left\{\log\det\left\{\mathbf{K}_k^{\mathrm{m}} + \mathbf{H}_k\mathbf{Q}^{\mathrm{m}}\mathbf{H}_k^H\right\}\right\} - \log\det\left\{\mathbf{K}_k^{\mathrm{m}}\right\}. \tag{8}$$

Denote multicast ergodic rate as $\min\limits_k R_k^{\mathrm{m}}$. By inserting the massive system model in (5) and $\det\{\mathbf{I}+\mathbf{MN}\} = \det\{\mathbf{I}+\mathbf{NM}\}$, the Sylvester's determinant identity, into (8), the multicast rate R_k^{m} in (8) becomes

$$R_k^{\mathrm{m}} = \mathsf{E}\left\{\log\det\left\{\overline{\mathbf{K}}_k^{\mathrm{m}} + \mathbf{G}_k\mathbf{V}^H\mathbf{Q}^{\mathrm{m}}\mathbf{V}\mathbf{G}_k^H\right\}\right\} - \log\det\left\{\overline{\mathbf{K}}_k^{\mathrm{m}}\right\}, \tag{9}$$

where $\overline{\mathbf{K}}_k^{\mathrm{m}}$ is defined as

$$\begin{aligned}\overline{\mathbf{K}}_k^{\mathrm{m}} &\triangleq \mathbf{U}_k^H\mathbf{K}_k^{\mathrm{m}}\mathbf{U}_k \\ &= \sigma^2\mathbf{I}_{N_k} + \sum_{k'\in\mathcal{K}} \mathsf{E}\left\{\mathbf{G}_k\mathbf{V}^H\mathbf{Q}_{k'}^{\mathrm{u}}\mathbf{V}\mathbf{G}_k^H\right\} \in \mathbb{C}^{N_k\times N_k}.\end{aligned} \tag{10}$$

Define a matrix-valued function $\mathbf{A}_k(\mathbf{X})$ by $\mathbf{A}_k(\mathbf{X}) \triangleq \mathsf{E}\left\{\mathbf{G}_k\mathbf{X}\mathbf{G}_k^H\right\}$. Since all the elements of $\mathbf{G}_k$ are zero-mean and independently distributed, the off-diagonal elements of $\mathbf{A}_k(\mathbf{X})$ are zero, so $\mathbf{A}_k(\mathbf{X})$ is a diagonal matrix-valued function of size $N_k \times N_k$, and its ith diagonal element is

$$[\mathbf{A}_k]_{i,i} = \mathrm{tr}\left\{\mathsf{diag}\left\{\left([\boldsymbol{\Omega}_k]_{i,:}\right)^T\right\}\mathbf{X}\right\}. \tag{11}$$

Then the terms $\mathsf{E}\left\{\mathbf{G}_k\mathbf{V}^H\mathbf{Q}_{k'}^{\mathrm{u}}\mathbf{V}\mathbf{G}_k^H\right\}$ in (10) can be rewritten as $\mathbf{A}_k\left(\mathbf{V}^H\mathbf{Q}_{k'}^{\mathrm{u}}\mathbf{V}\right)$.

For the unicast signal decoding, with SCI, the multicast signal is removed, so the interference only contains the unicast signal meant for other UTs. For the covariance matrix of the interference and noise at the kth UT, we have

$$\mathbf{K}_k^{\mathrm{u}} = \underbrace{\sigma^2 \mathbf{I}_{N_k}}_{\text{noise}} + \underbrace{\sum_{k' \neq k} \mathsf{E}\left\{\mathbf{H}_k \mathbf{Q}_{k'}^{\mathrm{u}} \mathbf{H}_k^H\right\}}_{\text{interference}} \in \mathbb{C}^{N_k \times N_k}. \tag{12}$$

Then we denote by R_k^{u} the ergodic rate of the kth UT during the unicast transmission

$$R_k^{\mathrm{u}} = \mathsf{E}\left\{\log\det\left\{\mathbf{K}_k^{\mathrm{u}} + \mathbf{H}_k \mathbf{Q}_k^{\mathrm{u}} \mathbf{H}_k^H\right\}\right\} - \log\det\left\{\mathbf{K}_k^{\mathrm{u}}\right\}. \tag{13}$$

By inserting the massive system model in (5) and the Sylvester's determinant identity into (13), the unicast rate R_k^{u} at the kth UT becomes

$$R_k^{\mathrm{u}} = \mathsf{E}\left\{\log\det\left\{\overline{\mathbf{K}}_k^{\mathrm{u}} + \mathbf{G}_k \mathbf{V}^H \mathbf{Q}_k^{\mathrm{u}} \mathbf{V} \mathbf{G}_k^H\right\}\right\} - \log\det\left\{\overline{\mathbf{K}}_k^{\mathrm{u}}\right\}, \tag{14}$$

where $\overline{\mathbf{K}}_k^{\mathrm{u}}$ is defined as

$$\overline{\mathbf{K}}_k^{\mathrm{u}} \triangleq \sigma^2 \mathbf{I}_{N_k} + \sum_{k' \neq k} \mathbf{A}_k\left(\mathbf{V}^H \mathbf{Q}_{k'}^{\mathrm{u}} \mathbf{V}\right) \in \mathbb{C}^{N_k \times N_k}, \tag{15}$$

and the definition of $\mathbf{A}_k(\mathbf{X})$ is given in (11).

Next, we consider the system power consumption. Apply the power consumption model the same as the one used in [29,35] as follows

$$P = \mu\left(\mathrm{tr}\left\{\mathbf{Q}^{\mathrm{m}}\right\} + \sum_{k \in \mathcal{K}} \mathrm{tr}\left\{\mathbf{Q}_k^{\mathrm{u}}\right\}\right) + MP_{\mathrm{c}} + P_{\mathrm{s}}, \tag{16}$$

where the constant-coefficient $\mu \geq 1$ accounts for the reciprocal of the transmit amplifier drain efficiency. $\mathrm{tr}\left\{\mathbf{Q}^{\mathrm{m}}\right\}$ means the multicast transmit power, and $\sum_{k \in \mathcal{K}} \mathrm{tr}\left\{\mathbf{Q}_k^{\mathrm{u}}\right\}$ denotes the total unicast transmit power. P_{c} stands for the constant circuit power consumption per antenna and is unaffected by the actual transmit power. P_{s} represents the BS static power consumption and is irrelevant to the number of antennas.

In the following, we formulate the EE optimization problem for NOUM massive MIMO system. We aim at identifying the optimal transmit covariance matrices $\mathbf{Q}^{\mathrm{m}}$ and $\mathbf{Q}_k^{\mathrm{u}}$ for multicast and unicast transmission that can maximize the system EE, respectively. We define a weight matrix $\mathbf{u} = [u_0, u_1, \ldots, u_K]$ with u_0 being the weight of multicast rate and u_k being the weight of kth unicast rate. Then we can denote by R the weighted sum rate as follows:

$$R \triangleq u_0 K(\min_k R_k^{\mathrm{m}}) + \sum_{k \in \mathcal{K}} u_k R_k^{\mathrm{u}}, \tag{17}$$

and the EE of the considered system with bandwidth W is given by

$$EE = \frac{WR}{P} = \frac{W\left(u_0 K(\min_k R_k^{\mathrm{m}}) + \sum_{k \in \mathcal{K}} u_k R_k^{\mathrm{u}}\right)}{\mu\left(\mathrm{tr}\left\{\mathbf{Q}^{\mathrm{m}}\right\} + \sum_{k \in \mathcal{K}} \mathrm{tr}\left\{\mathbf{Q}_k^{\mathrm{u}}\right\}\right) + MP_{\mathrm{c}} + P_{\mathrm{s}}}. \tag{18}$$

Therefore, the EE maximization problem is stated as

$$\begin{aligned} \max_{\mathbf{Q}^{\mathrm{m}},\mathbf{Q}_k^{\mathrm{u}},\forall k\in\mathcal{K}} \quad & \frac{W\left(u_0K(\min_k R_k^{\mathrm{m}})+\sum_{k\in\mathcal{K}}u_kR_k^{\mathrm{u}}\right)}{\mu\left(\mathrm{tr}\left\{\mathbf{Q}^{\mathrm{m}}\right\}+\sum_{k\in\mathcal{K}}\mathrm{tr}\left\{\mathbf{Q}_k^{\mathrm{u}}\right\}\right)+MP_{\mathrm{c}}+P_{\mathrm{s}}}, \\ \text{s.t.} \quad & \mathrm{tr}\left\{\mathbf{Q}^{\mathrm{m}}\right\}+\sum_{k\in\mathcal{K}}\mathrm{tr}\left\{\mathbf{Q}_k^{\mathrm{u}}\right\}\le P_{\max}, \\ & \mathbf{Q}^{\mathrm{m}}\succeq\mathbf{0},\ \mathbf{Q}_k^{\mathrm{u}}\succeq\mathbf{0}\ \ (\forall k\in\mathcal{K}), \end{aligned} \tag{19}$$

where $P_{\max}$ is the power budget at the BS.

