Energy-Efficient Power Allocation between Pilots and Data Symbols in Downlink OFDMA Systems

Zhikun Xu†, Chenyang Yang†, Geoffrey Ye Li‡, Shunqing Zhang§, Yan Chen§, and Shugong Xu§

†School of Electronics and Information Engineering, Beihang University, Beijing 100191, P. R. China
‡School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA
§GREAT Research Team, Huawei Technologies Co. Ltd., Shanghai, P. R. China

Email of corresponding author: xuzhikun@ee.buaa.edu.cn

Abstract—In this paper, power allocation between pilots and data symbols is investigated aiming at maximizing energy efficiency (EE) for downlink orthogonal frequency division multiple access (OFDMA) networks. We first derive an EE function when the channel estimation error is considered, which depends on the large-scale channel gains of multiple users, the allocated power to pilots and data symbols, and the circuit power consumption. Then an optimization problem is formulated to maximize the EE under overall transmit power constraint. The relationship between the power for pilots and data symbols is analyzed based on Karush-Kuhn-Tucker (KKT) conditions and the impacts of channel gains on both power allocation and the EE are studied. Exploiting the quasiconcavity property of the EE function, a bisection searching algorithm is developed to find the optimal power allocation. Simulation results demonstrate the performance gain of the proposed optimal power allocation scheme in terms of the EE and the required overall transmit power.

I. INTRODUCTION

Wireless communications are expected to play a more and more important role in energy consumption due to explosive growth of high-quality wireless services. Consequently, design of energy-efficient wireless systems has become an urgent task.

In wireless communication systems, pilots are usually inserted into data streams to facilitate channel estimation for coherent detection at receivers and preprocessing at transmitters. Pilot design has been studied from different aspects under various criteria. The number, positions, and power for pilot symbols are designed to maximize the capacity lower bound in [1] and to minimize the Cramer-Rao bound of channel estimation error in [2]. It is shown that the optimal placement of pilot symbols is periodical insertion in frequency domain. Power allocation for pilots is discussed for MIMO systems in [3] and for doubly selective fading channels in [4], respectively. A comprehensive overview of pilot-assisted transmission is provided in [5], from both information theory and signal processing point of view.

In contrast to the flourish on pilot design from the perspective of channel capacity and channel estimation errors, little attention has been paid to energy efficiency (EE). Different from the capacity oriented design, the overall transmit power needs to be optimized except for allocating resources between pilots and data symbols. Pilot design for maximizing EE is first studied in [6] for the single user case. By assuming the interference incurred by channel estimation errors as Gaussian, it is shown that the EE reduces to zero as SNR approaches zero and the maximum EE is achieved at a certain SNR. Power allocation between pilots and data symbols in multiple user case is more complicated. Different users suffer from different channel fading, which results in different requirements on the power for pilot symbols. On the other hand, all users may share common pilots to estimate channels, e.g. in downlink 3GPP LTE systems [7]. How to allocate power between pilots and data symbols to maximize the EE in the multiple user case is still unknown.

In this paper, we will investigate energy-efficient power allocation between pilots and data symbols in downlink orthogonal frequency division multiple access (OFDMA) systems. We consider minimum mean-square error (MMSE) channel estimation. Based on the correlation matrix of the channel estimation error, we will derive the ergodic capacity of each user and the EE function considering the circuit power consumption. The optimization problem is formulated to maximize EE with the overall transmit power constraint. The relationship between the optimal power for pilots and data symbols is derived based on Karush-Kuhn-Tucker (KKT) conditions and an efficient numerical algorithm is proposed. Simulation results show the performance gain of the proposed power allocation scheme.

II. SYSTEM AND CHANNEL MODEL

Consider a downlink OFDMA network as shown in Fig. 1. Pilot symbols are periodically placed in the frequency domain and shared by different users for channel estimation [1]. Denote $M$, $N$, and $K$ as the numbers of users, subcarriers, and pilot symbols, respectively. The number of subcarriers occupied by user $m$ is denoted as $N_m$. The index sets of subcarriers for pilots and data symbols for user $m$ are denoted as $S_p^m$ and $S_d^m$, respectively.

