Abstract—Energy efficiency is important for mobile devices because battery technology development has not kept up with the growing demand of ubiquitous broadband communications. This paper addresses optimal energy-efficient design for uplink (UL) MU-MIMO in a single cell environment. We assume mobile devices have dynamic antenna management capability and can turn off circuit operations when some antennas are not used to save power. We show that some antennas should not be used even when their channel states are good because turning them on consumes too much circuit power. This indicates a single antenna system could perform better than a multi-antenna system in terms of energy efficiency. Based on theoretical analysis, we further develop low-complexity yet globally optimal algorithms that converge to the optimum exponentially.

Index Terms— energy efficiency, multi-user MIMO, power allocation, SDMA

I. INTRODUCTION

Energy efficiency is increasingly important for mobile devices because battery development has not kept up with the growing demand of ubiquitous multimedia communications [1], [2]. In addition to energy saving, energy-efficient communications have the benefit of reducing interference to other co-channel users as well as lessening environmental impacts, e.g., heat dissipation and electronic pollution. Therefore, recent research has focused on energy-efficient wireless communication techniques [3]–[9]. On the other hand, multiple-input multiple-output (MIMO) has been a key technology for wireless systems because of its potential to achieve high capacity [10]. In a multi-user scenario, a multi-user multiple-input multiple-output (MU-MIMO) systems can provide a substantial gain by allowing multiple users to communicate in the same frequency and time slots [11]. While there has been extensive research on improving the spectral efficiency of MU-MIMO, little effort in literature is focused on energy-efficient communications in MU-MIMO systems.

In our previous paper [12], we addressed the energy-efficient design of uplink (UL) MU-MIMO in a single cell environment. The proposed scheme balances the energy consumption of circuit operations and RF transmissions of all users to achieve the maximum network energy efficiency. In [12], we assume that all users consume a fixed amount of circuit power in addition to the RF power. In practice users may have improved antenna management capability and turn off part of the circuit operations when some antennas are not used to reduce the circuit power consumption. This motivates the work in this paper. In this paper we study energy-efficient UL MU-MIMO with dynamic antenna management. We show that in this case, the problem is non-concave and multiple local maximums may exist. However, we will develop algorithms that converge to the global optimum. The paper is organized as follows. In Section II, we formulate the problem. In Section III, we study the optimal solution and design low-complexity algorithms to achieve the optimal solution. Simulation results are provided in Section IV to demonstrate the performance improvement and the conclusion is given in Section V.

II. ENERGY-EFFICIENT MU-MIMO

In this section, we introduce energy-efficient MU-MIMO with dynamic antenna management. In the paper, denote matrices by capital boldface letters, vectors by lowercase boldface, and scalars by either upper or lowercase letters without boldface.

Consider an uplink MU-MIMO system where one access point (AP) is serving $K$ users. Both the AP and all users desire energy-efficient communications. The AP has $N$ antennas. User $i$ has $k_i$ antennas and $\sum_{i=1}^{K} k_i \leq N$. Assume block fading, that is, the channel state remains constant during each data frame. The channel between the AP and users are predetermined earlier through either training pilots as in a time-division duplex system or a feedback channel as in a frequency-division duplex system. In a flat-fading propagation environment, the received signal at the AP is denoted by

$$y = H \cdot Q \cdot P \cdot x + n = \sum_{i=1}^{K} H_i \cdot Q_i \cdot P_i \cdot x_i + n,$$  \hspace{1cm} (1)$$

where $y = [y_1, y_2, ..., y_N]^T$, $x_i = [x_{i1}, x_{i2}, ..., x_{ik_i}]^T$, consists of the transmitted signals of User $i$ and $E[|x_{ij}|^2] = 1$, where $E$ is the expectation. Here $^T$ is the transpose of a vector. $P_i = \text{diag}\{\sqrt{P_{i1}}, \sqrt{P_{i2}}, ..., \sqrt{P_{ik_i}}\}$ is the power allocation matrix of User $i$. $Q_i$ is the precoding matrix of User $i$. $H_i$ is the $N \times k_i$ channel matrix of User $i$ and is assumed to have rank $k_i$, which is generally true in a rich-scattering environment. $n$ is the length-$N$ noise vector, which is Gaussian distributed with a zero mean and a covariance matrix $\sigma^2 I_N$, where $I_N$ is the
identity matrix of size $N$.

