This is the published version of a paper published in *IEEE Transactions on Wireless Communications*.

Citation for the original published paper (version of record):

*IEEE Transactions on Wireless Communications*, 13(1): 382-393

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-134011
Energy Efficient Pilot and Link Adaptation for Mobile Users in TDD Multi-User MIMO Systems

Yunesung Kim, Guowang Miao, and Taewon Hwang

Abstract—In this paper, we develop an uplink pilot and downlink link adaptation approach to improve the energy efficiency (EE) of mobile users in time division duplexing (TDD) multiuser multiple input and multiple output (MU-MIMO) systems. Assuming reciprocity between uplink and downlink channels, the downlink transmission is based on uplink channel estimation. While more uplink pilot power ensures more accurate channel estimation and better downlink performance, it incurs higher energy consumption of mobile users. This paper reveals the relationship and tradeoff among pilot power, channel estimation, and downlink link adaptation that achieves the highest energy efficiency for mobile users. We show that the energy efficiency of different users can be decoupled because the downlink average throughput of each user is independent of the pilot powers of other users and energy-efficient design can be done on a per-user basis. Based on the analysis, we propose an uplink pilot and downlink link adaptation algorithm to improve the EE of mobile users. Simulation results are finally provided to demonstrate the significant gain in energy efficiency for mobile users.

Index Terms—Energy efficiency, rate adaptation, power allocation, TDD, multiuser MIMO.

I. INTRODUCTION

THE demand of cellular data traffic has grown significantly in recent years. To accommodate the need, cellular infrastructures are getting denser and denser and consuming more and more energy resulting in a large amount of carbon dioxide emission and high capital and operating expenditures [1]. On the other hand, mobile terminals also desire high energy efficiency (EE) because the development of battery technology has not kept up with the demand of mobile communications [2]. Thus, energy-efficient design is becoming more and more important for both mobile operators to fulfill their social responsibility in preserving environments and to minimize their costs and mobile terminals to extend their battery lives [3], [4].

In the past decades, significant efforts have been dedicated to improving the EE of wireless systems [5]-[18]. In [5], an adaptive modulation strategy that minimizes the total energy consumption for transmitting a given number of bits in a single input and single output (SISO) AWGN channel is investigated. It shows that using the lowest modulation order is not always energy efficient if circuit energy consumption is considered. Energy-efficient link adaptation for a single user multichannel system is studied in [6]-[9]. In [10], energy-efficient link adaptation and subcarrier allocation scheme is proposed for uplink OFDMA systems assuming flat fading channels. It is proved that, for a given channel gain and constant circuit energy consumption, there exists a unique optimal transmission rate that maximizes EE. That work is extended to frequency-selective channels in [11]. In [12], link adaptation for MIMO-OFDM wireless systems is formulated as a convex optimization problem and optimal transmission mode is chosen to maximize EE with quality of service (QoS) constraints. In [13], the problem of energy-efficient input covariance matrix is investigated when terminals have multiple antennas. In [14], an energy-efficient power allocation algorithm for a single antenna OFDM system is developed. That work is later extended to the power loading problem for a single-carrier MIMO-SVD system in [15]. In [16], the EE capacity for an uplink MU-MIMO system is defined and a low-complexity power allocation algorithm that achieves this capacity is developed. In [17], an energy-efficient waterfilling algorithm for the downlink MU-MIMO system is developed. In [18], assuming BS uses the zero-forcing precoder, the optimal power allocation that maximizes the EE in the downlink of a multiuser multicarrier system is studied.

These studies [5]-[18] assume the availability of perfect channel state information (CSI). However, in practice, it is impossible to obtain perfect CSI because of channel estimation error and CSI can not be obtained without additional cost. Hence, an energy-efficient system design should consider both energy consumption for channel estimation and the performance degradation as a result of imperfect channel estimation. In [19], an energy-efficient pilot design in a training-based downlink system is studied for a single user case and the optimal overall transmit power and the power allocation between pilots and data symbols are investigated. This idea is later extended to a downlink multiuser OFDMA system in [20]. Both [19] and [20] consider energy-efficient pilot power allocation for single-antenna systems. To the best of our knowledge, there has been no research in literature that investigates energy-efficient pilot power allocation for multiuser multiple-input and multiple-output (MU-MIMO) systems.

In this paper, we study the EE of users in a time division duplexing (TDD) MU-MIMO system, where each user sends an uplink pilot sequence for channel estimation by the BS assuming perfect reciprocity between uplink and downlink channels. Based on the estimate, the BS performs zero-forcing (ZF) beamforming and transmits data to users. With higher
pilot power, higher downlink rate can be achieved because the BS can perform ZF beamforming with higher accuracy and the interference between users can be suppressed. However, higher pilot power indicates higher user power consumption. This paper will find the optimal uplink pilot power for each user. Our modeling considers channel estimation error and we show that the average throughput of each user is independent from the pilot power of others. The EE is defined as the average throughput per total energy consumed by the user and we find that the objective function is not quasi-concave in general. However, since the variables are uncoupled and the objective function is quasi-concave with respect to each variable in practice, we propose an iterative algorithm to find the optimal uplink pilot power and downlink transmission rate that maximizes the EE of all the users in the network.

The rest of the paper is organized as follows. In Section II, we describe the system model. In Section III, we analyze the downlink average throughput of the TDD MU-MIMO system based on the ZF precoder. In Section IV, we define the downlink EE of each user and propose an algorithm that finds the optimal uplink pilot power and downlink transmission rate to maximize the EE of each user. In Section V, simulation results are provided to demonstrate the performance of the proposed algorithm and the paper is concluded in Section VI.

Notations: \((\cdot)^T\) and \((\cdot)^H\) denote transpose and Hermitian transpose, respectively. \(I_N\) denotes an \(N\times N\) identity matrix.

II. SYSTEM MODEL

Consider the TDD MU-MIMO system shown in Fig. 1, where a BS is serving \(K\) users. The BS has \(M\) antennas and each user has one antenna. We consider zero-forcing (ZF) precoding at the BS because it is a practical low-complexity linear precoding scheme and it performs optimal among all the linear precoders at high SNR. Moreover, the SINR analysis under imperfect CSI at the transmitter is tractable when ZF precoder is employed. Due to these nice properties, ZF precoder has been frequently adopted in the system model of the papers on limited feedback such as [22], [25]-[27]. Following typical assumptions in MU-MIMO research, e.g., [21]-[22], we consider a narrowband system with flat fading channels. By assuming flat fading channel, the discussion on the tradeoff between the uplink pilot power and the downlink rate of users in a multiuser MIMO system can be simplified. However, the discussion for narrowband channels can be extended to frequency-selective channels by employing orthogonal frequency division multiplexing (OFDM) because in OFDM systems the wideband channel is divided into many narrowband sub-bands, each experiencing flat fading. Denote \(\sqrt{\beta_k} h_k^T\) to be the downlink channel from the BS to the \(k\)th user, where \(\beta_k\) models large-scale fading that incorporates path-loss and shadowing and \(h_k^T \sim C\mathcal{N}(0, I_M)\), a \(1 \times M\) vector, models small-scale fading. The received signal at the \(k\)th user is

\[
    r_k = \sqrt{\beta_k} h_k^T x + n_k
\]

where \(x\) and \(n_k\) \(\sim C\mathcal{N}(0, \sigma^2)\) are the \(M \times 1\) transmitted signal vector and complex additive white Gaussian noise, respectively. Assume ideal channel reciprocity and the uplink channels are the same as the downlink channels. In addition, assume block fading and the channel is constant in each frame. The large-scale fading coefficient \(\beta_k\) is known and the small-scale fading vector \(h_k\) needs to be estimated.

