Energy-Efficient Link Adaptation in Frequency-Selective Channels*

Guowang Miao[†,♭] , *Student Member, IEEE*, Nageen Himayat[††], *Member, IEEE*, and Geoffrey Ye Li[†], *Fellow, IEEE*

[†] School of ECE, Georgia Institute of Technology

[††] Communications Technology Lab., Intel Corporation

***Abstract*—Energy efficiency is becoming increasingly important for small form factor mobile devices, as battery technology has not kept up with the growing requirements stemming from ubiquitous multimedia applications. This paper addresses link adaptive transmission for maximizing energy efficiency, as measured by the "throughput per Joule" metric. In contrast to the existing water-filling power allocation schemes that maximize throughput subject to a fixed overall transmit power constraint, our scheme maximizes energy efficiency by adapting both overall transmit power and its allocation, according to the channel states and the circuit power consumed. We demonstrate the existence of a unique globally optimal link adaptation solution and develop iterative algorithms to obtain it. We further consider the special case of flat-fading channels to develop an upper bound on energy efficiency and to characterize its variation with bandwidth, channel gain and circuit power. Our results for OFDM systems demonstrate improved energy savings with energy optimal link adaptation as well as illustrate the fundamental tradeoff between energy-efficient and spectrum-efficient transmission.**

Index Terms– energy efficiency, link adaptation, frequency selective channel, energy aware, OFDM

I. Introduction

The quality of wireless channel varies with time and frequency. Therefore, link adaptation can be used to improve transmission performance. With link adaptation, modulation order, coding rate, and transmit power can be selected according to *channel state information* (CSI).

Earlier research on link adaptation focuses on power allocation to improve channel capacity subject to overall power constraint. Optimal power allocation for frequency-selective channels has been investigated in [1]. Here the highest data rate on a band-limited channel is achieved when the total received signal power at each frequency, consisting of channel noise and desired signal component, is a constant. The terminology, adaptive modulation, was first used in [2] even though work on adaptive modulation [3] had been reported before.

In addition to throughput improvement, energy efficiency is becoming increasingly important for mobile communications because of the slow progress of battery technology [4] and growing requirements of anytime and anywhere multimedia applications. With sufficient battery power, link adaptation can be geared toward peak performance delivery. However, with limited battery capacity, link adaptation could be adapted toward energy conservation to minimize battery drain. Energy-efficient communication also has the desirable benefit of reducing interference to other co-channel users as well as lessening environmental impacts, e.g. heat dissipation and electronic pollution. Hence, recent research has focused on energy-efficient link adaptation techniques [5]–[7]. It is shown in [5] that when the transmission bandwidth approaches infinity, the minimum received signal energy per bit for reliable communication over additive white Gaussian noise (AWGN) channels, approaches -1.59 dB. For band-limited transmission, the lowest order modulation should be used [6]. However, the investigation in [5], [6] does not account for additional circuit power consumed during transmission. Energy dissipation of both transmitter circuits and radio-frequency output is investigated in [8], where the modulation level is adapted to minimize the energy consumption according to the simulation observations. In [7], these ideas are extended to a detailed analysis of circuit and transmit powers for both adaptive *multiple quadrature amplitude modulation* (M-QAM) and *multiple frequency shift keying* (MFSK) in AWGN channels for short range energy-efficient communications.

Orthogonal frequency division multiplexing (OFDM) has emerged as a primary modulation scheme for the next-generation broadband wireless standards [9], [10]. The power allocation and bit-loading algorithms for OFDM are summarized in Chapter 3 of [11]. While extensive research has been conducted to improve throughput [11], [12], limited work has been done to address the energy-efficient communication for OFDM systems. In this paper, we address the energy-efficient link adaptation for frequency-selective fading channels. We account for both circuit and transmit powers when designing link adaptation schemes and emphasize energy efficiency over peak rates or throughput. The proposed link adaptation balances circuit power consumption and transmission power to achieve the maximum energy efficiency, which is defined as the number of bits transmitted per Joule of energy. In contrast to the existing water-filling power allocation schemes that maximize throughput subject to overall transmit power constraints, our scheme adapts both the overall transmit power and its allocation according to the states of all subchannels and the circuit power consumption to maximize the energy efficiency.