$$\mathbf{x} = [x_1, x_2, \ldots, x_K]^T,$$
$$\mathbf{P} = \text{diag}\{\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_K\},$$
$$\mathbf{Q} = \text{diag}\{\mathbf{Q}_1, \mathbf{Q}_2, \ldots, \mathbf{Q}_K\},$$
and
$$\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_K].$$

Assume each user has local channel knowledge $\mathbf{H}_i$ and sets the precoder $\mathbf{Q}_i$ to be the matrix of the right singular vectors of $\mathbf{H}_i$. The AP has full channel information $\mathbf{H}$ and uses the zero-forcing receiver. As shown in [12], the transmissions of different users can be uncoupled and the AP can detect each symbol independently. The achieved signal-to-noise ratio (SNR) of all the symbols for User $i$ is

$$\eta_i = \left[ \frac{p_1\lambda_2^2}{\sigma^2}, \frac{p_2\lambda_2^2}{\sigma^2}, \ldots, \frac{p_i\lambda_2^2}{\sigma^2} \right]^T.$$  \hspace{1cm} (2)

Correspondingly denote the data rate vector of User $i$, $\mathbf{R}_i = [r_{i1}, r_{i2}, \ldots, r_{ik}]$. The overall data rate is

$$R_i = \sum_{k=1}^{K} r_{ik}.  \hspace{1cm} (3)$$

Denote $B$ as the system bandwidth. The achievable data rate $r_{ik}$ is determined by [13]

$$r_{ik} = B \log_2(1 + \frac{\eta_{ik}}{\Gamma}),  \hspace{1cm} (4)$$

where $\eta_{ik} = \frac{p_{ik}\lambda_2^2}{\sigma^2}$ and $\Gamma$ is the SNR gap that defines the gap between the channel capacity and a practical coding and modulation scheme. Denote the overall transmission power of User $i$ to be $P_{Ti}$ and

$$P_{Ti} = \sum_{k=1}^{K} p_{ik} \zeta,  \hspace{1cm} (5)$$

where $\zeta \in [0, 1]$ is the power amplifier efficiency.

In addition to transmission power, mobile devices also incur additional circuit power consumption owing to inevitable electronic operations which is relatively independent of the radio frequency (RF) transmission [6]. This portion of energy consumption excludes that of the power amplifier and is independent of the transmission state. With dynamic antenna management, circuit power is a function of the set of antennas that are turned on. In the following, for notation simplicity, denote the circuit power of User $i$ by $P_{C_i}(k_i^o)$, where $k_i^o$ is the number of antennas that have positive power allocation. $P_{C_i}(k_i^o)$ is increasing in $k_i^o$. Please note the result in this section also applies in generic cases that $P_{C_i}$ is a function of the set of antennas that are turned on. A simple example is

$$P_{C_i}(k_i^o) = k_i^o P_{\alpha} + I(k_i^o) P_{\beta},  \hspace{1cm} (6)$$

where $P_{\alpha}$ is the extra antenna-related circuit power consumption when one more antenna is turned on and $P_{\beta}$ is the power consumption of circuit components that are independent of the number of antennas turned on. When $k_i^o = 0$, the user can be turned off completely to avoid any circuit power consumption and $P_{C_i}(0) = 0$. The indicator function $I(A)$ is

$$I(A) = \begin{cases} 1 & \text{if } A > 0, \\ 0 & \text{otherwise}. \end{cases} \hspace{1cm} (7)$$

So the total power consumption of User $i$ is

$$P_i = P_{Ti} + P_{C_i}(k_i^o).  \hspace{1cm} (8)$$

The AP also consumes electronic circuit energy to receive and decode signals. Denote the receiver circuit power to be $P_r$, which models the average energy consumption of AP device electronics, such as mixers, filters, and analog-to-digital converters.