A. Communication Procedure

As shown in Fig. 2, the system consists of three phases: 1) uplink channel estimation, 2) downlink effective channel estimation, and 3) downlink data transmission. Each frame has \(T\) symbols. We allocate \(\tau_{\text{up}}\) symbols for uplink channel estimation, \(\tau_{\text{dn}}\) symbols for downlink effective channel estimation, and the remaining \(\tau_{\text{tr}} = T - \tau_{\text{up}} - \tau_{\text{dn}}\) symbols for downlink data transmission.

1) Uplink Channel Estimation: To estimate \(h_k\) at the BS, the \(k\)th user transmits a \(1 \times \tau_{\text{up}}\) orthogonal pilot sequence vector \(\sqrt{p_k} \psi_k^T\) with \(\|\psi_k^T\|^2 = \tau_{\text{up}}\), where \(p_k\) denotes the transmit power of the uplink pilot symbols. The \(M \times \tau_{\text{up}}\) received signal matrix at the BS can be written as

\[
    Y = H \left( A^T \Psi^T \right) + N
\]

where \(H = [h_1, \ldots, h_K]\), \(A = \text{diag}(\beta_1 p_1, \beta_2 p_2, \ldots, \beta_K p_K)\), \(\Psi = [\psi_1^T, \ldots, \psi_K^T]\), and \(N\) is a \(M \times \tau_{\text{up}}\) noise matrix with its element in the \(i\)th row and the \(j\)th column \(n_{ij} \sim C\mathcal{N}(0, \sigma^2)\). Due to the orthogonality of the pilot sequences, \(\Psi^T \Psi = \tau_{\text{up}} I_{\tau_{\text{up}}}\). As shown in [23], the MMSE estimate of \(H\) can be written as

\[
    \hat{H} = Y \Psi^T A^T \left( \tau_{\text{up}} A + \sigma^2 I_K \right)^{-1}.
\]

The channel estimation error of the MMSE channel estimator is

\[
    W = H - \hat{H} = E \sigma \left( \tau_{\text{up}} A + \sigma^2 I_K \right)^{-1/2}
\]

where \(E\) is a \(M \times K\) matrix with its element in the \(i\)th row and the \(j\)th column \(e_{ij} \sim C\mathcal{N}(0, 1)\). Here, the estimation error \(W\) is independent of estimated channel \(\hat{H}\). Using the estimated channel, the BS designs its ZF precoder \(A = [\hat{a}_1, \ldots, \hat{a}_K]\).
where \( \hat{a}_k \) is the \( k \)th normalized column vector of \((\mathbf{H}^T)^\dagger = \mathbf{H}^* (\mathbf{H}^T \mathbf{H})^{-1} \).

2) Downlink Effective Channel Estimation: Assume coherent symbol detection and each user needs to know its effective downlink channel \( \mathbf{h}_k^T \hat{a}_k \). Therefore, before the BS transmits data symbols, it sends \( K \) distinct \( 1 \times T_{\text{ch}} \) orthogonal pilot sequences. For example, in the LTE-Advanced system, precoded downlink pilot signal called user equipment specific reference signal (UE-RS) is employed to enable each user to estimate its effective downlink channel [24]. The accuracy of the estimated effective downlink channel depends on the downlink pilot power of the BS. Since this paper investigates the EE of users, we assume the downlink pilot power of the BS is sufficiently high and each user has perfect knowledge of its downlink effective channel. The focus of this paper is on phases 1 and 3 and on how to determine the uplink pilot power and the downlink rate of users to help them achieve the highest EE.

3) Downlink Data Transmission: Then, the BS transmits the precoded data symbol vector

\[
x = \sum_{k=1}^{K} \hat{a}_k u_k
\]

where \( u_k \) is the message symbol for the \( k \)th user with \( \mathbb{E}[|u_k|^2] = p^m / K \).

From (1) and (3), the received signal at the \( k \)th user can be rewritten as

\[
r_k = \sqrt{\beta_k} \mathbf{h}_k^T \hat{a}_k u_k + \sum_{j=1, j \neq k}^{K} \sqrt{\beta_j} \mathbf{h}_k^T \hat{a}_j u_j + n_k.
\]

The SINR of the \( k \)th user can be written as

\[
S_k = \frac{\beta_k p_k^m}{\sum_{j=1, j \neq k}^{K} \beta_j p_j^m |\mathbf{h}_k^T \hat{a}_j|^2 + \sigma^2}.
\]

(4)

From (2), \( \mathbf{h}_k \) can be written as

\[
\mathbf{h}_k = \hat{\mathbf{h}}_k + \sqrt{\frac{\sigma^2}{\tau_{\text{up}} \beta_k p_k + \sigma^2}} \mathbf{e}_k
\]

where \( \hat{\mathbf{h}}_k \) and \( \mathbf{e}_k \) are the \( k \)th columns of \( \mathbf{H} \) and \( \mathbf{E} \), respectively. Using (5) and the nulling property of a ZF precoder, i.e.,

\[
\hat{\mathbf{h}}_k^T \hat{a}_j = 0
\]

for all \( l \neq j \), we have

\[
|\mathbf{h}_k^T \hat{a}_j|^2 = \frac{\sigma^2}{\tau_{\text{up}} \beta_k p_k + \sigma^2} |\mathbf{e}_k^T \hat{a}_j|^2
\]

for all \( j \neq k \). From (7), we can see that the power of the interuser interference experienced by user \( k \) depends only on its own pilot power \( p_k \) when the ZF precoder is used at the BS. It can be intuitively explained as follows. From the nulling property of a ZF precoder in (6), we see that the BS designs \( \hat{a}_j \) such that it lies in \( \mathcal{N}(\text{sp}\{\mathbf{h}_l\}_{l \neq j}) \), which denotes the null space spanned by \( \{\mathbf{h}_l\}_{l \neq j} \). If user \( k \) (\( k \neq j \)) uses a higher pilot power \( p_k \) to make its estimated channel \( \mathbf{h}_k \) more accurate (closer to \( \hat{\mathbf{h}}_k \), then \( \hat{a}_j \), which is perfectly orthogonal to \( \hat{\mathbf{h}}_k \), becomes closer to the null space of \( \mathbf{h}_k \), and the magnitude of the inner product \( \mathbf{h}_k^T \hat{a}_j \), which is due to the channel estimation error and causes interference to user \( k \), decreases. Furthermore, the channel estimation error of each user, as shown in (2), is independent of the pilot powers of other users. Therefore, the interference power experienced by user \( k \) depends only on its pilot power \( p_k \). Define \( X_k = |\mathbf{h}_k^T \hat{a}_k|^2 \) and \( Y_{k,j} = |\mathbf{e}_k^T \hat{a}_j|^2 \).