We assume that users undergo frequency-selective block fading channels. Specifically, the channel of user $m$ has $L$
The estimation error at subcarrier \( k \) can be obtained as \( \{ \mathbf{g}_{m,l} \}_{l=1}^{L} \). The noise at the receiver of each user is assumed to be additive white Gaussian with zero mean and variance \( \sigma^2 \). Denote \( \alpha \) and \( \beta_m \) as the transmit power for pilot symbols and for data symbols of user \( m \), respectively. We assume that the power is uniformly distributed among subcarriers, and then the power values per subcarrier for pilot symbols and data symbols are \( \alpha / K \) and \( \beta_m / N_m \), respectively.

### III. Problem Formulation

#### A. Performance of Channel Estimation

The frequency-domain pilot vector received by user \( m \) is
\[
y_{p,m} = \mathbf{S}_p \mathbf{h}^f_m + \mathbf{n}_{p,m},
\]
where \( \mathbf{S}_p \) denotes a \( K \times K \) diagonal matrix with pilot symbols \( \{ s_k \}_{k=1}^{K} \) as the diagonal entries, \( \mathbf{n}_{p,m} \) is the noise vector at the subcarriers occupied by pilot symbols, and \( \mathbf{h}^f_m \) is the channel frequency response (CFR) vector at the subcarriers occupied by pilot symbols.

Denote \( \mathbf{h}^f_m \) as the channel impulse response vector of user \( m \), and then we have
\[
\mathbf{h}^f_m = \mathbf{S}_p \mathbf{h}^f_m = \sqrt{N} \mathbf{\Psi}_p \mathbf{F}_L \mathbf{h}^f_m,
\]
where \( \mathbf{h}^f_m \) is the CFR vector of user \( m \) over all subcarriers and \( \mathbf{F}_L \) is an \( N \times L \) truncated normalized FFT matrix as follows,
\[
\mathbf{F}_L = \left[ \frac{1}{\sqrt{N}} e^{-j2\pi nl/N} \right]_{n,l=0}^{N-1,L-1}.
\]
Upon substituting (2), we can rewritten (1) as follows,
\[
y_{p,m} = \sqrt{N} \mathbf{S}_p \mathbf{\Psi}_p \mathbf{F}_L \mathbf{h}^f_m + \mathbf{n}_{p,m} = \mathbf{\Phi}_m \mathbf{h}^f_m + \mathbf{n}_{p,m},
\]
where \( \mathbf{\Phi}_m = \sqrt{N} \mathbf{S}_p \mathbf{\Psi}_p \mathbf{F}_L \).

When the MMSE channel estimator is used, the correlation matrix of estimation error can be obtained as [8, Ch.8]
\[
\mathbf{R}_{\Delta \mathbf{h}} = \mathbb{E}[(\mathbf{h}^f_m - \hat{\mathbf{h}}^f_m)(\mathbf{h}^f_m - \hat{\mathbf{h}}^f_m)^H] = \left( \mathbf{R}_{\mathbf{h}^f_m} + \frac{1}{\sigma^2} \mathbf{\Phi}_m \mathbf{\Phi}_m^H \right)^{-1},
\]
where \( \hat{\mathbf{h}}^f_m \) is the estimation of \( \mathbf{h}^f_m \). \( \mathbf{R}_{\mathbf{x}} \equiv \mathbb{E}[\mathbf{x}\mathbf{x}^H] \) and \( \mathbb{E}[\cdot] \) is the expectation operation. It can be readily derived that \( \mathbf{\Phi}_m \mathbf{\Phi}_m = \alpha \mathbf{I}_K \), where \( \mathbf{I}_K \) denotes a \( K \times K \) identity matrix.