The MU-MIMO system wants to maximize the system energy efficiency, i.e.,

$$U(P) = \sum_i w_i R_i \max_\mathbf{P} \frac{\sum_{i=1}^{K} w_i R_i}{\alpha \sum_{i} (P_{Ti} + P_{C_i}(k_i^o)) + \beta P_r},  \hspace{1cm} (9)$$

where $w_i \in (0, 1]$ is the predetermined weight for User $i$, indicating the corresponding priority in resource allocation. For a special case, $w_i = 1$ for all $i$ and $U$ is the total number of bits sent per Joule of energy consumption. The weights $\alpha, \beta \in [0, 1]$ characterize the priorities of transmitter and receiver power consumptions. If the receiver power is not considered, $\alpha = 1$ and $\beta = 0$. The unit of energy efficiency is bits per Joule. With metric (9), the energy used for sending each information bit is minimized. The energy efficiency capacity of MU-MIMO is defined as

$$U^* = \max_\mathbf{P} \hat{U}(P) = \max_\mathbf{P} \frac{\sum_{i=1}^{K} w_i R_i}{\alpha \sum_{i} (P_{Ti} + P_{C_i}(k_i^o)) + \beta P_r},  \hspace{1cm} (10)$$

and the optimal energy-efficient power allocation achieving the energy efficiency capacity is

$$\mathbf{P}^* = \arg \max_\mathbf{P} \hat{U}(P) = \arg \max_\mathbf{P} \frac{\sum_{i=1}^{K} w_i R_i}{\alpha \sum_{i} (P_{Ti} + P_{C_i}(k_i^o)) + \beta P_r},  \hspace{1cm} (11)$$

where $k_i^o = \sum_{k=1}^{K} I(p_{ik})$. Note that when $K = 1$, (10) and (11) give the energy efficiency capacity and the optimal power allocation for a point-to-point MIMO system. Therefore the results in this paper are also applicable to MIMO systems.

III. ENERGY-EFFICIENT ANTENNA MANAGEMENT AND POWER ALLOCATION

Different from the work in [12], the energy efficiency function $\hat{U}$ is no longer continuous or quasi-concave in $\mathbf{P}$ in (9). So it is difficult to find the optimal solution.

A. Properties of the Globally Optimal Solution

Observe Antenna $j$ of User $i$ and denote it to be Antenna $(i,j)$. Assume the power on all other antennas have been optimally allocated and denote $P_{ij}^{(0)}(p_{ij})$ to be the power allocation that equals the optimal power allocation except that the power on Antenna $(i,j)$ is $p_{ij}$. Antenna $(i,j)$ may have two states, on or off. If it is on, the energy efficiency is

$$\hat{U}\left(P_{ij}^{(0)}(p_{ij})\right) = \frac{R_{ij}^{(0)} + w_i B \log_2 \left(1 + \frac{p_{ij}\lambda_2^2}{\Gamma \sigma^2}\right)}{P_{ij}^{(0)} + p_{ij} \frac{\zeta}{k} + P_{C_i}(k_i^o)} \hspace{1cm} (12)$$

where $P_{ij}^{(0)}$ is the power on Antenna $(i,j)$ when it is off.
where \( R_{ij}^o = \sum_{(u, k, u \neq i, k \neq j)} w_i B \log_2 \left( 1 + \frac{p_{uk} \lambda_{ik}}{\Gamma \sigma^2} \right) \) and 
\[ P_{ij}^o = \alpha \sum_{u \neq i} P_{Ci}(k_u^o) + \beta P_r + \frac{\alpha}{\zeta} \sum_{(u, k, u \neq i, k \neq j)} p_{uk}. \]
It can be shown that \( \hat{U} \) is strictly quasi-concave in \( p_{ij} \). Hence there is a unique \( p_{ij} \) that maximizes \( \hat{U} \) if \( p_{ij} > 0 \), i.e., Antenna \((i, j)\) should be turned on. In the following we study the condition that Antenna \((i, j)\) should be turned on.