Then, using (7), (4) can be expressed as

\[
S_k = \frac{p_k \beta_k \mathbf{h}_k^T \hat{a}_k}{\tau_{\text{up}} p_k + 1} \sum_{j=1, j \neq k}^{K} Y_{k,j} + 1
\]

(8)

where \( \rho_k = \frac{\beta_k}{\sigma^2} \) is the channel to noise ratio of the \( k \)th user.

III. AVERAGE THROUGHPUT ANALYSIS

In this section, we derive the SINR distribution and the downlink average throughput of each user in the TDD MU-MIMO system. For analytic simplicity, similar to [25]-[27], we assume \( M = K \) for the rest of the paper. In a multiuser MIMO scenario, if the number of users \( K_u \) who want to access the channel is larger than the number of BS antennas, i.e., \( K_u > M \), a multiuser scheduler can be used to select only \( K = M \) users out of those \( K_u \) users to be serviced at the same time using multiuser MIMO [22]. This is usually the case in cellular communications. For example in LTE-A Release 10, the BS may have up to four antennas but there are usually more than ten or dozens of active users. So scheduling is always used. Therefore, the case of \( K = M \) has a significant meaning in a multiuser MIMO scenario.

A. SINR Distribution

To find the distribution of \( S_k \), we will calculate the distributions of \( X_k \) and \( Y_{k,j} \), respectively. First, we show that \( X_k = |\mathbf{h}_k^T \hat{a}_k|^2 \) is independent of all pilot powers \( p_j \) when \( M = K \). From (6), \( \hat{a}_k \in \mathcal{N}(\text{sp}\{\mathbf{h}_j\}_{j \neq k}) \), whose dimension reduces to one when \( M = K \). This implies that when \( M = K \), \( \hat{a}_k \) is uniquely determined by \( \mathbf{h}_k \). \( \hat{a}_k \), which is independent of \( p_k \) but dependent on \( \mathbf{h}_k \), i.e., \( \mathbf{h}_k \sim \mathcal{CN}(0, \mathbf{I}_M) \) is a Gaussian random vector and \( \hat{a}_k \) is an \( M \times 1 \) random unit-norm vector independent of \( \mathbf{h}_k \), by lemma 1,

\[
X_k = |\mathbf{h}_k^T \hat{a}_k|^2 \sim \text{Exp}(1)
\]

(9)

where \( \text{Exp}(m) \) denotes the exponential distribution with mean \( m \). It is clear from (9) that the distribution of \( X_k = |\mathbf{h}_k^T \hat{a}_k|^2 \) is independent of the other users’ pilot powers \( \{p_l\}_{l \neq k} \).

Lemma 1: Consider an \( M \times 1 \) Gaussian random vector \( \mathbf{g} \sim \mathcal{CN}(0, \mathbf{I}_M) \) and an \( M \times 1 \) unit-norm random vector \( \mathbf{u} \) independent of \( \mathbf{g} \). Then,

\[
|\mathbf{g}^T \mathbf{u}|^2 \sim \text{Exp}(1).
\]

Similarly, using lemma 1, we show that

\[
Y_{k,j} = |\mathbf{e}_k^T \hat{a}_j|^2 \sim \text{Exp}(1)
\]

(10)
where we use the fact that \( e_k \sim CN(0, I_M) \) and \( \mathbf{a}_j \) is a unit norm vector independent of \( e_k \). Define

\[
    Z_k = \sum_{j=1, j\neq k}^{K} Y_{k,j}.
\]

(11)

Then, (8) can be rewritten as

\[
    S_k = \frac{\rho_k \rho_k \rho_k}{\rho_k \rho_k + \tau Z_k + 1} X_k.
\]

(12)

From (12) and our discussion above, we can see that the SINR of the \( k \)th user \( S_k \) depends only on its pilot power \( p_k \) and is independent of the other users’ pilot powers \( \{p_j\}_{j\neq k} \) when \( M = K \). From (12), the probability density function (PDF) of \( S_k \) can be written as

\[
    f_{S_k}(s) = \int_0^\infty f_{S_k|Z_k}(s|z)f_{Z_k}(z)dz
\]

(13)

where \( f_{Z_k}(z) \) is the PDF of \( Z_k \). From (9) and (12), we have

\[
    f_{S_k|Z_k}(s|z) = \theta_se^{-\theta_s s} \quad (s \geq 0)
\]

(14)

where

\[
    \theta_z = \frac{1}{\tau \rho_k \rho_k + 1} z + \frac{K}{\rho_k \rho_k}.
\]

Next, we calculate the PDF of \( Z_k \) in (11). As shown in Appendix A, \( \{Y_{k,j}\}_{j=1}^{K} \) are not independent. However, we approximate the PDF of \( Z_k \) by treating \( \{Y_{k,j}\}_{j=1}^{K} \) as independent (exponential) random variables. Then, the approximated PDF of \( Z_k \) is a chi-square distribution with \( 2(K - 1) \) degrees of freedom [28], that is,

\[
    f_{Z_k}^{app}(z) = \frac{z^{K-2}e^{-z}}{\Gamma(K-1)} \quad (z \geq 0).
\]

(15)

Then, the approximated PDF of \( S_k \) is given by

\[
    f_{S_k}^{app}(s) = \int_0^\infty f_{S_k|Z_k}(s|z)f_{Z_k}^{app}(z)dz.
\]

(16)

We show in Appendix B that although the approximation error of \( f_{Z_k}^{app}(z) \) is not negligible, the approximation error of \( f_{S_k}^{app}(s) \) is small for a moderate number of \( M = K \). As shown in Appendix C, we can calculate the CDF of \( S_k \) from (16)

\[
    F_{S_k}(s) \approx 1 - e^{-\frac{\tau \rho_k p_k + 1}{\tau \rho_k p_k + 1 + s} K - 1}.
\]

(17)

Fig. 3 and Fig. 4 show that the accuracy of our CDF approximation in (17) is reasonable for all practical ranges of \( \rho_k \), \( p_k \), and \( M = K \).

B. Average Throughput

We consider the average throughput

\[
    \bar{R}_k = r_k (1 - F_{S_k}^{out})
\]

(18)

where \( r_k \) is the downlink transmission rate of the \( k \)th user and

\[
    F_{S_k}^{out} = P_{r}\{\log_2(1 + S_k) < r_k\}
\]

(19)

\[
    = F_{S_k}(2^{r_k} - 1)
\]

is the the outage probability of the \( k \)th user. Using (17) and (19), (18) can be rewritten as

\[
    \bar{R}_k(r_k, p_k) \approx \bar{R}_k^{out}(r_k) \cdot f_k(r_k, p_k)
\]

(20)

where

\[
    \bar{R}_k^{out}(r_k) \equiv r_k e^{-\frac{\tau \rho_k p_k + 1}{\tau \rho_k p_k + 2^{r_k}-1}}
\]

(21)

is the average throughput of the \( k \)th user when the BS has perfect channel knowledge and

\[
    f_k(r_k, p_k) \equiv \left(\frac{\tau \rho_k p_k + 1}{\tau \rho_k p_k + 2^{r_k}}\right)^{K-1}
\]

(22)

represents the throughput-loss factor due to the imperfect channel knowledge at the BS. As shown in Appendix D,
\( f_k(r_k, p_k) \) is a strictly increasing function of \( p_k \) and
\[
\frac{1}{2(K-1)}r_k \leq f_k(r_k, p_k) \leq 1.
\]
Also, it is easy to see that \( f_k(r_k, p_k) \) is decreasing in \( r_k \).