According to our channel assumption, the correlation matrix of \( \mathbf{h}^f_m \) is diagonal, i.e.,
\[
\mathbf{R}_{\mathbf{h}^f_m} = \text{diag}\{\gamma_{m,1}, \ldots, \gamma_{m,L}\}.
\]
Consequently, (4) becomes
\[
\mathbf{R}_{\Delta \mathbf{h}} = \text{diag}\left\{ \frac{1}{\gamma_{m,1} + \sigma^2}, \ldots, \frac{1}{\gamma_{m,L} + \sigma^2} \right\}.
\]

Denote \( \hat{\mathbf{h}}^f_{m,i} \) and \( \Delta \mathbf{h}^f_{m,i} \) as the channel estimation and estimation error at subcarrier \( i \) of user \( m \), respectively. From (5) and (6), their variances can be respectively obtained as
\[
\mathbb{E}[|\hat{\mathbf{h}}^f_{m,i}|^2] = \sum_{l=1}^{L} \frac{\gamma_{m,l}\alpha}{\gamma_{m,l} + \sigma^2} \quad \text{and} \quad \mathbb{E}[|\Delta \mathbf{h}^f_{m,i}|^2] = \sum_{l=1}^{L} \frac{\gamma_{m,l}\sigma^2}{\gamma_{m,l} + \sigma^2}.
\]

#### B. Ergodic Capacity of Each User

The frequency-domain data vector received by user \( m \) can be expressed as
\[
y_d = \mathbf{H}^f_m \mathbf{d} + \mathbf{n}_d = \hat{\mathbf{H}}^f_m \mathbf{d} + \Delta \mathbf{H}^f_m \mathbf{d} + \mathbf{n}_d,
\]
where \( \mathbf{d} \) is the data vector, \( \mathbf{n}_d \) is the noise vector, \( \mathbf{H}^f_m \triangleq \text{diag}\{\{\mathbf{h}^f_{m,i}\}_{i=1}^{L}\} \) is an \( N_m \times N_m \) diagonal matrix whose diagonal elements are the CFRs at the subcarriers occupied by user \( m \), \( \hat{\mathbf{H}}^f_m \) and \( \Delta \mathbf{H}^f_m \) respectively denote its estimation and estimation error, and \( \mathbf{v}_m \triangleq \Delta \mathbf{H}^f_m \mathbf{d} + \mathbf{n}_d \) denotes the total signal distortion.

If we treat the term \( \Delta \mathbf{H}^f_m \mathbf{d} \) as independent Gaussian noise and the data vector is subject to independent Gaussian distribution as in [1], the lower bound of ergodic channel capacity of user \( m \) can be obtained as
\[
C_m = \Delta f \ln \det(\mathbf{I}_{N_m} + \mathbf{R}_{\mathbf{v}_m}^{-1} \hat{\mathbf{H}}^f_m \mathbf{R}_d \hat{\mathbf{H}}^f_m H),
\]
where \( \Delta f \) is the subcarrier spacing, \( \mathbf{R}_d \) and \( \mathbf{R}_{\mathbf{v}_m} \) are the correlation matrices of the data vector and the signal distortion vector, respectively.

Since the transmit power for user \( m \) is equally allocated to its occupied subcarriers, the correlation matrix of the data vector can be expressed as
\[
\mathbf{R}_d = \frac{\beta_m}{N_m} \mathbf{I}_{N_m}.
\]

Considering the independency between data and noise, from (8) and (11) the correlation matrix of the signal distortion vector can be derived as
\[
\mathbf{R}_{\mathbf{v}_m} = \mathbb{E}[\Delta \mathbf{H}^f_m \mathbf{d} \mathbf{d}^H \Delta \mathbf{H}^f_m^H] + \mathbb{E}[\mathbf{n}_d \mathbf{n}_d^H]
\]
\[
= \frac{\beta_m}{N_m} \sum_{l=1}^{L} \frac{\gamma_{m,l}\sigma^2}{\gamma_{m,l} + \sigma^2} \mathbf{I}_{N_m}.
\]

To simplify the analysis of the impact of power allocation
on capacity, we model the channel estimate as
\[ \hat{h}_{m,i}^f = \sqrt{\mathbb{E}[|h_{m,i}^f|^2]} g, \]  
where \( g \triangleq \frac{\hat{h}_{m,i}^f}{\sqrt{\mathbb{E}[|h_{m,i}^f|^2]}} \) is a complex random variable with zero mean and unit variance.

