IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 8, 2019, accepted December 30, 2019, date of publication January 6, 2020, date of current version January 16, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2964393

Energy-Efficient Channel Estimation

GUANDONG WEI[1], KAI YANG[1], (Member, IEEE), XIAOZHENG GAO[1], YE YU[1], JINSONG WU[2], (Senior Member, IEEE), AND JIANPING AN[1], (Member, IEEE)

[1]School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China
[2]Department of Electrical Engineering, Universidad de Chile, Santiago 1058, Chile

Corresponding author: Kai Yang (yangkai@ieee.org)

This work was supported in part by the National Natural Science Foundation of China under Grant 61771054 and Grant 61601025.

ABSTRACT We investigate the energy-efficient channel estimation in wireless networks, where one pilot is inserted for every several data slots to estimate the channel coefficients. Both the channel state information estimation error and the time-varying model are considered. We first formulate the energy-efficient channel estimation problem into a mixed integer nonlinear programming (MINLP) problem, where the variables include the transmitted power and the number of data slots. Due to the NP-hard nature, we degenerate the MINLP problem into a series of non-concave optimization problems without integer variables. Then we solve these problems using successive convex approximation, geometric programming, and the Dinkelbach algorithm to obtain a point satisfying the Karush-Kuhn-Tucker (KKT) conditions. Furthermore, we develop a low-complexity sub-optimal scheme through binary variable relaxation to obtain a solution, and the convergence point satisfies the KKT conditions of the relaxed non-concave problem. Simulation results demonstrate the convergence and effectiveness of our proposed schemes.

INDEX TERMS Energy efficiency, channel estimation, power control.

I. INTRODUCTION

Due to the rapid development of the Internet of Things (IoT) [1]–[3], the impact of wireless communication systems on the environment cannot be ignored [4], and the design of energy-efficient wireless communication systems has become an urgent task [5]–[8]. In wireless networks, channel state information (CSI) is used to improve energy-efficiency (EE), and pilots are inserted into data streams to get the CSI [9]. It is noticeable that the pilot itself requires additional power consumption, and perfect CSI estimation is unachievable in practice due to the CSI estimation errors [10]. Better CSI quality, namely lower CSI error, can provide a higher data rate [11]–[13], but better CSI also requires higher pilot power consumption [14], [15]. Since EE could be defined as the ratio of data rate to energy consumption, an energy efficient pilot design has to make a good tradeoff between energy consumption and CSI quality [16].

The pilot design relies on the established channel model. Block-fading feedback channel model is widely used for the channel estimation in a given frequency, where the channel gains remain invariant over the channel coherence time. In [17], the authors discussed the optimal power allocation from the EE viewpoint with considering channel estimation expense, and made the adjustment of the pilot power and the data power. In [18], the authors investigated power allocation between pilots and data with channel estimation error. In [19], the authors jointly optimized the number of activated antennas, power allocation, and pilot assignment. In [20], the number of subcarriers was taken into account to optimize the EE of the orthogonal frequency-division multiple access (OFDMA) system, where one of the subcarriers was pilot subcarrier and the others were data subcarriers. It is noticeable that [17]–[20] made the adjustment optimization of pilot power and data power. However, the channel gains in these works changed randomly across different coherence time slots. Thus, pilots had to be sent in each time slot, which limited the potentials of EE optimization.

In this paper, we propose an improved block-fading feedback channel model. The channel varies in different time slots. There are non-zero correlation coefficients between different time slots, and the correlation coefficients decrease as time interval increrases. Pilots and data are sent in the first time slot and the next several slots, respectively. The length of time slots is flexible. On the one hand, a shorter length of time slots results in a higher pilot energy consumption proportion, which will decrease the EE. On the other hand, a longer length of time slots results in worse CSI quality, which will reduce

The associate editor coordinating the review of this manuscript and approving it for publication was Jie Tang.

the average data rate and the EE. In this case, an optimal length of time slots should be derived. We try to solve the EE-maximization power control problem jointly considering the CSI errors and the channel coefficients in time-varying channels, and we try to optimize pilot transmission power, data transmission power and the number of data slots at the same time.

Similar to [21], the downlink channel is estimated by the uplink feedback channel in our model. Compared with [22], the pilot power and the transmission power in each data slot are coupled together in our objective function, which is more complex and harder to solve. Our optimization goal is the EE of the system instead of the frame error probability. The pilot length in our paper is a variable, so the objective function is completely different from [23]. Compared with [24], our paper focuses on how to optimize the energy efficiency of the system through resource allocation and pilot design with channel estimation information, rather than calculating the upper channel capacity limit in the case of multiple antenna arrays. We would like to mention that the pilots and data are in different slots rather than within the channel coherence time, and we may reduce the additional pilot power cost and obtain higher EE, which is different from [25]–[27]. In addition, we note that the analytical expression of the channel data rate is more complicated in our paper. In this paper, due to the variable coupling, the fractional form of the EE, and the integer variables, the formulated optimization problem is a mixed–integer nonlinear programming (MINLP) problem [28]. Moreover, the number of independent variables will change with the integer variables, which is hard to solve. To overcome these difficulities, we investigate MINLP EE-maximization schemes for wireless communication system pilot design based on successive convex approximation (SCA), geometric programming (GP), and the Dinkelbach methods. We further present a low-complexity sub-optimal power control schemes based on the integer relaxation to give sub-optimal solutions. The main contributions of the paper can be summarized as follows.

1) We model the time-varying channel, then calculate the CSI errors with different correlation coefficients to obtain the analytical expression of the channel data rate, which depends on the number of data slots, the pilot transmission power, and the data transmission power. Then, the EE-maximization problem is formulated as an MINLP problem.
2) We design two schemes, namely sub-optimal and low-complexity sub-optimal schemes, for this MINLP problem, aiming to approximate the solution to meet the Karush-Kuhn-Tucker (KKT) conditions with acceptable costs. In addition, we analyze the performance of our schemes, including the complexity and convergence.
3) We investigate the following network performance through simulation results, including the convergence rates of sub-optimal and low-complexity sub-optimal power control schemes, the gap between two solutions, the impact of the length of time slots, and the impact of the maximum allowed transmission power. The gap between two power control schemes is small. With the increase of the number of data slots, the EE increases and the average throughput decreases.

The rest parts of the paper are organized as follows. In Section II, we introduce the system model with time-varying channel, and then calculate the CSI error. In Section III, we formulate the EE-maximization problem and its equivalent problem, which is used to develop sub-optimal and low-complexity sub-optimal schemes, respectively. In Section IV, we present sub-optimal power control algorithm. A low-complexity sub-optimal power control algorithm is designed in Section V. Simulation studies are shown in Section VI. Finally, Section VII concludes the paper.

Notations: $\boldsymbol{a}$ is a vector, a is a scalar, $a(x)$ means a as a function of x, $a(x_0)$ (where x_0 is a given value) means the value of function a with given independent variable value x_0, $\mathcal{CN}\left(m, \sigma^2\right)$ is a complex Gaussian scalar with mean m and covariance σ^2, and $E[*]$ represents expectation.

II. SYSTEM MODEL

Let us consider a system consisting of one single antenna source node and one single antenna destination node,[1] where one pilot symbol is inserted for every N data slots to estimate the channel coefficients, as shown in Figure 1. Here, we will present the channel model, system throughput, and power consumption model consequently to facilitate the following analysis.

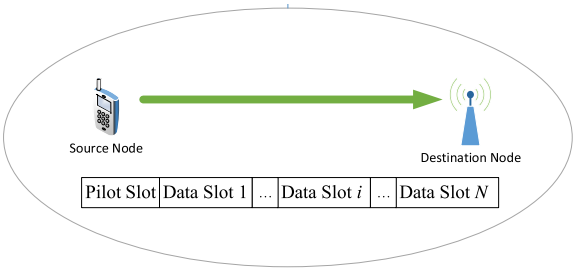

FIGURE 1. Illustration of the pilot and data slots.

A. CHANNEL MODEL

Considering a time-varying Rayleigh block-fading channel [29], the received signal can be written as

$$y_i = \sqrt{G} h_i x_i + n_i, \quad i \in \{0, 1, \ldots, N\}, \tag{1}$$

where x_i is the transmission signal with transmission power P_i, $h_i \sim \mathcal{CN}\left(0, \sigma_h^2\right)$ is the channel coefficient of the i-th time slot with the variance of σ_h^2, $n_i \sim \mathcal{CN}\left(0, \sigma_n^2\right)$ represents the

[1] When the source node and/or the destination node are equipped with multiple antennas, the formulated problem has almost the same form as that in the single antenna scenario, indicating that our proposed strategy in this paper can be extended to the multi-antenna scenario.

additional white Gaussian noise with the variance of σ_n^2, and $G = (4\pi d_0/\lambda)^{-2}(d/d_0)^{-\nu}$ is the pathloss. Here, d_0, d, λ, and ν are the reference distance, the distance between the source and the destination, the wavelength, and the path loss exponent, respectively.

To model the relationship between h_i and h_0, we employ the time-varying channel feedback error model to express the channel coefficient as [21]

$$h_i = \rho_i h_0 + \left(1-\rho_i^2\right)^{\frac{1}{2}} e, \tag{2}$$

where ρ_i is the cross correlation between h_i and h_0, and $e \sim \mathcal{CN}\left(0, \sigma_e^2\right)$ is the error with the variance σ_e^2.

In the pilot slot, a pilot x_0 is sent from the source node to the destination node and the received signal is $y_0 = h_0 x_0 + n$. The destination node estimates the channel. Denote $\hat{h}_0$ as the MMSE estimation of h_0, and it can be expressed as [30], [31]

$$\hat{h}_0 = \frac{\sigma_h^2 x_0^H y_0}{\sigma_h^2 G P_0 + \sigma_n^2}, \tag{3}$$

where P_0 is the transmitted power of x_0. Then $\hat{h}_0$ is transmitted back from the destination node to the source node through the feedback channel. As the channel is estimated in the pilot slot, we take $\hat{h}_0$ as the estimation of the channel coefficient for all data slots, i.e.,

$$\hat{h}_i = \hat{h}_0, \tag{4}$$

where $\hat{h}_i$ represents the estimation of h_i.

B. SYSTEM THROUGHPUT

With $\hat{h}_i$, the data rate of data slot i, $i \in \{1, 2, \ldots, N\}$, is give by [20], [21]

$$R_i = B\log_2\left(1 + \frac{|\hat{h}_i|^2 G P_0}{\sigma^2_{h_i-\hat{h}_i} G P_0 + \sigma_n^2}\right), \tag{5}$$

where

$$\sigma^2_{h_i-\hat{h}_i} = E\left[\left(h_i-\hat{h}_i\right)\left(h_i-\hat{h}_i\right)^H\right] = \frac{a_i G P_i + b_i}{G P_i \sigma_h^2 + \sigma_n^2}, \tag{6}$$

with

$$h_i - \hat{h}_i = \rho_i h_0 + \left(1-\rho_i^2\right)^{\frac{1}{2}} e - \hat{h}_0, \tag{7}$$

$$a_i = \left(1-\rho_i^2\right)\sigma_h^2\sigma_e^2 + (1-\rho_i)^2\sigma_h^4, \tag{8}$$

$$b_i = \rho_i^2\sigma_h^2\sigma_n^2 + \left(1-\rho_i^2\right)\sigma_n^2\sigma_e^2. \tag{9}$$

The detailed calculation of (6) is presented in Appendix A.