The partial derivative of \( \hat{U} \) with respect to \( p_{ij} \) is
\[
\frac{\partial \hat{U}}{\partial p_{ij}} = \frac{f(p_{ij})}{(1 + \frac{p_{ij} \lambda_{ij}}{\Gamma \sigma^2})(\xi p_{ij} + P_{ij}^o + P_{Ci}(k_{ij}^o))^2},
\]
where 
\[ f(p_{ij}) = w_i B \log_2 \left( 1 + \frac{p_{ij} \lambda_{ij}}{\Gamma \sigma^2} \right) - p_{ij} B \log_2 (1 + \frac{p_{ij} \lambda_{ij}}{\Gamma \sigma^2}). \]
Because \( \hat{U} \) is strictly quasi-concave, if there exists a \( p_{ij} > 0 \) such that \( \frac{\partial \hat{U}}{\partial p_{ij}} = 0 \), it is unique. This further indicates that if \( p_{ij} > 0 \), there is only one \( p_{ij} \) such that \( f(p_{ij}) = 0 \). In Appendix I, we prove that only when
\[
\frac{\lambda_{ij}^2}{\sigma^2} > \frac{R_{ij}^o \alpha \Gamma \ln 2}{(P_{ij}^o + P_{Ci}(k_{ij}^o)) w_i B \xi},
\]
where \( k_{ij}^o \) is the number of antennas when Antenna \((i, j)\) is turned on, will there exist a \( p_{ij} > 0 \) such that \( f(p_{ij}) = 0 \). This is also the condition for Antenna \((i, j)\) to be turned on.

Readily, by setting \( \frac{\partial \hat{U}}{\partial p_{ij}} \) to be zero, we have the following necessary condition of globally optimal energy-efficient power allocation.

**Theorem 1.** With improved circuit management, the optimal energy-efficient power allocation \( \mathbf{P}^* \) achieving the energy capacity satisfies, for antennas that are turned on,
\[
p_{ij}^* = \frac{w_i B \xi}{\alpha \Gamma \ln 2} - \frac{\Gamma \sigma^2}{\lambda_{ij}^2},
\]
and these antennas have channel conditions
\[
\frac{\lambda_{ij}^2}{\sigma^2 \Gamma} > \frac{R_{ij}^o \alpha \Gamma \ln 2}{(P_{ij}^o + P_{Ci}(k_{ij}^o)) w_i B \xi},
\]
where \( k_{ij}^o \) is the number of antennas of User \( i \) when Antenna \((i, j)\) is turned on. Correspondingly, the energy efficiency capacity is
\[
U^* = \hat{U}(\mathbf{P}^*).
\]

According to Theorem 1, whether or not Antenna \((i, j)\) should be turned on is determined by multiple factors. \( \frac{\lambda_{ij}^2}{\sigma^2 \Gamma} \) characterizes the channel condition of Antenna \((i, j)\) and determines the effective receiver SNR once the power is allocated. If it is above the threshold \( \frac{R_{ij}^o \alpha \Gamma \ln 2}{(P_{ij}^o + P_{Ci}(k_{ij}^o)) w_i B \xi} \), Antenna \((i, j)\) should be used since it improves the overall network energy efficiency. The threshold decreases with the power amplifier efficiency, \( \xi \), indicating that with improved power amplifier efficiency, worse channel conditions can be used for data transmission on Antenna \((i, j)\) to achieve energy efficiency capacity. The threshold is also determined by the states of all other antennas. The energy efficiency achieved by all other antennas is \( \frac{R_{ij}^o}{P_{ij}^o + P_{Ci}(k_{ij}^o)} \) assuming Antenna \((i, j)\) has an infinite small amount of power allocated. The threshold increases with \( \frac{R_{ij}^o \alpha \Gamma \ln 2}{(P_{ij}^o + P_{Ci}(k_{ij}^o)) w_i B \xi} \), which indicates that if the energy efficiency achieved by all other antennas increases, the channel state for Antenna \((i, j)\) should also be better such that turning it on improves the overall network energy efficiency.