* This work was supported by Intel Corp. and the U.S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-20-0011.

♭ Corresponding author. Email: gmiao3@gatech.edu. Address: School of Electrical and Computer Engineering Georgia Institute of Technology, Atlanta, Georgia, 30332–0250

We demonstrate the existence of a unique globally optimal link adaptation solution and provide iterative algorithms to obtain this optimum. While the usefulness of our technique is illustrated using frequency selective OFDM as an example in this paper, the solution developed is applicable to more general transmission scenarios where transmission occurs over resources experiencing different channel conditions.

The rest of the paper is organized as follows. In Section III, we investigate optimal conditions for energy-efficient transmission and develop algorithms to obtain the globally optimal solution. In Section III-B, we consider a special case when the channel is with flat fading. We also consider energy-efficient link adaptation when the user has either data rate requirement or peak power limit in Section IV. As an example of energy-efficient link adaptation, we apply the energy-efficient scheme in OFDM systems and provide simulation results to demonstrate energy efficiency improvement in Section VI. Finally, we conclude the paper in Section VII.

II. Problem Formulation

In this section, we formulate the problem of energy-efficient link adaptation.

Assume that K subchannels are used for transmission, each with a different channel gain. An example of this scenario is OFDM transmission over frequency-selective channels. Assume block fading [13], [14], that is, the channel state remains constant during each data frame and is independent from one to another. Denote the data rate on Subchannel i as r_i and the data rate vector on all subchannels as

$$\mathbf{R} = [r_1, r_2, \cdots, r_K]^T, \tag{1}$$

where $[]^T$ is the transpose of a vector. The data rate vector, $\mathbf{R}$, depends on the channel state, coding, and power allocation. Correspondingly, the overall data rate is

$$R = \sum_{i=1}^{K} r_i. \tag{2}$$

For a given channel state, the transmit power on each subchannel is determined by the requirement of reliable data transmission. If we denote W as the subchannel bandwidth, N_o the power spectral density, g_i the power gain, and P_{T_i} the allocated transmit power on Subchannel i, the channel output *signal-to-noise ratio* (SNR) will be

$$\eta_i = \frac{P_{T_i} g_i}{N_o W} \tag{3}$$

and the achievable data transmission rate r_i is determined by [15]

$$r_i = W \log(1 + \frac{\eta_i}{\Gamma}), \tag{4}$$

where Γ is the SNR gap that defines the gap between the channel capacity and a practical coding and modulation scheme. The SNR gap depends on the coding and modulation scheme used and on the target probability of error. For a coded *quadrature amplitude modulation* (QAM) system, the gap is given by [15]

$$\Gamma = 9.8 + \gamma_m - \gamma_c \text{ (dB)}, \tag{5}$$

where γ_m is the system design margin and γ_c is the coding gain. For Shannon capacity [16], $\Gamma = 0$ dB. Denote the overall transmit power as $P_T(\mathbf{R})$ and

$$P_T(\mathbf{R}) = \frac{\sum_{i=1}^{K} P_{T_i}}{\zeta} = \sum_{i=1}^{K} (e^{\frac{r_i}{W}} - 1) \frac{N_o W \Gamma}{g_i \zeta}, \tag{6}$$

where $\zeta \in [0, 1]$ is the power amplifier efficiency and depends on the design and implementation of the transmitter. $P_T(\mathbf{R})$ is strictly convex and monotonically increasing in $\mathbf{R}$. In fact, the developed theory and approaches can be used for any $P_T(\mathbf{R})$ that is strictly convex and monotonically increasing in $\mathbf{R}$ with $P_T(\mathbf{0}) = 0$, where $\mathbf{0} = [0, 0, \cdots, 0]^T$.