Substituting (6) into (5), the total system throughput is

$$\begin{aligned} R &= \sum_{i=1}^{N} R_i \\ &= \sum_{i=1}^{N} B\log_2\left(1+\frac{|\hat{h}|^2 G P_0\left(G P \sigma_h^2+\sigma_n^2\right)}{\left(a_i G P_i+b_i\right) G P_i+\sigma_h^2\sigma_n^2 G P_0+\sigma_n^4}\right). \end{aligned} \tag{10}$$

C. POWER CONSUMPTION MODEL

In our system, the total power consumption P includes the transmission power and the static power. The static power includes the power of circuits incurred by active circuit blocks and signal processing [32]. Note that a part of the static power can be employed to capture an effect of the computation power. We denote P_0, P_i, P_s as the power of the pilot, the transmission power in the i-th time slot, and the static power of the source node, respectively. Therefore, the total power consumption of the system can be given by

$$P = P_s + P_0 + \sum_{i=1}^{N} P_i. \tag{11}$$

The total EE η is defined as the amount of the transmitted data per unit energy consumed. Based on (10) and (11), we have

$$\begin{aligned} \eta\left(\boldsymbol{P}_{[N]}, N\right) &= \frac{R}{P} \\ &= \frac{B\sum_{i=1}^{N}\log_2\left(1+\frac{|\hat{h}|^2 G P_0\left(G P_i \sigma_h^2+\sigma_n^2\right)}{\left(a_i G P_i+b_i\right) G P_i+\sigma_h^2\sigma_n^2 G P_0+\sigma_n^4}\right)}{P_s + P_0 + \sum_{i=1}^{N} P_i}, \end{aligned} \tag{12}$$

where $\boldsymbol{P}_{[N]} = [P_0, P_1, \ldots, P_N]$.

III. PROBLEM FORMULATION

In this section, we will formulate the energy-efficient power control problem, where the number of data slots is flexible. Taking the minimum data rate requirement into consideration, the EE-maximization problem can be formulated as

$$\begin{aligned} \mathcal{P}1: \max_{\boldsymbol{P}_{[N]}, N} \quad & \eta\left(\boldsymbol{P}_{[N]}, N\right) \\ s.t.\ & C1: R_i \geq R_{\text{th}}, \quad i \in \{1, 2, \ldots, N\}, \\ & C2: P_0, P_i \leq P_{\max}, \quad i \in \{1, 2, \ldots, N\}, \\ & C3: N \leq L, \end{aligned} \tag{13}$$

where R_{th} represents the minimum data rate requirement, $P_{\max}$ represents the maximum allowed transmission power, and L represents the maximum number of data slots. Due to the integer variable N, the fractional form of EE, and the non-concavity of R_i, $\mathcal{P}1$ is a mixed integer non-concave fractional programming (FP) problem.

We can notice that the form of the objective function in $\mathcal{P}1$ depends on variable N. In particular, there are $N+2$ independent variables in $\mathcal{P}1$, namely $P_0, P_1, \ldots, P_N$ and N, and there are N and $N+2$ items in the numerator and denominator of the objective function, respectively. Both the number of independent variables and the number of items in the objective function change with N. For the convenience of analysis, we introduce binary variables x_i, $i \in \{1, 2, \ldots, L\}$, which is given by

$$x_i = \begin{cases} 1, & 1 \leq i \leq N \\ 0, & \text{else.} \end{cases} \tag{14}$$

Using x_i, $\mathcal{P}1$ can be rewritten into an equivalent form. Define

$$\tilde{R}_i = B\log_2\left(1 + x_i\frac{|\hat{h}|^2 GP_0}{\sigma^2_{h_i-\hat{h}}GP_0 + \sigma_n^2}\right), \tag{15}$$

$$\tilde{P}_i = x_i P_i, \quad i \in \{1, 2, \ldots, L\}, \tag{16}$$

$$\boldsymbol{x} = [x_1, x_2, \ldots, x_L]. \tag{17}$$

Accordingly, we have

$$\tilde{R} = \sum_{i=1}^{L}\tilde{R}_i, \tag{18}$$

$$\tilde{P} = P_s + P_0 + \sum_{i=1}^{L}\tilde{P}_i, \tag{19}$$

$$\begin{aligned}\tilde{\eta}\left(\boldsymbol{P}, \boldsymbol{x}\right) &= \frac{\tilde{R}}{\tilde{P}} \\ &= \frac{B\sum_{i=1}^{L}\log_2\left(1 + \frac{x_i|\hat{h}|^2 GP_0\left(GP_i\sigma_h^2+\sigma_n^2\right)}{\left(a_iGP_i+b_i\right)GP_i+\sigma_h^2\sigma_n^2GP_0+\sigma_n^4}\right)}{P_s + P_0 + \sum_{i=1}^{L}x_iP_i}.\end{aligned} \tag{20}$$

Therefore, $\mathcal{P}1$ is transformed into the following problem.

$$\begin{aligned}\mathcal{P}2: \ &\max_{\boldsymbol{P},\boldsymbol{x}} \ \tilde{\eta}\left(\boldsymbol{P}, \boldsymbol{x}\right) \\ &s.t.\ (14), \\ &\quad C4: \tilde{R}_i \geq x_i R_{\text{th}}, \quad i \in \{1, 2, \ldots, L\}, \\ &\quad C5: P_0, \tilde{P}_i \leq P_{\max}, \quad i \in \{1, 2, \ldots, L\},\end{aligned}$$

where $\boldsymbol{P} = [P_0, P_1, \ldots, P_L]$. Since $x_i \in \{0, 1\}$, the elements in $\boldsymbol{P}$ are continuous variables, and $\tilde{\eta}$ is a nonlinear function, $\mathcal{P}2$ is an MINLP problem which has $2L + 1$ independent variables. We will prove that $\mathcal{P}1$ and $\mathcal{P}2$ are equivalent, hence the maximum EE can be reached through either $\mathcal{P}1$ or $\mathcal{P}2$.

Proposition 1: $\mathcal{P}1$ and $\mathcal{P}2$ are equivalent.

Proof: The proof is presented in Appendix B. □

Both $\mathcal{P}1$ and $\mathcal{P}2$ are MINLP problems, which are NP-hard [33], and there exists no general efficient method to solve the large-scale nonconvex problems [34]. Therefore, we will investigate the sub-optimal power control scheme from $\mathcal{P}1$ in Section IV and the low-complexity sub-optimal power control scheme from $\mathcal{P}2$ in Section V.

IV. POWER CONTROL SCHEME

In this section, we design a three-loop scheme to obtain a point satisfing the KKT conditions of $\mathcal{P}1$, where the optimal N, $\boldsymbol{P}_{[N]}$, and η are derived in the outer, middle, and inner loops based on traversing, SCA and GP, and the Dinkelbach method in subsections IV-A, IV-B, and IV-C, respectively.

A. OUTER LOOP

To make the MINLP problem tractable, we fix the integer value N. Denote $R_{[N]}$ and $P_{[N]}$ as S and P with given N respectively, and

$$\eta_{[N]}\left(\boldsymbol{P}_{[N]}\right) = \frac{R_{[N]}}{P_{[N]}}. \tag{21}$$

As shown in Algorithm 1, let N change from 1 to L, $\mathcal{P}1$ is reduced to a series of sub-problems $\mathcal{P}3$.

$$\begin{aligned}\mathcal{P}3: \ &\max_{\boldsymbol{P}_{[N]}} \ \eta_{[N]}\left(\boldsymbol{P}_{[N]}\right) \\ &s.t. C1 \text{ and } C2.\end{aligned} \tag{22}$$

$\mathcal{P}1$ can be solved through solving $\mathcal{P}3$ as follows

$$\underset{\boldsymbol{P}_{[N]},N}{\text{argmax}}\ \eta\left(\boldsymbol{P}_{[N]}, N\right) = \underset{N}{\text{argmax}}\left[\underset{\boldsymbol{P}_{[N]}}{\text{argmax}} \ \eta_{[N]}\left(\boldsymbol{P}_{[N]}\right)\right]. \tag{23}$$

While N is fixed, $\mathcal{P}3$ is an optimization problem with no integer variable, but it is still a non-concave problem due to the fractional form of EE and the non-concavity of R_i. This means $\mathcal{P}3$ is still hard to solve.

Algorithm 1 Outer Loop of Traversing the Variable N

1: Set $N \leftarrow 0$, $N_{optimal} \leftarrow 1$ and $\eta \leftarrow 0$.
2: **while** $N < L$ **do**
3: $N \leftarrow N + 1$.
4: Solve $\mathcal{P}3$ with given N to obtain $\boldsymbol{P}_{[N]}$ and $\eta_{[N]}$.
5: **if** $\eta_{[N]} > \eta$ **then**
6: $\eta \leftarrow \eta_{[N]}$, $N_{optimal} \leftarrow N$.
7: **end if**
8: **end while**
9: $\boldsymbol{P}_{optimal} \leftarrow \boldsymbol{P}_{[N_{optimal}]}$.

B. MIDDLE LOOP

Due to the non-concavity of $\mathcal{P}3$, we will adopt SCA and GP based method. SCA method is used to approximate a difficult non-convex problem with a series of convex problems to obtain a KKT point of the original problem [35]–[37]. GP method can transform a positive polynomial form optimization problem into a convex problem. In the middle loop, we approximate $\mathcal{P}3$ to a series of problems that can be reduced into concave fractional programming problems through exponential variable substitutions [38]. The solution sequence converges to a point that satisfies the KKT conditions of $\mathcal{P}3$.

In the following part, the Jensen's inequality [38] is used to construct a series of convex problems to approximate $\mathcal{P}3$. Define

$$\begin{aligned}f_i\left(P_0, P_i\right) &= \left(a_iGP_i + b_i\right)GP_i + \sigma_h^2\sigma_n^2GP_0 + \sigma_n^4 \\ &= a_iG^2P_0P_i + b_iGP_i + \sigma_h^2\sigma_n^2GP_0 + \sigma_n^4, \\ g_i\left(P_0, P_i\right) &= |\hat{h}|^2GP_0\left(GP_i\sigma_h^2 + \sigma_n^2\right) \\ &\quad + \left(a_iGP_i + b_i\right)GP_i + \sigma_h^2\sigma_n^2GP_0 + \sigma_n^4, \\ \alpha_{0i} &= \left(a_i + |\hat{h}|^2\sigma_h^2\right)G^2P_0^{(0)}P_i^{(0)} \\ &\quad + \left(b_i + |\hat{h}|^2\sigma_n^2\right)GP_i^{(0)} + \sigma_h^2\sigma_n^2GP_0^{(0)} + \sigma_n^4,\end{aligned} \tag{24}$$

$$\alpha_{1i} = \frac{\left(a_i + |\hat{h}|^2\sigma_h^2\right) G^2 P_0^{(0)} P_i^{(0)}}{\alpha_{0i}},$$
$$\alpha_{2i} = \frac{\left(b_i + |\hat{h}|^2\sigma_n^2\right) G P_i^{(0)}}{\alpha_{0i}},$$
$$\alpha_{3i} = \frac{\sigma_h^2\sigma_n^2 G P_0^{(0)} P_i^{(0)}}{\alpha_{0i}},$$
$$\alpha_{4i} = \frac{\sigma_n^4}{\alpha_{0i}}, \tag{25}$$

and

$$\gamma_i(P_0, P_i) = \left[\frac{\left(a_i + |\hat{h}|^2\sigma_h^2\right) G^2 P_0 P_i}{\alpha_{1i}}\right]^{\alpha_{1i}} \cdot \left[\frac{\left(b_i + |\hat{h}|^2\sigma_n^2\right) G P_i}{\alpha_{2i}}\right]^{\alpha_{2i}} \cdot \left[\frac{\sigma_h^2\sigma_n^2 G P_0}{\alpha_{3i}}\right]^{\alpha_{3i}} \cdot \left[\frac{\sigma_n^4}{\alpha_{4i}}\right]^{\alpha_{4i}}, \tag{26}$$

where $\boldsymbol{P}_{[N]}^{(0)} = [P_0^{(0)}, P_1^{(0)}, \ldots, P_N^{(0)}]$ is the initial value of $\boldsymbol{P}_{[N]}$. Based on the Jensen's inequality, we have the following lemma.