IV. ENERGY-EFFICIENT ADAPTATION OF UPLINK PILOT POWER AND DOWNLINK RATE

We are interested in the trade-off between the downlink rate that each user achieves and the uplink pilot power that the user consumes in the TDD MU-MIMO system. Therefore, we define the EE of each user as
\[
\eta_k = \frac{\tau_k \bar{R}_k}{\tau_{up} p_k + E_{cir}}
\]
where \( E_{cir} \triangleq T \rho_{cir} \) is the circuit energy consumption during a frame and \( p_{cir} \) is the circuit power of a mobile user which includes power consumption in a mixer, a frequency synthesizer, low noise amplifiers (LNA), analog-to-digital (A/D) converters, and filters, etc.

Since the EE of a user \( \eta_k \) does not depend on the other users’ pilot power \( \{p_j\}_{j \neq k} \), each user can find optimal \( p_k \) to maximize its EE individually. Therefore, we formulate the following EE optimization problem.

(P1) maximize \( \eta_k(r_k, p_k) \)
subject to \( p_k \leq p_{max} \),
\( r_k \geq r_{min} \)
where
\[
\eta_k(r_k, p_k) = \frac{\tau_k r_k e^{-\frac{\tau_k p_k + 1}{\tau_{up} p_k + E_{cir}}}}{r_k} + \frac{1}{r_k}
\]
and \( p_{max} \) is the maximum pilot power and \( r_{min} \) is the minimum downlink transmission rate.

We can show that \( \eta_k(r_k, p_k) \) is strictly quasi-concave in \( r_k \). Also, we can show that \( \eta_k(r_k, p_k) \) is strictly quasi-concave in \( p_k \) if practical values are used for system parameters \( E_{cir} \) and \( \rho_{cir} \). Now, we prove the strictly quasi-concavity of \( \eta_k \) in each coordinate.

A. Quasi-concavity of \( \eta_k(p_k) \)

We use the following lemma in [29] to check the quasi-concavity of \( \eta_k(p_k) \).

\textbf{Lemma 2:} A continuous function \( f : \mathbb{R} \to \mathbb{R} \) is strictly quasi-concave if and only if at least one of the following conditions holds:
- \( f \) is strictly decreasing,
- \( f \) is strictly increasing,
- There is a point \( c \in \text{dom } f \) such that for \( t \leq c \) (and \( t \in \text{dom } f \), \( f \) is strictly increasing, and for \( t \geq c \) (and \( t \in \text{dom } f \), \( f \) is strictly decreasing.

Consider the first-order derivative of \( \eta_k(p_k) \)
\[
\frac{\partial \eta_k(p_k)}{\partial p_k} = \frac{\tau_{up} \rho_{cir} p_k + 1}{\tau_{up} p_k + E_{cir}}
\]
where
\[
h_k(p_k) = -\tau_{up} \rho_{cir}^2 p_k - (K - (K - 2)2^{r_k}) \tau_{up} p_k^2 + (2^{r_k} - 1)(K - 1)p_k E_{cir} - 2^{r_k}
\]
Since the sign of \( h_k(p_k) \) is equal to that of \( \frac{\partial \eta_k}{\partial p_k} \), we only need to consider \( h_k(p_k) \) to characterize the shape of \( \eta_k(p_k) \). Note that the two roots of \( h_k(p_k) \) are
\[
\omega_1 \triangleq \frac{(K - 2)2^{r_k} - K - \sqrt{D_k}}{2p_{up} \tau_{up}}
\]
\( \omega_2 \triangleq \frac{(K - 2)2^{r_k} - K + \sqrt{D_k}}{2p_{up} \tau_{up}} \)
where
\[
D_k = (2^{r_k} - 1)(4(K - 1)\rho_k E_{cir} + 2^{r_k}(K - 2)^2 - K^2)
\]
is the discriminant of \( h_k(p_k) \). We consider the following cases of \( D_k \):
- Case I \( (D_k > 0) \):\nIn this case, \( h_k(p_k) \) can be written as
\[
h_k(p_k) = a(p_k - \omega_1)(p_k - \omega_2)
\]
where \( a = -\rho_{cir}^2 \tau_{up} < 0 \) and \( \omega_1 \) and \( \omega_2 \) are real numbers. Therefore, \( \eta_k(p_k) \) is strictly decreasing for \( p_k \in (-\infty, \omega_1) \), strictly increasing for \( p_k \in (\omega_1, \omega_2) \), and strictly decreasing for \( p_k \in (\omega_2, \infty) \).

Depending on the location of \( \omega_1 \) and \( \omega_2 \), we consider the following three cases.
- Case I-a (\( \omega_1 < \omega_2 < 0 \)): \( \eta_k(p_k) \) is strictly decreasing for \( p_k \in (0, \infty) \). \( \eta_k(p_k) \) is strictly quasi-concave.
- Case I-b (\( \omega_1 < 0 < \omega_2 \)): \( \eta_k(p_k) \) is strictly increasing for \( p_k \in (0, \omega_1) \) and decreasing for \( p_k \in (\omega_1, \omega_2) \). \( \eta_k(p_k) \) is strictly quasi-concave.
- Case I-c (\( 0 < \omega_1 < \omega_2 \)): \( \eta_k(p_k) \) is strictly decreasing for \( p_k \in (0, \omega_1) \), strictly increasing for \( p_k \in (\omega_1, \omega_2) \), and strictly decreasing for \( p_k \in (\omega_2, \infty) \). \( \eta_k(p_k) \) is not strictly quasi-concave.
- Case II (\( D_k \leq 0 \)):\nIn this case, \( h_k(p_k) \leq 0 \) for all \( p_k \geq 0 \). Therefore, \( \eta_k(p_k) \) is a strictly decreasing function of \( p_k \) and it is strictly quasi-concave.

\textbf{Remark:} From (28) and (29), the condition for Case I-a (\( \omega_1 < \omega_2 < 0 \)) is equivalent to
\[
\max \left( \frac{K}{2}, (K - 1)\rho_k E_{cir} \right) < 1 + \frac{1}{2^{r_k} - 1}.
\]
Since \( (K - 1)\rho_k E_{cir} \) is usually larger than \( K/2 \), (31) can be rewritten as
\[
r_k < \log_2 \left( 1 + \frac{1}{(K - 1)\rho_k E_{cir} - 1} \right)
\]
which implies that Case I-a happens when \( r_k \) is very low. Similarly, the condition for Case I-b (\( \omega_1 < 0 < \omega_2 \)) is equivalent to
\[
r_k > \log_2 \left( 1 + \frac{1}{(K - 1)\rho_k E_{cir} - 1} \right)
\]
which implies that Case I-b happens when \( r_k \) is not very low. However, as shown in Appendix E, Case I-c and Case II hardly occur under practical system parameters. Thus \( \eta_k(p_k) \) is strictly quasi-concave in \( p_k \) if we use practical system parameters.