**B. Algorithm Development**

Theorem 1 only gives the necessary conditions of globally optimal energy efficient power allocation. There may be multiple power allocation schemes that satisfy Theorem 1 because there may be multiple local maximums of the energy efficiency function \( \hat{U} \). An example is given in Fig. 1, where we assume one user with two antennas is communicating to the AP, i.e., a MIMO system. The circuit power of the user is assumed to be \( P_{Ci} = k_{ii}^o + I(k_{ii}^o) \). Observing Fig. 1, \( \hat{U} \) has three local maximums, each of which satisfies Theorem 1. When both antennas are turned on, there is a unique power allocation that maximizes \( \hat{U} \). When the state of one antenna switches from on to off, i.e., \( p_{11} \) or \( p_{12} \) goes to zero, the energy efficiency \( \hat{U} \) increases abruptly because of the reduction of circuit power. In this case, we need to compare the three local maximums to determine the energy efficiency capacity and the optimal power allocation. Therefore, to find the optimal power allocation, we can use the EMMPA algorithm, which is globally optimal when circuit power is fixed, in [12] to determine the corresponding optimal power allocation and the maximum energy efficiency achieved when Antenna \((1, 1)\) is on, \((1, 2)\) is on, or both are on. The antenna configuration and power allocation that achieve the highest energy efficiency is the optimal one.

1) **An Exhaustive Search Algorithm:** For a generic case, denote the antenna configuration to be a binary vector, \( e \), of length \( \sum_i k_i \), in which 1 indicates the corresponding antenna is on and 0 otherwise. Given a certain antenna configuration, the circuit power of all users is determined and the power allocation is the same as that in [12]. So the EMMPA algorithm in [12] can be used to determine the corresponding optimal power allocation for that antenna configuration. Therefore we
need to determine the globally optimal antenna configuration and use EMMPA to find the corresponding power allocation.

One simple approach is that we can exhaustively search all antenna configurations and use EMMPA to determine the maximum energy efficiency achieved for each configuration. The antennas configuration and corresponding power allocation that achieves the highest energy efficiency is the globally optimal one. We call this as the exhaustive search power allocation (ESPA) algorithm and summarize it in Table I.

### Table I: ESPA Algorithm

<table>
<thead>
<tr>
<th>Algorithm Exhaustive search power allocation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ( U_{\text{max}} \leftarrow 0; \ P^* \leftarrow 0; )</td>
</tr>
<tr>
<td>2. for all antenna configurations</td>
</tr>
<tr>
<td>3. Calculate the circuit power for each user;</td>
</tr>
<tr>
<td>4. Use EMMPA to find the optimal power allocation ( P ) and the corresponding EE ( U );</td>
</tr>
<tr>
<td>5. if ( U &gt; U_{\text{max}} ) and all antennas turned on have positive power allocation</td>
</tr>
<tr>
<td>6. ( c \leftarrow \text{current antenna configuration}; )</td>
</tr>
<tr>
<td>7. ( U_{\text{max}} \leftarrow U ) and ( P^* \leftarrow P; )</td>
</tr>
<tr>
<td>8. return ( c, P^* ), and ( U_{\text{max}} ).</td>
</tr>
</tbody>
</table>

The complexity of ESPA grows exponentially with the total number of antennas at the user side and it can be easily shown that the convergence rate is characterized by Proposition 1.

**Proposition 1.** ESPA converges to the globally optimal power allocation. The globally optimal antenna configuration, as well as the power allocation for antennas that are turned on, \( P_{ik} = \left[ \mu - \frac{RC_{ik}}{N_{ik}} \right]^+, \) where \( \mu \) satisfies \( \mu - \frac{RC_{ik}}{\sigma^2} \leq \epsilon, \) can be found within at most \( \left\lceil \log_2 \left( \frac{\sigma^2}{\epsilon} \right) \right\rceil (2\sum_i k_i - 1) \) iterations.

2) A Quadratic-Complexity Algorithm: For a small number of users and antennas, ESPA is effective in finding the globally optimal solution. When there are many users in the system and each has many antennas, ESPA has high complexity. In the following, we further develop a low complexity algorithm. This algorithm consists of two steps. In the first step, the linear-complexity EMMPA is used to filter out antennas that are a subset of antennas that should be turned off in the globally optimal antenna configuration. These antennas are turned off because of their poor channel conditions. In the second step, the remaining antennas are examined and some may be turned off to achieve higher energy efficiency because they consume relatively large amounts of circuit power. We will show that while the proposed algorithm has quadratic complexity, it is also globally optimal.