Lemma 1:

$$1)\ g_i(P_0, P_i) \geq \gamma_i(P_0, P_i), \tag{27}$$
$$2)\ g_i\left(P_0^{(0)}, P_i^{(0)}\right) = \gamma_i\left(P_0^{(0)}, P_i^{(0)}\right), \tag{28}$$
$$3)\ \nabla g_i\left(P_0^{(0)}, P_i^{(0)}\right) = \nabla\gamma_i\left(P_0^{(0)}, P_i^{(0)}\right), \tag{29}$$

where $\nabla = \left(\frac{\partial}{\partial P_0}, \frac{\partial}{\partial P_1}, \ldots, \frac{\partial}{\partial P_N}\right)$.

Proof: The proof is presented in Appendix C. □

With (24) and (25), R_i can be written as

$$R_i(P_0, P_i) = -B\log_2\left(\frac{f_i(P_0, P_i)}{g_i(P_0, P_i)}\right). \tag{30}$$

Denote

$$\zeta_i(P_0, P_i) = -B\log_2\left(\frac{f_i(P_0, P_i)}{\gamma_i(P_0, P_i)}\right). \tag{31}$$

From Lemma 1, we can see that ζ_i is a lower bound of R_i, and we will use ζ_i to approximate R_i. However, $\zeta_i(P_0, P_i)$ is still a non-concave function, so we will adopt GP based method to transform ζ_i into a concave function. It is noticeable that $\gamma_i(P_0, P_i)$ is a monomial, and $f_i(P_0, P_i)$ and $g_i(P_0, P_i)$ are polynomials. Therefore $\frac{f_i(P_0,P_i)}{g_i(P_0,P_i)}$ is fractional and $\frac{f_i(P_0,P_i)}{\gamma_i(P_0,P_i)}$ is a polynomial. Since ζ_i has the form of log(polynomial), We choose the following variable substitution

$$\boldsymbol{P}_{[N]} = P_{\max}[e^{p_0}, e^{p_1}, \ldots, e^{p_N}], \tag{32}$$

where $p_0, p_1, \ldots, p_N \in (-\infty, 0]$. Now, $\zeta_i(P_{\max}e^{p_0}, P_{\max}e^{p_i})$ is a concave function due to the negative log-sum-exp form [38]. Based on Lemma 1, we have the following corollary.

Corollary 1:

$$1)\ R_i\left(P_{\max}e^{p_0}, P_{\max}e^{p_i}\right) \geq \zeta_i\left(P_{\max}e^{p_0}, P_{\max}e^{p_i}\right), \tag{33}$$
$$2)\ R_i\left(P_{\max}e^{p_0^{(0)}}, P_{\max}e^{p_i^{(0)}}\right) = \zeta_i\left(P_{\max}e^{p_0^{(0)}}, P_{\max}e^{p_i^{(0)}}\right), \tag{34}$$
$$3)\ \nabla R_i\left(P_{\max}e^{p_0^{(0)}}, P_{\max}e^{p_i^{(0)}}\right) = \nabla\zeta_i\left(P_{\max}e^{p_0^{(0)}}, P_{\max}e^{p_i^{(0)}}\right), \tag{35}$$

where $(P_{\max}e^{p_0^{(0)}}, P_{\max}e^{p_i^{(0)}}) = (P_0^{(0)}, P_i^{(0)})$ and $\nabla = \left(\frac{\partial}{\partial p_0}, \frac{\partial}{\partial p_1}, \ldots, \frac{\partial}{\partial p_N}\right)$.

We replace R_i with ζ_i, and have the following problem.

$$\mathcal{P}4: \max_{\boldsymbol{p}_{[N]}} \ \underline{\eta}_{[N]}\left(\boldsymbol{p}_{[N]}\right)$$
$$s.t.\ \zeta_i\left(P_{\max}e^{p_0}, P_{\max}e^{p_i}\right) \geq R_{\text{th}}, \quad i \in \{1, 2, \ldots, N\},$$
$$p_0, p_1, \ldots, p_N \leq 0, \tag{36}$$

where

$$\boldsymbol{p}_{[N]} = [p_0, p_1, \ldots, p_N], \tag{37}$$
$$\zeta_{[N]}\left(\boldsymbol{p}_{[N]}\right) = \sum_{i=1}^{N}\zeta_i\left(P_{\max}e^{p_0}, P_{\max}e^{p_i}\right), \tag{38}$$
$$\pi_{[N]}\left(\boldsymbol{p}_{[N]}\right) = P_s + P_{\max}\sum_{i=0}^{N}e^{p_i}, \tag{39}$$
$$\underline{\eta}_{[N]}\left(\boldsymbol{p}_{[N]}\right) = \frac{\zeta_{[N]}\left(\boldsymbol{p}_{[N]}\right)}{\pi_{[N]}\left(\boldsymbol{p}_{[N]}\right)}. \tag{40}$$

Because of the concavity of $\zeta_i(P_{\max}e^{p_0}, P_{\max}e^{p_i})$, $\zeta_{[N]}\left(\boldsymbol{p}_{[N]}\right)$ is also a concave function. Thus, $\pi_{[N]}$ is a convex function. Therefore, $\underline{\eta}_{[N]}$ is a fractional function with a concave numerator and a convex denominator.

Algorithm 2 presents an algorithm to solve $\mathcal{P}3$ via solving $\mathcal{P}4$ iteratively. We can prove that the optimal objective value of $\mathcal{P}4$ provides a lower bound for the optimal objective value of $\mathcal{P}3$. We assume that $\boldsymbol{p}_{[N]}^*$ is the optimal solution to $\mathcal{P}4$. Recalling (33), we have

$$\max_{\boldsymbol{p}_{[N]}} \underline{\eta}_{[N]} = \underline{\eta}_{[N]}(\boldsymbol{p}_{[N]}^*) \leq \eta_{[N]}(\boldsymbol{p}_{[N]}^*) \leq \max_{\boldsymbol{p}_{[N]}} \eta_{[N]}. \tag{41}$$

Lemma 2: $\eta^{(k)}$ in Algorithm 2 is a monotonically increasing sequence.

Proof: The proof is presented in Appendix D. □

Lemma 2 and inequality (41) jointly show that the sequence $\eta^{(k)}$ increases in each iteration, and it is upper-bounded by $\max_{\boldsymbol{p}_{[N]}} \eta_{[N]}$, so Algorithm 2 converges, as follows.

$$\underline{\eta}^{(0)} \leq \cdots \leq \underline{\eta}^{(k-1)} \leq \underline{\eta}^{(k)} \leq \underline{\eta}^{(k+1)} \cdots \leq \max_{\boldsymbol{p}_{[N]}} \underline{\eta}_{[N]}. \tag{42}$$

Lemma 3: The convergence point of Algorithm 2 satisfies the KKT conditions of the original problem $\mathcal{P}3$.

Proof: The proof is presented in Appendix E. □

Based on Lemmas 2 and 3, we have the following corollary.

Algorithm 2 Medium Loop of Solving $\mathcal{P}3$ by Successive Convex Approximation

Set tolerance δ.
Initialize $k \leftarrow 0$, $\boldsymbol{p}^{(k)} \leftarrow \mathbf{0}$, and $\underline{\eta}^{(k)} \leftarrow 0$.
$k \leftarrow k+1$.
Update α_{1i}, α_{2i}, α_{3i} and α_{4i} with $\boldsymbol{p}^{(k-1)}$, $i \in \{1, 2, \ldots, N\}$.
Solve problem $\mathcal{P}4$ to derive the optimal solution $\boldsymbol{p}^{(k)}$ and $\underline{\eta}^{(k)}$.
while $|\underline{\eta}^{(k)} - \underline{\eta}^{(k-1)}| \geq \delta$ **do**
  $k \leftarrow k+1$.
  Update α_{1i}, α_{2i}, α_{3i} and α_{4i} with $\boldsymbol{p}^{(k-1)}$, $i \in \{1, 2, \ldots, N\}$.
  Solve $\mathcal{P}4$ to derive the optimal solution $\boldsymbol{p}^{(k)}$ and $\underline{\eta}^{(k)}$.
end while
$\eta_{optimal} \leftarrow \underline{\eta}^{(k)}$.
$\boldsymbol{P}_{optimal} \leftarrow P_{\max}[e^{p_0}, e^{p_1}, \ldots, e^{p_N}]$.

Algorithm 3 Inner Loop of Deriving λ in $\mathcal{P}5$ based on the Dinkelbach Method

Set tolerance ϵ.
Initialize $k \leftarrow 0$ and $\lambda^{(k)} \leftarrow 0$.
Solve the $\mathcal{P}5$ with $\lambda^{(k)}$ to obtain the optimal solution $\boldsymbol{p}^{(k)}$.
while $|F\left(\lambda^{(k)}\right)| > \epsilon$ **do**
  $k \leftarrow k+1$.
  $\lambda^{(k)} \leftarrow \frac{\zeta_{[N]}}{\pi_{[N]}}$ with $\boldsymbol{p}^{(k-1)}$.
  Solve $\mathcal{P}5$ with $\lambda^{(k)}$ to obtain $\boldsymbol{P}^{(k)}$.
end while
$\lambda^* \leftarrow \lambda^{(k)}$.

Corollary 2: Algorithm 2 converges to a point satisfying the KKT conditions of $\mathcal{P}3$, and a locally optimal solution of $\mathcal{P}3$ is obtained.