3.2. Optimal Transmit Directions

The problem in (19) aims at designing large-dimensional complex matrices $\mathbf{Q}^{\mathrm{m}}$ and $\mathbf{Q}_k^{\mathrm{u}}$ $(\forall k)$, and the computational complexity can be very high. To simplify this problem, first, we decompose the transmit covariance matrices as $\mathbf{Q}^{\mathrm{m}}=\mathbf{\Phi}^{\mathrm{m}}\mathbf{\Lambda}^{\mathrm{m}}\left(\mathbf{\Phi}^{\mathrm{m}}\right)^H$ and $\mathbf{Q}_k^{\mathrm{u}}=\mathbf{\Phi}_k^{\mathrm{u}}\mathbf{\Lambda}_k^{\mathrm{u}}\left(\mathbf{\Phi}_k^{\mathrm{u}}\right)^H$, respectively. $\mathbf{\Phi}^{\mathrm{m}}$ and $\mathbf{\Phi}_k^{\mathrm{u}}$ are constituted by the eigenvectors of $\mathbf{Q}^{\mathrm{m}}$ and $\mathbf{Q}_k^{\mathrm{u}}$, respectively, which represent the directions of the transmitted signals. Meanwhile, $\mathbf{\Lambda}^{\mathrm{m}}$ and $\mathbf{\Lambda}_k^{\mathrm{u}}$ are diagonal matrices with their diagonal elements constituted by the eigenvalues of $\mathbf{Q}^{\mathrm{m}}$ and $\mathbf{Q}_k^{\mathrm{u}}$, respectively, which denote the allocated power over the corresponding directions.

The following theorem determines the values of the eigenvectors of $\mathbf{Q}^{\mathrm{m}}$ and $\mathbf{Q}_k^{\mathrm{u}}$.

Theorem 1. *The optimal multicast and unicast transmit covariance matrices of problem (19) is*

$$\mathbf{Q}^{\mathrm{m,opt}}=\mathbf{V}\mathbf{\Lambda}^{\mathrm{m}}\mathbf{V}^H,\quad \mathbf{Q}_k^{\mathrm{u,opt}}=\mathbf{V}\mathbf{\Lambda}_k^{\mathrm{u}}\mathbf{V}^H,\forall k, \tag{20}$$

where $\mathbf{\Lambda}^{\mathrm{m}}$ and $\mathbf{\Lambda}_k^{\mathrm{u}}$ $(\forall k)$ are both diagonal matrices, and the matrix $\mathbf{V}$ equals to the eigenvector matrices of the correlation matrices between the BS and all UTs and only depends on the BS array topology. The eigenvectors of $\mathbf{Q}^{\mathrm{m}}$ and $\mathbf{Q}_k^{\mathrm{u}}$ are given by the columns of the matrix $\mathbf{V}$,

Proof. Please refer to the Appendix A. □

Theorem 1 above indicates that when solving problem (19), since the eigenvectors are deterministic, we only have to determine the power allocation matrix denoted by $\mathbf{\Lambda}\triangleq\left\{\mathbf{\Lambda}^{\mathrm{m}},\mathbf{\Lambda}_1^{\mathrm{u}},\mathbf{\Lambda}_2^{\mathrm{u}},\ldots,\mathbf{\Lambda}_K^{\mathrm{u}}\right\}$, which reduces the number of variables to be optimized and the computational complexity significantly. Therefore, the large-dimensional complex-matrix-valued EE maximization problem can be transformed into a real-vector-valued power allocation problem in the beam domain.

Rewrite $\overline{\mathbf{K}}_k^{\mathrm{m}}$ and $\overline{\mathbf{K}}_k^{\mathrm{u}}$ as follows

$$\overline{\mathbf{K}}_k^{\mathrm{m}}\left(\mathbf{\Lambda}\right)\triangleq\sigma^2\mathbf{I}_{N_k}+\sum_{k'\in\mathcal{K}}\mathbf{A}_k\left(\mathbf{\Lambda}_{k'}^{\mathrm{u}}\right), \tag{21}$$

$$\overline{\mathbf{K}}_k^{\mathrm{u}}\left(\mathbf{\Lambda}\right)\triangleq\sigma^2\mathbf{I}_{N_k}+\sum_{k'\neq k}\mathbf{A}_k\left(\mathbf{\Lambda}_{k'}^{\mathrm{u}}\right), \tag{22}$$

and without loss of optimality, we can simplify the problem in (19) into the problem below

$$\begin{aligned} \max_{\mathbf{\Lambda}} \quad & \frac{W\left(u_0K\left(\min_k R_k^{\mathrm{m}}\left(\mathbf{\Lambda}\right)\right)+\sum_{k\in\mathcal{K}}u_kR_k^{\mathrm{u}}\left(\mathbf{\Lambda}\right)\right)}{\mu\left(\mathrm{tr}\left\{\mathbf{\Lambda}^{\mathrm{m}}\right\}+\sum_{k\in\mathcal{K}}\mathrm{tr}\left\{\mathbf{\Lambda}_k^{\mathrm{u}}\right\}\right)+MP_{\mathrm{c}}+P_{\mathrm{s}}}, \\ \text{s.t.} \quad & \mathrm{tr}\left\{\mathbf{\Lambda}^{\mathrm{m}}\right\}+\sum_{k\in\mathcal{K}}\mathrm{tr}\left\{\mathbf{\Lambda}_k^{\mathrm{u}}\right\}\le P_{\max}, \\ & \mathbf{\Lambda}^{\mathrm{m}}\succeq\mathbf{0},\ \mathbf{\Lambda}^{\mathrm{m}}\ \text{diagonal},\ \mathbf{\Lambda}_k^{\mathrm{u}}\succeq\mathbf{0},\ \mathbf{\Lambda}_k^{\mathrm{u}}\ \text{diagonal}\ (\forall k\in\mathcal{K}), \end{aligned} \tag{23}$$

with

$$R_k^{\mathrm{m}}(\mathbf{\Lambda}) = \underbrace{\mathsf{E}\left\{\log\det\left\{\overline{\mathbf{K}}_k^{\mathrm{m}}(\mathbf{\Lambda}) + \mathbf{G}_k\mathbf{\Lambda}^{\mathrm{m}}\mathbf{G}_k^H\right\}\right\}}_{\triangleq s_k^+(\mathbf{\Lambda})} - \underbrace{\log\det\left\{\overline{\mathbf{K}}_k^{\mathrm{m}}(\mathbf{\Lambda})\right\}}_{\triangleq s_k^-(\mathbf{\Lambda})}, \tag{24}$$

$$R_k^{\mathrm{u}}(\mathbf{\Lambda}) = \underbrace{\mathsf{E}\left\{\log\det\left\{\overline{\mathbf{K}}_k^{\mathrm{u}}(\mathbf{\Lambda}) + \mathbf{G}_k\mathbf{\Lambda}_k^{\mathrm{u}}\mathbf{G}_k^H\right\}\right\}}_{\triangleq t_k^+(\mathbf{\Lambda})} - \underbrace{\log\det\left\{\overline{\mathbf{K}}_k^{\mathrm{u}}(\mathbf{\Lambda})\right\}}_{\triangleq t_k^-(\mathbf{\Lambda})}. \tag{25}$$

