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

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) multi-user 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 multicarrier 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-frocing 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

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

2   IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, ACCEPTED FOR PUBLICATION

Fig. 1.   TDD multiuser MIMO system model.

Fig. 2.   Downlink frame structure of TDD system.

orthogonal frequency division multiplexing (OFDM) because in OFDM systems the wideband channel is divided into many narrowband sub-bands, each experiencing flat fading. Denote $\sqrt{\beta_k}\mathbf{h}_k^T$ to be the downlink channel from the BS to the kth user, where β_k models large-scale fading that incorporates path-loss and shadowing and $\mathbf{h}_k^T \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$, a $1 \times M$ vector, models small-scale fading. The received signal at the kth user is

$$r_k = \sqrt{\beta_k}\mathbf{h}_k^T\mathbf{x} + n_k \qquad (1)$$

where $\mathbf{x}$ and $n_k \sim \mathcal{CN}(0, \sigma^2)$ 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 β_k is known and the small-scale fading vector $\mathbf{h}_k$ 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 τ_{up} symbols for uplink channel estimation, τ_{dn} symbols for downlink effective channel estimation, and the remaining $\tau_{\text{tr}} \triangleq T - \tau_{\text{up}} - \tau_{\text{dn}}$ symbols for downlink data transmission.

1) Uplink Channel Estimation: To estimate $\mathbf{h}_k$ at the BS, the kth user transmits a $1 \times \tau_{\text{up}}$ orthogonal pilot sequence vector $\sqrt{p_k}\boldsymbol{\psi}_k^T$ with $\|\boldsymbol{\psi}_k^T\|^2 = \tau_{\text{up}}$, 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

$$\mathbf{Y} = \mathbf{H}\left(\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Psi}^T\right) + \mathbf{N}$$

where $\mathbf{H} = [\mathbf{h}_1, \cdots, \mathbf{h}_K]$, $\boldsymbol{\Lambda} = \text{diag}(\beta_1 p_1, \beta_2 p_2, \cdots, \beta_K p_K)$, $\boldsymbol{\Psi} = [\boldsymbol{\psi}_1, \cdots, \boldsymbol{\psi}_K]$, and $\mathbf{N}$ is a $M \times \tau_{\text{up}}$ noise matrix with its element in the ith row and the jth column $n_{ij} \sim \mathcal{CN}(0, \sigma^2)$. Due to the orthogonality of the pilot sequences, $\boldsymbol{\Psi}^T\boldsymbol{\Psi}^* = \tau_{\text{up}}\mathbf{I}_{\tau_{\text{up}}}$. As shown in [23], the MMSE estimate of $\mathbf{H}$ can be written as

$$\hat{\mathbf{H}} = \mathbf{Y}\boldsymbol{\Psi}^*\boldsymbol{\Lambda}^{\frac{1}{2}}\left(\tau_{\text{up}}\boldsymbol{\Lambda} + \sigma^2\mathbf{I}_K\right)^{-1}.$$

The channel estimation error of the MMSE channel estimator is

$$\mathbf{W} \triangleq \mathbf{H} - \hat{\mathbf{H}}$$
$$= \mathbf{E}\sigma(\tau_{\text{up}}\boldsymbol{\Lambda} + \sigma^2\mathbf{I}_K)^{-\frac{1}{2}} \qquad (2)$$

where $\mathbf{E}$ is a $M \times K$ matrix with its element in the ith row and the jth colum $e_{ij} \sim \mathcal{CN}(0, 1)$. Here, the estimation error $\mathbf{W}$ is independent of estimated channel $\hat{\mathbf{H}}$. Using the estimated channel, the BS designs its ZF precoder $\hat{\mathbf{A}} = [\hat{\mathbf{a}}_1 \cdots \hat{\mathbf{a}}_K]$

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 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. $\mathbf{I}_N$ denotes an N by 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

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

KIM *et al.*: ENERGY EFFICIENT PILOT AND LINK ADAPTATION FOR MOBILE USERS IN TDD MULTI-USER MIMO SYSTEMS 3

where $\hat{\mathbf{a}}_k$ is the kth normalized column vector of $(\hat{\mathbf{H}}^T)^\dagger = \hat{\mathbf{H}}^* \left(\hat{\mathbf{H}}^T \hat{\mathbf{H}}^* \right)^{-1}$.

2) Downlink Effective Channel Estimation: Assume co-herent symbol detection and each user needs to know its effective downlink channel $\mathbf{h}_k^T \hat{\mathbf{a}}_k$. Therefore, before the BS transmits data symbols, it sends K distinct $1 \times \tau_{\text{dn}}$ 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

$$\mathbf{x} = \sum_{k=1}^{K} \hat{\mathbf{a}}_k u_k \tag{3}$$

where u_k is the message symbol for the kth user with $\mathbb{E}[|u_k|^2] = p^{\text{dn}}/K$.

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

$$r_k = \underbrace{\sqrt{\beta_k} \mathbf{h}_k^T \hat{\mathbf{a}}_k u_k}_{\text{signal}} + \underbrace{\sum_{j=1,j\neq k}^{K} \sqrt{\beta_k} \mathbf{h}_k^T \hat{\mathbf{a}}_j u_j}_{\text{interference}} + \underbrace{n_k}_{\text{noise}}.$$

The SINR of the kth user can be written as

$$S_k = \frac{\frac{\beta_k p^{\text{dn}}}{K} |\mathbf{h}_k^T \hat{\mathbf{a}}_k|^2}{\sum_{j=1,j\neq k}^{K} \frac{\beta_k p^{\text{dn}}}{K} |\mathbf{h}_k^T \hat{\mathbf{a}}_j|^2 + \sigma^2}. \tag{4}$$

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

$$\mathbf{h}_k = \hat{\mathbf{h}}_k + \sqrt{\frac{\sigma^2}{\tau_{\text{up}} \beta_k p_k + \sigma^2}} \mathbf{e}_k \tag{5}$$

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

$$\hat{\mathbf{h}}_l^T \hat{\mathbf{a}}_j = 0 \tag{6}$$

for all $l \neq j$, we have

$$|\mathbf{h}_k^T \hat{\mathbf{a}}_j|^2 = \frac{\sigma^2}{\tau_{\text{up}} \beta_k p_k + \sigma^2} |\mathbf{e}_k^T \hat{\mathbf{a}}_j|^2 \tag{7}$$

for all $j \neq k$. From (7), we can see that the power of the inter-user 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 $\hat{\mathbf{a}}_j$ such that it lies in $\mathcal{N}(\text{sp}(\{\hat{\mathbf{h}}_l\}_{l \neq j}))$, which denotes the null space spanned by $\{\hat{\mathbf{h}}_l\}_{l \neq j}$. If user k ($k \neq j$) uses a higher pilot power p_k to make its estimated channel $\hat{\mathbf{h}}_k$ more accurate

(closer to $\mathbf{h}_k$), then $\hat{\mathbf{a}}_j$, 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 $\mathbf{h}_k^T \hat{\mathbf{a}}_j$, 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 $X_k \triangleq |\mathbf{h}_k^T \hat{\mathbf{a}}_k|^2$ and $Y_{k,j} \triangleq |\mathbf{e}_k^T \hat{\mathbf{a}}_j|^2$. Then, using (7), (4) can be expressed as

$$S_k = \frac{\frac{\rho_k p^{\text{dn}}}{K} X_k}{\frac{\frac{\rho_k p^{\text{dn}}}{K}}{\tau_{\text{up}} \rho_k p_k + 1} \sum_{j=1,j\neq k}^{K} Y_{k,j} + 1} \tag{8}$$

where $\rho_k \triangleq \frac{\beta_k}{\sigma^2}$ is the channel to noise ratio of the kth 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 dis-tributions of X_k and $Y_{k,j}$, respectively. First, we show that $X_k = |\mathbf{h}_k^T \hat{\mathbf{a}}_k|^2$ is independent of all pilot powers p_l's when $M = K$. From (6), $\hat{\mathbf{a}}_k \in \mathcal{N}(\text{sp}\{\hat{\mathbf{h}}_j\}_{j \neq k})$, whose dimension reduces to one when $M = K$. This implies that when $M = K$, $\hat{\mathbf{a}}_k$ is uniquely determined by $\hat{\mathbf{h}}_j$'s ($j \neq k$) and independent of $\hat{\mathbf{h}}_k$. So, $\hat{\mathbf{a}}_k$ is independent of p_k. Moreover, since $\hat{\mathbf{h}}_j$'s ($j \neq k$) are statistically independent of $\mathbf{h}_k$, $\hat{\mathbf{a}}_k$ is also independent of $\mathbf{h}_k$. Since $\mathbf{h}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ is a Gaussian random vector and $\hat{\mathbf{a}}_k$ is an $M \times 1$ random unit-norm vector independent of $\mathbf{h}_k$, by lemma 1,

$$X_k = |\mathbf{h}_k^T \hat{\mathbf{a}}_k|^2 \sim \text{Exp}(1) \tag{9}$$

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

Lemma 1: Consider an $M \times 1$ Gaussian random vector $\mathbf{g} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ 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

$$Y_{k,j} = |\mathbf{e}_k^T \hat{\mathbf{a}}_j|^2 \sim \text{Exp}(1) \tag{10}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