C. INNER LOOP

We notice that $\underline{\eta}_{[N]}$ has a concave numerator and a convex denominator, and all the constraints are convex. Accordingly, $\mathcal{P}4$ can be globally solved by the Dinkelbach's algorithm [39]. We have the equivalent Dinkelbach problem as follows.

$$\begin{aligned}\mathcal{P}5: \max_{\boldsymbol{p}_{[N]}} \ & \zeta_{[N]}\left(\boldsymbol{p}_{[N]}\right) - \lambda\pi_{[N]}\left(\boldsymbol{p}_{[N]}\right) \\ s.t. \ & \zeta_i\left(P_{\max}e^{p_0}, P_{\max}e^{p_i}\right) \geq R_{\text{th}}, \quad i \in \{1, 2, \ldots, N\}, \\ & p_0, p_1, \ldots, p_N \leq 0, \end{aligned} \tag{43}$$

where λ is a non-negative parameter. Due to the concavity of $\zeta_{[N]}$ and the convexity of $\pi_{[N]}$, $\zeta_{[N]} - \lambda\pi_{[N]}$ is a concave function and $\mathcal{P}5$ is a concave problem. Therefore, it can be easily solved through convex optimization methods.

Through solving the concave problem $\mathcal{P}5$, we can solve FP $\mathcal{P}4$, as shown in Algorithm 3, where

$$F(\lambda) = \max_{\boldsymbol{p}_{[N]}} \{\zeta_{[N]}\left(\boldsymbol{p}_{[N]}\right) - \lambda\pi_{[N]}\left(\boldsymbol{p}_{[N]}\right)\}. \tag{44}$$

Since both $\zeta_{[N]}$ and $\pi_{[N]}$ are positive, $F(\lambda)$ is continuous and strictly monotonic decreasing in λ and has an unique root. Within the Dinkelbach method, the value of λ increases in each iteration until converges to $\lambda^\star$, where $\lambda^\star$ in Algorithm 3 is the root of $F(\lambda)$ [40]. Also, $\lambda^\star$ is equal to the optimal EE [40], [41], i.e.,

$$\lambda^\star = \max_{\boldsymbol{p}_{[N]}} \ \underline{\eta}_{[N]}\left(\boldsymbol{p}_{[N]}\right). \tag{45}$$

So far, we have solved the original problem $\mathcal{P}1$ entirely.

D. COMPLEXITY ANALYSIS

Here, we regard the tolerance ϵ as a constant, and the complexity is considered as a function of the size of the optimization problem. The sub-optimal scheme includes three loops, i.e., traversing the variable N in outer loop, the SCA method in middle loop, and the Dinkelbach method in inner loop. The complexity of the outer loop is $O(L)$. The simulation results in section VI show that the convergence rate of successive convex is very fast and almost stable with L. According to [40], updating λ based on the Dinkelbach method converges at a superlinear convergence rate. The complexity of solving convex problem $\mathcal{P}5$ on the Dinkelbach method is $O\left(L^{3.5}\right)$, based on interior point method [42]. The accurate analytical complexities of the SCA and the Dinkelbach method are challenging to derive, and therefore we illustrate the average number of the required iterations by simulations in Section VI.

V. FAST SUB-OPTIMAL POWER CONTROL SCHEME

We design a three-loop scheme to obtain a solution satisfying the KKT conditions of $\mathcal{P}1$ in the previous section. However, the complexity is too high due to the traversal of the variable N in the outer loop. In this section, we will investigate an integer relaxation based low-complexity scheme to obtain a point satisfying the KKT conditions of the relaxed $\mathcal{P}2$. We first relax the binary variables in $\mathcal{P}2$ to transform $\mathcal{P}2$ into a problem without integer variables, and then solve the problem via optimization to obtain the optimal N. With the optimal N, we solve $\mathcal{P}3$ to get the final results.

A. INTEGER RELAXATION

Relaxation is an important approach to solve MINLP problems. The original integer variables in (14), i.e., $x_1, x_2, \ldots, x_L$, are binary variables, which can be relaxed into continuous variables, i.e., $\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_L$. Accordingly, we define $\tilde{\boldsymbol{x}} = [\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_L]$, and we have

$$1 \geq \tilde{x}_1 \geq \tilde{x}_2 \geq \cdots \geq \tilde{x}_L \geq 0. \tag{46}$$

Constraint (46) can be further expressed as the following linear constraints.

$$\begin{aligned} & 0 \leq \tilde{x}_i \leq 1, \quad i \in \{1, 2, \ldots, L\}, \\ & -\tilde{x}_1 + \tilde{x}_2 \leq 0, \\ & -\tilde{x}_2 + \tilde{x}_3 \leq 0, \\ & \cdots \\ & -\tilde{x}_{L-1} + \tilde{x}_L \leq 0. \end{aligned} \tag{47}$$

Recalling that $\boldsymbol{x}$ represents the number of data slots, in order to get a reasonable result, we need to decode $\tilde{\boldsymbol{x}}$ into the binary $\boldsymbol{x}$ in $\mathcal{P}2$. Let the optimal results be $\tilde{\boldsymbol{x}^*} = [\tilde{x}_1^*, \tilde{x}_2^*, \ldots, \tilde{x}_L^*]$, and set a as the decision threshold, i.e.,

$$x_i = \begin{cases} 1, & \tilde{x}_i^* \geq a \\ 0, & \text{else}, \end{cases} \quad i \in \{1, 2, \ldots, L\}. \tag{48}$$

After $\boldsymbol{x}$ is fixed (equivalent to fixed N), $\mathcal{P}2$ can be solved in the same way as $\mathcal{P}3$.

B. OUTER LOOP

We replace $\boldsymbol{x}$ and constraints (14) in $\mathcal{P}2$ with $\tilde{\boldsymbol{x}}$ and constraints (47) respectively, and $\mathcal{P}2$ becomes a non-concave optimization problem without integer variables, which is similar to $\mathcal{P}3$. Therefore, we can use the similar scheme as in Section IV-B to solve it. We define

$$\begin{aligned}
\tilde{f}_i\left(P_0, P_i, \tilde{x}_i\right) &= \left(a_i G P_i + b_i\right) G P_i + \sigma_h^2 \sigma_n^2 G P_0 + \sigma_n^4 \\
&= a_i G^2 P_0 P_i + b_i G P_i + \sigma_h^2 \sigma_n^2 G P_0 + \sigma_n^4, \quad (49) \\
\tilde{g}_i\left(P_0, P_i, \tilde{x}_i\right) &= \tilde{x}_i |\hat{h}|^2 G P_0 \left(G P_i \sigma_h^2 + \sigma_n^2\right) \\
&\quad + \left(a_i G P_i + b_i\right) G P_i + \sigma_h^2 \sigma_n^2 G P_0 + \sigma_n^4, \quad (50) \\
\tilde{\alpha}_{0i} &= \left(a_i + \tilde{x}_i |\hat{h}|^2 \sigma_h^2\right) G^2 P_0^{(0)} P_i^{(0)} \\
&\quad + \left(b_i + \tilde{x}_i |\hat{h}|^2 \sigma_n^2\right) G P_i^{(0)} + \sigma_h^2 \sigma_n^2 G P_0^{(0)} + \sigma_n^4, \\
\tilde{\alpha}_{1i} &= \frac{\left(a_i + \tilde{x}_i |\hat{h}|^2 \sigma_h^2\right) G^2 P_0^{(0)} P_i^{(0)}}{\tilde{\alpha}_{0i}}, \\
\tilde{\alpha}_{2i} &= \frac{\left(b_i + \tilde{x}_i |\hat{h}|^2 \sigma_n^2\right) G P_i^{(0)}}{\tilde{\alpha}_{0i}}, \\
\tilde{\alpha}_{3i} &= \frac{\sigma_h^2 \sigma_n^2 G P_0^{(0)} P_i^{(0)}}{\tilde{\alpha}_{0i}}, \\
\tilde{\alpha}_{4i} &= \frac{\sigma_n^4}{\tilde{\alpha}_{0i}}, \quad (51)
\end{aligned}$$

and

$$\begin{aligned}
\tilde{\gamma}_i\left(P_0, P_i, \tilde{x}_i\right) = &\left[\frac{\left(a_i + \tilde{x}_i |\hat{h}|^2 \sigma_h^2\right) G^2 P_0 P_i}{\alpha_{1i}}\right]^{\alpha_{1i}} \\
&\cdot \left[\frac{\left(b_i + \tilde{x}_i |\hat{h}|^2 \sigma_n^2\right) G P_i}{\alpha_{2i}}\right]^{\alpha_{2i}} \\
&\cdot \left[\frac{\sigma_h^2 \sigma_n^2 G P_0}{\alpha_{3i}}\right]^{\alpha_{3i}} \cdot \left[\frac{\sigma_n^4}{\alpha_{4i}}\right]^{\alpha_{4i}}, \quad (52)
\end{aligned}$$

where $\boldsymbol{P}^{(0)} = [P_0^{(0)}, P_1^{(0)}, \ldots, P_L^{(0)}]$ is the initial value of $\boldsymbol{P}$, and $\tilde{\boldsymbol{x}^{(0)}} = [\tilde{x}_1^{(0)}, \tilde{x}_2^{(0)}, \ldots, \tilde{x}_L^{(0)}]$ is the initial value of $\tilde{\boldsymbol{x}}$. Based on (49) and (50), R_i in (15) can be further expressed as

$$\tilde{R}_i\left(P_0, P_i, \tilde{x}_i\right) = -B \log_2 \left(\frac{\tilde{f}_i\left(P_0, P_i, \tilde{x}_i\right)}{\tilde{g}_i\left(P_0, P_i, \tilde{x}_i\right)}\right). \tag{53}$$

Thus, we can construct a lower bound of $\tilde{R}_i$ using (49) and (52), which is given by

$$\tilde{\zeta}_i\left(P_0, P_i, \tilde{x}_i\right) = -B \log_2 \left(\frac{\tilde{f}_i\left(P_0, P_i, \tilde{x}_i\right)}{\tilde{\gamma}_i\left(P_0, P_i, \tilde{x}_i\right)}\right). \tag{54}$$

$\tilde{R}_i$ is lower bounded by $\tilde{\zeta}_i$ due to Jensen's inequality, similar to Section IV-B. We also have the similar conclusions, listed as follows.

$$1)\ \tilde{R}_i\left(P_0, P_i, \tilde{x}_i\right) \geq \tilde{\zeta}_i\left(P_0, P_i, \tilde{x}_i\right), \tag{55}$$

$$2)\ \tilde{R}_i\left(P_0^{(0)}, P_i^{(0)}, \tilde{x}_i^{(0)}\right) = \tilde{\zeta}_i\left(P_0^{(0)}, P_i^{(0)}, \tilde{x}_i^{(0)}\right), \tag{56}$$

$$3)\ \nabla \tilde{R}_i\left(P_0^{(0)}, P_i^{(0)}, \tilde{x}_i^{(0)}\right) = \nabla \tilde{\zeta}_i\left(P_0^{(0)}, P_i^{(0)}, \tilde{x}_i^{(0)}\right), \tag{57}$$

where $\nabla = \left(\frac{\partial}{\partial P_0}, \frac{\partial}{\partial P_1}, \ldots, \frac{\partial}{\partial P_L}, \frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_L}\right)$. The above conclusions can be proved in the same way as that in Appendix C. Noticing that $\frac{\tilde{f}_i(P_0, P_i, \tilde{x}_i)}{\tilde{\gamma}_i(P_0, P_i, \tilde{x}_i)}$ is a polynomial, we can transform $\tilde{\zeta}_i\left(P_0, P_i, \tilde{x}_i\right)$ into a concave function by the following exponential variable substitution,

$$\begin{aligned}
\tilde{\boldsymbol{x}} &= [e^{y_1}, e^{y_2}, \ldots, e^{y_L}], \\
\boldsymbol{P} &= P_{\max}[e^{p_0}, e^{p_1}, \ldots, e^{p_L}]. \quad (58)
\end{aligned}$$