Denote the lower bound of $R_k^{\mathrm{m}}(\mathbf{\Lambda})$ $(\forall k)$ as an auxiliary variable η, the problem in (23) can be equivalently expressed as

$$\begin{aligned} \max_{\mathbf{\Lambda}} \quad & \frac{W\left(u_0K\eta + \sum_{k\in\mathcal{K}} u_k R_k^{\mathrm{u}}(\mathbf{\Lambda})\right)}{\mu\left(\mathrm{tr}\left\{\mathbf{\Lambda}^{\mathrm{m}}\right\} + \sum_{k\in\mathcal{K}}\mathrm{tr}\left\{\mathbf{\Lambda}_k^{\mathrm{u}}\right\}\right) + MP_{\mathrm{c}} + P_{\mathrm{s}}}, \\ \text{s.t.} \quad & R_k^{\mathrm{m}}(\mathbf{\Lambda}) \geq \eta \ \ (\forall k\in\mathcal{K}), \\ & \mathrm{tr}\left\{\mathbf{\Lambda}^{\mathrm{m}}\right\} + \sum_{k\in\mathcal{K}}\mathrm{tr}\left\{\mathbf{\Lambda}_k^{\mathrm{u}}\right\} \leq P_{\max}, \\ & \mathbf{\Lambda}^{\mathrm{m}} \succeq \mathbf{0},\ \mathbf{\Lambda}^{\mathrm{m}}\ \text{diagonal},\ \mathbf{\Lambda}_k^{\mathrm{u}} \succeq \mathbf{0},\ \mathbf{\Lambda}_k^{\mathrm{u}}\ \text{diagonal}\ \ (\forall k\in\mathcal{K}). \end{aligned} \tag{26}$$

3.3. Energy-Efficient Power Allocation for NOUM Transmission

By observing problem (26), we can conclude that the numerator of the objective function is a difference of concave functions. We adopt the MM algorithm to deal with the problem. It is an iteration optimization process, where during each iteration, we replace the objective function with its lower bound function.

In this problem, we substitute $s_k^-(\mathbf{\Lambda})$ in (24) and $t_k^-(\mathbf{\Lambda})$ in (25) with their first-order Taylor expansions, respectively, to transfer the numerator of the objective function into a concave function, which leads to a concave-linear fractional program. We can solve problem (26) by solving a series of substitution problems iteratively. Then at the pth iteration, $\mathbf{\Lambda}_{(p)} = \left\{\mathbf{\Lambda}_{(p)}^{m}, \mathbf{\Lambda}_{1,(p)}^{\mathrm{u}}, \ldots, \mathbf{\Lambda}_{K,(p)}^{\mathrm{u}}\right\}$, and the sub-problem is

$$\begin{aligned} & \mathbf{\Lambda}_{(p+1)} = \\ & \arg\max_{\mathbf{\Lambda}} \ \frac{W\left(u_0K\eta + \sum_{k\in\mathcal{K}} u_k\left(t_k^+(\mathbf{\Lambda}) - t_k^-\left(\mathbf{\Lambda}_{(p)}\right) - \sum_{a\neq k}\mathrm{tr}\left\{\left(\frac{\partial t_k^-\left(\mathbf{\Lambda}_{(p)}\right)}{\partial\mathbf{\Lambda}_a^{\mathrm{u}}}\right)^T\left(\mathbf{\Lambda}_a^{\mathrm{u}} - \mathbf{\Lambda}_{a,(p)}^{\mathrm{u}}\right)\right\}\right)\right)}{\mu\left(\mathrm{tr}\left\{\mathbf{\Lambda}^{\mathrm{m}}\right\} + \sum_{k\in\mathcal{K}}\mathrm{tr}\left\{\mathbf{\Lambda}_k^{\mathrm{u}}\right\}\right) + MP_{\mathrm{c}} + P_{\mathrm{s}}}, \\ & \text{s.t.} \quad s_k^+(\mathbf{\Lambda}) - s_k^-\left(\mathbf{\Lambda}_{(p)}\right) - \sum_{a\in\mathcal{K}}\mathrm{tr}\left\{\left(\frac{\partial s_k^-\left(\mathbf{\Lambda}_{(p)}\right)}{\partial\mathbf{\Lambda}_a^{\mathrm{u}}}\right)^T\left(\mathbf{\Lambda}_a^{\mathrm{u}} - \mathbf{\Lambda}_{a,(p)}^{\mathrm{u}}\right)\right\} - \eta \geq 0\ \ (\forall k\in\mathcal{K}), \\ & \qquad \mathrm{tr}\left\{\mathbf{\Lambda}^{\mathrm{m}}\right\} + \sum_{k\in\mathcal{K}}\mathrm{tr}\left\{\mathbf{\Lambda}_k^{\mathrm{u}}\right\} \leq P_{\max}, \\ & \qquad \mathbf{\Lambda}^{\mathrm{m}} \succeq \mathbf{0},\ \mathbf{\Lambda}^{\mathrm{m}}\ \text{diagonal},\ \mathbf{\Lambda}_k^{\mathrm{u}} \succeq \mathbf{0},\ \mathbf{\Lambda}_k^{\mathrm{u}}\ \text{diagonal}\ \ (\forall k\in\mathcal{K}), \end{aligned} \tag{27}$$

where the gradients of $s_k^-\left(\mathbf{\Lambda}_{(p)}\right)$ and $t_k^-\left(\mathbf{\Lambda}_{(p)}\right)$ with respect to $\mathbf{\Lambda}_a^{\mathrm{u}}$ are defined by $\Delta s_k^{(p)}$ and $\Delta t_k^{(p)}$, respectively, with their diagonal elements being

$$\left[\Delta s_k^{(p)}\right]_{i,i} = \left[\frac{\partial s_k^-\left(\mathbf{\Lambda}_{(p)}\right)}{\partial\mathbf{\Lambda}_a^{\mathrm{u}}}\right]_{i,i} = \sum_{n=1}^{N_k}\frac{\left[\mathbf{\Omega}_k\right]_{n,i}}{\sigma^2 + \sum_{k'\in\mathcal{K}}\sum_{m=1}^{M}\left[\mathbf{\Omega}_k\right]_{n,m}\left[\mathbf{\Lambda}_{k',(p)}^{\mathrm{u}}\right]_{m,m}}, \tag{28}$$

$$\left[\Delta t_k^{(p)}\right]_{i,i} = \left[\frac{\partial t_k^-\left(\mathbf{\Lambda}_{(p)}\right)}{\partial \mathbf{\Lambda}_a^{\mathrm{u}}}\right]_{i,i} = \begin{cases} \sum\limits_{n=1}^{N_k} \dfrac{[\mathbf{\Omega}_k]_{n,i}}{\sigma^2 + \sum\limits_{k' \neq k} \sum\limits_{m=1}^{M} [\mathbf{\Omega}_k]_{n,m} \left[\mathbf{\Lambda}_{k',(p)}^{\mathrm{u}}\right]_{m,m}}, & a \neq k, \\ 0, & a = k, \end{cases} \tag{29}$$

respectively.