### B. Quasi-concavity of \( \eta_k(r_k) \)

To check the quasi-concavity of \( \eta_k(r_k) \), we consider its first-order derivative

\[
\frac{\partial \eta_k(r_k)}{\partial r_k} = \eta_k(r_k) g_k(r_k) \tag{33}
\]

where

\[
g_k(r_k) \triangleq \frac{1}{r_k} - \frac{K}{\rho_k p_k^2 r_k^2} \ln 2 - \frac{(K - 1) 2^{r_k} \ln 2}{\tau_p \rho_k p_k + 2 r_k} \tag{34}
\]

Since the sign of \( g_k(r_k) \) is equal to that of \( \frac{\partial \eta_k}{\partial r_k} \), we only need to consider \( g_k(r_k) \) to characterize the shape of \( \eta_k(r_k) \). As shown in Appendix F, \( \eta_k(r_k) \) is strictly increasing for all \( r_k \in (0, \nu_k) \) and strictly decreasing for all \( r_k \in (\nu_k, \infty) \) where \( \nu_k \) is the unique solution of \( g_k(r_k) = 0 \) or

\[
K^2 r_k + (K - 1) \rho_k p_k^2 + K \tau_p \rho_k p_k = \nu_k^2 2^{r_k} \ln 2. \tag{35}
\]

Unfortunately, \( \nu_k \) cannot be expressed in a closed form. However, it can be easily found using a numerical algorithm, e.g. bisection method. From lemma 2, we can see that \( \eta_k(r_k) \) is strictly quasi-concave.

### C. Proposed Algorithm

Since \( \eta_k \) is strictly quasi-concave in each coordinate, we use the cyclic coordinated search method [32], which alternatively updates \( r_k \) and \( p_k \) by solving the following two subproblems.

- **Subproblem A**: Optimize \( p_k \) for a given \( r_k \), i.e.,
  \[
  \max_{p_k} \eta_k(p_k) \quad \text{subject to} \quad p_k \leq p_{\max}.
  \]

  From the four cases in subsection IV-A, we can see that the solution of subproblem A depends on the values of \( \omega_1 \) and \( \omega_2 \). In summary, the solution of subproblem A is

  \[
  p_k^\star(r_k) = \begin{cases} 
  0 & (\text{Case I-a or Case II}) \\
  \min_{\omega_2, p_{\max}} & (\text{Case I-b}) \\
  \max_{\{0, \min(\omega_2, p_{\max})\}} \eta_k(p_k) & (\text{Case I-c})
  \end{cases} \tag{36}
  \]

- **Subproblem B**: Optimize \( r_k \) for a given \( p_k \), i.e.,
  \[
  \max_{r_k} \eta_k(r_k) \quad \text{subject to} \quad r_k \geq r_{\min}.
  \]

  From subsection IV-B, \( \nu_k \) is the unique maximum point of \( \eta_k(r_k) \). Since the feasible set of subproblem B is \( r_k \geq r_{\min} \), its unique solution is

  \[
  r_k^\star(p_k) = \max(\nu_k, r_{\min}). \tag{37}
  \]

Using the solutions of subproblem A and B, we obtain the following cyclic coordinated search algorithm.

### Algorithm 1

**Initialization**: Choose \( r_k^{(1)} \geq r_{\min} \) and a tolerance \( \epsilon > 0 \).

**Iterations**: \( i \geq 1 \)

1. Calculate \( p_k^{(i+1)} = p_k^\star(r_k^{(i)}) \).
2. Calculate \( r_k^{(i+1)} = r_k^\star(p_k^{(i+1)}) \).
3. Set \( x^{(i+1)} = [r_k^{(i+1)}, p_k^{(i+1)}] \).
4. If \( ||x^{(i+1)} - x^{(i)}|| \leq \epsilon \) then stop; else set \( i = i + 1 \) and repeat next iteration.

### D. Convergence Property

Now, we establish the convergence of the proposed algorithm under the assumption of using practical values for the system parameters, which excludes Case I-c and Case II. As proved in Appendix G, we have the following properties of \( p_k^\star(r_k) \) and \( r_k^\star(p_k) \):

1. Excluding Case I-c and Case II, \( p_k^\star(r_k) \) is increasing in \( r_k \).
2. \( r_k^\star(p_k) \) is increasing in \( p_k \).
3. \( r_k^{(i)} \) is upper bounded by \( p_{\max} \).

Let \( \{x^{(i)} = [r_k^{(i)}, p_k^{(i)}]\}_{i=1}^{\infty} \) be the sequence of points (vectors) generated by Algorithm 1 and \( A(\cdot) \) be the mapping from \( x^{(i)} \) to \( x^{(i+1)} \), that is, \( x^{(i+1)} = A(x^{(i)}) \). The following theorem guarantees the existence of a fixed point of Algorithm 1.

**Theorem 1**: Algorithm 1 always has a fixed point \( \bar{x} \in \mathcal{F} \) defined as

\[
\lim_{i \to \infty} x^{(i)} = \bar{x} \in \mathcal{F}.
\]

**Proof**: See Appendix H.

Define the set of fixed points of Algorithm 1 as

\[
\mathcal{S}_k \triangleq \{x \in \mathcal{F} | x = A(x)\}.
\]

From theorem 1, we know that the solution set \( \mathcal{S}_k \neq \emptyset \) and the proposed algorithm always converges. Also, any solution of Algorithm 1 has the following property.

**Theorem 2**: Any fixed point of Algorithm 1 satisfies the first-order necessary condition (FONC). In other words, for any \( \bar{x} \in \mathcal{S}_k \),

\[
d^T \nabla \eta_k(\bar{x}) \leq 0
\]

where \( d \) is any feasible direction at \( \bar{x} \).

**Proof**: See Appendix I.

From theorem 2, we see that any fixed point of Algorithm 1 is a local maximum.

### V. Simulation Results

In this section, we provide simulation results to demonstrate the performance of the proposed algorithm.

The system parameters are listed in Table I. Since \( W = 10 \) kHz and \( N_0 = -174 \) dBm/Hz, the noise power \( \sigma^2 = N_0 W = -134 \) dBm. We use the channel parameters for the macrocell system in [30]. Then, the large-scale fading coefficient of the \( k \)th user is

\[
\beta_k \text{ (dB)} = -L(l_k) \text{ (dB)} + G_T \text{ (dBi)} + G_R \text{ (dBi)} + \sigma^2_0 \text{ (dB)} \tag{38}
\]

where \( l_k \) is the distance between the BS and the \( k \)th user in km,

\[
L(l_k) = 128.1 + 37.6 \log_{10}(l_k) \tag{39}
\]
is pathloss, $G_T = 14$ dB is the BS antenna gain, $G_R = 0$ dB is the user antenna gain, and $\sigma^2_\Omega$ (dB) is shadowing modeled as a Gaussian random variable with standard deviation of 8 dB, i.e., $\sigma^2_\Omega$ (dB) $\sim \mathcal{N}(0, 8^2)$.

The typical value of the circuit power of a mobile terminal is $p_{cir} = 20$ dBm (100 mW) [31]. We consider the circuit power in the range of $15$ dBm $\leq p_{cir} \leq 31$ dBm. Then, since $T = 30$, the circuit energy is in the range of 0 dB $\leq E_{cir} \leq 15$ dB. Also, we assume that $\tau_{up} = \tau_{dn} = K$.