4   IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, ACCEPTED FOR PUBLICATION

where we use the fact that $\mathbf{e}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ and $\hat{\mathbf{a}}_j$ is a unit norm vector independent of $\mathbf{e}_k$. Define

$$Z_k \triangleq \sum_{j=1, j \neq k}^{K} Y_{k,j}. \tag{11}$$

Then, (8) can be rewritten as

$$S_k = \frac{\frac{\rho_k p^{\text{dn}}}{K} X_k}{\frac{\rho_k p^{\text{dn}}}{\tau_{\text{up}} \rho_k p_k + 1} Z_k + 1}. \tag{12}$$

From (12) and our discussion above, we can see that the SINR of the kth 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

$$f_{S_k}(s) = \int_0^\infty f_{S_k | Z_k}(s|z) f_{Z_k}(z) dz \tag{13}$$

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

$$f_{S_k | Z_k}(s|z) = \theta_z e^{-\theta_z s} \quad (s \geq 0) \tag{14}$$

where

$$\theta_z = \frac{1}{\tau_{\text{up}} \rho_k p_k + 1} z + \frac{K}{\rho_k p^{\text{dn}}}.$$

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,

$$f_{Z_k}^{\text{app}}(z) = \frac{z^{K-2} e^{-z}}{\Gamma(K-1)} \quad (z \geq 0). \tag{15}$$

Then, the approximated PDF of S_k is given by

$$f_{S_k}^{\text{app}}(s) = \int_0^\infty f_{S_k | Z_k}(s|z) f_{Z_k}^{\text{app}}(z) dz. \tag{16}$$

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

$$F_{S_k}(s) \approx 1 - e^{-\frac{K}{\rho_k p^{\text{dn}}} s} \left(\frac{\tau_{\text{up}} \rho_k p_k + 1}{\tau_{\text{up}} \rho_k p_k + 1 + s} \right)^{K-1}. \tag{17}$$

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

B. Average Throughput

We consider the average throughput

$$\bar{R}_k = r_k \left(1 - P_k^{\text{out}}\right) \tag{18}$$

where r_k is the downlink transmission rate of the kth user and

$$\begin{aligned} P_k^{\text{out}} &= \Pr\{\log_2\left(1 + S_k\right) < r_k\} \\ &= F_{S_k}(2^{r_k} - 1) \end{aligned} \tag{19}$$
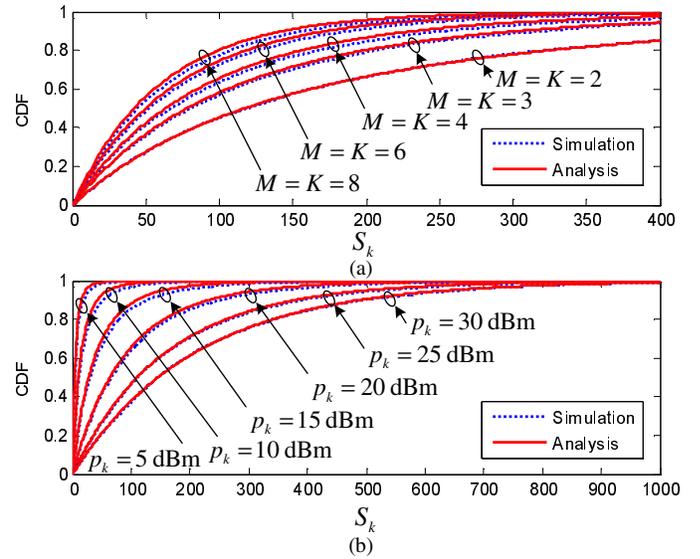
Fig. 3. Accuracy of the approximated CDF of S_k, $F_{S_k}^{\text{app}}(s)$, when $\rho_k = 30$ dB. (a) $M = K = 2, 3, 4, 6, 8$ and $p_k = 20$ dBm (b) $p_k = 5, 10, 15, 20, 25, 30$ dBm and $M = K = 4$.
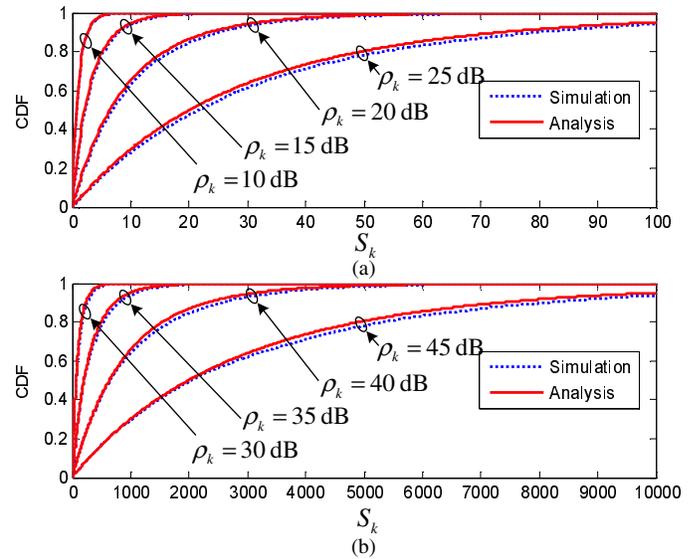
Fig. 4. Accuracy of the approximated CDF of S_k, $F_{S_k}^{\text{app}}(s)$, when $M = K = 4$ and $p_k = 20$ dBm. (a) $\rho_k = 10, 15, 20, 25$ dB (b) $\rho_k = 30, 35, 40, 45$ dBm.

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

$$\bar{R}_k(r_k, p_k) \approx \bar{R}_k^{\text{pf}}(r_k) \cdot f_k(r_k, p_k) \tag{20}$$

where

$$\bar{R}_k^{\text{pf}}(r_k) \triangleq r_k e^{-\frac{K}{\rho_k p^{\text{dn}}}(2^{r_k} - 1)} \tag{21}$$

is the average throughput of the kth user when the BS has perfect channel knowledge and

$$f_k(r_k, p_k) \triangleq \left(\frac{\tau_{\text{up}} \rho_k p_k + 1}{\tau_{\text{up}} \rho_k p_k + 2^{r_k}} \right)^{K-1} \tag{22}$$

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

$$\frac{1}{2^{(K-1)r_k}} \leq f_k(r_k, p_k) \leq 1. \tag{23}$$

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

$$\eta_k \triangleq \frac{\tau_{tr}\bar{R}_k}{\tau_{up}p_k + E_{cir}} \tag{24}$$

where E_{cir}[1] $\triangleq Tp_{cir}$ 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 η_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.

$$(\text{P1}) \quad \underset{r_k, p_k}{\text{maximize}} \ \eta_k(r_k, p_k)$$
$$\text{subject to } p_k \leq p_{\max},$$
$$r_k \geq r_{\min}$$

where

$$\eta_k(r_k, p_k) = \frac{\tau_{tr}r_k e^{-\frac{K}{\rho_k p^{dn}}}(2^{r_k}-1)\left(\frac{\tau_{up}\rho_k p_k+1}{\tau_{up}\rho_k p_k+2^{r_k}}\right)^{K-1}}{\tau_{up}p_k + E_{cir}} \tag{25}$$

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 ρ_k. Now, we prove the strictly quasi-concavity of η_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 $\eta_k(p_k)$.