Since $t_k^-\left(\mathbf{\Lambda}_{(p)}\right)$ and $\mathbf{\Lambda}_{a,(p)}^{\mathrm{u}}$ in (27) are constant in each iteration, we can ignore them and obtain an equivalent optimization problem as

$$\begin{aligned} &\mathbf{\Lambda}_{(p+1)} = \\ &\underset{\mathbf{\Lambda}}{\arg\max} \quad \frac{W\left(u_0 K\eta + \sum\limits_{k\in\mathcal{K}} u_k \left(t_k^+(\mathbf{\Lambda}) - \sum\limits_{a\neq k} \mathrm{tr}\left\{\left(\frac{\partial t_k^-\left(\mathbf{\Lambda}_{(p)}\right)}{\partial \mathbf{\Lambda}_a^{\mathrm{u}}}\right)^T \mathbf{\Lambda}_a^{\mathrm{u}}\right\}\right)\right)}{\mu\left(\mathrm{tr}\left\{\mathbf{\Lambda}^{\mathrm{m}}\right\} + \sum\limits_{k\in\mathcal{K}} \mathrm{tr}\left\{\mathbf{\Lambda}_k^{\mathrm{u}}\right\}\right) + MP_{\mathrm{c}} + P_{\mathrm{s}}}, \\ &\quad\ \mathrm{s.t.} \quad s_k^+(\mathbf{\Lambda}) - s_k^-\left(\mathbf{\Lambda}_{(p)}\right) - \sum_{a\in\mathcal{K}} \mathrm{tr}\left\{\left(\frac{\partial s_k^-\left(\mathbf{\Lambda}_{(p)}\right)}{\partial \mathbf{\Lambda}_a^{\mathrm{u}}}\right)^T \left(\mathbf{\Lambda}_a^{\mathrm{u}} - \mathbf{\Lambda}_{a,(p)}^{\mathrm{u}}\right)\right\} - \eta \geq 0 \ \ (\forall k \in \mathcal{K}), \\ &\qquad\quad\ \mathrm{tr}\left\{\mathbf{\Lambda}^{\mathrm{m}}\right\} + \sum_{k\in\mathcal{K}} \mathrm{tr}\left\{\mathbf{\Lambda}_k^{\mathrm{u}}\right\} \leq P_{\max}, \\ &\qquad\quad\ \mathbf{\Lambda}^{\mathrm{m}} \succeq \mathbf{0},\ \mathbf{\Lambda}^{\mathrm{m}} \text{ diagonal},\ \mathbf{\Lambda}_k^{\mathrm{u}} \succeq \mathbf{0},\ \mathbf{\Lambda}_k^{\mathrm{u}} \text{ diagonal } \ (\forall k \in \mathcal{K}). \end{aligned} \tag{30}$$

Although the numerator of the objective function and constraint of the transformed sub-problem (30) are concave, the computational complexity can still be quite high if the expectation operation is manipulated using Monte-Carlo methods. Via applying the large-dimensional random matrix theory in [36,37], we further reduce the optimization complexity by substituting the minuends of $R_k^{\mathrm{m}}(\mathbf{\Lambda})$ and $R_k^{\mathrm{u}}(\mathbf{\Lambda})$ with their DEs, respectively.

First, we define a diagonal matrix-valued function $\mathbf{Y}_k(\mathbf{X})$ of size $M \times M$, and its ith diagonal element is

$$[\mathbf{Y}_k(\mathbf{X})]_{i,i} = \mathrm{tr}\left\{\mathrm{diag}\left\{[\mathbf{\Omega}_k]_{:,i}\right\}\mathbf{X}\right\}. \tag{31}$$

Then, we can write the DE of $s_k^+(\mathbf{\Lambda})$ as

$$\begin{aligned} S_k^+(\mathbf{\Lambda}) &= \log\det\left\{\mathbf{I}_M + \mathbf{\Gamma}_k^{\mathrm{m}}\mathbf{\Lambda}^{\mathrm{m}}\right\} \\ &\quad + \log\det\left\{\widetilde{\mathbf{\Gamma}}_k^{\mathrm{m}} + \overline{\mathbf{K}}_k^{\mathrm{m}}(\mathbf{\Lambda})\right\} - \mathrm{tr}\left\{\mathbf{I}_{N_k} - \left(\widetilde{\mathbf{\Phi}}_k^{\mathrm{m}}\right)^{-1}\right\}, \end{aligned} \tag{32}$$

where $\mathbf{\Gamma}_k^{\mathrm{m}}$, $\widetilde{\mathbf{\Gamma}}_k^{\mathrm{m}}$ and $\widetilde{\mathbf{\Phi}}_k^{\mathrm{m}}$ are given by

$$\begin{aligned} \mathbf{\Gamma}_k^{\mathrm{m}} &= \mathbf{Y}_k\left(\left(\widetilde{\mathbf{\Phi}}_k^{\mathrm{m}}\overline{\mathbf{K}}_k^{\mathrm{m}}(\mathbf{\Lambda})\right)^{-1}\right) \in \mathbb{C}^{M\times M}, \\ \widetilde{\mathbf{\Gamma}}_k^{\mathrm{m}} &= \mathbf{A}_k\left(\mathbf{\Lambda}^{\mathrm{m}}\left(\mathbf{I}_M + \mathbf{\Lambda}^{\mathrm{m}}\mathbf{\Gamma}_k^{\mathrm{m}}\right)^{-1}\right) \in \mathbb{C}^{N_k\times N_k}, \\ \widetilde{\mathbf{\Phi}}_k^{\mathrm{m}} &= \mathbf{I}_{N_k} + \widetilde{\mathbf{\Gamma}}_k^{\mathrm{m}}\left(\overline{\mathbf{K}}_k^{\mathrm{m}}(\mathbf{\Lambda})\right)^{-1} \in \mathbb{C}^{N_k\times N_k}, \end{aligned} \tag{33}$$

and the definition of $\mathbf{A}_k(\mathbf{X})$ is given in (11).

Likewise, we have the DE of $t_k^+(\mathbf{\Lambda})$ as

$$\begin{aligned} T_k^+(\mathbf{\Lambda}) &= \log\det\left\{\mathbf{I}_M + \mathbf{\Gamma}_k^{\mathrm{u}}\mathbf{\Lambda}_k^{\mathrm{u}}\right\} \\ &\quad + \log\det\left\{\widetilde{\mathbf{\Gamma}}_k^{\mathrm{u}} + \overline{\mathbf{K}}_k^{\mathrm{u}}(\mathbf{\Lambda})\right\} - \mathrm{tr}\left\{\mathbf{I}_{N_k} - \left(\widetilde{\mathbf{\Phi}}_k^{\mathrm{u}}\right)^{-1}\right\}, \end{aligned} \tag{34}$$

where $\mathbf{\Gamma}_k^{\mathrm{u}}$, $\widetilde{\mathbf{\Gamma}}_k^{\mathrm{u}}$ and $\widetilde{\mathbf{\Phi}}_k^{\mathrm{u}}$ are given by

$$\begin{aligned}
\mathbf{\Gamma}_k^{\mathrm{u}} &= \mathbf{Y}_k\left(\left(\widetilde{\mathbf{\Phi}}_k^{\mathrm{u}}\overline{\mathbf{K}}_k^{\mathrm{u}}\left(\mathbf{\Lambda}\right)\right)^{-1}\right) \in \mathbb{C}^{M\times M},\\
\widetilde{\mathbf{\Gamma}}_k^{\mathrm{u}} &= \mathbf{A}_k\left(\mathbf{\Lambda}_k^{\mathrm{u}}\left(\mathbf{I}_M + \mathbf{\Lambda}_k^{\mathrm{u}}\mathbf{\Gamma}_k^{\mathrm{u}}\right)^{-1}\right) \in \mathbb{C}^{N_k\times N_k},\\
\widetilde{\mathbf{\Phi}}_k^{\mathrm{u}} &= \mathbf{I}_{N_k} + \widetilde{\mathbf{\Gamma}}_k^{\mathrm{u}}\left(\overline{\mathbf{K}}_k^{\mathrm{u}}\left(\mathbf{\Lambda}\right)\right)^{-1} \in \mathbb{C}^{N_k\times N_k}.
\end{aligned} \tag{35}$$