Then, we replace $\tilde{R}_i$ with $\tilde{\zeta}_i$ and apply exponential variable substitution (58) to obtain $\mathcal{P}6$.

$$\begin{aligned}
\mathcal{P}6: \ &\max_{\boldsymbol{p}, \boldsymbol{y}} \ \underline{\eta}\left(\boldsymbol{p}, \boldsymbol{y}\right) \\
&s.t.\ \zeta_i\left(P_{\max} e^{p_0}, P_{\max} e^{p_i}, e^{y_i}\right) \geq R_{\text{th}}, \quad i \in \{1, 2, \ldots, L\}, \\
&\quad p_0, p_1, \ldots, p_L \leq 0, \\
&\quad y_1, \ldots, y_L \leq 0, \quad (59)
\end{aligned}$$

where

$$\boldsymbol{p} = [p_0, p_1, \ldots, p_L], \tag{60}$$

$$\boldsymbol{y} = [y_1, \ldots, p_L], \tag{61}$$

$$\zeta\left(\boldsymbol{p}, \boldsymbol{y}\right) = \sum_{i=1}^{L} \zeta_i\left(P_{\max} e^{p_0}, P_{\max} e^{p_i}, e^{y_i}\right), \tag{62}$$

$$\pi\left(\boldsymbol{p}, \boldsymbol{y}\right) = P_s + P_{\max} e^{p_0} + P_{\max} \sum_{i=1}^{L} e^{p_i} e^{y_i}, \tag{63}$$

and

$$\begin{aligned}
\underline{\eta}\left(\boldsymbol{p}, \boldsymbol{y}\right) &= \frac{\zeta\left(\boldsymbol{p}, \boldsymbol{y}\right)}{\pi\left(\boldsymbol{p}, \boldsymbol{y}\right)} \\
&= \frac{-B \sum_{i=1}^{L} \log_2 \left(\frac{\tilde{f}_i\left(P_{\max} e^{p_0}, P_{\max} e^{p_i}, e^{y_i}\right)}{\tilde{\gamma}_i\left(P_{\max} e^{p_0}, P_{\max} e^{p_i}, e^{y_i}\right)}\right)}{P_s + P_{\max} e^{p_0} + P_{\max} \sum_{i=1}^{L} e^{p_i} e^{y_i}}. \quad (64)
\end{aligned}$$

Here, ζ is a concave function and π is a convex function because of the concavity of ζ_i and the convexity of exponential functions. Assume that $\boldsymbol{p}^*$ and $\boldsymbol{y}^*$ is the optimal solution to $\mathcal{P}6$. The optimal objective value of $\mathcal{P}6$ gives a lower bound for the optimal objective value of $\mathcal{P}2$ due to the fact that $\tilde{\zeta}_i$ is a lower bound of $\tilde{R}_i$, as follows.

$$\max_{\boldsymbol{p}, \boldsymbol{y}} \ \underline{\eta} = \underline{\eta}(\boldsymbol{p}^*, \boldsymbol{y}^*) \leq \eta(\boldsymbol{p}^*, \boldsymbol{y}^*) \leq \max_{\boldsymbol{p}, \boldsymbol{y}} \ \eta. \tag{65}$$

Algorithm 4 presents an algorithm to solve $\mathcal{P}2$ via solving $\mathcal{P}6$ iteratively.

Algorithm 4 Outer Loop of Solving $\mathcal{P}2$ by Successive Convex Approximation

Set tolerance δ and threshold a.
Initialize $k \leftarrow 0, \boldsymbol{p}^{(k)} \leftarrow \mathbf{0}, \boldsymbol{y}^{(k)} \leftarrow \mathbf{0}$, and $\underline{\eta}^{(k)} \leftarrow 0$.
Update $\tilde{\alpha}_{1i}$, $\tilde{\alpha}_{2i}$, $\tilde{\alpha}_{3i}$ and $\tilde{\alpha}_{4i}$ with $\boldsymbol{p}^{(k)}$ and $\boldsymbol{y}^{(k)}$, $i \in \{1, 2, \ldots, L\}$.
Set $k \leftarrow 1$.
Solve $\mathcal{P}6$ to derive the optimal solution $\boldsymbol{p}^{(k)}, \boldsymbol{y}^{(k)}$ and $\underline{\eta}^{(k)}$.
while $|\underline{\eta}^{(k)} - \underline{\eta}^{(k-1)}| \geq \delta$ **do**
  $k \leftarrow k + 1$.
  Update $\tilde{\alpha}_{1i}, \tilde{\alpha}_{2i}, \tilde{\alpha}_{3i}$ and $\tilde{\alpha}_{4i}$ with $\boldsymbol{p}^{(k-1)}$ and $\boldsymbol{y}^{(k-1)}$, $i \in \{1, 2, \ldots, L\}$.
  Solve $\mathcal{P}6$ to derive the optimal solution $\boldsymbol{p}^{(k)}, \boldsymbol{y}^{(k)}$ and $\underline{\eta}^{(k)}$.
end while
$x_i^* \leftarrow 0, i \in \{1, 2, \ldots, L\}$.
$i \leftarrow 1$.
while $i \leq L$ **do**
  if $e^{y_i} \geq a$ **then**
    $x_i^* \leftarrow 1$.
  end if
  $i \leftarrow i + 1$.
end while
Solve problem $\mathcal{P}2$ with $\boldsymbol{x} \leftarrow \boldsymbol{x}^*$ to get the optimal $\boldsymbol{P}$ and η.

Similar to Lemmas 2 and 3, we can prove that the sequence $\eta^{(k)}$ in Algorithm 4 increases and is upper-bounded by $\max\limits_{\boldsymbol{p},\boldsymbol{y}} \ \eta$, so Algorithm 4 converges. Furthermore, the convergence point of Algorithm 4 satisfies the KKT conditions of the relaxed non-concave problem.

C. INNER LOOP

The objective function in $\mathcal{P}6$ has a concave numerator and a convex denominator, and all the constraints are convex. This means that $\mathcal{P}4$ can be globally solved through the Dinkelbach's algorithm, and the Dinkelbach problem is a concave problem which can be easily solved by the convex optimization, as follows.

$$\begin{aligned} \mathcal{P}7: \max_{\boldsymbol{p},\boldsymbol{y}} \ & \zeta\left(\boldsymbol{p},\boldsymbol{y}\right) - \lambda\pi\left(\boldsymbol{p},\boldsymbol{y}\right) \\ s.t. \ & \zeta_i\left(P_{\max}e^{p_0}, P_{\max}e^{p_i}, e^{y_i}\right) \geq R_{\text{th}}, \quad i \in \{1, 2, \ldots, L\}, \\ & p_0, p_1, \ldots, p_L \leq 0, \\ & y_1, \ldots, y_L \leq 0. \end{aligned} \tag{66}$$

Algorithm 5 shows how to solve $\mathcal{P}4$ through the Dinkelbach method. Sequence $\lambda^{(k)}$ in Algorithm 5 will keep increasing until converging to $\max\limits_{\boldsymbol{p},\boldsymbol{y}} \ \eta$ [40].

D. COMPLEXITY ANALYSIS

Simulations in Section VI show that the integer relaxation variables will converge in several iterations. Similar to Section IV, the average numbers of the required iterations for the SCA and the Dinkelbach are shown through the simulation results in Section VI.

Algorithm 5 Inner Loop of Deriving λ in $\mathcal{P}7$ based on the Dinkelbach Method

Set tolerance ϵ.
Initialize $k \leftarrow 0$ and $\lambda^{(k)} \leftarrow 0$.
Solve the concave problem $\mathcal{P}7$ with $\lambda^{(k)}$ to obtain the optimal solution $\boldsymbol{p}^{(k)}$ and $\boldsymbol{y}^{(k)}$.
while $|F\left(\lambda^{(k)}\right)| > \epsilon$ **do**
  $k \leftarrow k + 1$
  $\lambda^{(k)} \leftarrow \frac{\zeta}{\pi}$ with $\boldsymbol{p}^{(k-1)}$ and $\boldsymbol{y}^{(k-1)}$
  Solve $\mathcal{P}7$ with $\lambda^{(k)}$ to obtain $\boldsymbol{p}^{(k)}$ and $\boldsymbol{y}^{(k)}$.
end while
$\lambda^* \leftarrow \lambda^{(k)}$

VI. SIMULATION RESULTS

In this section, we provide some simulation results to demonstrate the performance of our proposed energy-efficient power control schemes in Sections IV and V. Here, we consider a two-node communicate link where the distance between the source node and the destination node d is 200 meters (m), the path loss is expressed as $G_{dB} = 140.7 + 37.6\log_{10}(d/1000)$ in dB and $G = 10^{-G_{dB}/10}$, and the noise power spectral density is -174dBm/Hz [9]. The bandwidth of the channel is 180 kHz. $\sigma_h^2 = G$, $\sigma_e^2 = 0.5\ G$ [21], and the cross correlation between h_i and h_0 is considered as $\rho_i = \exp(-i)$. The number of the maximum data slots L is 128. The static power P_s is 500 milli-Watts (mW) [39], the maximum allowed transmission power $P_{\max}$ is 300 mW (24.7 dBm), and the minimum data rate requirement R_{th} is 18 kilobits per second ($kbps$). The tolerance is set as $\delta = \epsilon = 10^{-6}$ and the decision threshold a in Algorithm 4 is $0.5 * x_1^*$.

Figure 2 shows that the parameters of the objective function in Algorithm 2 converge very fast, namely in less than 5 iterations. Moreover, the convergence rate of Algorithm 2 is almost independent of the size of the problem. As L increases,

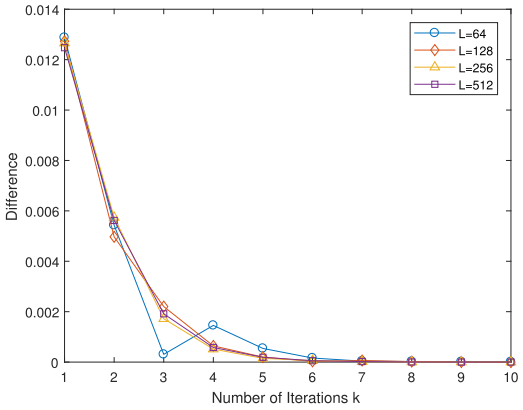

FIGURE 2. The difference between $\alpha_{0i}^{(k)}$ and $\alpha_{0i}^{(k-1)}$ Versus Iteration times in terms of $\left|\sum_i \alpha_{0i}^{(k)} - \alpha_{0i}^{(k-1)}\right| \Big/ \sum_i \alpha_{0i}^{(k-1)}$.

the convergence rate of Algorithm 2 will almost not change. The convergence of η in Algorithm 2 is shown in Figure 3. Figure 3 shows that the convergence of Algorithm 2 is independent of variable L as well. The optimal value tends to be stable after two iterations.