Substituting (7), (11), (12), and (13) into (10), we have
\[ C_m = \Delta f N_m \mathbb{E}[\log_2(1 + \frac{\alpha \gamma_{m,l}}{N_m + \beta_m \sum_{l=1}^{L} \gamma_{m,l} / (\rho + \sum_{m=1}^{M} \beta_m)})]. \]  
(14)

Since the power of received pilot symbols is usually much larger than the noise power, i.e., \( \alpha \gamma_{m,i} \gg \sigma^2 \), we have \( \alpha \gamma_{m,i} + \sigma^2 \approx \alpha \gamma_{m,i} \). Then the lower capacity bound of user \( m \) can be approximated as
\[ C_m \approx \Delta f N_m \mathbb{E}[\log_2(1 + \frac{\alpha \gamma_{m,l}}{N_m \alpha + L \beta m / (\rho + \sum_{m=1}^{M} \beta_m)})]. \]  
(15)

To calculate the approximate ergodic capacity, we need to know the distribution of \( |g|^2 \), which is hard to find in general. To deal with this issue, we further approximate the ergodic capacity using Jensen’s inequality as follows,
\[ C_m \approx \Delta f N_m \mathbb{E}[\log_2(1 + \frac{\alpha \beta m \sum_{l=1}^{L} \gamma_{m,l}}{(N_m \alpha + L \beta m) / (\rho + \sum_{m=1}^{M} \beta_m)})]. \]  
(16)

We will show through simulations later that these approximations have little impact on the considered optimization problem. Yet with such an approximate capacity, the optimal power allocation only depends on the summation of channel power profile of each user, i.e., large-scale channel gains, and does not rely on the distribution of small-scale channel fading. More importantly, the ergodic capacity function is converted from a nonconcave function to a concave one with respect to \( \alpha \) and \( \{\beta_m\}_{m=1}^{M} \), which is beneficial for the optimization.

C. Problem Formulation to Maximize the Energy Efficiency

The linear power model is usually employed to describe the overall consumed power [9], which is
\[ P_{tot} = \rho P_t + P_c, \]
where \( \rho > 1 \) denotes the reciprocal of the efficiency of power amplifier, \( P_t \) is the transmit power, and \( P_c \) is the circuit power. Considering that the transmit power consists of of those for pilot and data symbols, the overall power consumption is rewritten as
\[ P_{tot} = \rho (\alpha + \sum_{m=1}^{M} \beta_m) + P_c. \]  
(17)

The EE is defined as the overall number of bits per unit energy [10] and is equivalent to the sum capacity of multiple users per unit power. From (16) and (17), the expression of EE is
\[ \eta = \frac{\sum_{m=1}^{M} C_m}{P_{tot}} = \frac{\Delta f \sum_{m=1}^{M} N_m \log_2(1 + \frac{\psi_m \alpha \beta m}{N_m \alpha + L \beta m})}{\rho (\alpha + \sum_{m=1}^{M} \beta_m) + P_c}. \]  
(18)

where \( \psi_m \triangleq \sum_{l=1}^{L} \gamma_{m,l}/\sigma^2 \).

The optimization problem of power allocation to maximize EE under transmit power constraint is finally formulated as follows,
\[ \max_{\alpha, \{\beta_m\}_{m=1}^{M}} \Delta f \sum_{m=1}^{M} N_m \log_2(1 + \frac{\psi_m \alpha \beta m}{N_m \alpha + L \beta m}) \]
\[ \text{s. t.} \quad \alpha + \sum_{m=1}^{M} \beta_m \leq P_{max}, \]
\[ \alpha \geq 0, \quad \beta_m \geq 0, \quad m = 1, 2, \ldots, M, \]  
(19)

where \( P_{max} \) is the maximum transmit power.

Since the target application of our work is for best-effort services pursuing high EE, the minimum data rate requirements of multiple users are not considered.

IV. OPTIMAL POWER ALLOCATION

In this section, we will first analyze the relationship between the optimal power for pilots and for data symbols based on KKT conditions, which provides useful insight into the problem. Then we will present an efficient algorithm to find the solution for problem (19).