With the DEs of $s_k^{+}\left(\mathbf{\Lambda}\right)$ and $t_k^{+}\left(\mathbf{\Lambda}\right)$ defined above, the optimization problem in (30) becomes

$$\begin{aligned}
&\mathbf{\Lambda}_{(p+1)} = \\
&\underset{\mathbf{\Lambda}}{\arg\max} \quad \frac{W\left(u_0K\eta + \sum_{k\in\mathcal{K}} u_k\left(T_k^{+}\left(\mathbf{\Lambda}\right) - \sum_{a\neq k}\mathrm{tr}\left\{\left(\frac{\partial t_k^{-}\left(\mathbf{\Lambda}_{(p)}\right)}{\partial\mathbf{\Lambda}_a^{\mathrm{u}}}\right)^T\mathbf{\Lambda}_a^{\mathrm{u}}\right\}\right)\right)}{\mu\left(\mathrm{tr}\left\{\mathbf{\Lambda}^{\mathrm{m}}\right\} + \sum_{k\in\mathcal{K}}\mathrm{tr}\left\{\mathbf{\Lambda}_k^{\mathrm{u}}\right\}\right) + MP_{\mathrm{c}} + P_{\mathrm{s}}},\\
&\quad \text{s.t.} \quad S_k^{+}\left(\mathbf{\Lambda}\right) - s_k^{-}\left(\mathbf{\Lambda}_{(p)}\right) - \sum_{a\in\mathcal{K}}\mathrm{tr}\left\{\left(\frac{\partial s_k^{-}\left(\mathbf{\Lambda}_{(p)}\right)}{\partial\mathbf{\Lambda}_a^{\mathrm{u}}}\right)^T\left(\mathbf{\Lambda}_a^{\mathrm{u}} - \mathbf{\Lambda}_{a,(p)}^{\mathrm{u}}\right)\right\} - \eta \geq 0 \;\left(\forall k\in\mathcal{K}\right),\\
&\qquad\quad \mathrm{tr}\left\{\mathbf{\Lambda}^{\mathrm{m}}\right\} + \sum_{k\in\mathcal{K}}\mathrm{tr}\left\{\mathbf{\Lambda}_k^{\mathrm{u}}\right\} \leq P_{\max},\\
&\qquad\quad \mathbf{\Lambda}^{\mathrm{m}} \succeq \mathbf{0},\; \mathbf{\Lambda}^{\mathrm{m}}\text{ diagonal},\; \mathbf{\Lambda}_k^{\mathrm{u}} \succeq \mathbf{0},\; \mathbf{\Lambda}_k^{\mathrm{u}}\text{ diagonal }\;\left(\forall k\in\mathcal{K}\right).
\end{aligned} \tag{36}$$

We can observe from the optimization problem in (36) that the denominator and numerator of the objective function are linear and concave functions of $\mathbf{\Lambda}$, respectively. We invoke Dinkelbach's transform [38] to deal with this concave-linear program. We can obtain the solution to (36) via solving a series of problems below

$$\begin{aligned}
&\left\{\mathbf{\Lambda}_{(p)}^{(q+1)}, \eta^{(q+1)}\right\} = \\
&\underset{\mathbf{\Lambda}}{\arg\max} \quad W\left(u_0K\eta + \sum_{k\in\mathcal{K}} u_k\left(T_k^{+}\left(\mathbf{\Lambda}\right) - \sum_{a\neq k}\mathrm{tr}\left\{\left(\frac{\partial t_k^{-}\left(\mathbf{\Lambda}_{(p)}\right)}{\partial\mathbf{\Lambda}_a^{\mathrm{u}}}\right)^T\mathbf{\Lambda}_a^{\mathrm{u}}\right\}\right)\right) - \chi_{(p)}^{(q)}P\left(\mathbf{\Lambda}\right),\\
&\quad \text{s.t.} \quad S_k^{+}\left(\mathbf{\Lambda}\right) - s_k^{-}\left(\mathbf{\Lambda}_{(p)}\right) - \sum_{a\in\mathcal{K}}\mathrm{tr}\left\{\left(\frac{\partial s_k^{-}\left(\mathbf{\Lambda}_{(p)}\right)}{\partial\mathbf{\Lambda}_a^{\mathrm{u}}}\right)^T\left(\mathbf{\Lambda}_a^{\mathrm{u}} - \mathbf{\Lambda}_{a,(p)}^{\mathrm{u}}\right)\right\} - \eta \geq 0 \;\left(\forall k\in\mathcal{K}\right),\\
&\qquad\quad \mathrm{tr}\left\{\mathbf{\Lambda}^{\mathrm{m}}\right\} + \sum_{k\in\mathcal{K}}\mathrm{tr}\left\{\mathbf{\Lambda}_k^{\mathrm{u}}\right\} \leq P_{\max},\\
&\qquad\quad \mathbf{\Lambda}^{\mathrm{m}} \succeq \mathbf{0},\; \mathbf{\Lambda}^{\mathrm{m}}\text{ diagonal},\; \mathbf{\Lambda}_k^{\mathrm{u}} \succeq \mathbf{0},\; \mathbf{\Lambda}_k^{\mathrm{u}}\text{ diagonal }\;\left(\forall k\in\mathcal{K}\right),
\end{aligned} \tag{37}$$

where $P\left(\mathbf{\Lambda}\right) = \mu\left(\mathrm{tr}\left\{\mathbf{\Lambda}^{\mathrm{m}}\right\} + \sum_{k\in\mathcal{K}}\mathrm{tr}\left\{\mathbf{\Lambda}_k^{\mathrm{u}}\right\}\right) + MP_{\mathrm{c}} + P_{\mathrm{s}}$, q is the iteration index, and $\chi_{(p)}^{(q)}$ is the auxiliary variable. During each iteration, we update $\chi_{(p)}^{(q)}$ using the following equation

$$\chi_{(p)}^{(q)} = \frac{W\left(u_0K\eta^{(q)} + \sum_{k\in\mathcal{K}} u_k\left(T_k^{+}\left(\mathbf{\Lambda}_{(p)}^{(q)}\right) - \sum_{a\neq k}\mathrm{tr}\left\{\left(\frac{\partial t_k^{-}\left(\mathbf{\Lambda}_{(p)}\right)}{\partial\mathbf{\Lambda}_a^{\mathrm{u}}}\right)^T\mathbf{\Lambda}_{a,(p)}^{\mathrm{u},(q)}\right\}\right)\right)}{P\left(\mathbf{\Lambda}_{(p)}^{(q)}\right)}. \tag{38}$$

From the analysis above, we can observe that the proposed EE optimization algorithm involves two-layer iterations. During the outer iteration, via invoking the MM algorithm, we replace the numerator of the objective function in (26) with its lower bound function, thus making the numerator concave. The MM-based algorithm is guaranteed to converge to the locally optimal solution [39–41]; in the inner iteration, we transform the fractional problem in (36) into solvable convex optimization problems in (37) via Dinkelbach's transform, which can derive the global optimum solution to (36) with guaranteed convergence [42]. After several iterations, we can obtain the optimal beam domain power allocation matrix $\mathbf{\Lambda}$. Please note that $\mathbf{\Lambda}$ is locally optimal due to the local optimality of MM algorithm. We present our algorithm in Algorithm 1.