In addition to transmit power, mobile devices also incur additional circuit power during transmissions which is relatively independent of the transmission rate [8], [17]. While the transmit power models all the power used for reliable data transmission, we let the circuit power represents the average energy consumption of device electronics, such as mixers, filters, and digital-to-analog converters, and this portion of energy consumption excludes that of the power amplifier and is independent of the transmission state. If we denote the circuit power as P_C, the overall power consumption given a data rate vector will be

$$P(\mathbf{R}) = P_C + P_T(\mathbf{R}). \tag{7}$$

For energy-efficient communications, it is desirable to maximize the amount of data sent with a given amount of energy. Hence, given any amount of energy $\triangle e$ consumed in a duration, $\triangle t$, i.e. $\triangle e = \triangle t(P_C + P_T(\mathbf{R}))$, the mobile wants to send a maximum amount of data by choosing the data rate vector to maximize

$$\frac{R \triangle t}{\triangle e}, \tag{8}$$

which is equivalent to maximizing

$$U(\mathbf{R}) = \frac{R}{\triangle e / \triangle t} = \frac{R}{P_C + P_T(\mathbf{R})}. \tag{9}$$

$U(\mathbf{R})$ is called energy efficiency. The unit of the energy efficiency is bits per Joule, which has been frequently used in literature for energy-efficient communications [5], [6], [18]–[20]. The optimal energy-efficient link adaptation achieves maximum energy efficiency, i.e.

$$\mathbf{R}^* = \arg\max_{\mathbf{R}} U(\mathbf{R}) = \arg\max_{\mathbf{R}} \frac{R}{P_C + P_T(\mathbf{R})}. \tag{10}$$

Note that if we fix the overall transmit power, the objective of Equation (10) is equivalent to maximizing the overall throughput and the existing water-filling power allocation approach [1] gives the solution. However, besides adapting the power distributions on all subchannels, the overall transmit power can also be adapted according to the states of all subchannels to maximize the energy efficiency. Hence, the solution to Equation (10) is in general different from existing power allocation schemes that maximize throughput with power constraints.

III. Principles of Energy-Efficient Link Adaptation

In the following, we demonstrate that a unique globally optimal data rate vector always exists and give the necessary and sufficient conditions for a data rate vector to be globally optimal.

A. Conditions of Optimality

The concept of quasiconcavity will be used in our discussion and is defined as [21].

Definition 1. *A function f, which maps from a convex set of real n-dimensional vectors, $\mathcal{D}$, to a real number, is called strictly quasiconcave if for any $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{D}$ and $\mathbf{x}_1 \neq \mathbf{x}_2$,*

$$f(\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) > \min\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}, \tag{11}$$

for any $0 < \lambda < 1$.

Any strictly monotonic function is quasiconcave. Besides, any strictly concave function is also strictly quasiconcave but the reverse is not generally true. An example is the Gaussian function, which is strictly quasiconcave but not concave.

It is proved in Appendix I that $U(\mathbf{R})$ has the following properties.

Lemma 1. *If $P_T(\mathbf{R})$ is strictly convex in $\mathbf{R}$, $U(\mathbf{R})$ is strictly quasiconcave. Furthermore, $U(\mathbf{R})$ is either strictly decreasing or first strictly increasing and then strictly decreasing in any r_i of $\mathbf{R}$, i.e. the local maximum of $U(\mathbf{R})$ for each r_i exists at either 0 or a positive finite value.*

For strictly quasiconcave functions, if a local maximum exists, it is also globally optimal [21]. Hence, a unique globally optimal transmission rate vector always exists and its characteristics are summarized in Theorem 1 according to the proofs in Appendix I.

Theorem 1. *If $P_T(\mathbf{R})$ is strictly convex, there exists a unique globally optimal transmission data rate vector $\mathbf{R}^* = [r_1^*, r_2^*, \cdots, r_K^*]^T$ for (10), where r_i^* is given by*

(i) when $\frac{P_C + P_T(\mathbf{R}_i^{(0)})}{R_i^{(0)}} \geq \left.\frac{\partial P_T(\mathbf{R})}{\partial r_i}\right|_{\mathbf{R}=\mathbf{R}_i^{(0)}}$, $\left.\frac{\partial U(\mathbf{R})}{\partial r_i}\right|_{\mathbf{R}=\mathbf{R}^*} = 0$, i.e. $\frac{1}{\frac{\partial P_T(\mathbf{R}^*)}{\partial r_i^*}} = \frac{R^*}{P_C + P_T(\mathbf{R}^*)} = U(\mathbf{R}^*)$;