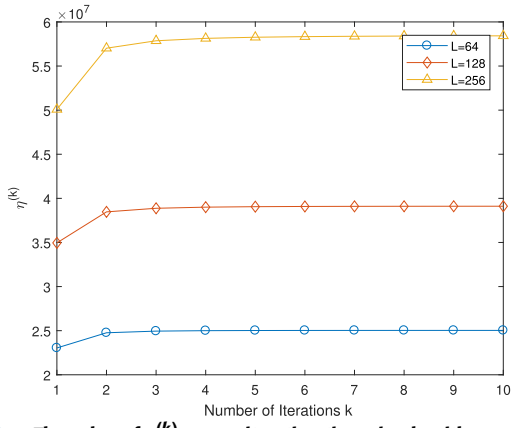

FIGURE 3. The value of $\eta^{(k)}$ versus iteration times in algorithm 3.

Figure 4 shows the convergence of relaxed $\tilde{\boldsymbol{x}}$. We use $\frac{1}{L}\sum_i \min\{\tilde{x}_i/\tilde{x}_1, 1-\tilde{x}_i/\tilde{x}_1\}$ to see if $\tilde{\boldsymbol{x}}$ converges to zero or the first component of $\tilde{\boldsymbol{x}}$ (which is also the component with the largest value). We can see that as the number of iterations increases, most of the elements in $\tilde{\boldsymbol{x}}$ are distributed around 0 or $\tilde{x}_1$, which provides a basis for us to set a decision threshold to choose the optimal N. Moreover, the convergence rate of the integer relaxation is also almost independent of L.

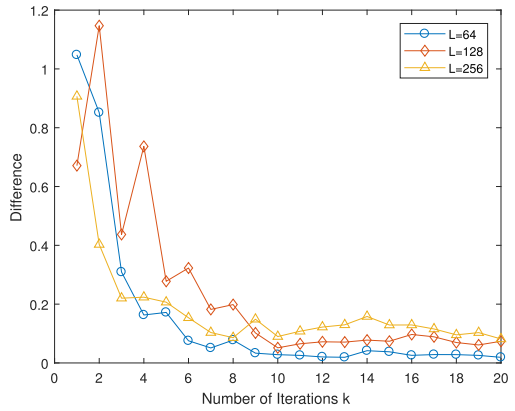

FIGURE 4. The convergence of relaxed variables $\tilde{x}$ versus iteration times in terms of $\frac{1}{L}\sum_i \min\{\tilde{x}_i/\tilde{x}_1, 1-\tilde{x}_i/\tilde{x}_1\}$. $d = 120$ m.

From Figures 2, 3 and 4, we notice that the complexity of the SCA and the integer relaxation are almost independent of value L, which is related to the size of the original problem. As we set the tolerance ϵ as a constant, the complexity is considered as a function of size. Therefore, the complexity of arithmetic operations of interior point method is $O\left(L^{3.5}\right)$, and the complexity of the traversal of variable N is $O(L)$, as mentioned in Sections IV-D and V-D.

We consider the impacts of cell radius in Figures 5 and 6, where $d = 200$ m and $d = 500$ m, respectively. We can see that in these figures, the curves of $d = 500$ m are much lower than the curves of $d = 200$ m, which indicates that a

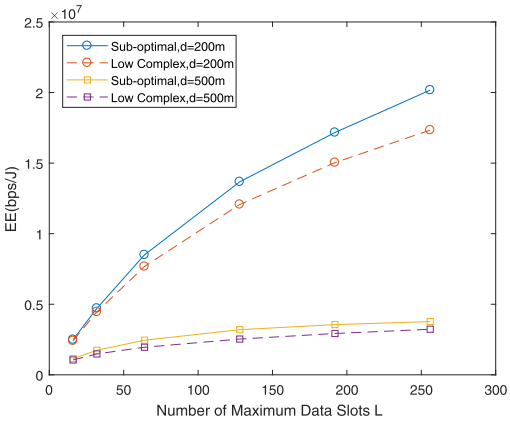

FIGURE 5. Impact of number of maximum data slots L on the EE.

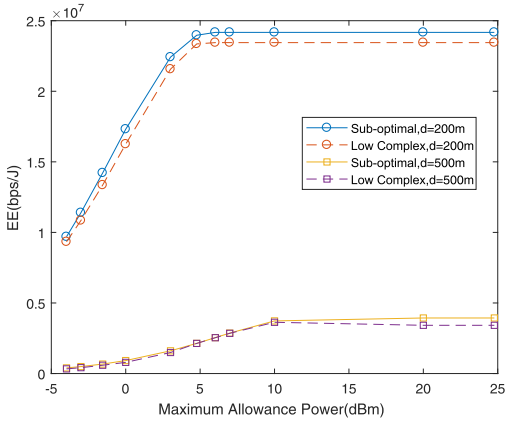

FIGURE 6. Impact of maximum transmission power $P_{\max}$ on EE.

larger d results in a lower EE and a lower average throughput. This is due to the fact that the longer distance requires higher transmission power to offset the larger path loss.

Figure 5 shows the impacts of the maximum number of time slots L on the EE. We can see that the achieved EEs of sub-optimal schemes are higher than those of low-complexity sub-optimal schemes, but they are very close. This is due to the fact that the low-complexity sub-optimal solution is obtained by solving the $\mathcal{P}2$ again after N is determined in the Algorithm 4. We also observe that EE increases with the increase of L, which shows that it is preferred to use less pilots from the EE perspective.

The impacts of the maximum transmission power $P_{\max}$ on EE are shown in Figure 6. We can see that EE firstly increases and then remains almost stable with the increase of $P_{\max}$. We note that the increasing speed of data rate is always lower than that of the transmission power due to the logarithmic function, the system EE strictly decreases when the transmission power increases, and the maximum EE should be achieved when the data rate is R_{th}, if the static power is zero. Under these circumstances, the average throughput should remain constant, i.e., R_{th}. However, in practice the static power is always larger than zero. If the transmission power is extremely low, the static power becomes a dominant influence and the system EE will increase as the transmission power increases. On the other hand, if the transmission power is very high, the influence of the static power is negligi-

ble and the system EE will decrease as the transmission power increases just like the case of zero static power. From Figure 6, we find that there exists a break point for the transmission power from the EE perspective, i.e. around 5 dBm (curves of $d = 200m$) or 10 dBm (curves of $d = 500m$), and the EE would be degraded when the transmitted power is higher than the break point.

Figure 7 shows the impacts of the maximum number of time slots L and the minimum date rate R_{th} on the average throughput. We consider three different minimum data rates R_{th} in this figure. We can see that the average throughput decreases with the increase of L until reaching the minimum date rate R_{th}, which is due to the average throughput loss from the worse channel estimation of additional data slot. A data rate in every slot must be higher than R_{th}, and the average throughput becomes stable after reaching R_{th}. We can also notice that in the high throughput area all curves are similar. However, in the low average throughput area, before reaching the minimum date rate, the curves of lower minimum date rate are above the curves of higher minimum date rate, from the curves of $R_{th} = 0.1 * B$ and the curves of $R_{th} = 0.6 * B$ shown in Figure 7. Since the data rates in additional slots have reached R_{th} and raised the total power, to achieve a higher EE, the power values of the previous slots are reduced and the throughput is reduced.

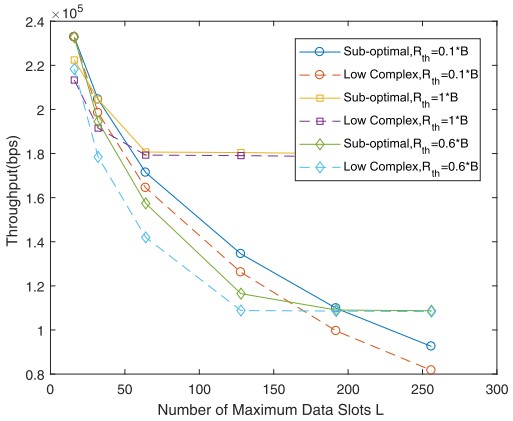

FIGURE 7. Impact of the number of the maximum data slots L on the average throughput.

From Figures 5 and 7, it is noticed that we should carefully choose the number of maximum data slots to make a good tradeoff between system EE and average throughput. From Figures 5, 6 and 7, we can see that the EE performance of the low-complexity sub-optimal scheme is alomst the same as that of the sub-optimal scheme. Furthermore, as the low-complexity sub-optimal one is faster than the sub-optimal one, when L is large, the low-complexity sub-optimal scheme is a better choice for practical applications. If we need a better EE performance, and the maximum data slots L is small, the sub-optimal scheme is more suitable.

The comparison of our proposed sub-optimal scheme with the global optimal solution is shown in Figure 8. We employ exclusive search to find the maximum objective function value, which is a global optimal solution. However, the complexity of exclusive search increases exponentially with the problem size, which is unaffordable. Due to the high complexity, we set $L = 4$ to get a global optimal solution in a reasonable time. It can be observed that the performance of our proposed sub-optimal scheme is very close to the global optimal solution.

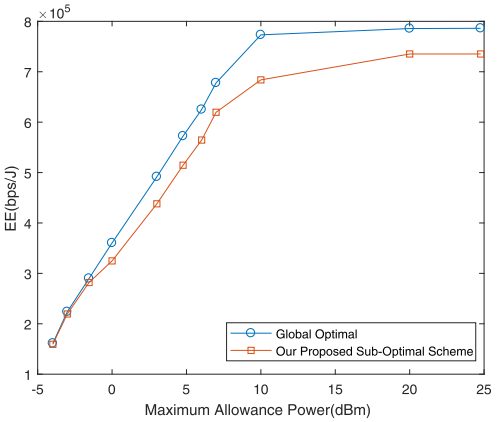

FIGURE 8. Energy efficiency of our proposed sub-optimal scheme and the global optimal solution.

VII. CONCLUSION

In this paper, we have investigated the energy-efficient power control schemes for wireless communication networks, where both the CSI error and the cross correlation of the time-varying channel have been considered. Based on the data rate expression we calculated, an MINLP problem is formulated to maximum EE, and we use SCA and the Dinkelbach methods to give a solution satisfying the KKT conditions. We have further developed the low-complexity sub-optimal scheme through relaxing binary variables to reduce the complexity. Numerical simulations have been conducted to demonstrate the convergence and the effectiveness of our proposed schemes. Simulation results have shown that our proposed schemes converge rapidly, and the EE loss caused by the integer relaxation is very minor. We have illustrated the impact of various network parameters, including the number of maximum data slots, the maximum allowed transmission power, and the cell radius, on the system performance, which could provide some useful guidance for future pilot designs.