Lemma 2: A continuous function $f : \mathbb{R} \rightarrow \mathbb{R}$ 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.

---

[1] In more detail, the circuit energy consumption can be written as $E_{cir} = \tau_{up}p_{cir,up} + \tau_{dn}p_{cir,dn} + \tau_{tr}p_{cir,tr}$ where $p_{cir,up}$, $p_{cir,dn}$, and $p_{cir,tr}$ are the circuit powers during uplink channel estimation phase, downlink effective channel estimation phase, downlink data transmission phase, respectively.

Consider the first-order derivative of $\eta_k(p_k)$

$$\frac{\partial \eta_k(p_k)}{\partial p_k} = \eta_k \frac{\tau_{up} \cdot h_k(p_k)}{(\tau_{up}\rho_k p_k + 2^{r_k})(\tau_{up}\rho_k p_k + 1)(\tau_{up}p_k + E_{cir})} \tag{26}$$

where

$$h_k(p_k) = -\tau_{up}^2\rho_k^2 p_k^2 - (K - (K-2)2^{r_k})\tau_{up}\rho_k p_k + (2^{r_k}-1)(K-1)\rho_k E_{cir} - 2^{r_k}. \tag{27}$$

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

$$\omega_1 \triangleq \frac{(K-2)2^{r_k} - K - \sqrt{D_k}}{2\rho_k\tau_{up}} \tag{28}$$

$$\omega_2 \triangleq \frac{(K-2)2^{r_k} - K + \sqrt{D_k}}{2\rho_k\tau_{up}} \tag{29}$$

where

$$D_k = (2^{r_k}-1)\left(4(K-1)\rho_k E_{cir} + 2^{r_k}(K-2)^2 - K^2\right) \tag{30}$$

is the discriminant of $h_k(p_k)$. We consider the following cases of D_k.

• Case I ($D_k > 0$)
In this case, (27) can be written as

$$h_k(p_k) = a(p_k - \omega_1)(p_k - \omega_2)$$

where $a = -\rho_k^2\tau_{up}^2 < 0$ and ω_1 and ω_2 ($\omega_1 < \omega_2$) 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 ω_1 and ω_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 $p_k \in (\omega_1, \infty)$. $\eta_k(p_k)$ 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 ($D_k \leq 0$)
In 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 thus it is strictly quasi-concave.

Remark: From (28) and (29), the condition for Case I-a ($\omega_1 < \omega_2 < 0$) is equivalent to

$$\max\left(\frac{K}{2}, (K-1)\rho_k E_{cir}\right) < 1 + \frac{1}{2^{r_k}-1}. \tag{31}$$

Since $(K-1)\rho_k E_{cir}$ is usually larger than $\frac{K}{2}$, (31) can be rewritten as

$$r_k < \log_2\left(1 + \frac{1}{(K-1)\rho_k E_{cir} - 1}\right). \tag{32}$$

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

$$r_k > \log_2\left(1 + \frac{1}{(K-1)\rho_k E_{cir} - 1}\right)$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

6   IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, ACCEPTED FOR PUBLICATION

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

$$g_k(r_k) \triangleq \frac{1}{r_k} - \frac{K}{\rho_k p^{\mathrm{dn}}} 2^{r_k} \ln 2 - \frac{(K-1)2^{r_k} \ln 2}{\tau_{\mathrm{up}} \rho_k p_k + 2^{r_k}}. \tag{34}$$

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 ν_k is the unique solution of $g_k(r_k) = 0$ or

$$\frac{\rho_k p^{\mathrm{dn}} 2^{\nu_k} + \tau_{\mathrm{up}} \rho_k^2 p^{\mathrm{dn}} p_k}{K 2^{\nu_k} + (K-1)\rho_k p^{\mathrm{dn}} + K \tau_{\mathrm{up}} \rho_k p_k} = \nu_k 2^{\nu_k} \ln 2. \tag{35}$$

Unfortunately, ν_k can not 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 η_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.,

$$\underset{p_k}{\text{maximize }} \eta_k(p_k)$$
$$\text{subject to } p_k \leq p_{\max}.$$

From the four cases in subsection IV-A, we can see that the solution of subproblem A depends on the values of ω_1 and ω_2. In summary, the solution of subproblem A is

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

• *Subproblem B:* Optimize r_k for a given p_k, i.e.,

$$\underset{r_k}{\text{maximize }} \eta_k(r_k)$$
$$\text{subject to } r_k \geq r_{\min}.$$

From subsection IV-B, ν_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

Initialize: 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 $\mathbf{x}^{(i+1)} = [r_k^{(i+1)}, p_k^{(i+1)}]$.
4) If $\|\mathbf{x}^{(i+1)} - \mathbf{x}^{(i)}\| \leq \epsilon$ 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 $r_{\max} \triangleq r_k^\star(p_{\max})$.
Let $\{\mathbf{x}^{(i)} = [r_k^{(i)}, p_k^{(i)}]\}_{i=1}^\infty$ be the sequence of points (vectors) generated by Algorithm 1 and $A(\cdot)$ be the mapping from $\mathbf{x}^{(i)}$ to $\mathbf{x}^{(i+1)}$, that is, $\mathbf{x}^{(i+1)} = A(\mathbf{x}^{(i)})$. The following theorem guarantees the existence of a fixed point of Algorithm 1.

Theorem 1: Algorithm 1 always has a fixed point $\bar{\mathbf{x}} \in \mathcal{F} \triangleq \{(r_k^{(i)}, p_k^{(i)}) | 0 \leq p_k \leq p_{\max}, r_{\min} \leq r_k \leq r_{\max}\}$, that is,

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

Proof: See Appendix H ∎
Define the set of fixed points of Algorithm 1 as

$$\mathcal{S}_k \triangleq \{\mathbf{x} \in \mathcal{F} \,|\, \mathbf{x} = A(\mathbf{x})\}.$$

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 $\bar{\mathbf{x}} \in \mathcal{S}_k$,

$$\mathbf{d}^T \nabla \eta_k(\bar{\mathbf{x}}) \leq 0$$

where $\mathbf{d}$ is any feasible direction at $\bar{\mathbf{x}}$.
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 kth user is

$$\beta_k \text{ (dB)} = -L(l_k) \text{ (dB)} + G_T \text{ (dBi)} + G_R \text{ (dBi)} + \sigma_\Omega^2 \text{ (dB)} \tag{38}$$

where l_k is the distance between the BS and the kth user in km,

$$L(l_k) = 128.1 + 37.6 \log_{10}(l_k) \tag{39}$$

TABLE I
SYSTEM PARAMETERS

System bandwidth W	10 kHz
Thermal noise N_0	-174 dBm/Hz
Frame length T	30 symbols
Total BS Tx power p_{dn}	30 dBm
Maximum Tx power of user p_{max}	25 dBm
Path loss L	$128.1 + 37.6 \log_{10}(l)$ dB, l in km
Shadowing standard deviation	8 dB
Inter-site distance (ISD) I	1732 m
BS antenna gain G_T	14 dBi
User antenna gain G_R	0 dBi

is pathloss, $G_T = 14$ dBi is the BS antenna gain, $G_R = 0$ dBi is the user antenna gain, and σ_Ω^2 (dB) is shadowing modeled as a Gaussian random variable with standard deviation of 8 dB, i.e., σ_Ω^2 (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.,

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

2) The nth interval is characterized by its mid-point, i.e.,

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

3) For each $r_k^{(n)}$, calculate $p_k^{(n)} = p_k^\star(r_k^{(n)})$.
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 of users. It is clear that this one-dimensional exhaustive search algorithm can find a point arbitrarily close to the global optimal point with a sufficiently large N_r. In simulation, we partition $[r_{min}, r_{max}]$ into N_r intervals of equal length 0.01. Also, we consider the SE maximization problem under the same constraints, i.e.,

(P2) $\quad \underset{r_k, p_k}{\text{maximize}} \quad \bar{R}_k(r_k, p_k) = \bar{R}_k^{pf}(r_k) f_k(r_k, p_k)$

$\quad \text{subject to} \quad r_k \geq r_{min},$

$\qquad\qquad\qquad p_k \leq p_{max}.$

Since $f_k(r_k, p_k)$ is a strictly increasing function of p_k, we can easily see that optimal solution to (P2) is the solution of subproblem B at p_{max}, i.e., $r_k^\star(p_{max})$.