(ii) when $\frac{P_C + P_T(\mathbf{R}_i^{(0)})}{R_i^{(0)}} < \left.\frac{\partial P_T(\mathbf{R})}{\partial r_i}\right|_{\mathbf{R}=\mathbf{R}_i^{(0)}}$, $r_i^* = 0$,

where $\mathbf{R}_i^{(0)} = [r_1^*, r_2^*, \cdots, r_{i-1}^*, 0, r_{i+1}^*, \cdots, r_K^*]$ and $R_i^{(0)} = \sum_{j \neq i} r_j^*$, i.e. the overall data rate on all other subchannels except i.

Theorem 1 has clear physical insights. $P_C + P_T(\mathbf{R}_i^{(0)})$ is the power consumption of both circuit and all other subchannels when Subchannel i is not used. $\frac{P_C + P_T(\mathbf{R}_i^{(0)})}{R_i^{(0)}}$ is the per-bit energy consumption when Subchannel i is not used and the overall per-bit energy consumption needs to be minimized for energy-efficient communications. $\left.\frac{\partial P_T(\mathbf{R})}{\partial r_i}\right|_{\mathbf{R}=\mathbf{R}_i^{(0)}}$ is the per-bit energy consumption transmitting infinitely small data rate on Subchannel i conditioned on the optimal status of all other subchannels. Hence, Subchannel i should not transmit anything when $\frac{P_C + P_T(\mathbf{R}_i^{(0)})}{R_i^{(0)}} < \left.\frac{\partial P_T(\mathbf{R})}{\partial r_i}\right|_{\mathbf{R}=\mathbf{R}_i^{(0)}}$. Otherwise, there should be a tradeoff between the desired data rate on Subchannel i and the incurred power consumption. The tradeoff closely depends on the power consumption of both circuits and transmission on all other subchannels and can be found through the unique zero derivative of $U(\mathbf{R})$ with respect to r_i.

To further understand Theorem 1, we consider an example when each subchannel achieves the Shannon capacity and the transmit power on each subchannel is given in (6) with $\Gamma = 0$ dB and $\zeta = 1$. The overall transmit power is

$$P_T(\mathbf{R}) = \sum_{k=1}^{K} (e^{\frac{r_k}{W}} - 1)\frac{N_o W}{g_k}. \tag{12}$$

According to Condition (i) of Theorem 1, when $r_k > 0$, we have

$$\frac{1}{\frac{\partial P_T(\mathbf{R})}{\partial r_k}} = \frac{1}{e^{\frac{r_k}{W}} \frac{N_o}{g_k}} = U(\mathbf{R}^*). \tag{13}$$

Hence, the transmit power on Subchannel k is

$$P_{T_n} = (e^{\frac{r_k}{W}} - 1)\frac{N_o W}{g_k} = \frac{W}{U(\mathbf{R}^*)} - \frac{N_o W}{g_k}, \tag{14}$$

which is a water-filling to level $\frac{W}{U(\mathbf{R}^*)}$. Since the water level is determined by the optimal energy efficiency, we refer to our scheme as *dynamic energy-efficient water-filling*. Note that while the absolute value of power allocation is determined by the maximum energy efficiency $U(\mathbf{R}^*)$, which relies on both the circuit power and channel state, the relative differences of power allocations on different subchannels depend only on the channel gains on those subchannels.

B. A Special Case: When the Channel is Flat Fading

To facilitate the understanding of the fundamental dependence of energy efficiency on the channel gain, circuit power, and bandwidth, we consider a special case that the channel is experiencing flat fading in this section. Hence, all subchannels are with the same channel gain and the same link adaptation is applied on all subchannels. The overall data rate is

$$R = Kr. \tag{15}$$

According to Theorem 1, the optimal transmission data rate follows immediately and is summarized by Theorem 2, where the upper bound is proved in Appendix II.