Algorithm 1 Energy-Efficient Power Allocation Algorithm in the Beam Domain for Massive MIMO NOUM Transmission

Input: Beam domain channel statistics $\mathbf{\Omega}_k$, initial power allocation matrix $\mathbf{\Lambda}_{(0)}$, outer iteration threshold ϵ_1 and inner iteration threshold ϵ_2

Output: Power allocation matrix $\mathbf{\Lambda}$ in the beam domain

Initialization: $\overline{EE}_{(-1)} = 0, p = 0$

Calculate

$$\overline{EE}_{(p)} = \frac{W\left(u_0 K\left(\min_k \left\{S_k^+\left(\mathbf{\Lambda}_{(p)}\right) - s_k^-\left(\mathbf{\Lambda}_{(p)}\right)\right\}\right) + \sum_{k\in\mathcal{K}} u_k\left(T_k^+\left(\mathbf{\Lambda}_{(p)}\right) - t_k^-\left(\mathbf{\Lambda}_{(p)}\right)\right)\right)}{P\left(\mathbf{\Lambda}_{(p)}\right)} \tag{39}$$

while $\left|\overline{EE}_{(p)} - \overline{EE}_{(p-1)}\right| \geq \epsilon_1$ **do**

  Initialization: $q = 0$, let $\mathbf{\Lambda}_{(p)}^{(q)} = \mathbf{\Lambda}_{(p)}$, calculate $\chi_{(p)}^{(q)}$ with (38)

  while $\left|\chi_{(p)}^{(q)} - \chi_{(p)}^{(q-1)}\right| \geq \epsilon_2$ **do**

    Let $q = q + 1$

    Calculate $\mathbf{\Lambda}_{(p)}^{(q)}$ via solving problem (37) with $\chi_{(p)}^{(q-1)}$

    Calculate $\chi_{(p)}^{(q)}$ using (38)

  end while

  Let $p = p + 1$

  Let $\mathbf{\Lambda}_{(p)} = \mathbf{\Lambda}_{(p-1)}^{(q)}$

  Calculate $\overline{EE}_{(p)}$ with (39)

end while

return $\mathbf{\Lambda} = \mathbf{\Lambda}_{(p)}$

4. Numerical Results

We provide numerical simulation results to demonstrate the performance of the EE optimization algorithm proposed above for NOUM transmission massive MIMO scenario with statistical CSI. Table 1 illustrates how the numerical simulation parameters are set.

First of all, in Figure 2, we illustrate the convergence performance by showing the iteration process of our EE optimization algorithm under different transmit power budgets $P_{\max}$. The horizontal ordinate is the outer iteration index. As we can see, the EE converges after only a few iterations. Also, we can observe that in the lower power budget regime, the EE performance convergences faster than that in the higher power budget regime.

Table 1. Simulation parameters.

Parameter	Value
Scenario	Suburban macro
Channel model	3GPP SCM
Pathloss	-120 dB $(\forall k)$
Array topology	ULA with antenna spacing half wavelength
Noise variance	$\sigma^2 = -131$ dBm
Number of UTs	$K = 8$
Number of BS antennas	$M = 128$
Number of UT antennas	$N_k = 4\ (\forall k)$
Transmission bandwidth	W= 10 MHz
Amplifier drain efficiency	$\mu = 5$
Circuit power consumption per antenna	$P_c = 30$ dBm
Static power consumption	$P_s = 40$ dBm
Weights	$u_0 = 0.7, u_k = 0.3\ (\forall k)$

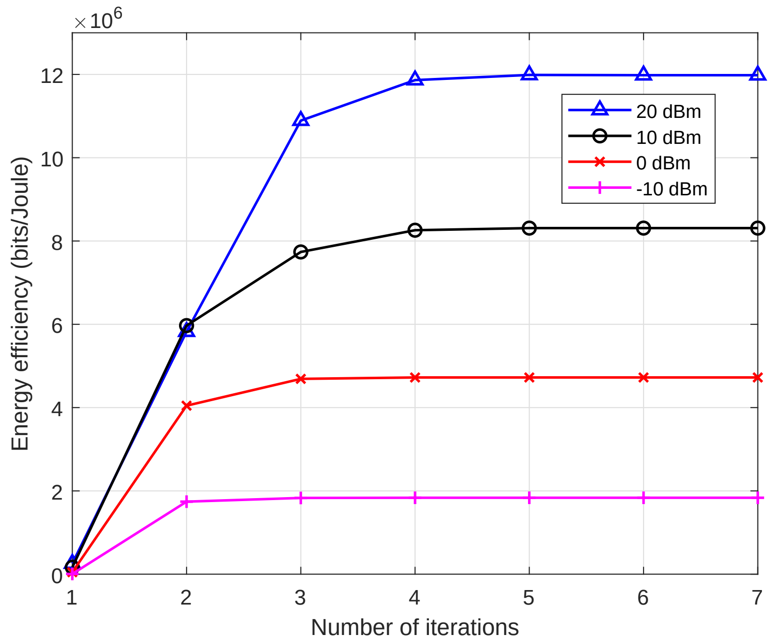

Figure 2. The convergence performance of the proposed EE optimization algorithm for different power budgets $P_{\max}$.

Then, we evaluate the EE of the NOUM transmission versus the power budget $P_{\max}$ under different numbers the antennas M at the BS in Figure 3. As we can see, the EE performance decreases when the BS antenna number M increases for the reason that in the power consumption model we adopted in (16), the total circuit power consumption grows linearly with M, the BS antenna number.

Next, the comparison of the EE performance of the power allocation algorithm proposed above with the rate maximization approach [11] is shown in Figure 4. We notice that the EE performance of the two approaches are similar at low transmit power budget regime. However, when the transmit power budget gets high, the EE performance of the rate maximization approach decreases, while that of our EE maximization approach remains high. This indicates that the rate maximization approach can achieve almost EE optimal when $P_{\max}$ is low. However, our EE maximization approach outperforms the rate maximization one at high transmit power budget regime.

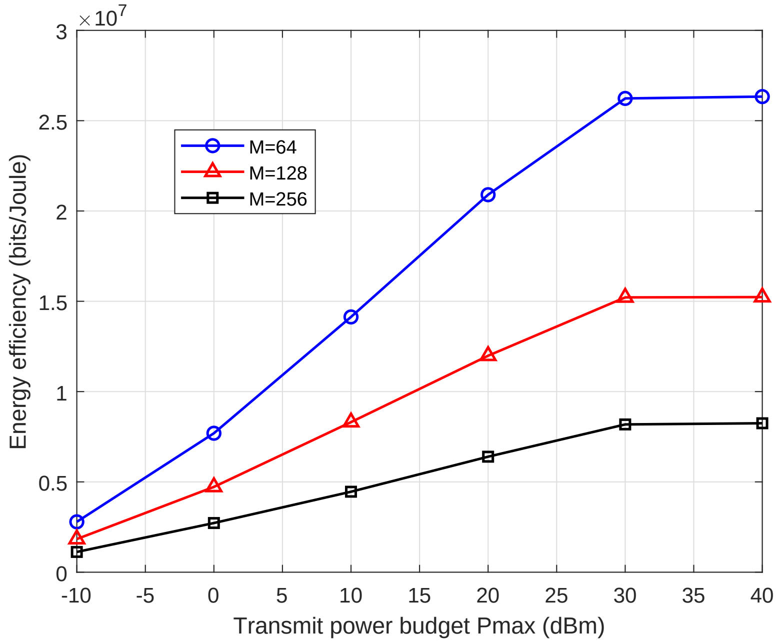

Figure 3. The EE performance of the NOUM transmission versus the power budget $P_{\max}$ for different numbers of BS antennas M.

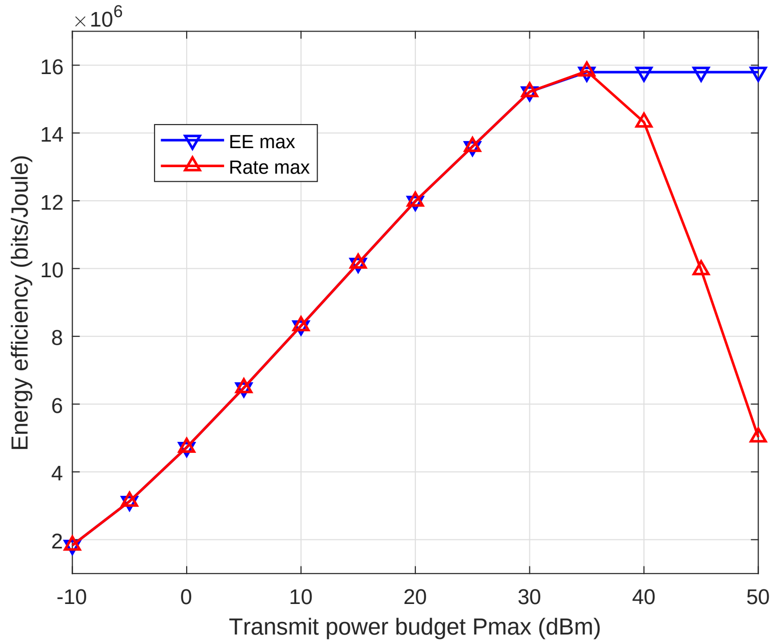

Figure 4. The EE performance of the proposed beam domain power allocation algorithm compared with the rate maximization approach.