Here, we briefly compare the complexity of the proposed algorithm and that of the exhaustive search. The proposed algorithm calculates at each iteration: i) $p_k^\star(r_k)$ using (36) and ii) $r_k^\star(p_k)$ using (37). In calculating (36), ω_2 is given in a closed-form expression by (29). But, to calculate ν_k
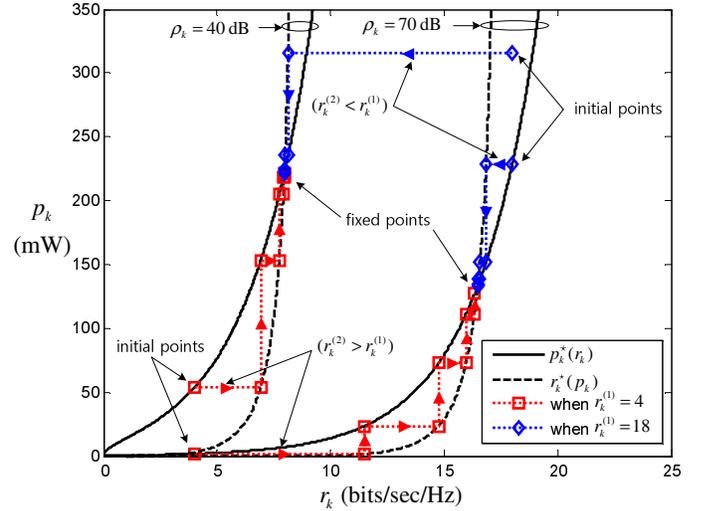
Fig. 5. Convergence of Algorithm 1 for two different initial values $r_k^{(1)} = 4, 18$. ($\rho_k = 40, 70$ dB, $E_{cir} = 10$ dB, $M = K = 4$, $r_{min} = 1$).

that appears in (37), we solve (35) for ν_k using the bisection algorithm. Since calculating (29) and (35) require only several multiplications and divisions, the main complexity of the proposed algorithm is that of the bisection algorithm to calculate ν_k. The complexity of the bisection is $O(\log_2 N_r)$, where N_r is the number of subintervals in $[r_{max}, r_{min}]$. In contrast, the complexity of the one-dimensional exhaustive search to calculate ν_k is $O(N_r)$. The complexity of the complete exhaustive search to find $p^\star$ and $r^\star$ is $O(N_r N_p)$, where N_p is the number of subintervals in $[0, p_{max}]$. Let I denote the number of iterations required for convergence. Then, from above the total complexity of Algorithm 1 is $I \times O(\log_2 N_r)$. It is usually difficult to find the number of iterations required for convergence I in an analytical way for iterative algorithms. According to our simulation results (see Fig. 4), usually $I = 4$ or 5 iterations is sufficient for the convergence of Algorithm 1.

Fig. 5 shows the convergence of Algorithm 1 for two different initial values $r_k^{(1)} = 4, 18$. We can check that when $r_k^{(2)} > r_k^{(1)}$ ($r_k^{(2)} < r_k^{(1)}$), $r_k^{(i)}$ and $p_k^{(i)}$ are increasing (decreasing) sequences for all $i = 2, 3, \cdots$ as expected in the increasing properties of $p_k^\star(r_k)$ and $r_k^\star(p_k)$. As shown in the figure, for two different starting points, the proposed algorithm converges to the same fixed point within a few iterations. Fig. 6 shows optimal uplink pilot power and downlink transmission rate pairs $\{(p_k^{op}, r_k^{op})\}$'s found by Algorithm 1 and 2, respectively. As shown in the figure, Algorithm 1 and Algorithm 2 yield the same solutions for all system parameters ρ_k and E_{cir}. The reason is that although the objective function $\eta_k(r_k, p_k)$ is not jointly quasi-concave in general, experimentally, we can verify that $\eta_k(r_k, p_k)$ 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 ρ_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,
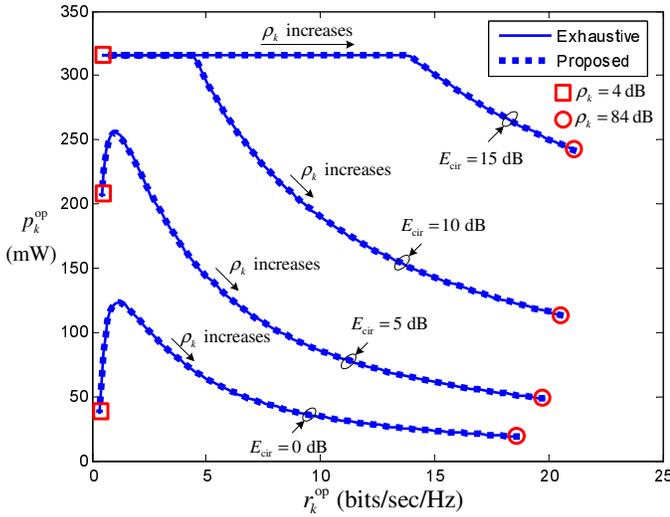
Fig. 6. Optimal uplink pilot power and downlink transmission rate $(r_k^{\text{op}}, p_k^{\text{op}})$ for all 4 dB $\leq \rho_k \leq 84$ dB ($E_{\text{cir}} = 0, 5, 10, 15$ dB, $M = K = 4$, $r_{\min} = 0$).

Fig. 7. EE of the kth user for 0 dBm $\leq p^{\text{dn}} \leq 50$ dBm. ($\rho_1 = 60$ dB, $\rho_2 = 50$ dB, $\rho_3 = 40$ dB, $\rho_4 = 30$ dB, $E_{\text{cir}} = 5$ dB, $M = K = 4$, $r_{\min} = 1$).

Fig. 8. SE of the kth user for 0 dBm $\leq p^{\text{dn}} \leq 50$ dBm. ($\rho_1 = 60$ dB, $\rho_2 = 50$ dB, $\rho_3 = 40$ dB, $\rho_4 = 30$ dB, $E_{\text{cir}} = 5$ dB, $M = K = 4$, $r_{\min} = 1$).

when E_{cir} is sufficiently large, the EE maximization algorithm is the same as the SE maximization algorithm. Second, as ρ_k increases, r_k^{op} increases too. Since ρ_k represents the channel to noise ratio, it is natural. The relationship between ρ_k and p_k^{op} is more interesting. As shown in the figure, p_k^{op} is increasing in ρ_k when ρ_k is low while p_k^{op} is decreasing in ρ_k when ρ_k is sufficiently high.

Fig. 7 and Fig. 8 show the EE and the SE of the kth user for 0 dBm $\leq p^{\text{dn}} \leq 50$ dBm and $E_{\text{cir}} = 5$ 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 ρ_k increases, the EE gap between Algorithm 1 and SE maximization algorithm increases too because in the high ρ_k regime, p_k^{op} of Algorithm 1 is a decreasing function of ρ_k as shown in Fig. 6, but p_k^{op} of SE maximization algorithm is unchanged and $p_k^{\text{op}} = p_{\max}$.