Theorem 2. *If $P_T(R)$ is monotonically increasing and strictly convex in R, there exists a unique globally optimal transmission data rate to maximize energy efficiency and is given by*

$$R^* = \frac{P_C + P_T(R^*)}{P_T'(R^*)}, \tag{16}$$

where $P_T'(\cdot)$ is the first order derivative of function $P_T(\cdot)$. Besides, energy efficiency is upper bounded by $\frac{1}{P_T'(0)}$.

When Shannon capacity is achieved in AWGN channels, the upper bound is $\frac{g}{N_o}$.

In the following, we investigate some basic properties of energy-efficient link adaptation. Propositions 1, 2, and 3

summarize the impact of channel gain, circuit power, and the number of subchannels on the optimal energy-efficient transmission, and are proved in Appendix III.

Proposition 1. *Both the data rate and energy efficiency increase with channel gain.*

Proposition 2. *The data rate increases with circuit power, while the energy efficiency decreases with it. With zero circuit power, the highest energy efficiency,* $\frac{1}{P_T'(0)}$*, is obtained by transmitting with infinite small data rate.*

From Proposition 2, when circuit power dominates power consumption, which is usually true with short-range communication, the highest data rate should be used to finish transmission as soon as possible, which has been commonly assumed by most MAC layer energy-efficient optimization schemes as describe in the introduction of this paper. However, when the circuit power is negligible, which is usually true with long-range communication like satellite communications, the lowest data rate should be used, which coincides with the results in [6] and [22].

Proposition 3. *The data rate on each subchannel decreases with increasing number of subchannels while the energy efficiency increases with it. With infinite number of subchannels, the highest energy efficiency,* $\frac{1}{P_T'(0)}$*, is obtained by transmitting with infinite small data rate.*

Propositions 1, 2, and 3 discover three ways to improve energy efficiency: increasing channel power gain, reducing circuit power, and allocating more subchannels. The energy-efficiency upper bound is achieved by transmitting with infinite small data rate when either circuit power is zero or infinite number of subchannels is assigned.

IV. Constrained Energy-Efficient Link Adaptation

In this section, we study energy-efficient link adaptation when user has either a data rate requirement or a peak power limit.

With a data rate requirement Γ, the energy-efficient link adaptation is given by

$$\widehat{\mathbf{R}}^* = \arg\max_{\mathbf{R}} \frac{R}{P_C + P_T(\mathbf{R})}, \tag{17a}$$

subject to

$$R \geq \Gamma. \tag{17b}$$

If the optimal data rate vector without constraint in (10) satisfies $R^* \geq \Gamma$, it is also the solution to Problem (17), i.e. $\widehat{\mathbf{R}}^* = \mathbf{R}^*$. Otherwise, Problem (17) is equivalent to

$$\widehat{\mathbf{R}}^* = \arg\max_{\mathbf{\Gamma}} \frac{\Gamma}{P_C + P_T(\mathbf{R})} = \arg\min_{\mathbf{R}} P_T(\mathbf{R}), \tag{18a}$$

subject to

$$R = \Gamma. \tag{18b}$$

Since $P_T(\mathbf{R})$ is strictly convex, a unique globally optimal $\widehat{\mathbf{R}}^*$ exists. Denote

$$f_k(r_k) = \frac{\partial P_T(\mathbf{R})}{\partial r_k} \tag{19}$$

and its inverse function to be $f_k^{-1}()$. Then $\widehat{\mathbf{R}}^*$ can be easily obtained via the Lagrangian technique [23] and is

$$\widehat{r}_k^* = \max\left\{f_k^{-1}(\lambda), 0\right\} \tag{20}$$

for $k = 1, \cdots, K$, where λ is determined by

$$\sum_{k=1}^{K} \widehat{r}_k^* = \Gamma. \tag{21}$$

When the channel capacity is achieved on each subchannel, the corresponding optimal power allocation is a water-filling allocation, which achieves the sum channel capacity Γ.