APPENDIXES

APPENDIX A

DERIVATION PROCESS OF CHANNEL COEFFICIENT COVARIANCE

According to (7), $\sigma^2_{h_i-\hat{h}_i}$ is defined as

$$\begin{aligned}\sigma^2_{h_i-\hat{h}_i} &= E\left[\left(h_i-\hat{h}_i\right)\left(h_i-\hat{h}_i\right)^H\right] \\ &= E\left[\left(\rho_i h_0+\left(1-\rho_i^2\right)^{\frac{1}{2}} e-\hat{h}_0\right)\right. \\ &\quad \left.\left(\rho_i h_0+\left(1-\rho_i^2\right)^{\frac{1}{2}} e-\hat{h}_0\right)^H\right]. \end{aligned} \quad (67)$$

Recalling (3), we have

$$\hat{h}_0 = \frac{\sigma_h^2 x_0^H y_0}{\sigma_h^2 G P_0 + \sigma_n^2}. \tag{68}$$

Since h and e are statistically independent, we have $E[he^H] = E[eh^H] = 0$, therefore

$$\begin{aligned}\sigma^2_{h_i-\hat{h}_i} &= E\left[\left(\rho_i h + \left(1-\rho_i^2\right)^{\frac{1}{2}} e - \frac{\sigma_h^2 x_0^H y_0}{\sigma_h^2 G P_0 + \sigma_n^2}\right)\right. \\ &\quad \left.\times \left(\rho_i h + \left(1-\rho_i^2\right)^{\frac{1}{2}} e - \frac{\sigma_h^2 x_0^H y_0}{\sigma_h^2 G P_0 + \sigma_n^2}\right)^H\right] \\ &= E\left[\rho_i^2 h h^H + \left(1-\rho_i^2\right) e e^H \right. \\ &\quad \left. + \frac{\sigma_h^4 x_0^H y_0 y_0^H x_0}{\left(\sigma_h^2 G P_0 + \sigma_n^2\right)^2} - \frac{2\rho_i \sigma_h^2 h x_0 x_0^H h^H}{\sigma_h^2 G P_0 + \sigma_n^2}\right]. \end{aligned} \tag{69}$$

Recalling that $E[hh^H] = \sigma_h^2$, $E[ee^H] = \sigma_e^2$, $E[x_0\, x_0^H] = P_0$, $y_0 = \sqrt{G} h_0\, x_0 + n$, and $E[y_0\, y_0^H] = \sigma_h^2\, G P_0 + \sigma_n^2$, we have

$$\begin{aligned}\sigma^2_{h_i-\hat{h}_i} &= \rho_i^2 \sigma_h^2 + \left(1-\rho_i^2\right)\sigma_e^2 + \frac{\sigma_h^4 P_0 - 2\rho_i \sigma_h^4 P_0}{\sigma_h^2 G P_0 + \sigma_n^2} \\ &= \frac{a_i G P_i + b_i}{G P_i \sigma_h^2 + \sigma_n^2}, \end{aligned} \tag{70}$$

where $a_i = \left(1-\rho_i^2\right)\sigma_h^2 \sigma_e^2 + (1-\rho_i)^2 \sigma_h^4$ and $b_i = \rho_i^2 \sigma_h^2 \sigma_n^2 + \left(1-\rho_i^2\right)\sigma_n^2 \sigma_e^2$.

Note that when we consider the case of perfect correlation, i.e. $\rho_i = 1$, (70) degenerates into

$$\frac{\sigma_h^2 \sigma_n^2}{G P_i \sigma_h^2 + \sigma_n^2}, \tag{71}$$

which is exactly the same as that in [21].

APPENDIX B
PROOF OF PROPOSITION 1

If $i \le N$, we have $x_i = 1$. Hence, $\tilde{R}_i = x_i R_i = R_i$, $\tilde{P}_i = x_i P_i = P_i$ and the constraint $C4$ becomes $R_i \ge R_{\text{th}}$. If $N < i \le L$, such $x_i = 0$, we have $\tilde{R}_i = B \log_2(1+0) = 0, \tilde{P}_i = x_i P_i = 0$ and the constraint $C4$ becomes $0 \ge 0$. Therefore,

$$\begin{aligned}\tilde{R} &= \sum_{i=1}^{L} \tilde{R}_i = \sum_{i=1}^{N} R_i = R, \\ \tilde{P} &= \sum_{i=1}^{L} \tilde{P}_i = \sum_{i=1}^{N} P_i = P. \end{aligned} \tag{72}$$

This result means that $\tilde{\eta}(\boldsymbol{P}, \boldsymbol{x}) = \eta\left(\boldsymbol{P}_{[N]}, N\right)$ and constraint $C4$ is equivalent to constraint $C1$.

APPENDIX C
PROOF OF LEMMA 1

1) Assume that

$$\theta_1, \theta_2, \ldots, \theta_m \ge 0, \quad \sum_{j=1}^{m} \theta_j = 1, \tag{73}$$

$$x_1, x_2, \ldots, x_m \ge 0. \tag{74}$$

Recalling that $\ln(x)$ is a concave function, based on Jensen's inequality, we have

$$\sum_{j=1}^{m} \theta_j \ln\left(\frac{x_j}{\theta_j}\right) \le \ln\left(\sum_{j=1}^{m} \theta_j \cdot \frac{x_j}{\theta_j}\right) = \ln\left(\sum_{j=1}^{m} x_j\right), \tag{75}$$

which is equivalent to

$$\sum_{j=1}^{m} x_j \ge \prod_{j=1}^{m} \left(\frac{x_j}{\theta_j}\right)^{\theta_j}. \tag{76}$$

Notice that in (26) we have $\alpha_{1i} + \alpha_{2i} + \alpha_{3i} + \alpha_{4i} = 1$, and we will use (76) to get the inequality we need. Let $m = 4$, $\left(a_i + |\hat{h}|^2 \sigma_h^2\right) G^2\, P_0\, P_i = x_1$, $\left(b_i + |\hat{h}|^2 \sigma_n^2\right) G P_i = x_2$, $\sigma_h^2 \sigma_n^2 G P_0 = x_3$, and $\sigma_n^4 = x_4$, then the left term in (76) becomes $g_i(P_0, P_i)$ and the right term becomes $\gamma_i(P_0, P_i)$. Specifically, we get

$$g_i(P_0, P_i) \ge \gamma_i(P_0, P_i).$$

2) Recalling that the Jensen's inequality (76) is tightened when $\frac{x_1}{\theta_1} = \frac{x_2}{\theta_2} = \cdots = \frac{x_m}{\theta_m}$, thus, in (50) and (52), when $\boldsymbol{P}_{[N]} = \boldsymbol{P}^{(0)}_{[N]}$, we have

$$\begin{aligned}\frac{\left(a_i + |\hat{h}|^2 \sigma_h^2\right) G^2 P_0 P_i}{\alpha_{1i}} &= \frac{\left(b_i + |\hat{h}|^2 \sigma_n^2\right) G P_i}{\alpha_{2i}} \\ &= \frac{\sigma_h^2 \sigma_n^2 G P_0}{\alpha_{3i}} = \frac{\sigma_n^4}{\alpha_{4i}} = \alpha_{0i}, \end{aligned} \tag{77}$$

which satisfies the equal condition. Therefore, the inequality (33) is tightened when $\boldsymbol{P}_{[N]} = \boldsymbol{P}^{(0)}_{[N]}$, namely

$$g_i\left(P_0^{(0)}, P_i^{(0)}\right) = \gamma_i\left(P_0^{(0)}, P_i^{(0)}\right). \tag{78}$$

3) From the chain rule, we have $\nabla R_i = R_i \nabla \ln R_i$ and $\nabla \zeta_i = \zeta_i \nabla \ln \zeta_i$. We will prove $\nabla R_i\left(P_0^{(0)}, P_i^{(0)}\right) = \nabla \zeta_i\left(P_0^{(0)}, P_i^{(0)}\right)$ by calculating all the first-order partial derivatives of $\ln R_i$ and $\ln \zeta_i$ as follows

$$\begin{aligned}\frac{\partial \ln \gamma_i}{\partial P_0}\left(P_0^{(0)}, P_i^{(0)}\right) &= \frac{\partial \ln g_i}{\partial P_0}\left(P_0^{(0)}, P_i^{(0)}\right) \\ &= \frac{\alpha_{1i} P_i^{(0)} + \alpha_{3i}}{P_0^{(0)}}, \\ \frac{\partial \ln \gamma_i}{\partial P_i}\left(P_0^{(0)}, P_i^{(0)}\right) &= \frac{\partial \ln g_i}{\partial P_i}\left(P_0^{(0)}, P_i^{(0)}\right) \\ &= \frac{\alpha_{1i} P_0^{(0)} + \alpha_{2i}}{P_i^{(0)}}, \\ \frac{\partial \ln \gamma_i}{\partial P_j}\left(P_0^{(0)}, P_i^{(0)}\right) &= 0, \quad i \ne j. \end{aligned}$$

Accordingly,

$$(\nabla \ln R_i)\left(P_0^{(0)}, P_i^{(0)}\right) = (\nabla \ln \zeta_i)\left(P_0^{(0)}, P_i^{(0)}\right). \tag{79}$$

Hence, from (78) and (79), we have

$$\nabla R_i\left(P_0^{(0)}, P_i^{(0)}\right) = \nabla \zeta_i\left(P_0^{(0)}, P_i^{(0)}\right). \tag{80}$$

APPENDIX D
PROOF OF LEMMA 2

In iteration k, $\boldsymbol{p}^{(k-1)}$ is the initial value of $\boldsymbol{p}$. According to (33) and (34), we have

$$\underline{\eta}^{(k-1)} \leq \eta_{[N]}(\boldsymbol{p}^{(k-1)}), \tag{81}$$

$$\eta_{[N]}(\boldsymbol{p}^{(k-1)}) = \underline{\eta}(\boldsymbol{p}^{(k-1)}), \tag{82}$$

$$\underline{\eta}(\boldsymbol{p}^{(k-1)}) \leq \max_{\boldsymbol{p}_{[N]}} \underline{\eta}_{[N]} = \underline{\eta}^{(k)}. \tag{83}$$

Hence, $\underline{\eta}^{(k-1)} \leq \underline{\eta}^{(k)}$.

APPENDIX E
PROOF OF LEMMA 3

Let sequence $\boldsymbol{p}^{(k)}$ converge to

$$\boldsymbol{p}^* = [p_0^*, p_1^*, \ldots, p_N^*].$$

Hence, $\boldsymbol{P}_{[N]}^{(0)}$ in (26) equals to $P_{\max}[e^{p_0^*}, e^{p_1^*}, \ldots, e^{p_N^*}]$. According to (35), we have

$$\nabla R_i \left(P_{\max} e^{p_0^{(0)}}, P_{\max} e^{p_i^{(0)}}\right) = \nabla \zeta_i \left(P_{\max} e^{p_0^{(0)}}, P_{\max} e^{p_i^{(0)}}\right).$$

Consequently, the objective function and each condition have the same gradient as the original problem $\mathcal{P}3$ at point $\boldsymbol{p}^*$ when Algorithm 2 converges. Moreover, $\boldsymbol{p}^*$ satisfies the KKT conditions of $\mathcal{P}4$, i.e.,

$$\nabla \underline{\eta}_{[N]}(\boldsymbol{p}^*) + \sum_{i=1}^{N} \mu_i \nabla \zeta_i - P_{\max} \sum_{j=0}^{N} \nu_j \nabla e^{p_j} = 0, \tag{84}$$

where μ_i and ν_j are the KKT multipliers. Considering Corollary 1, it also means that $\boldsymbol{p}^*$ satisfies the KKT conditions of $\mathcal{P}3$, i.e.,

$$\nabla \eta_{[N]}(\boldsymbol{p}^*) + \sum_{i=1}^{N} \mu_i \nabla R_i - P_{\max} \sum_{j=0}^{N} \nu_j \nabla e^{p_j} = 0. \tag{85}$$

REFERENCES

[1] J. An, K. Yang, J. Wu, N. Ye, S. Guo, and Z. Liao, "Achieving sustainable ultra-dense heterogeneous networks for 5G," *IEEE Commun. Mag.*, vol. 55, no. 12, pp. 84–90, Dec. 2017.