A. Relationship between Optimal Power for Pilots and Data Symbols

The Lagrange function of problem (19) is shown in (20) on the top of next page, where \( \lambda_1, \lambda_2 \), and \( \{v_m\}_{m=1}^{M} \) are Lagrange multipliers associated with inequality constraints. The KKT conditions are expressed as follows,
\[ \frac{\partial L}{\partial \alpha} = \frac{\partial L}{\partial \beta m} = 0, \]
\[ \frac{\partial L}{\partial \alpha} = \frac{\partial L}{\partial \beta m} \]
\[ \lambda_1 \geq 0, \quad \alpha + \sum_{m=1}^{M} \beta_m \leq P_{max}, \]
\[ \lambda_1 (\alpha + \sum_{m=1}^{M} \beta_m - P_{max}) = 0, \]
\[ \alpha \geq 0, \quad \lambda_2 \geq 0, \quad \lambda_2 \alpha = 0, \]
\[ \beta_m \geq 0, \quad v_m \geq 0, \quad v_m \beta_m = 0, \quad m = 1, \ldots, M. \]  
(21)

1The optimization problem can be readily changed by adding some data rate constraints from multiple users to guarantee user’s performance, which will be considered in our future work.
We can prove that the KKT conditions are the necessary condition of problem (19) by showing that all the constraints of the problem satisfy linear independence constraints qualification [11].

After some manipulations\(^2\), we can respectively derive the optimal power for pilots and the optimal power for data symbols of multiple users as follows,

\[
\alpha^* = \min_{\alpha} \frac{M}{\sum_{m=1}^{M} L^2 m^2} \frac{1}{N_m},
\]

and

\[
\beta_m^* = \begin{cases} 
0, & \psi_m \leq \mu \\
\frac{N_m \alpha^* (\sqrt{\log_2 (\rho \eta^* + \lambda_1 (\rho (\alpha^* + \sum_{m=1}^{M} \beta_m^* + P_c)) \cdot 2) + \psi_m \alpha^*)}}{2L (\log_2 (\rho \eta^* + \lambda_1 (\rho (\alpha^* + \sum_{m=1}^{M} \beta_m^* + P_c)))}, & \psi_m > \mu
\end{cases}
\]

where \(\mu \triangleq \frac{\ln 2}{2} (\rho \eta^* + \lambda_1 (\rho (\alpha^* + \sum_{m=1}^{M} \beta_m^* + P_c)))\) and \(\eta^*\) is the optimal EE.

It follows from (22) that the optimal power allocated to pilots depends on the weighted sum of the optimal power allocated to data symbols of multiple users, and increases with the power for data symbols. This implies that when more power is allocated to data transmission to achieve higher capacity, more power should be allocated for pilots to provide more accurate channel estimation.

It is shown from (23) that user \(m\) will transmit data only when its channel gain \(\psi_m\) exceeds a threshold \(\mu\). Furthermore, when \(\psi_m > \mu\), the power allocated for each subcarrier of user \(m\), i.e., \(\beta_m^* / N_m\), can be shown as a monotonously increasing function of \(\psi_m\). This implies that more power will be allocated to a user with higher channel gain.

If all the power values for \(\{\beta_m\}_{m=1}^{M}\) are zero, the EE in (18) will be zero. Therefore, at least one nonzero value exists among \(\{\beta_m\}_{m=1}^{M}\) when the optimal EE is achieved. From the nonzero condition in (23), we have \(\mu < \max\{\psi_1, \cdots, \psi_M\}\). After considering the expression of \(\beta_m\), we can obtain that

\[
\frac{\ln 2}{\Delta f} (\rho \eta^* + \lambda_1 (\rho (\alpha^* + \sum_{m=1}^{M} \beta_m^* + P_c))) < \max\{\psi_1, \cdots, \psi_M\}
\]

i.e.,

\[
\eta^* < \frac{1}{\rho} \frac{\Delta f}{\ln 2} \max\{\psi_1, \cdots, \psi_M\} = \lambda_1 (\rho (\alpha^* + \sum_{m=1}^{M} \beta_m^* + P_c)).
\]

Since \(\lambda_1 \geq 0\), the optimal EE is upper bounded by

\[
\eta^* < \frac{\Delta f}{\rho \ln 2} \max\{\psi_1, \cdots, \psi_M\}.
\]

\(^2\)Due to the lack of space, the derivation is omitted.

### B. Bisection Searching Algorithm

Although a numerical solution can be obtained by solving (22) and (23) iteratively, the convergence of this algorithm is not guaranteed. In the following, we derive an efficient algorithm.