Finally, in Figure 5, the EE performance of our power allocation approach and that of full CSI approach, which assumes instantaneous CSI is known at the BS, is compared. Since full CSI is an ideal case, it can achieve better EE performance than other imperfect CSI situation. However, the full CSI case suffers from pilot overhead. As Figure 5 illustrates, our proposed algorithm surpasses the full CSI approach with 3/7 pilot overhead [43] in the EE performance.

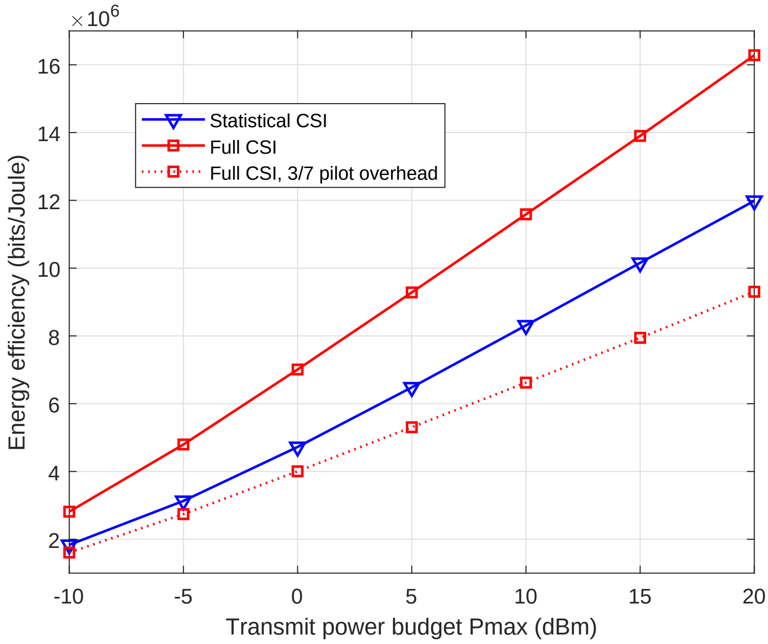

Figure 5. The comparison on the EE performance of proposed algorithm, full CSI case and full CSI with 3/7 overhead.

5. Conclusions

To conclude, we considered the EE optimization problem in NOUM transmission systems with statistical CSI available at the BS. We first formulated the EE maximization problem, and then determined the closed-form optimal eigenvectors of the multicast and unicast transmit covariance matrices for optimal EE, respectively. Next, with guaranteed convergence, we proposed a beam domain power allocation algorithm adopting the MM algorithm, DE and Dinkelbach's transform and derived the locally optimal power allocation strategy to achieve the EE optimization. Finally, with numerical results, we presented the performance gain of our EE maximization algorithm compared with the conventional approach.

Author Contributions: W.W. and Y.H. perceived the idea and wrote the manuscript. L.Y., J.X., J.L. and X.G. gave valuable suggestions on the structuring of the paper and assisted in the revising and proofreading.

Funding: This work was supported by the National Key R&D Program of China under Grant 2018YFB1801103, the National Natural Science Foundation of China under Grants 61631018, 61801114, 61761136016, 61771264, 61871465, and 61501113, the Natural Science Foundation of Jiangsu Province under Grant BK20170688, the Fundamental Research Funds for the Central Universities, and the Huawei Cooperation Project.

Acknowledgments: The authors would like to thank the Editor and the anonymous reviewers for their valuable comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proof of Theorem 1

Firstly, we can tell from (11) that the value of $\mathbf{A}_k(\mathbf{X})$ is only affected by the diagonal elements of $\mathbf{X}$, so $\overline{\mathbf{K}}_k^{\mathrm{m}}$ in (10) and $\overline{\mathbf{K}}_k^{\mathrm{u}}$ in (15) are irrelevant to the off-diagonal elements of $\mathbf{V}^H\mathbf{Q}_{k'}^{\mathrm{u}}\mathbf{V}$. Then, the power consumption in (16) has no relationship with the off-diagonal elements of $\mathbf{V}\mathbf{\Lambda}^{\mathrm{m}}\mathbf{V}^H$ or $\mathbf{V}\mathbf{\Lambda}_k^{\mathrm{u}}\mathbf{V}^H$. Moreover, applying the proof method similar to [44], we can conclude that $\mathbf{V}\mathbf{\Lambda}^{\mathrm{m}}\mathbf{V}^H$ and $\mathbf{V}\mathbf{\Lambda}_k^{\mathrm{u}}\mathbf{V}^H$ should be diagonal to maximize R_k^{m} and R_k^{u}, respectively. Therefore, to maximize the objective function in (19), $\mathbf{Q}^{\mathrm{m}}$ and $\mathbf{Q}_k^{\mathrm{u}}$ should be both diagonal matrices. This concludes the proof.