**Algorithm 2 One-dimensional Exhaustive Search**

1) We partition $[r_{min}, r_{max}]$ into $N_r > 0$ smaller intervals, i.e.,

$$r_{min} = \psi_k^0 < \psi_k^1 < \ldots < \psi_k^{(N_r)} = r_{max}.\$$

2) The $n$th interval is characterized by its mid-point, i.e.,

$$r_k^{(n)} = \frac{\psi_k^{(n)} + \psi_k^{(n-1)}}{2}, \quad n = 1, \ldots, N_r.$$

3) For each $r_k^{(n)}$, calculate $p_k^{(n)} = p_k^0(r_k^{(n)})$.

4) Find a point which maximizes $\eta_k(r_k, p_k)$ among $N_r$ candidate points, $(p_k^{(n)}, r_k^{(n)})$.

We compare three algorithms: 1) Algorithm 1, 2) a one-dimensional exhaustive search algorithm, 3) a spectral efficiency (SE) maximization algorithm. Using the closed-form solution of subproblem A, we employ the following one-dimensional exhaustive search algorithm to maximize the EE solution of subproblem A, we employ the following one-dimensional exhaustive search algorithm to maximize the EE solution of subproblem A, we employ the following one-dimensional exhaustive search algorithm to maximize the EE solution of subproblem A, we employ the following one-dimensional exhaustive search algorithm to maximize the EE solution of subproblem A.

The reason is that although the objective function is not jointly quasi-concave in general, experimentally, we can verify that it is jointly quasi-concave for nearly all practical parameters and thus the point satisfies FONC coincides with the globally optimal solution of (P1).

Also, from the figure, we know how system parameters $p_k$ and $E_{cir}$ affect optimal point $(p_k^{op}, r_k^{op})$. First, as $E_{cir}$ increases, $p_k^{op}$ increases and eventually reaches the $p_{max}$ which is the solution of the SE maximization scheme. This is clear because $E_{cir}$ can be viewed as a fixed cost of communication which is consumed even if we do not send any uplink pilot. Therefore,
when $E_{\text{cir}}$ is sufficiently large, the EE maximization algorithm is the same as the SE maximization algorithm. Second, as $\rho_k$ increases, $r_k^{\text{op}}$ increases too. Since $\rho_k$ represents the channel to noise ratio, it is natural. The relationship between $\rho_k$ and $p_k^{\text{op}}$ is more interesting. As shown in the figure, $p_k^{\text{op}}$ is increasing in $\rho_k$ when $p_k$ is low while $p_k^{\text{op}}$ is decreasing in $\rho_k$ when $\rho_k$ is sufficiently high.

Fig. 7 and Fig. 8 show the EE and the SE of the $k$th user for $0 \text{ dBm} \leq p_k^{\text{op}} \leq 50 \text{ dBm}$ and $E_{\text{cir}} = 5 \text{ dB}$, respectively. By optimizing the uplink pilot power $\{p_k\}_{k=1}^K$ and downlink transmission rate $\{r_k\}_{k=1}^K$ in terms of EE, Algorithm 1 enhances the EE of the users significantly compared to that obtained from the SE maximization algorithm at the expense of a relatively small SE loss. As $\rho_k$ increases, the EE gap between Algorithm 1 and SE maximization algorithm increases too because in the high $\rho_k$ regime, $p_k^{\text{op}}$ of Algorithm 1 is a decreasing function of $\rho_k$ as shown in Fig. 6, but $p_k^{\text{op}}$ of SE maximization algorithm is unchanged and $p_k^{\text{op}} = p_{\max}$. EE of users. Currently, our work considers the case of $M = K$ in a single-cell environment. Extending this work to the case of $M > K$ and to a multi-cell environment will be interesting future research topics.

**VI. CONCLUSION**

In this paper, we have investigated the EE of users in a TDD MU-MIMO system. We have derived the closed-form expression of the average throughput and shown that the average throughput of the $k$th user is independent of the uplink pilot powers of the other users. Therefore, each user can maximize its EE independently. Unfortunately, the EE $\eta_k(r_k, p_k)$ function is not quasi-concave in general. But, with practical system parameters, we have shown that the EE function is strictly quasi-concave with respect to each coordinate, $r_k$ and $p_k$. Therefore, we have proposed an iterative uplink pilot power and downlink transmission rate adaptation algorithm to maximize the EE of users. We have proved that for any arbitrary starting point, the algorithm converges to a point that satisfies the first-order necessary condition. Comprehensive simulation results have been provided to demonstrate how system parameters affect optimal settings as well as the performance gain. From the simulation results, we can see that the proposed algorithm converges to the globally optimal solution within a few iterations and significantly enhances the performance gain. From the simulation results, we can see that the proposed algorithm converges to the globally optimal solution within a few iterations and significantly enhances the performance gain. From the simulation results, we can see that the proposed algorithm converges to the globally optimal solution within a few iterations and significantly enhances the performance gain.

**APPENDIX A**

**PROOF OF DEPENDENCY OF $Y_{k,i}$ AND $Y_{k,j}$ ($i \neq j$)**

Consider $Y_{k,i}$ and $Y_{k,j}$ ($i \neq j$). Let $\theta_{i,j}$ denote the angle between $\hat{a}_i$ and $\hat{a}_j$, that is,

$$\theta_{i,j} \triangleq \cos^{-1} \left| \hat{a}_i^H \hat{a}_j \right| \in [0, \pi/2].$$

Then, $\hat{a}_i$ can be written as the sum of $\hat{a}_j$ ($j \neq i$) and one of its orthonormal vectors, that is,

$$\hat{a}_i = \cos \theta_{i,j} \hat{a}_j + \sin \theta_{i,j} \mathbf{g}_{i,j}$$

(40)
where $g_{i,j}$ is a unit norm vector isotropically distributed in the nullspace of $\mathbf{A}_i^H$ and is independent of $\sin \theta_{i,j}$. From (10) and (40), $Y_{k,i}$ can be rewritten as

$$
Y_{k,i} = \cos^2 \theta_{i,j} |e_k^T \mathbf{a}_i|^2 + \sin^2 \theta_{i,j} |e_k^T \mathbf{g}_{i,j}|^2
$$

$$= \cos^2 \theta_{i,j} Y_{k,j} + \sin^2 \theta_{i,j} Y_{k,j}^\perp
$$

where $Y_{k,j} \perp \triangleq |e_k^T \mathbf{g}_{i,j}|^2 \sim \text{Exp}(1)$. Since $\cos^2 \theta_{i,j}$, $Y_{k,i}$, and $Y_{k,j}^\perp$ are independent, from (41), we have

$$
\text{Cov}[Y_{k,i}, Y_{k,j}] = E[Y_{k,i} Y_{k,j}^\perp] - E[Y_{k,i}] E[Y_{k,j}]
= E[|\cos^2 \theta_{i,j} Y_{k,j}^\perp + |\sin^2 \theta_{i,j} Y_{k,j}^\perp|] - 1
= 2E[|\cos^2 \theta_{i,j}|] + E[|\sin^2 \theta_{i,j}|] - 1
= E[|\cos^2 \theta_{i,j}|].
$$

(42)

Since the columns of a ZF precoding matrix $\mathbf{a}_i$'s are not orthogonal in general, $\cos \theta_{k,i} \neq 0$. Therefore, from (42), it is clear that $Y_{k,i}$ and $Y_{k,j}^\perp$ are not independent.