Similarly, with a maximum transmit power constraint, the problem is to find

$$\widetilde{\mathbf{R}}^* = \arg\max_{\mathbf{R}} \frac{R}{P_C + P_T(\mathbf{R})}, \tag{22a}$$

subject to

$$P_T(\mathbf{R}) \leq P_m. \tag{22b}$$

If the optimal data rate vector without constraint in (10) satisfies $P_T(\mathbf{R}^*) \leq P_m$, it is also the solution to Problem (22), i.e. $\widetilde{\mathbf{R}}^* = \mathbf{R}^*$. Otherwise, via the the Lagrangian technique again, we have the unique optimal solution as follows

$$\widetilde{r}_k^* = \max\left\{f_k^{-1}(\lambda), 0\right\}, k = 1, \cdots, K, \tag{23}$$

where λ is determined by

$$P_T(\widetilde{\mathbf{R}}^*) = P_m. \tag{24}$$

When channel capacity is achieved on each subchannel, the power allocation is the classical water-filling where the water level is determined by P_m [1].

V. Algorithm Design

Theorem 1 provides the necessary and sufficient conditions for a rate vector to be the unique and globally optimum one. However, it is usually difficult to directly solve the joint nonlinear equations according to Theorem 1 to obtain the optimal vector $\mathbf{R}^*$. Therefore, we develop iterative methods to search the optimal $\mathbf{R}$ for maximizing $U(\mathbf{R})$. The global optimality of the proposed methods is guaranteed by the strict quasiconcavity of $U(\mathbf{R})$. In the following, we describe our low-complexity iterative algorithms.

A. Gradient Assisted Binary Search

When there is only one subchannel, Lemma 1 shows that function $U(\mathrm{r})$ has a unique r^* such that for any $r < r^*$, $\frac{dU(\mathrm{r})}{dr} > 0$, and for any $r > r^*$, $\frac{dU(\mathrm{r})}{dr} < 0$. Hence, we have the following lemma to seek two points r_1 and r_2 such that $r_1 \leq r^* \leq r_2$.

Proposition 4. *Let the initial setting* $r^{[0]} > 0$ *and set* $\alpha > 1$*. For any* $i \geq 0$*, let*

$$r^{[i+1]} = \begin{cases} \frac{r^{[i]}}{\alpha} & \left.\frac{dU(r)}{dr}\right|_{r^{[0]}} < 0 \\ \alpha r^{[i]} & otherwise \end{cases}. \tag{25}$$

Repeat (25) until $r^{[I]}$ such that $\frac{dU(r)}{dr}\Big|_{r^{[I]}}$ has a different sign from $\frac{dU(r)}{dr}\Big|_{r^{[0]}}$. Then r^* must be between $r^{[I]}$ and $r^{[I-1]}$.

To locate r^* between r_1 and r_2, let $\widehat{r} = \frac{r_1+r_2}{2}$. If $\frac{dU(r)}{dr}\Big|_{\widehat{r}} = 0$, r^* is found. If $\frac{dU(r)}{dr}\Big|_{\widehat{r}} < 0$, $r_1 < r^* < \widehat{r}$ and replace r_2 with $\widehat{r}$; otherwise, replace r_1 with $\widehat{r}$. This leads to the *gradient assisted binary search* (GABS) for maximizing $U(r)$, which is summarized in Table I.

Algorithm *GABS*(r_o)
($*$ algorithm for single-subchannel transmission. $*$)
Input: initial guess: $r_o > 0$
Output: optimal transmission rate: r^*
$r_1 = r_o$, $h_1 \leftarrow \frac{dU(r)}{dr}\Big|_{r_1}$, initialize $\alpha > 1$ (e.g.10)
if $h_1 < 0$
($*$ seek r_1 and r_2 such that $r_1 < r^* < r_2$ $*$)
then $r_2 \leftarrow r_1$, $r_1 \leftarrow \frac{r_1}{\alpha}$, and $h_1 \leftarrow \frac{dU(r)}{dr}\Big|_{r_1}$
while $h_1 < 0$
do $r_2 \leftarrow r_1$, $r_1 \leftarrow \frac{r_1}{\alpha}$, and $h_1 \leftarrow \frac{dU(r)}{dr}\Big|_{r_1}$
else $r_2 \leftarrow r_1 * \alpha$ and $h_2 \leftarrow \frac{dU(r)}{dr}\Big|_{r_2}$
while $h_2 > 0$
do $r_1 \leftarrow r_2$, $r_2 \leftarrow r_2 * \alpha$, and $h_2 \leftarrow \frac{dU(r)}{dr}\Big|_{r_2}$
while no convergence
($*$ seek r^* between r_1 and r_2 $*$)
do $\widehat{r} \leftarrow \frac{r_2+r_1}{2}$; $\widehat{h} \leftarrow \frac{dU(r)}{dr}\Big|_{\widehat{r}}$
if $\widehat{h} > 0$
then $r_1 = \widehat{r}$;
else $r_2 = \widehat{r}$
return $\widehat{r}$