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 kth 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
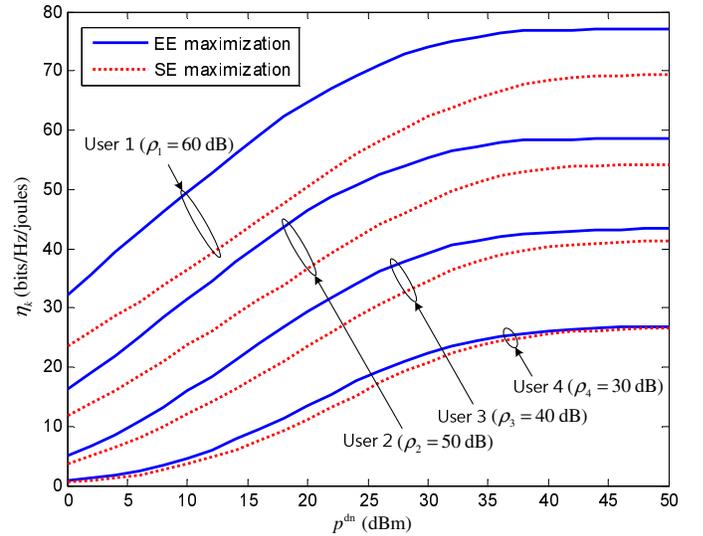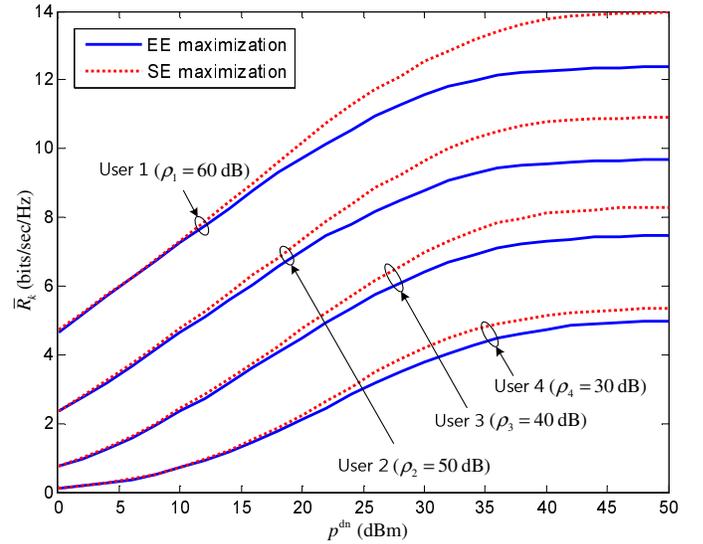
EE of users. Currently, our work considers the case of $M = K$ in a single-cell environment. Extending this work to the case of $M > K$ and to a multi-cell environment will be interesting future research topics.

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 $\hat{\mathbf{a}}_i$ and $\hat{\mathbf{a}}_j$, that is,

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

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

$$\hat{\mathbf{a}}_i = \cos \theta_{i,j} \hat{\mathbf{a}}_j + \sin \theta_{i,j} \mathbf{g}_{i,j} \qquad (40)$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

KIM *et al.*: ENERGY EFFICIENT PILOT AND LINK ADAPTATION FOR MOBILE USERS IN TDD MULTI-USER MIMO SYSTEMS 9

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

$$Y_{k,i} = \cos^2\theta_{i,j}|\mathbf{e}_k^T\hat{\mathbf{a}}_j|^2 + \sin^2\theta_{i,j}|\mathbf{e}_k^T\mathbf{g}_{i,j}|^2$$
$$= \cos^2\theta_{i,j}Y_{k,j} + \sin^2\theta_{i,j}Y_{k,j}^\perp \qquad (41)$$

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

$$\mathrm{Cov}[Y_{k,i}Y_{k,j}] = \mathbb{E}[Y_{k,i}Y_{k,j}] - \mathbb{E}[Y_{k,i}]\mathbb{E}[Y_{k,j}]$$
$$= \mathbb{E}[\cos^2\theta_{i,j}Y_{k,j}^2] + \mathbb{E}[\sin^2\theta_{i,j}Y_{k,j}^\perp Y_{k,j}] - 1$$
$$= 2\mathbb{E}[\cos^2\theta_{i,j}] + \mathbb{E}[\sin^2\theta_{i,j}] - 1$$
$$= \mathbb{E}[\cos^2\theta_{i,j}]. \qquad (42)$$

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

APPENDIX B
APPROXIMATION OF SINR PDF

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

$$e_S(s) = \int_0^\infty f_{S_k|Z_k}(s|z)e_Z(z)dz \qquad (43)$$

where $e_Z(z) \triangleq f_{Z_k}(z) - f_{Z_k}^{\mathrm{app}}(z)$. 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 $e_Z(z)$ increases as $M = K$ increases. Since from (14) $f_{S_k|Z_k}(s|z) = \theta_z e^{-\theta_z s}$ where $\theta_z = \frac{1}{\tau_{\mathrm{up}}\rho_k p_k + 1}z + \frac{K}{\rho_k p^{\mathrm{dn}}}$, for a sufficiently small θ_z or a sufficiently large θ_z, $f_{S_k|Z_k}(s|z)$ is small and the contribution of $e_Z(z)$ to $e_S(s)$ diminishes. As shown in Fig. 3-(b), even for intermediate values of θ_z ($\rho_k = 30$ dB, $p^{\mathrm{dn}} = 30$ dBm, $p_k = 15, 20$ dBm) that might maximize $f_{S_k|Z_k}(s|z)$, the approximation error $e_S(s)$ is small enough as long as $M = K$ is not too large. Fig. 9-(a) shows $e_Z(z)$ when $M = K = 4, 8$. We can see that the approximation error increases as $M = K$ increases. Also, $e_Z(z)$ is positive for $[0, a]$, $e_Z(z)$ is negative for $(a, b]$, $e_Z(z)$ is positive for $(b, c]$, and $e_Z(z)$ is negligible for $[c, \infty)$. From (11) and (15), we know that $e_Z(z)$ 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

$$e_S(s) \leq a_1\int_0^a e_Z(z)dz + a_2\int_a^b e_Z(z)dz + a_3\int_b^c e_Z(z)dz \qquad (44)$$

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

$$e_S(s) \geq b_1\int_0^a e_Z(z)dz + b_2\int_a^b e_Z(z)dz + b_3\int_b^c e_Z(z)dz \qquad (45)$$

where $b_1 = \min_{z\in[0,a]}\left(f_{S_k|Z_k}(s|z)\right)$, $b_2 = \max_{z\in(a,b]}\left(f_{S_k|Z_k}(s|z)\right)$, and $b_3 = \min_{z\in(b,c]}\left(f_{S_k|Z_k}(s|z)\right)$. Since we have the closed-form expression of $f_{S_k|Z_k}(s|z)$ in (14), we can always find the exact solution of $\max_{z\in[x_1,x_2]}\left(f_{S_k|Z_k}(s|z)\right)$ or $\min_{z\in[x_1,x_2]}\left(f_{S_k|Z_k}(s|z)\right)$
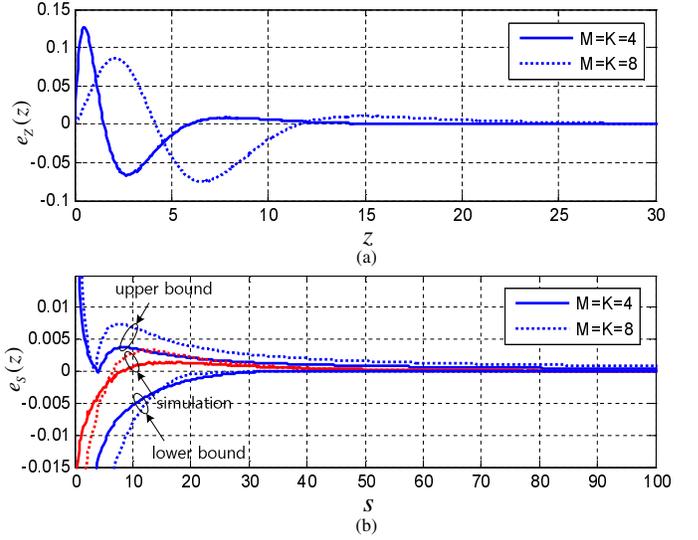
Fig. 9. (a) $e_Z(z)$ and (b) $e_S(s)$ and its upper and lower bounds in (44) and (45), respectively. ($M = K = 4, 8$, $\tau_{\mathrm{up}} = K$, $p^{\mathrm{dn}} = 30$ dBm, $\rho_k = 30$ dB, $p_k = 5$ dBm)