**APPENDIX B**

**APPROXIMATION OF SINR PDF**

The approximation error $e_S(s) \triangleq f_{S_k}(s) - f_{S_k}^{\text{app}}(s)$ can be written as

$$
e_S(s) = \int_{-\infty}^{\infty} f_{S_k|^z}(s) e^z dz
$$

(43)

where $e_S(z) \triangleq f_{S_k}(z) - f_{S_k}^{\text{app}}(z)$. The outline of our argument that $e_S(s)$ is reasonable for sufficiently small $M = K$ is as follows. We show experimentally in Fig. 9-(a) that the energy of $e_S(z)$ increases as $M = K$ increases. Since from (14) $f_{S_k|^z}(s) = \theta_2 e^{-\theta_2 s}$ where $\theta_2 = 1 + \frac{p_k^2}{\rho_k^2} + \frac{p^n}{\rho_n^2}$, for a sufficiently small $\theta_2$, $f_{S_k|^z}(s)$ is small and the contribution of $e_S(z)$ to $e_S(s)$ diminishes. As shown in Fig. 3-(b), even for intermediate values of $\theta_2$ ($\rho_k = 30$ dB, $p^{bn} = 30$ dBm, $p_k = 15, 20$ dBm) that might maximize $f_{S_k|^z}(s)$, the approximation error $e_S(s)$ is small enough as long as $M = K$ is not too large. Fig. 9-(a) shows $e_S(z)$ when $M = K = 4, 8$. We can see that the approximation error increases as $M = K$ increases. Also, $e_S(z)$ is positive for $[0, a]$, $e_S(z)$ is negative for $(a, b]$, $e_S(z)$ is positive for $(b, c]$, and $e_S(z)$ is negligible for $[c, \infty)$. From (11) and (15), we know that $e_S(z)$ depends only on $M = K$. In other words, $a, b, c > 0$ depends only on $M = K$ and is independent of other parameters. Therefore, (43) is upper-bounded by

$$
e_S(s) \leq a_1 \int_{a}^{b} e_S(z) dz + a_2 \int_{b}^{c} e_S(z) dz + a_3 \int_{c}^{\infty} e_S(z) dz
$$

(44)

where $a_1 = \max_{z \in [0,a]} f_{S_k|^z}(s)$, $a_2 = \min_{z \in (a,b]} (f_{S_k|^z}(s))$, and $a_3 = \max_{z \in (b,c]} (f_{S_k|^z}(s))$. Similarly, (43) is lower-bounded by

$$
e_S(s) \geq b_1 \int_{a}^{b} e_S(z) dz + b_2 \int_{b}^{c} e_S(z) dz + b_3 \int_{c}^{\infty} e_S(z) dz
$$

(45)

where $b_1 = \min_{z \in [0,a]} (f_{S_k|^z}(s))$, $b_2 = \min_{z \in (a,b]} (f_{S_k|^z}(s))$, and $b_3 = \min_{z \in [b,c]} (f_{S_k|^z}(s))$. Since we have the closed-form expression of $f_{S_k|^z}(s)$ in (14), we can always find the exact solution of $\max_{z \in [x_1,x_2]} (f_{S_k|^z}(s))$ or $\min_{z \in [x_1,x_2]} (f_{S_k|^z}(s))$ for an arbitrary interval $[x_1, x_2]$. Also, by numerical integration, $\int_{x_1}^{x_2} e_S(z) dz$ can be calculated for an arbitrary interval $[x_1, x_2]$. For example, when $M = K = 4$, from Fig. 9, we have $a_1 = 1.5, b = 5.9, c = 20$ and $\int_{1.5}^{5.9} e_S(z) dz = 0.1152$, $\int_{5.9}^{20} e_S(z) dz = 0.0442$. Using these values, we can easily evaluate the upper and lower bounds of $e_S(s)$ given by (44) and (45), respectively.

Fig. 9-(b) shows $e_S(s)$ and its upper and lower bounds given by (44) and (45) when $M = K = 4, 8$. Though the input error $e_S(z)$ is not negligible, the resulting error $e_S(s)$ is relatively small. For example, when $M = K = 8$, the maximum value of $e_S(s)$ is close to 0.1. But, the resulting error $e_S(s) \leq 0.005$ for nearly all $s$. The reason is twofold. First, as $s \rightarrow \infty$, $f_{S_k|^z}(s) = \theta_2 e^{-\theta_2 s}$ converges to zero very fast and it alleviates the error propagation. Second, even for small $s$, since $e(z)$ in $[0, a]$ and $(a, b]$ have different signs and thus the resulting errors in these two intervals cancel each other. Therefore, we conclude that the accuracy of our SINR PDF approximation is reasonable for small $M = K$.

**APPENDIX C**

**PROOF OF EQUATION (17)**

From (14) and (15), (16) can be rewritten as

$$
f_{S_k}^{\text{app}}(s) = \int_0^\infty \left( \frac{1}{\mu_k} \right) e^{-\left( \frac{1}{\mu_k^2} + \frac{K}{\rho_k^{pbn}} \right) s} dz
$$

$$= \frac{1}{\mu_k} e^{-\frac{K}{\rho_k^{pbn}}} \int_0^\infty \frac{K-2}{\Gamma(K-1)} z^{K-2} e^{-\left( \frac{K}{\rho_k^{pbn}} + \frac{1}{\mu_k^2} \right) z} dz
$$

(46)

where

$$
\mu_k = \tau_k \rho_k p_k + 1.
$$
Using the following integration in [33]
\[ \int_0^{\infty} x^n e^{-ax} \, dx = \frac{\Gamma(n+1)}{a^{n+1}}, \]
(46) can be written as
\[
f_{S_k}(s) \approx \frac{K-1}{\mu_k} e^{-\frac{k}{\rho_k p_m}} s \left( \frac{1}{\mu_k} s + 1 \right)^{-K} + \frac{K}{\rho_k p_m} e^{-\frac{k}{\rho_k p_m}} s \left( \frac{1}{\mu_k} s + 1 \right)^{-(K-1)}. \tag{47}
\]
From (47), the approximated CDF of \( S_k \) can be written as
\[
F_{S_k}(s) = \int_0^{\infty} f_{S_k}(x) \, dx \\
= \frac{K-1}{\mu_k} \int_0^{s} e^{-\frac{k}{\rho_k p_m} x} \left( \frac{1}{\mu_k} x + 1 \right)^{-K} \, dx \\
+ \frac{K}{\rho_k p_m} \int_0^{s} e^{-\frac{k}{\rho_k p_m} x} \left( \frac{1}{\mu_k} x + 1 \right)^{-(K-1)} \, dx \\
= 1 - e^{-\frac{k}{\rho_k p_m} s} \left( \tau_{apf} \rho_k p_k + 1 \right) K^{-1}. \tag{50}
\]

APPENDIX D
PROPERTY OF \( f_k(r_k, p_k) \)

\( f_k(r_k, p_k) \) is a strictly increasing function of \( p_k \) because
\[
\frac{\partial f_k}{\partial p_k} = f_k(r_k, p_k) \frac{(K-1)\tau_{apf} \rho_k p_k (2^{r_k} - 1)}{(\tau_{apf} \rho_k p_k + 1)(\tau_{apf} \rho_k p_k + 1)} > 0
\]
for all \( p_k \geq 0 \). Also, from (22), we have
\[
\lim_{p_k \to 0} f_k(p_k, r_k) = \frac{1}{2(K-1)r_k}, \\
\lim_{p_k \to \infty} f_k(p_k, r_k) = 1.
\]
Therefore, we obtain (23).