In the first step, assume all antennas are turned on and the circuit power, \( P_{C_i} (k_i^*) \), of all users can be determined. According to Theorem 1 in [12], antennas with channel conditions \( w_i B C_{ik} \leq \sigma^2 \) will have zero power allocation and we denote these antennas by set \( S^{(1)} \). If any antenna is turned off, the circuit power of the corresponding user is reduced. According to Proposition 1 in [12], the optimal power allocated on each spatial channel, if nonzero, will be reduced. Therefore, any antenna in \( S^{(1)} \) will not be allocated positive power if any antenna is turned off. This indicates that \( S^{(1)} \) belongs to the set of antennas that should be turned off in the globally optimal antenna configuration. \( S^{(1)} \) can be determined by EMMPA. In the second round, we turn off all antennas in \( S^{(1)} \) and calculate the circuit power, \( P_{C_i} (k_i^*) \), of all users. Then we use EMMPA again to determine \( S^{(2)} \), the set of antennas that should be turned off in this round. Similarly, \( S^{(2)} \) also belongs to the set of antennas that should be turned off in the globally optimal antenna configuration. In addition, \( S^{(2)} \) is a super set of \( S^{(1)} \). We iterate this process until a round that all antennas turned on have positive power allocation. This Iterative EMMPA algorithm is summarized in Table II, whose property is given in Proposition 2 based on the above analysis.

### Table II: Iterative EMMPA Algorithm

<table>
<thead>
<tr>
<th>Algorithm Iterative EMMPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Let ( S^{(0)} ) be an empty set.</td>
</tr>
<tr>
<td>2. Assume all antennas are turned on and calculate the circuit power, ( P_{C_i} (k_i^*) ), of all users.</td>
</tr>
<tr>
<td>3. Use EMMPA to determine ( S^{(1)} ) and the corresponding optimal power allocation ( P^* ).</td>
</tr>
<tr>
<td>4. while ( S^{(m)} ) differs from ( S^{(m-1)} ) do</td>
</tr>
<tr>
<td>5. Turn off all antennas in ( S^{(m)} ) and calculate the circuit power, ( P_{C_i} (k_i^*) ), of all users.</td>
</tr>
<tr>
<td>6. ( m \leftarrow m + 1 ).</td>
</tr>
<tr>
<td>7. Use EMMPA to determine ( S^{(m)} ) and the corresponding optimal power allocation ( P^* ).</td>
</tr>
<tr>
<td>8. return ( S^{(m)}, P^* ).</td>
</tr>
</tbody>
</table>

**Proposition 2.** The output of the iterative EMMPA algorithm, \( S^{(m)} \), is a subset of antennas that should be turned off in the globally optimal antenna configuration. \( S^{(m)} \) can be found within at most \( \sum_i k_i - 1 \) rounds.

We can look at the iterative EMMPA algorithm from another perspective. In each round, some additional antennas are turned off to achieve higher energy efficiency. Therefore, \( U^* \) increases in each round. More antennas may be turned off in each round because although in the first several rounds, their channel conditions may not fulfill \( w_i B C_{ik} \leq \sigma^2 \), they fulfill this condition in a later round owing to the increase of \( U^* \). Finally, only the antennas with the best channel conditions are kept on for data transmission.

The iterative EMMPA algorithm finds a subset of antennas that should be turned off in the globally optimal antenna configuration. However some more antennas may still need to be turned off to achieve the energy efficiency capacity. An example has been given in Fig. 1. In this example, both antennas have good channel states and the iterative EMMPA algorithm determines that both antennas should be turned on. However, turning both antennas on may not be globally optimal because one of them may consume too much circuit power. Therefore, the remaining antennas that are kept on by the iterative EMMPA algorithm should be further examined.