for an arbitrary interval $[x_1, x_2]$. Also, by numerical integration, $\int_{x_1}^{x_2} e_Z(z)dz$ can be calculated for an arbitrary interval $[x_1, x_2]$. For example, when $M = K = 4$, from Fig. 9, we have $a = 1.5$, $b = 5.9$, $c = 20$ and $\int_0^{1.5} e_Z(z)dz = 0.1152$, $\int_{1.5}^{5.9} e_Z(z)dz = -0.1598$, $\int_{5.9}^{20} e_Z(z)dz = 0.0442$. 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 $e_Z(s)$ is not negligible, the resulting error $e_S(s)$ is relatively small. For example, when $M = K = 8$, the maximum value of $e_Z(s)$ 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 $s \to \infty$, $f_{S_k|Z_k}(s|z) = \theta_z e^{-\theta_z s}$ 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

$$f_{S_k}^{\mathrm{app}}(s) = \int_0^\infty \left(\frac{1}{\mu_k}z + \frac{K}{\rho_k p^{\mathrm{dn}}}\right)e^{-\left(\frac{1}{\mu_k}z + \frac{K}{\rho_k p^{\mathrm{dn}}}\right)s}$$
$$\cdot z_k^{K-2}\frac{e^{-z}}{\Gamma(K-1)}dz$$
$$= \frac{1}{\mu_k} \cdot \frac{e^{-\frac{K}{\rho_k p^{\mathrm{dn}}}s}}{\Gamma(K-1)}\int_0^\infty z^{K-1}e^{-\left(\frac{s}{\mu_k}+1\right)z}dz$$
$$+ \frac{K}{\rho_k p^{\mathrm{dn}}} \cdot \frac{e^{-\frac{K}{\rho_k p^{\mathrm{dn}}}s}}{\Gamma(K-1)}\int_0^\infty z^{K-2}e^{-\left(\frac{s}{\mu_k}+1\right)z}dz \qquad (46)$$

where

$$\mu_k = \tau_{\mathrm{up}}\rho_k p_k + 1.$$

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

$$
f_{S_k}^{\text{app}}(s) \approx \frac{K-1}{\mu_k} e^{-\frac{K}{\rho_k p^{\text{dn}}} s} \left(\frac{1}{\mu_k} s + 1 \right)^{-K}
$$
$$
+ \frac{K}{\rho_k p^{\text{dn}}} e^{-\frac{K}{\rho_k p^{\text{dn}}} s} \left(\frac{1}{\mu_k} s + 1 \right)^{-(K-1)}. \quad (47)
$$

From (47), the approximated CDF of S_k can be written as

$$
F_{S_k}^{\text{app}}(s) = \int_0^\infty f_{S_k}^{\text{app}}(x) dx
$$
$$
= \frac{K-1}{\mu_k} \int_0^s e^{-\frac{K}{\rho_k p^{\text{dn}}} x} \left(\frac{1}{\mu_k} x + 1 \right)^{-K} dx
$$
$$
+ \frac{K}{\rho_k p^{\text{dn}}} \int_0^s e^{-\frac{K}{\rho_k p^{\text{dn}}} x} \left(\frac{1}{\mu_k} x + 1 \right)^{-(K-1)} dx
$$
$$
= 1 - e^{-\frac{K}{\rho_k p^{\text{dn}}} s} \left(\frac{\tau_{\text{up}} \rho_k p_k + 1}{\tau_{\text{up}} \rho_k p_k + 1 + s} \right)^{K-1}.
$$

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

$$
\frac{\partial f_k}{\partial p_k} = f_k(r_k, p_k) \frac{(K-1)\tau_{\text{up}} \rho_k (2^{r_k} - 1)}{(\tau_{\text{up}} \rho_k p_k + 2^{r_k})(\tau_{\text{up}} \rho_k p_k + 1)} > 0
$$

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

$$
p_{\text{cir}} < \frac{K}{2T\rho_k(K-1)}
$$

or

$$
p_{\text{cir}}(\text{dBm}) < \frac{K}{2(K-1)}(\text{dB}) - T(\text{dB}) + \sigma^2(\text{dBm}) - \beta_k(\text{dB}).
$$
(48)

We use the practical system parameters in Table I and choose $K = 4$. Then, (48) can be written as

$$
p_{\text{cir}}(\text{dBm}) < -150.5 - \beta_k(\text{dB}). \quad (49)
$$

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

$$
L(l_k) \leq 125.75 \text{ dB}. \quad (50)
$$

Since $\sigma_\Omega^2(\text{dB}) \sim \mathcal{N}(0, 8^2)$, $\sigma_\Omega^2(\text{dB}) \geq -2.33 \times 8$ dB with probability of 0.99. From (38) and (50), we have

$$
\beta_k(\text{dB}) \geq -130.39 \text{ dB} \quad (51)
$$

TABLE II
POSSIBLE COMBINATIONS OF $p_k^\star(r_{k,1})$ AND $p_k^\star(r_{k,2})$

	$p_k^\star(r_{k,1})$	$p_k^\star(r_{k,2})$
Case 1	0	0
Case 2	0	$\min(\omega_2(r_{k,2}), p_{\max})$
Case 3	$\min(\omega_2(r_{k,1}), p_{\max})$	0
Case 4	$\min(\omega_2(r_{k,1}), p_{\max})$	$\min(\omega_2(r_{k,2}), p_{\max})$

with probability of 0.99 even at the user located farthest from the BS, i.e., $l_k = 0.866$ km. From (51), the righthand side of (49) is no more than -20.14 dBm with probability of 0.99. However, p_{cir} is much larger than -20.14 dBm (9.68 μW). 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

$$
4(K-1)Tp_{\text{cir}}\rho_k + 2^{r_k}(K-2)^2 \leq K^2. \quad (52)
$$

For the same system parameters as above and $p_{\text{cir}} = 20$ dBm (100 mW), (52) becomes $36\rho_k + 4 \cdot 2^{r_k} \leq 16$, which is satisfied for only impractically small ρ_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

$$
\frac{\partial g_k}{\partial r_k} = -\frac{1}{r_k^2} - \frac{K}{\rho_k p^{\text{dn}}} 2^{r_k} (\ln 2)^2 - \frac{(K-1)2^{r_k}(\ln 2)^2 \tau_{\text{up}} \rho_k p_k}{(\tau_{\text{up}} \rho_k p_k + 2^{r_k})^2}
$$
$$
< 0. \quad (53)
$$

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 $p_k^\star(r_k)$ AND $r_k^\star(p_k)$

A. Proof of Increasing Property of $p_k^\star(r_k)$

For $r_{k,2} > r_{k,1}$, we prove $p_k^\star(r_{k,2}) \geq p_k^\star(r_{k,1})$. 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 4. From (29), $\omega_2(r_k)$ is a strictly increasing function in r_k. Therefore, in Case 4, we have

$$
p_k^\star(r_{k,2}) = \min(\omega_2(r_{k,2}), p_{\max}) \geq \min(\omega_2(r_{k,1}), p_{\max}) = p_k^\star(r_{k,1}).
$$

Finally, we show that Case 3 is impossible. For Case 3, $p_k^\star(r_{k,1}) = \min(\omega_2(r_{k,1}), p_{\max})$, which implies $\omega_1(r_{k,1}) <$

$0 < \omega_2(r_{k,1})$ (Case I-b). This condition is equivalent to $\omega_1(r_{k,1})\omega_2(r_{k,1}) < 0$. Then, we have

$$(K-1)\rho_k E_{\text{cir}} \underset{(a)}{>} 1 + \frac{1}{2^{r_{k,1}} - 1} \underset{(b)}{>} 1 + \frac{1}{2^{r_{k,2}} - 1} \quad (54)$$

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 $\omega_1(r_{k,2})\omega_2(r_{k,2}) < 0$ (Case I-b) and in this case $p_k^\star(r_{k,2}) = \min(\omega_2(r_{k,2}), p_{\max})$, which contradicts $p_k^\star(r_{k,2}) = 0$ for case 3.