APPENDIX E
REMARK ON CASE I-c AND CASE II

The condition for Case I-c (\( 0 < \omega_1 < \omega_2 \)) is equivalent to
\[ p_{\text{cir}} < \frac{K}{2T p_k (K-1)} \]
and for Case II (\( D_k > 0 \)),
\[ p_{\text{cir}}(\text{dBm}) < \frac{K}{2(K-1)}(\text{dB}) - T(\text{dB}) + \sigma^2(\text{dBm}) - \beta_k(\text{dB}). \tag{48} \]

We use the practical system parameters in Table I and choose \( K = 4 \). Then, (48) can be written as
\[ p_{\text{cir}}(\text{dBm}) < -150.5 - \beta_k(\text{dB}). \tag{49} \]

Since the inter-site distance (ISD) of a typical macrocell is ISD = 1732 m [30], the maximum distance between the BS and a user is ISD/2 = 0.866 km, that is, \( l_k \leq 0.866 \) km. Thus, the path loss \( L(l_k) \) in (39) is upper bounded by
\[ L(l_k) \leq 125.75 \text{ dB}. \tag{50} \]

Since \( \sigma_{11}^2(\text{dB}) \sim N(0, 8^2) \), \( \sigma_{22}^2(\text{dB}) \geq -2.33 \times 8 \) dB with probability of 0.99. From (38) and (50), we have
\[ \beta_k(\text{dB}) \geq -130.39 \text{ dB}. \tag{51} \]

with probability of 0.99 even at the user located farthest from the BS, i.e., \( l_k = 0.866 \) km. From (51), the right-hand side of (49) is no more than \(-20.14 \) dBm with probability of 0.99. However, \( p_{\text{cir}} \) is much larger than \(-20.14 \) dBm (9.68 \mu W). For example, in [31], the typical value of the circuit power of mobile user is \( p_{\text{cir}} = 20 \) dBm (100 mW). Therefore, Case I-c hardly occurs under practical system parameters.

From (30), the condition for Case II (\( D_k \leq 0 \)) is equivalent to
\[ 4(K-1)Tp_{\text{cir}} \rho_k + 2^{-1} \epsilon (K-2)^2 \leq K^2. \tag{52} \]
For the same system parameters as above and \( p_{\text{cir}} = 20 \) dBm (100 mW), (52) becomes \( 36 \rho_k + 4.2^2 \leq 16 \), which is satisfied for only impractically small \( p_k \) and \( r_k \). Therefore, we conclude that Case II hardly occurs in practice.

APPENDIX F
PROOF OF QUASI-CONCAVITY OF \( \eta_k(r_k) \)

From (34), we have
\[
\frac{\partial g_k}{\partial r_k} = -\frac{1}{r_k^2} - \frac{K}{\rho_k p_m} 2^{r_k} (\ln 2)^2 - \frac{(K-1)2^{r_k} (\ln 2)^2 \tau_{apf} \rho_k p_k}{(\tau_{apf} \rho_k p_k + 1)^2} < 0. \tag{53} \]
which implies that \( g_k(r_k) \) is strictly decreasing over \( r_k \in (0, \infty) \). Since
\[
\lim_{r_k \to 0} g_k(r_k) = \infty, \\
\lim_{r_k \to \infty} g_k(r_k) = -\infty
\]
and \( g_k(r_k) \) is continuous, we know that there exists a unique \( \nu_k \in (0, \infty) \) such that \( g_k(\nu_k) = 0 \). Since the sign of \( g_k(r_k) \) is same as that of \( \eta_k(r_k) \), we conclude that \( \eta_k(r_k) \) is strictly increasing over \( r_k \in (0, \nu_k) \) and strictly decreasing over \( r_k \in (\nu_k, \infty) \).

APPENDIX G
PROOF OF PROPERTY OF \( p_k^*(r_{k, 1}) \) AND \( r_k^*(p_k) \)

A. Proof of Increasing Property of \( p_k^*(r_{k, 1}) \)

For \( r_{k,2} > r_{k,1} \), we prove \( p_k^*(r_{k,2}) \geq p_k^*(r_{k,1}) \). Since we exclude Case I-c and Case II from (36), we only need to consider the four cases listed in Table II.

In Case 1 and Case 2, the proof is trivial. Consider Case 3. From (29), \( \omega_2(r_k) \) is a strictly increasing function in \( r_k \). Therefore, in Case 4, we have
\[
p_k^*(r_{k,1}) = \min(\omega_2(r_{k,1}), p_{\text{max}}) \geq \min(\omega_2(r_{k,1}), p_{\text{max}}) = p_k^*(r_{k,1}).
\]

Finally, we show that Case 3 is impossible. For Case 3, \( p_k^*(r_{k,1}) = \min(\omega_2(r_{k,1}), p_{\text{max}}) \), which implies \( \omega_1(r_{k,1}) <
where (c) follows from (20) and (24).

where (a) follows from \( \omega_1(r_{k,1})/\omega_2(r_{k,1}) < 0 \) and (27), and (b) follows from \( r_{k,2} > r_{k,1} \). From (54), we have

\[
\omega_1(r_{k,1})/\omega_2(r_{k,2}) < 0 \tag{Case I-b}
\]

\( \text{From steps 1 and 2 of Algorithm 1, we have}
\]

\[
\max_{t_1: x_k + t_1 u_1 \in \mathcal{F}} \eta_k(x_k + t_1 u_1) = \eta_k(x_k)
\]

\[
\max_{t_2: x_k + t_2 u_2 \in \mathcal{F}} \eta_k(x_k + t_2 u_2) = \eta_k(x_k).
\]

\[
\eta_k(x_k + t_1 u_1) < \eta_k(x_k) \tag{62}
\]

\[
\eta_k(x_k + t_2 u_2) < \eta_k(x_k) \tag{63}
\]

\[
\frac{\partial \eta_k}{\partial p_k} \bigg|_{x_k = \bar{x}_k} = 0.
\]

\[
\frac{\partial \eta_k}{\partial p_k} \bigg|_{x_k = \bar{x}_k} < 0.
\]

\[
\frac{\partial \eta_k}{\partial p_k} \bigg|_{x_k = \bar{x}_k} > 0.
\]

\[
\nabla \eta_k(\bar{x}) = \left[ \frac{\partial \eta_k}{\partial p_k} \bigg|_{x_k = \bar{x}_k}, \frac{\partial \eta_k}{\partial r_k} \bigg|_{x_k = \bar{x}_k} \right]^T
\]

\[
\text{we have obtained the theorem.}
\]
REFERENCES