We can show that the EE function in (18) is quasiconcave in variables \(\alpha\) and \(\{\beta_m\}_{m=1}^{M}\). In addition, all the constraints are linear. Therefore, (19) is a quasiconcave optimization problem, whose solution can be obtained by solving a set of concave feasibility problems [12].

Based on the expression of EE, we define a function as

\[
\phi_t(\alpha, \beta_1, \cdots, \beta_M) \triangleq \Delta f \sum_{m=1}^{M} N_m \log_2 (1 + f(\alpha, \beta_m)) - t (\rho (\alpha + \sum_{m=1}^{M} \beta_m) + P_c),
\]

where \(t\) is a parameter. Since the sum capacity is a concave function with respect to \(\alpha\) and \(\{\beta_m\}_{m=1}^{M}\), \(\phi_t(\alpha, \beta_1, \cdots, \beta_M)\) is also concave. For different values of \(t\), a set of concave feasibility problems is formulated as follows,

\[
\begin{aligned}
\text{find} & \quad \alpha, \{\beta_m\}_{m=1}^{M} \\
\text{s. t.} & \quad \phi_t(\alpha, \{\beta_m\}_{m=1}^{M}) \geq 0, \\
& \quad \alpha + \sum_{m=1}^{M} \beta_m \leq P_{\max}, \\
& \quad \alpha \geq 0, \\
& \quad \beta_m \geq 0, \quad m = 1, 2, \cdots, M.
\end{aligned}
\]

For a given \(t\), if this concave problem is feasible, we have \(\eta^* \geq t\) from constraint (24a). Otherwise, if the problem is infeasible, \(\eta^* < t\). Based on this observation, we propose a bisection searching algorithm to solve problem (19).

The algorithm starts with an interval \([\eta_l, \eta_u]\), which contains the optimal value \(\eta^*\). Then problem (24) is solved at its midpoint, \(t = (\eta_l + \eta_u) / 2\), where interior point methods can be applied [12]. If the problem is feasible, the optimal EE \(\eta^*\)
Table II
List of Simulation Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subcarrier spacing, $\Delta f$</td>
<td>15 kHz</td>
</tr>
<tr>
<td>Number of subcarriers, $N$</td>
<td>1024</td>
</tr>
<tr>
<td>Number of pilots, $K$</td>
<td>16</td>
</tr>
<tr>
<td>Number of users, $M$</td>
<td>40</td>
</tr>
<tr>
<td>Number of subcarriers occupied by user $m$, $N_m$</td>
<td>25</td>
</tr>
<tr>
<td>Cell radius, $R$</td>
<td>200-2000 m</td>
</tr>
<tr>
<td>Distribution of $M$ users</td>
<td>Independently and uniformly distributed in the cell</td>
</tr>
<tr>
<td>Channel length of User $m$, $L$</td>
<td>8</td>
</tr>
<tr>
<td>Power spectral density of noise</td>
<td>-174 dBm/Hz</td>
</tr>
<tr>
<td>Variance of Shadowing</td>
<td>8 dB</td>
</tr>
<tr>
<td>Noise amplifier gain</td>
<td>7 dB</td>
</tr>
<tr>
<td>Minimum distance from base station to users, $d_{\text{min}}$</td>
<td>35 m</td>
</tr>
<tr>
<td>Path loss (dB)</td>
<td>$35 + 38 \log_{10} d$</td>
</tr>
<tr>
<td>Circuit power consumption, $P_c$</td>
<td>58 W</td>
</tr>
<tr>
<td>Maximum transmit power, $P_{\text{max}}$</td>
<td>40 W</td>
</tr>
<tr>
<td>Efficiency of Power Amplifier</td>
<td>38%</td>
</tr>
</tbody>
</table>

lies in the upper half of the interval. Otherwise, the optimal EE $\eta^*$ lies in the lower half of the interval. Update the interval accordingly and repeat the procedures until the duration of the interval is small enough. The initial interval can be set as the lower and upper bound of $\eta^*$, respectively. Based on the EE upper bound derived in Subsection A, we can set $\eta_l = 0$ and $\eta_u = \frac{\Delta f}{\rho L_m^2} \max\{\psi_1, \ldots, \psi_M\}$. The detailed algorithm is shown in Table I.