In the second step, we determine which remaining antennas should be turned off. Denote \( \bar{S}^{(0)} \) to be the set of the remaining antennas and \( U^{(0)} \) to be the highest energy efficiency achieved when all the remaining antennas are turned on. Denote \( U_{ij}^{(0)} \) to
be the highest energy efficiency achieved when only Antenna \((i, j)\) in \(S^{(0)}\) is turned off. If turning off any antenna in \(S^{(0)}\) will not improve the energy efficiency, i.e.,

\[
\max_{(i, j) \in S^{(0)}} U_{i,j}^{(0)} \leq U^{(0)},
\]

no antennas in \(S^{(0)}\) should be turned off and the selection process is done. Otherwise, we turn off the antenna that results in the highest energy efficiency, i.e., Antenna \((k, l)\), where

\[
(k, l) = \arg\max_{(i, j) \in S^{(0)}} U_{i,j}^{(0)}.
\]

Then a higher energy efficiency \(U_{k,l}^{(0)}\) is achieved. In the second round, let \(U^{(1)} = U_{k,l}^{(0)}\). Denote the set of remaining antennas that are still on to be \(S^{(1)}\). The above selection process can be repeated until in round \(m\), no antennas should be turned off, i.e.,

\[
\max_{(i, j) \in S^{(m)}} U_{i,j}^{(m)} \leq U^{(m)}.
\]

Note that \(m\) is always smaller than the number of antennas in \(S^{(0)}\) because in each round, one antenna will be turned off and finally, at least one antenna in \(S^{(0)}\) must be kept on to achieve non-zero energy efficiency.

The whole algorithm for energy-efficient MU-MIMO with improved circuit management is named improved EMMPA and summarized in Table III. The global optimality of improved EMMPA is verified later in Section IV.

**TABLE III: Improved EMMPA Algorithm**

**Algorithm Improved EMMPA**

1. Use iterative EMMPA to determine \(S^{(0)}\) and \(U^{(0)}\).
2. \(m \leftarrow 0\) and \(U_{\text{max}} \leftarrow 0\).
3. repeat
4. for Antenna \((i, j)\) in \(S^{(m)}\)
5. do Turn on only antennas in \(S^{(m)}\) excluding \((i, j)\) and calculate circuit power of all users.
6. Use EMMPA to determine \(U_{i,j}^{(m)}\) and the corresponding power allocation \(P^*\).
7. if \(U_{i,j}^{(m)} > U_{\text{max}}\)
8. then \(U_{\text{max}} \leftarrow U_{i,j}^{(m)}\) and \((k, l) \leftarrow (i, j)\).
9. \(P^* = P\).
10. if \(U_{\text{max}} > U^{(m)}\)
11. then \(S^{(m+1)} \leftarrow S^{(m)}\) excluding \((k, l)\)
12. \(U^{(m+1)} = U_{\text{max}}\).
13. \(m \leftarrow m + 1\).
14. until \(U_{\text{max}} \leq U^{(m+1)}\).
15. return \(S^{(m)}\) and \(P^*\).

The complexity of the improved EMMPA grows quadratically with the total number of antennas at the user side and it can be easily shown that the convergence rate is characterized by Proposition 3.

**Proposition 3.** The output of the improved EMMPA is obtained within at most

\[
\left\lfloor \log_2 \left( \frac{(\alpha - 1)P^*}{\epsilon} \right) \right\rfloor + \frac{1}{2} \left( \sum_i k_i - 1 \right) \sum_i k_i
\]

iterations.

**IV. SIMULATION RESULTS FOR ENERGY-EFFICIENT MU-MIMO**

The proposed energy-efficient MU-MIMO can be applied in different types of wireless networks to improve the network energy efficiency. In this section, we provide simulation results for a single-cell cellular network to demonstrate the performance of energy-efficient MU-MIMO. System parameters are the same as those in [12]. In each trial, users are dropped uniformly within 250 meters from the AP. The performance below is the average over all trials. Assume \(P_{C_1}(k_i^o) = P_{k_i}^o + P_\beta I(k_i^o)\) mW for all users.