B. Proof of Increasing Property of $r_k^\star(p_k)$

First, we show that $\frac{\partial \nu_k}{\partial p_k} > 0$ for all $p_k < \infty$. By differentiating (35) with respect to p_k, we have

$$\frac{\partial \nu_k}{\partial p_k} = \frac{\tau \rho_k a_k}{2^{\nu_k} b_k + c_k} \quad (55)$$

where

$$a_k = \frac{1}{\nu_k} - A2^{\nu_k}, \quad (56)$$

$$b_k = \frac{1}{\nu_k^2} - \ln 2 \frac{1}{\nu_k} + (K-1)(\ln 2)^2 \quad (57)$$

and

$$c_k = A2^{\nu_k} \ln 2(2^{\nu_k+1} + \tau \rho_k p_k) + \frac{\tau \rho_k p_k}{\nu_k^2} \quad (58)$$

with

$$A = \frac{K \ln 2}{\rho_k p^{\text{dn}}}.$$

Define

$$\nu_{\max} = \arg \max_{r_k} \bar{R}_k^{\text{pf}}(r_k).$$

Then, for $r_k > \nu_{\max}$, we have

$$\bar{R}_k^{\text{pf}}(r_k) f_k(r_k, p_k) \leq \bar{R}_k^{\text{pf}}(\nu_{\max}) f_k(\nu_{\max}, p_k) \quad (59)$$

where we use the facts that i) $\bar{R}_k^{\text{pf}}(r_k) \leq \bar{R}_k^{\text{pf}}(\nu_{\max})$ and ii) $f_k(r_k, p_k)$ is strictly decreasing in r_k. Therefore, for any given $p_k < \infty$,

$$\nu_k \underset{(c)}{=} \arg \max_{r_k} \bar{R}_k^{\text{pf}}(r_k) f_k(r_k, p_k)$$
$$< \nu_{\max}. \quad (60)$$

where (c) follows from (20) and (24).

Since $f_k(r_k, p_k) \to 1$ as $p_k \to \infty$, we see that $\nu_k \to \nu_{\max}$ as $p_k \to \infty$. Therefore, $\nu_{\max}$ needs to satisfy (35) when $p_k \to \infty$, that is,

$$\frac{1}{\nu_{\max}} - A2^{\nu_{\max}} = 0. \quad (61)$$

From (56) we can easily see that a_k is a strictly decreasing function of ν_k. Also, from (61) we have $a_k = 0$ when $\nu_k = \nu_{\max}$. But, $\nu_k < \nu_{\max}$ from (60). Therefore, we have $a_k > 0$, which implies that the numerator of (55) is positive for all $p_k < \infty$.

Now, we show that the denominator of (55) is positive for all $p_k < \infty$. Since $c_k > 0$ from (58), we only need to show that $b_k > 0$. From (57), b_k is a second order polynomial of $\frac{1}{\nu_k} > 0$ with a positive second order coefficient. Since the discriminant of (57) is $(\ln 2)^2(5 - 4K) < 0$ for all $K \geq 2$, we obtain $b_k > 0$.

From above, (55) is positive and ν_k is an increasing function of p_k. Therefore, it is clear that $r_k^\star(p_k) = \max(\nu_k, r_{\min})$ in (37) is also an increasing function of p_k.

C. Proof of the Upper Limit of $r_k^{(i)}$

Since $p_k \leq p_{\max}$ and $r_k^\star(p_k)$ is an increasing function of p_k, $r_k^{(i)}$ is upper bounded by $r_{\max} \triangleq r_k^\star(p_{\max})$.

APPENDIX H
PROOF OF THEOREM 1

Suppose $r_k^{(2)} \geq r_k^{(1)}$. Since $p_k^{(i+1)} = p_k^\star(r_k^{(i)})$ and $p_k^\star(r_k)$ is a increasing function of r_k, we have $p_k^{(3)} \geq p_k^{(2)}$. Then, since $r_k^{(i+1)} = r_k^\star(p_k^{(i+1)})$ and $r_k^\star(p_k)$ is a increasing function of p_k, we have $r_k^{(3)} \geq r_k^{(2)}$, which also implies $p_k^{(4)} \geq p_k^{(3)}$. From above,

$$r_k^{(i+1)} \geq r_k^{(i)} \quad \text{and} \quad p_k^{(i+1)} \geq p_k^{(i)}$$

for $i \geq 2$. Since $r_k^{(i)}$ and $p_k^{(i)}$ are increasing and bounded sequences in $\mathcal{F}$, by the monotone convergence theorem [34], $r_k^{(i)}$ and $p_k^{(i)}$ converge. Therefore, $\mathbf{x}^{(i)} = [r_k^{(i)}, p_k^{(i)}]$ converges to $\bar{\mathbf{x}} = [\bar{r}_k, \bar{p}_k] \in \mathcal{F}$. For the case $r_k^{(2)} < r_k^{(1)}$, theorem 1 can be similarly proved.

APPENDIX I
PROOF OF THEOREM 2

Since $\bar{\mathbf{x}} \in \mathcal{S}_k$, $A(\bar{\mathbf{x}}) = \bar{\mathbf{x}}$. From steps 1 and 2 of Algorithm 1, we have

$$\max_{t_1 : \bar{\mathbf{x}}_k + t_1 \mathbf{u}_1 \in \mathcal{F}} \eta_k(\bar{\mathbf{x}}_k + t_1 \mathbf{u}_1) = \eta_k(\bar{\mathbf{x}}_k)$$
$$\max_{t_2 : \bar{\mathbf{x}}_k + t_2 \mathbf{u}_2 \in \mathcal{F}} \eta_k(\bar{\mathbf{x}}_k + t_2 \mathbf{u}_2) = \eta_k(\bar{\mathbf{x}}_k).$$

where $\mathbf{u}_1 = [0, 1]$ and $\mathbf{u}_2 = [1, 0]$. Since $\eta_k(r_k, p_k)$ is strictly quasi-concave with respect to each coordinate r_k and p_k, optimal t_1 and t_2 are zero. For any $t_1 \neq 0$ such that $\bar{\mathbf{x}}_k + t_1 \mathbf{u}_1 \in \mathcal{F}$,

$$\eta_k(\bar{\mathbf{x}}_k + t_1 \mathbf{u}_1) < \eta_k(\bar{\mathbf{x}}_k). \quad (62)$$

For any $t_2 \neq 0$ such that $\bar{\mathbf{x}}_k + t_2 \mathbf{u}_2 \in \mathcal{F}$,

$$\eta_k(\bar{\mathbf{x}}_k + t_2 \mathbf{u}_2) < \eta_k(\bar{\mathbf{x}}_k). \quad (63)$$

Clearly, $0 \leq \bar{p}_k \leq p_{\max}$. This can be split into three cases.
- Case a: $\bar{\mathbf{x}}_k = [\bar{r}_k, \bar{p}_k]$ is an interior point, $0 < \bar{p}_k < p_{\max}$. From (62), we have

$$\left. \frac{\partial \eta_k}{\partial p_k} \right|_{\mathbf{x}_k = \bar{\mathbf{x}}_k} = 0.$$

- Case b: $\bar{\mathbf{x}}_k = [\bar{r}_k, \bar{p}_k]$ is a boundary point, $\bar{p}_k = 0$. From (62), we have

$$\left. \frac{\partial \eta_k}{\partial p_k} \right|_{\mathbf{x}_k = \bar{\mathbf{x}}_k} < 0.$$

- Case c: $\bar{\mathbf{x}}_k = [\bar{r}_k, \bar{p}_k]$ is a boundary point, $\bar{p}_k = p_{\max}$. From (62), we have

$$\left. \frac{\partial \eta_k}{\partial p_k} \right|_{\mathbf{x}_k = \bar{\mathbf{x}}_k} > 0.$$

For $r_{\min} \leq \bar{r}_k$, we can obtain similar results from (63). Since

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

we have obtained the theorem.

REFERENCES

[1] Z. Hasan, H. Boostanimehr, and V. K. Bhargava, "Green cellular networks: a survey, some research issues and challenges," *IEEE Commun. Surveys & Tutorials,* vol. 13, no. 4, pp. 524–540, Fourth Quarter 2011.