The convergence of this algorithm is guaranteed by applying bisection searching and by solving the concave feasibility problem in each step. Since the interval is divided into two parts in each iteration, the interval duration after $k$ iterations becomes $2^{-k}(\eta_u - \eta_l)$. Then $\lceil \log_2((\eta_u - \eta_l)/\epsilon) \rceil$ iterations are required before the algorithm terminates, where $\epsilon$ is EE accuracy.

V. SIMULATION RESULTS

In this section, we evaluate performances of the proposed scheme in terms of the EE, the overall transmit power, and the ratio of the power for pilots to the overall transmit power. System parameters are listed in Table II. The maximum transmit power, the circuit power and the efficiency of power amplifier are configured as in [13].

In order to show the accuracy of the approximation on the lower capacity bound in Section III, we compare the EE when the approximate lower capacity bound in (16) is used for power optimization with the EE when the accurate lower capacity bound in (14) is used. Since the EE function does not have a good property for designing efficient algorithm when using (14), we employ the exhaustive searching method to find the optimal power allocation. Due to its high complexity, only one user is considered. Fig. 2 shows the EE versus the cell size for both cases. It can be observed that the EE gap between these two cases is very small.

We evaluate the performance of the proposed algorithm (“Optimal PA Scheme” in the legend) by comparing it with equal power allocation and full power allocation. In the equal power allocation scheme, the power at each subcarrier for data transmission is equal, i.e., $\beta_{n1} = \beta_{n2} = \cdots = \beta_{nM} = \beta$. In the full power allocation scheme, the maximum transmit power is employed for pilot and data transmission, i.e., constraint (19a) becomes $\alpha + \sum_{m=1}^{M} \beta_m = P_{\text{max}}$. It is readily proved that optimization problems in these two schemes are still quasiconcave, and thus similar algorithms to that in Table I can be used to find the optimal solutions.

Figure 3 shows the EE of different power allocation schemes versus the cell size. It is shown that the EE decreases with the cell size, which implies that microcell deployment is more energy efficient. The optimal power allocation scheme outperforms other two schemes. When the cell size is small, the performance of equal power allocation scheme is close to that of the optimal power allocation scheme and much better than that of full power allocation scheme. When the cell size

![Fig. 2. EE vs. cell size when the capacity expressions in (14) and (16) are used for designing power allocation](image)

![Fig. 3. EE vs. cell size](image)
is large, the performance of full power allocation scheme is close to that of the optimal power allocation scheme and better than that of equal power allocation scheme.

The required overall transmit powers of three power allocation schemes are compared in Fig. 4. For the optimal and equal power allocation schemes, when the cell size is small, the transmit powers are similar and much lower than the maximum power. As the cell size grows, the transmit powers for these two schemes increase and the optimal power allocation scheme always consumes the least power.

Figure 5 shows the ratio of the power for pilots to the overall transmit power. We can see that the ratio increases with the cell size when the optimal or full power allocation scheme is used, which implies that channel estimation takes up more resources as the cell size gets larger. It can also be observed that the ratio with the full power allocation scheme is less than that with the optimal power allocation scheme. This is because more overall transmit power is consumed in the full power allocation scheme. In the equal power allocation scheme, the ratio is a constant with the value \( \frac{\sqrt{L}}{\sum_{m=1}^{M} N_m + \sqrt{L}} \).

**VI. CONCLUSION**

In this paper, power allocation between pilots and data symbols aiming at maximizing the EE of training-based downlink OFDMA systems was studied. We derived the EE function when channel estimation error existed, which only depended on the statistical channel information of multiple users. From the KKT conditions, we analyzed the relationship between the optimal power for pilots and data symbols. By exploiting the quasiconcavity property of the EE function, a bisection searching algorithm was developed to find the solution of optimal power allocation. Simulation results show that the optimal power allocation scheme outperforms the equal and full power allocation schemes in terms of both EE and overall transmit power consumption. For the optimal power allocation scheme, the EE decreases with the cell size while both the overall transmit power and the ratio of the power for pilots to the overall transmit power increases with the cell size.

**REFERENCES**