References

1. Lecompte, D.; Gabin, F. Evolved multimedia broadcast/multicast service (eMBMS) in LTE-advanced: Overview and Rel-11 enhancements. *IEEE Commun. Mag.* **2012**, *50*, 68–74. [CrossRef]
2. Wang, W.; Liu, A.; Zhang, Q.; You, L.; Gao, X.; Zheng, G. Robust multigroup multicast transmission for frame-based multi-beam satellite systems. *IEEE Access* **2018**, *6*, 46074–46083. [CrossRef]
3. You, L.; Liu, A.; Wang, W.; Gao, X. Outage constrained robust multigroup multicast beamforming for multi-beam satellite communication systems. *IEEE Wirel. Commun. Lett.* **2019**, *8*, 352–355. [CrossRef]
4. You, L.; Wang, J.; Wang, W.; Gao, X. Secure multicast transmission for massive MIMO with statistical channel state information. *IEEE Signal Process. Lett.* **2019**, *26*, 803–807. [CrossRef]
5. Chen, E.; Tao, M.; Liu, Y. Joint base station clustering and beamforming for non-orthogonal multicast and unicast transmission with backhaul constraints. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 6265–6279. [CrossRef]
6. Sadeghi, M.; Björnson, E.; Larsson, E.G.; Yuen, C.; Marzetta, T. Joint unicast and multi-group multicast transmission in massive MIMO systems. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 6375–6388. [CrossRef]
7. Wang, J.; Xu, H.; Zhu, B.; Fan, L.; Zhou, A. Hybrid beamforming design for mmWave joint unicast and multicast transmission. *IEEE Commun. Lett.* **2018**, *22*, 2012–2015. [CrossRef]
8. Ngo, H.Q.; Larsson, E.G.; Marzetta, T.L. Energy and spectral efficiency of very large multiuser MIMO systems. *IEEE Trans. Commun.* **2013**, *61*, 1436–1449.
9. Hoydis, J.; ten Brink, S.; Debbah, M. Massive MIMO in the UL/DL of cellular networks: How many antennas do we need? *IEEE J. Sel. Areas Commun.* **2013**, *31*, 160–171. [CrossRef]
10. Sadeghi, M.; Björnson, E.; Larsson, E.G.; Yuen, C.; Marzetta, T.L. Max–min fair transmit precoding for multi-group multicasting in massive MIMO. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 1358–1373. [CrossRef]
11. You, L.; Xiong, J.; Li, K.; Wang, W.; Gao, X. Non-orthogonal unicast and multicast transmission for massive MIMO with statistical channel state information. *IEEE Access* **2018**, *6*, 66841–66849. [CrossRef]
12. Alibakhshikenari, M.; Virdee, B.S.; See, C.H.; Abd-Alhameed, R.A.; Falcone, F.; Limiti, E. High-isolation leaky-wave array antenna based on CRLH-metamaterial implemented on SIW with 30o frequency beam-scanning capability at millimetre-waves. *Electronics* **2019**, *8*, 642. [CrossRef]
13. Song, Y.J.; Sarabandi, K. Equivalent circuit model for metamaterial-based electromagnetic band-gap isolator. *IEEE Antennas Wirel. Propag. Lett.* **2012**, *11*, 1366–1369. [CrossRef]
14. Alibakhshikenari, M.; Khalily, M.; Virdee, B.S.; See, C.H.; Abd-Alhameed, R.A.; Limiti, E. Mutual-coupling isolation using embedded metamaterial EM bandgap decoupling slab for densely packed array antennas. *IEEE Access* **2019**, *7*, 51827–51840. [CrossRef]
15. Farsi, S.; Aliakbarian, H.; Schreurs, D.; Nauwelaers, B.; Vandenbosch, G.A.E. Mutual coupling reduction between planar antennas by using a simple microstrip U-section. *IEEE Antennas Wirel. Propag. Lett.* **2012**, *11*, 1501–1503. [CrossRef]
16. Alibakhshikenari, M.; Khalily, M.; Virdee, B.S.; See, C.H.; Abd-Alhameed, R.A.; Limiti, E. Mutual coupling suppression between two closely placed microstrip patches using EM-bandgap metamaterial fractal loading. *IEEE Access* **2019**, *7*, 23606–23614. [CrossRef]
17. Singh, K.; Gupta, A.; Ratnarajah, T. QoS-driven energy-efficient resource allocation in multiuser amplify-and-forward relay networks. *IEEE Trans. Signal Inf. Process. Over Netw.* **2017**, *3*, 771–786. [CrossRef]
18. Singh, K.; Gupta, A.; Ratnarajah, T. QoS-driven resource allocation and EE-balancing for multiuser two-way amplify-and-forward relay networks. *IEEE Trans. Wirel. Commun.* **2017**, *16*, 3189–3204. [CrossRef]
19. Singh, K.; Gupta, A.; Ratnarajah, T. A utility-based joint subcarrier and power allocation for green communications in multi-user two-way regenerative relay networks. *IEEE Trans. Commun.* **2017**, *65*, 3705–3722. [CrossRef]
20. Li, G.Y.; Xu, Z.; Xiong, C.; Yang, C.; Zhang, S.; Chen, Y.; Xu, S. Energy-efficient wireless communications: tutorial, survey, and open issues. *IEEE Wirel. Commun.* **2011**, *18*, 28–35. [CrossRef]
21. Khan, T.A.; Yazdan, A.; Heath, R.W. Optimization of power transfer efficiency and energy efficiency for wireless-powered systems with massive MIMO. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 7159–7172. [CrossRef]
22. Björnson, E.; Sanguinetti, L.; Hoydis, J.; Debbah, M. Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer? *IEEE Trans. Wirel. Commun.* **2015**, *14*, 3059–3075. [CrossRef]

23. Tervo, O.; Tran, L.; Chatzinotas, S.; Juntti, M.; Ottersten, B. Energy-efficient joint unicast and multicast beamforming with multi-antenna user terminals. In Proceedings of the 2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Sapporo, Japan, 3–6 July 2017; pp. 1–5.
24. He, S.; Huang, Y.; Jin, S.; Yang, L. Energy efficient coordinated beamforming design in multi-cell multicast networks. *IEEE Commun. Lett.* **2015**, *19*, 985–988. [CrossRef]
25. Tervo, O.; Tran, L.; Pennanen, H.; Chatzinotas, S.; Ottersten, B.; Juntti, M. Energy-efficient multicell multigroup multicasting with joint beamforming and antenna selection. *IEEE Trans. Signal Process.* **2018**, *66*, 4904–4919. [CrossRef]
26. You, L.; Gao, X.; Xia, X.; Ma, N.; Peng, Y. Pilot reuse for massive MIMO transmission over spatially correlated rayleigh fading channels. *IEEE Trans. Wirel. Commun.* **2015**, *14*, 3352–3366. [CrossRef]
27. You, L.; Gao, X.; Swindlehurst, A.L.; Zhong, W. Channel acquisition for massive MIMO-OFDM with adjustable phase shift pilots. *IEEE Trans. Signal Process.* **2016**, *64*, 1461–1476. [CrossRef]
28. You, L.; Chen, X.; Wang, W.; Gao, X. Coordinated multicast precoding for multi-cell massive MIMO transmission exploiting statistical channel state information. *Electronics* **2018**, *7*, 338. [CrossRef]
29. You, L.; Wang, W.; Gao, X. Energy-efficient multicast precoding for massive MIMO transmission with statistical CSI. *Energies* **2018**, *11*, 3175. [CrossRef]
30. Weichselberger, W.; Herdin, M.; Ozcelik, H.; Bonek, E. A stochastic MIMO channel model with joint correlation of both link ends. *IEEE Trans. Wirel. Commun.* **2006**, *5*, 90–100. [CrossRef]
31. Gao, X.; Jiang, B.; Li, X.; Gershman, A.B.; McKay, M.R. Statistical eigenmode transmission over jointly correlated MIMO channels. *IEEE Trans. Inf. Theory* **2009**, *55*, 3735–3750. [CrossRef]
32. You, L.; Gao, X.; Li, G.Y.; Xia, X.; Ma, N. BDMA for millimeter-wave/terahertz massive MIMO transmission with per-beam synchronization. *IEEE J. Sel. Areas Commun.* **2017**, *35*, 1550–1563. [CrossRef]
33. Sun, C.; Gao, X.; Jin, S.; Matthaiou, M.; Ding, Z.; Xiao, C. Beam division multiple access transmission for massive MIMO communications. *IEEE Trans. Commun.* **2015**, *63*, 2170–2184. [CrossRef]
34. Wen, C.; Jin, S.; Wong, K.; Chen, J.; Ting, P. Channel estimation for massive MIMO using Gaussian-mixture Bayesian learning. *IEEE Trans. Wirel. Commun.* **2015**, *14*, 1356–1368. [CrossRef]
35. Ng, D.W.K.; Lo, E.S.; Schober, R. Energy-efficient resource allocation in OFDMA systems with large numbers of base station antennas. *IEEE Trans. Wirel. Commun.* **2012**, *11*, 3292–3304. [CrossRef]
36. Lu, A.; Gao, X.; Xiao, C. Free deterministic equivalents for the analysis of MIMO multiple access channel. *IEEE Trans. Inf. Theory* **2016**, *62*, 4604–4629. [CrossRef]
37. Couillet, R.; Debbah, M. *Random Matrix Methods for Wireless Communications*; Cambridge University Press: Cambridge, UK, 2011. [CrossRef]
38. Dinkelbach, W. On nonlinear fractional programming. *Manag. Sci.* **1967**, *13*, 492–498. [CrossRef]
39. Lipp, T.; Boyd, S. Variations and extension of the convex-concave procedure. *Optim. Eng.* **2016**, *17*, 263–287. [CrossRef]
40. Yuille, A.L.; Rangarajan, A. The concave-convex procedure. *Neural Comput.* **2003**, *15*, 915–936. [CrossRef]
41. Hunter, D.R.; Lange, K. A tutorial on MM algorithms. *Am. Stat.* **2004**, *58*, 30–37. [CrossRef]
42. Zappone, A.; Jorswieck, E. Energy efficiency in wireless networks via fractional programming theory. *Found. Trends® Commun. Inf. Theory* **2015**, *11*, 185–396. [CrossRef]
43. Marzetta, T.L. Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas. *IEEE Trans. Wirel. Commun.* **2010**, *9*, 3590–3600. [CrossRef]
44. Tulino, A.M.; Lozano, A.; Verdú, S. Capacity-achieving input covariance for single-user multi-antenna channels. *IEEE Trans. Wirel. Commun.* **2006**, *5*, 662–671. [CrossRef]

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).