[2] G. Miao, N. Himayat, G. Y. Li, and A. Swami, "Cross-layer optimization for energy-efficient wireless communications: a survey," *Wireless Commun. and Mobile Comput.* vol. 9, no. 4. pp. 529–542, 2009.

[3] G. Y. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient wireless communications: tutorial, survey, and open issues," *IEEE Wireless Commun.*, vol. 18, no. 6, pp. 28–35, Dec. 2011.

[4] Y. Chen, S. Zhang, S. Xu, and G. Y. Li, "Fundamental trade-offs on green wireless networks," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 30–37, June 2011.

[5] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349–2360, Sept. 2005.

[6] C. Isheden and G. P. Fettweis, "Energy-efficient multi-carrier link adaptation with sum rate-dependent circuit power," in *Proc. 2010 IEEE GLOBECOM*, pp. 1–6.

[7] C. Isheden and G. P. Fettweis, "Energy-efficient link adaptation on parallel channels," in *Proc. 2011 Euripean Signal Process. Conf.*, pp. 874–878.

[8] G. Miao, N. Himayat, and G. Y. Li, "Energy-efficient transmission in frequency-selective channels," in *Proc. 2008 IEEE GLOBECOM*, pp. 1–5.

[9] G. Miao, N. Himayat, and G. Y. Li, "Energy-efficient link adaptation in frequency-selective channels," *IEEE Trans. Commun.*, vol. 58, no. 2, pp. 545–554, 2010.

[10] G. Miao, N. Himayat, G. Y. Li, and D. Bormann, "Energy-efficient design in wireless OFDMA," in *Proc. 2008 IEEE ICC*, pp. 3307–3312.

[11] G. Miao, N. Himayat, and G. Y. Li, "Low-complexity energy-efficient OFDMA," in *Proc. 2009 IEEE ICC*, pp. 1–5.

[12] H. Kim and B. Daneshrad, "Energy-constrained link adaptation for MIMO OFDM wireless communication systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2820–2832, Sept. 2010.

[13] E. V. Belmega and S. Lasaulce, "Energy-efficient precoding for multiple antenna terminals," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 329–340, Jan. 2011.

[14] R. S. Prabhu and B. Daneshrad, "An energy-efficient water-filling algorithm for OFDM systems," in *Proc. 2010 IEEE ICC*, pp. 1–5.

[15] R. S. Prabhu and B. Daneshrad, "Energy-efficient power loading for a MIMO-SVD system and its performance in flat fading," in *Proc. 2010 IEEE GLOBECOM*, pp. 1–5.

[16] G. Miao and J. Zhang, "On optimal energy-efficient multi-user MIMO," in *Proc. 2011 IEEE GLOBECOM*, pp. 1–6.

[17] J. Xu, L. Qiu, and S. Zhang, "Energy-efficient iterative waterfilling for the MIMO broadcasting channels," in *Proc. 2012 IEEE WCNC*, pp. 1–5.

[18] C. Hellings, N. Damak, and W. Utschick, "Energy-efficient zero-forcing with user selection in parallel vector broadcast channels," in *Proc. 2012 International ITG Workshop on Smart Antennas*, pp. 168–175.

[19] M. C. Gursoy, "On the capacity and energy efficiency of training-base transmissions over fading channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4543–4567, Oct. 2009.

[20] Z. Xu, G. Y. Li, C. Y. yang, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient power allocation for pilots in training-based downlink OFDMA systems," *IEEE Trans. Commun.*, vol. 60, no. 10, pp. 3047–3058, Oct. 2012.

[21] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channels with partial side information," *IEEE Trans. Inf. Theory.*, vol. 51, no. 2, Feb. 2005.

[22] T. Yoo, N. Jindal, and A. Goldsmith, "Multi-antenna downlink channels with limited feedback and user selection," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1478–1491, Sept. 2007.

[23] H. Poor, *An Introduction to Signal Detection and Estimation*. Springer-Verlag, 1994.

[24] Y. Nam, Y. Akimoto, Y. Kim, M. Lee, K. Bhattad, and A. Ekpenyong, "Evolution of reference signals for LTE-Advanced systems," *IEEE Commun. Mag.*, vol. 50, no. 2, pp. 132–138, Feb. 2012.

[25] N. Jindal, "MIMO broadcast channels with finite rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5059, Nov 2006.

[26] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, June 2010.

[27] M. Kobayashi, N. Jindal, and G. Caire, "Training and feedback optimization for multiuser MIMO downlink," *IEEE Trans. Commun.*, vol. 59, no. 8, pp. 2228–2240, Aug. 2011.

[28] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes,* 4th ed. McGraw-Hill, Inc., 2002.

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

[30] 3GPP TR 25.814 v7.1.0, "Physical Layer Aspect for Evolved Universal Terrestrial Radio Access (UTRA)."

[31] H. Kim, C. Chae, G. Veciana, and R. W. Heath, Jr., "A cross-layer approach to energy efficiency for adaptive MIMO systems exploiting spare capacity," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4264–4275, Aug. 2009.

[32] D. G. Luenberger, *Linear and Nonlinear Programming,* 2nd ed. Springer-Verlag, 2003.

[33] M. R. Spiegel and J. Liu, *Mathematical Handbook of Formulas and Tables*, 2nd ed. McGraw-Hill, 1999.

[34] W. Rudin, *Principles of Mathematical Analysis*. McGraw-Hill, 1976.

Yunesung Kim received his B.S. and M.S. degrees in Electrical and Electronic Engineering from Yonsei University, Seoul, Korea, in 2011 and 2013, respectively. He has been working on standardization and commercialization of the 3GPP LTE-Advanced and ETSI ISG ORI. His research interests include energy efficient communication, massive MIMO, mmwave beamforming, and CPRI compression for the next generation wireless systems.

Guowang Miao received his B.S. and M.S. degrees in electronic engineering in 2003 and 2006 from Tsinghua University, Beijing, China, and his M.S. and Ph.D. degrees in electrical and computer engineering in 2009 from Georgia Institute of Technology, Atlanta, GA, USA. He joined Dallas Telecom Lab of Samsung, Texas, in Jan 2010 as a Senior Algorithms and Standards Engineer and also a 3GPP RAN 1 delegate and worked on 5G technologies and 3GPP Long Term Evolution - Advanced (LTE-A) Standard, with a focus on both physical and MAC layers. In 2011, he won an Individual Gold Award from Samsung Telecom America for his contribution in LTE-A standardization. Starting in Fall 2011, he is an assistant professor in the Department of Communications Systems, KTH - The Royal Institute of Technology, Stockholm, Sweden. His research interest is in the design and optimization of wireless communications and networking. His paper on energy-efficient link adaptation is the most cited paper in IEEE TRANSACTIONS ON COMMUNICATIONS since 2010. He also has a top cited paper from recent years in *Wireless Communications and Mobile Computing*. Besides dozens of journal and conference papers, he has more than ten patents either approved or filed and several of them have been adopted in 4G cellular network standards. He serves as a technical reviewer of more than twenty international journals and conferences. He has been a technical program committee member of many international conferences. He is also an associate editor or on the editorial board of several international journals. He was an exemplary reviewer for IEEE COMMUNICATIONS LETTERS in 2011, representing fewer than 3% of all reviewers according to EiC.

Taewon Hwang received his B.S. and M.S. degrees in Electrical and Electronic Engineering from Yonsei University, Seoul, Korea in 1993 and in Electrical and Computer Engineering from Georgia Institute of Technology, Atlanta, Georgia in 1995, respectively. From 1995 to 2000, he was a Member of Technical Staff at Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea. He received his Ph.D. degree from the Department of Electrical and Computer Engineering, Georgia Institute of Technology in 2005. Since 2006, he has been with the School of Electrical and Electronic Engineering at Yonsei University, Seoul, Korea, where he is currently an Associate Professor. His research interests include energy-efficient communications, cognitive radio, and MIMO communications. He has been a technical program committee member of many international conferences. Also, he is currently serving as an Editor of the IEEE JOURNAL ON SELECTED AREAS ON COMMUNICATIONS: COGNITIVE RADIO SERIES.