First we verify the global optimality of improved EMMPA. Assume \(P_\beta = 20\) mW and \(P_\alpha\) is varied to observe the impact of circuit power consumption of individual antennas on the suboptimality gap of improved EMMPA and iterative EMMPA algorithms. Fig. 2 gives the normalized energy efficiency of improved and iterative EMMPA when the AP has 16 antennas and each user has 4 antennas. The average energy efficiency achieved by improved and iterative EMMPA is normalized by that of ESPA, which is globally optimal. From Fig. 2, improved EMMPA performs exactly the same as ESPA. In addition to the same average energy efficiency, improved EMMPA always obtains the same instantaneous power allocation and energy efficiency as ESPA according to our observation of simulation results. Therefore, improved EMMPA is also globally optimal. When \(P_\alpha\) is small, iterative EMMPA performs very closely to the global optimum because channel state is the main factor that determines which antennas should be turned on. When \(P_\alpha\) is larger, the impact of individual antenna circuit power consumption grows and since iterative EMMPA does not consider this impact, it has larger performance loss. Therefore improved EMMPA is needed to take this impact into account to achieve the globally optimal performance.

Fig. 3 gives the average energy efficiency capacity when the AP has 16 antennas while each user has two antennas. The number of users is varied from one to eight. \(P_\beta = 20\) mW and \(P_\alpha\) is varied to observe the impact of antenna circuit power on energy efficiency capacity. From Fig. 3 we can see more users always help improve the network energy efficiency because of increased multi-user diversity. Fig. 4 illustrates the average energy efficiency capacity when the AP has 16 antennas and two users are accessing the AP. The number of antennas of each user is varied from one to eight and \(P_\beta = 20\) mW. We can see that when individual antennas consume a small amount of circuit power, i.e., \(P_\alpha\) is small, more antennas always improve the network energy efficiency. However when individual antennas consume too much circuit power and \(P_\alpha\) is large, e.g., \(P_\alpha = 40\) dBm, more antennas do not help improve the network energy efficiency because most antennas should be turned off in this case to reduce circuit power consumption.

**V. CONCLUSION**

In this paper, we investigated the optimal energy-efficient MU-MIMO with dynamic antenna management where users can choose to turn off electronic circuit operations when some antennas are not used. Our result showed that some antennas
should not be used even when they have good channel states because turning them on consumes too much circuit power. We developed globally optimal algorithms to determine the set of antennas that should be kept on and the corresponding power allocation. Comprehensive simulation results were provided to demonstrate the performance and the significant gain in energy efficiency for a single-cell cellular network.

VI. ACKNOWLEDGEMENT

Part of the work was done when the author was with Samsung Telecom America. The author would like to thank C. Zhang, L. Liu, and Y. Nam for helpful discussions.

APPENDIX I

CONDITION OF AN ANTENNA BEING TURNED ON

Proof: \[ \frac{\partial f(p_{ij})}{\partial p_{ij}} = -\frac{\lambda^2_{ij}}{\sigma^2} \frac{\theta}{\zeta} \left( R_{ij}^o + \ln(1 + \frac{p_{ij}\lambda^2_{ij}}{\sigma^2}) \right) < 0. \]

Therefore, \( f(p_{ij}) \) is strictly decreasing. It is easy to see that as \( p_{ij} \to +\infty, f(p_{ij}) \to -\infty. \) If when \( p_{ij} \to 0, f(p_{ij}) > 0, \) then there exists a unique \( p_{ij} > 0 \) such that \( f(p_{ij}) = 0. \) Otherwise, Antenna \((i,j)\) should be turned off.

\[
\lim_{p_{ij} \to 0} f(p_{ij}) = w_i \frac{\lambda^2_{ij}}{\ln 2} \frac{\sigma^2}{\Gamma^2} (P^o_{ij} + P_{Ci}(\lambda^2_{ij})) - \frac{\alpha}{\zeta} R^o_{ij} > 0.
\]

Therefore, Antenna \((i,j)\) should be turned on when \( \frac{\lambda^2_{ij}}{\sigma^2} > \frac{R^o_{ij} \ln 2}{(P^o_{ij} + P_{Ci}(\lambda^2_{ij})) w_i B \zeta} \).

REFERENCES