[2] D. Feng, L. Lu, Y. Yuan-Wu, G. Li, S. Li, and G. Feng, "Device-to-device communications in cellular networks," *IEEE Commun. Mag.*, vol. 52, no. 4, pp. 49–55, Apr. 2014.

[3] X. Gao, P. Wang, D. Niyato, K. Yang, and J. An, "Auction-based time scheduling for backscatter-aided RF-powered cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, pp. 1684–1697, Mar. 2019.

[4] X. Xia, K. Xu, Y. Wang, and Y. Xu, "A 5G-enabling technology: Benefits, feasibility, and limitations of in-band full-duplex mMIMO," *IEEE Veh. Technol. Mag.*, vol. 13, no. 3, pp. 81–90, Sep. 2018.

[5] G. 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.

[6] J. Tang, A. Shojaeifard, D. K. C. So, K.-K. Wong, and N. Zhao, "Energy efficiency optimization for CoMP-SWIPT heterogeneous networks," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6368–6383, Dec. 2018.

[7] K. Yang, N. Yang, N. Ye, M. Jia, Z. Gao, and R. Fan, "Non-orthogonal multiple access: Achieving sustainable future radio access," *IEEE Commun. Mag.*, vol. 57, no. 2, pp. 116–121, Feb. 2019.

[8] J. Tang, D. K. C. So, N. Zhao, A. Shojaeifard, and K.-K. Wong, "Energy efficiency optimization with SWIPT in MIMO broadcast channels for Internet of Things," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2605–2619, Aug. 2018.

[9] B. Sklar, *Digital Communications: Fundamentals and Applications*. Upper Saddle River, NJ, USA: Prentice-Hall, 1988.

[10] D. Singh Gurjar and P. K. Upadhyay, "Impact of channel estimation error on zero-forcing-based multiple-input-multiple-output two-way relaying," *IET Signal Process.*, vol. 10, no. 3, pp. 210–217, May 2016.

[11] O. Simeone and U. Spagnolini, "Lower bound on training-based channel estimation error for frequency-selective block-fading Rayleigh MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 11, pp. 3265–3277, Nov. 2004.

[12] G. Wang, F. Gao, and C. Tellambura, "Superimposed pilots aided joint CFO and channel estimation for ZP-OFDM modulated two–way relay networks," in *Proc. IEEE 72nd Veh. Technol. Conf. Fall*, Sep. 2010, pp. 394–398.

[13] F. S. Tabataba, P. Sadeghi, C. Hucher, and M. R. Pakravan, "Impact of channel estimation errors and power allocation on analog network coding and routing in two–way relaying," *IEEE Trans. Veh. Technol.*, vol. 61, no. 7, pp. 3223–3239, Sep. 2012.

[14] A. Vosoughi and Y. Jia, "How does channel estimation error affect average sum-rate in two-way amplify-and-forward relay networks?" *IEEE Trans. Wireless Commun.*, vol. 11, no. 5, pp. 1676–1687, May 2012.

[15] S. N. Islam, "Achievable rate and error performance of an amplify and forward multi-way relay network in the presence of imperfect channel estimation," *IET Commun.*, vol. 10, no. 3, pp. 272–282, Feb. 2016.

[16] N. Zhao, F. R. Yu, and H. Sun, "Adaptive energy-efficient power allocation in green interference-alignment-based wireless networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 9, pp. 4268–4281, Sep. 2015.

[17] L. Chen and G. Wei, "Energy-efficient power allocation for training-based multiple-input multiple-output system with and without feedback," *IET Commun.*, vol. 7, no. 15, pp. 1697–1707, Oct. 2013.

[18] Z. Xu, G. Y. Li, C. 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.

[19] T. M. Nguyen and L. B. Le, "Joint pilot assignment and resource allocation in multicell massive MIMO network: Throughput and energy efficiency maximization," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2015, pp. 393–398.

[20] F. Liu, Q. Yang, P. Gong, and K. S. Kwak, "Subcarrier and power allocation for multi-user OFDMA wireless networks under imperfect channel state information," *IET Commun.*, vol. 10, no. 8, pp. 873–881, May 2016.

[21] O. Amin, E. Bedeer, M. H. Ahmed, O. A. Dobre, and M.-S. Alouini, "Opportunistic energy-aware amplify-and-forward cooperative systems with imperfect CSI," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 4875–4886, Jul. 2016.

[22] Y. Kim, G. Miao, and T. Hwang, "Energy efficient pilot and link adaptation for mobile users in TDD multi-user MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 382–393, Jan. 2014.

[23] A. Z. Ghanavati and D. C. Lee, "Optimal energy allocation between pilot and data symbols for minimizing frame error probability under imperfect fading channel information," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2014, pp. 1990–1995.

[24] E. Bjornson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7112–7139, Nov. 2014.

[25] A. Almamori and S. Mohan, "Improved MMSE channel estimation in massive MIMO system with a method for the prediction of channel correlation matrix," in *Proc. IEEE 8th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Jan. 2018, pp. 670–672.

[26] E. Bjornson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3059–3075, Jun. 2015.

[27] C. Pradhan, H. Chen, Y. Li, and B. Vucetic, "Joint beamwidth and energy optimization for multi–user millimeter wave communications," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2018, pp. 1–6.

[28] Y. Yu, X. Bu, K. Yang, Z. Wu, and Z. Han, "Green large-scale fog computing resource allocation using joint benders decomposition, Dinkelbach algorithm, ADMM, and branch-and-bound," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4106–4117, Jun. 2019.

[29] X. Xia, Y. Xu, K. Xu, and W. Ma, "On the design of relay selection strategy for two-way amplify-and-forward mobile relaying," *IET Commun.*, vol. 7, no. 17, pp. 1948–1957, Nov. 2013.

[30] M. C. Gursoy, "On the capacity and energy efficiency of training-based transmissions over fading channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4543–4567, Oct. 2009.
[31] H. Poor, *An Introduction to Signal Detection and Estimation*. New York, NY, USA: Springer-Verlag, 1994.
[32] S. Cui, A. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349–2360, Sep. 2005.
[33] S. Burer and A. N. Letchford, "Non-convex mixed-integer nonlinear programming: A survey," *Surv. Oper. Res. Manage. Sci.*, vol. 17, no. 2, pp. 97–106, Jul. 2012.
[34] M. Locatelli and F. Schoen, "Efficient algorithms for large scale global optimization: Lennard-jones clusters," *Comput. Optim. Appl.*, vol. 26, no. 2, pp. 173–190, Nov. 2003.
[35] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Oper. Res.*, vol. 26, no. 4, pp. 681–683, Aug. 1978.
[36] A. Zappone, L. Sanguinetti, G. Bacci, E. Jorswieck, and M. Debbah, "Energy-efficient power control: A look at 5G wireless technologies," *IEEE Trans. Signal Process.*, vol. 64, no. 7, pp. 1668–1683, Apr. 2016.
[37] X. Gao, D. Niyato, P. Wang, K. Yang, and J. An, "Contract design for time resource assignment and pricing in backscatter-assisted RF-powered networks," *IEEE Wireless Commun. Lett.*, to be published, doi: 10.1109/lwc.2019.2940942.
[38] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[39] K. Yang, S. Martin, C. Xing, J. Wu, and R. Fan, "Energy-efficient power control for device-to-device communications," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3208–3220, Dec. 2016.
[40] W. Dinkelbach, "On nonlinear fractional programming," *Manage. Sci.*, vol. 13, no. 7, pp. 492–498, Mar. 1967.
[41] C. Isheden, Z. Chong, E. Jorswieck, and G. Fettweis, "Framework for link-level energy efficiency optimization with informed transmitter," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2946–2957, Aug. 2012.
[42] A. Nemirovski. (1996). *Interior Point Polynomial Time Methods in Convex Programming*. [Online]. Available: https://www2.isye.gatech.edu/~nemirovs/Lect_IPM.pdf

GUANDONG WEI received the B.S. degree in information engineering from the Beijing Institute of Technology, Beijing, China, in 2019, where he is currently pursuing the master's degree with the School of Information and Electronics. His research interests include green communications and fog computing.

KAI YANG received the B.E. and Ph.D. degrees from the National University of Defense Technology and the Beijing Institute of Technology, China, in 2005 and 2010, respectively, both in communications engineering. From January 2010 to July 2010, he was with the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University. From 2010 to 2013, he was with Alcatel-Lucent Shanghai Bell, Shanghai, China. In 2013, he joined the Laboratoire de Recherche en Informatique, University Paris Sud 11, Orsay, France. He is currently with the School of Information and Electronics, Beijing Institute of Technology, Beijing, China. His current research interests include convex optimization, cooperative communications, MIMO systems, resource allocation, and interference mitigation.

XIAOZHENG GAO received the B.S. degree in information engineering from the Beijing Institute of Technology, Beijing, China, in 2014, where he is currently pursuing the Ph.D. degree with the School of Information and Electronics. He was a Visiting Student with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. His current research interests include cellular communications, backscatter communications, and energy harvesting.

YE YU received the B.Sc. degree in information engineering from the Beijing Institute of Technology, Beijing, China, in 2014, where he is currently pursuing the Ph.D. degree with the School of Information and Electronics. From September 2016 to September 2018, he was a Visiting Student with the Computer Science Department, University of Houston, Houston, TX, USA. His research interests include green communications, network virtualization, heterogeneous networks, and fog computing.

JINSONG WU (Senior Member, IEEE) received the Ph.D. degree from the Department of Electrical and Computer Engineering, Queen's University, Canada. He won both 2017 and 2019 IEEE System Journal Best Paper Awards. His coauthored paper has won the 2018 IEEE TCGCC Best Magazine Paper Award. He received the IEEE Green Communications and Computing Technical Committee 2017 Excellent Services Award for Excellent Technical Leadership and Services in the Green Communications and Computing Community. He is an elected Vice-Chair, Technical Activities, the IEEE Environmental Engineering Initiative, a pan-IEEE effort under the IEEE Technical Activities Board (TAB). He was the Founder and Founding Chair of the IEEE Technical Committee on Green Communications and Computing (TCGCC). He is also the Co-Founder and founding Vice-Chair of the IEEE Technical Committee on Big Data (TCBD). He was the leading Editor and a coauthor of the comprehensive book, entitled *"Green Communications: Theoretical Fundamentals, Algorithms, and Applications"* (CRC Press), in September 2012.

JIANPING AN received the Ph.D. degree from the Beijing Institute of Technology, China, in 1996. He joined the School of Information and Electronics, Beijing Institute of Technology, in 1995, where he is currently a Full Professor. He is currently the Dean of the School of Information and Electronics, Beijing Institute of Technology. His research interests are in the field of digital signal processing, cognitive radio, wireless networks, and high-dynamic broadband wireless transmission technology.