TABLE I: Gradient assisted binary search

B. Binary Search Assisted Ascent

To find the optimal data rate vector for the multiple subchannel case, we design a gradient ascent method to produce a maximizing sequence $\mathbf{R}^{[i]}$, $n = 0, 1, \cdots$, and

$$\mathbf{R}^{[i+1]} = \left[\mathbf{R}^{[i]} + \mu \nabla U(\mathbf{R}^{[i]})\right]^+, \tag{26}$$

where $[\mathbf{R}]^+$ sets the negative part of the vector $\mathbf{R}$ to be zero, $\mu > 0$ is the search step size, and $\nabla U(\mathbf{R}^{[i]})$ is the gradient at iteration i. With sufficiently small step size, $U(\mathbf{R}^{[i+1]})$ will be always bigger than $U(\mathbf{R}^{[i]})$ except when $\nabla U(\mathbf{R}^{[i]}) = 0$ that indicates the optimality of $\mathbf{R}^{[i]}$ [23]. However, small step size leads to slow convergence. Besides, each element of the gradient depends on the corresponding subchannel power gain, which potentially differs from each other by orders of magnitude. Hence, a line search of the optimal step size needs to cover a large range to assure global convergence on all subchannels, which is computationally expensive. Therefore, at each $\mathbf{R}^{[i]}$, an efficient algorithm is needed to find the optimal step size. Denote

$$f_i(\mu) = U(\left[\mathbf{R}^{[i]} + \mu \nabla U(\mathbf{R}^{[i]})\right]^+). \tag{27}$$

Similar to the proof of Lemma 1, it is easy to show that $g_i(\mu)$ is also strictly quasiconcave in μ and has a unique globally maximum μ^* such that for any $\mu < \mu^*$, $\frac{df_i(\mu)}{d\mu} > 0$, and for any $\mu > \mu^*$, $\frac{df_i(\mu)}{d\mu} < 0$. Let $\nabla U(\mathbf{R}^{[i]}) = [\widehat{g}_1, \widehat{g}_2, \cdots, \widehat{g}_K]$. Replace $\frac{dU(r)}{dr}$ in GABS to be

$$\frac{df_i(\mu)}{d\mu} = [\nabla U(\mathbf{R}^{[i+1]})]^T \widetilde{\mathbf{G}}[i], \tag{28}$$

where $\widetilde{\mathbf{G}}[i] = \frac{d\left[\mathbf{R}^{[i]} + \mu \nabla U(\mathbf{R}^{[i]})\right]^+}{d\mu} = [\widetilde{g}_1, \widetilde{g}_2, \cdots, \widetilde{g}_K]$, in which $\widetilde{g}_k = \widehat{g}_k$ if the kth component of $\mathbf{R}^{[i]} + \mu \nabla U(\mathbf{R}^{[i]})$ is positive and $\widetilde{g}_k = 0$ otherwise. Then GABS can be used for quick location of the optimal step size. This leads to the *binary search assisted ascent* (BSAA) algorithm in Table II.

Algorithm *BSAA*($\mathbf{R}_o$)
($*$ algorithm for multi-subchannel transmission. $*$)
Input: initial guess: $\mathbf{R}_o$(default transmission rate can be used)
Output: optimal transmission rate vector: $\mathbf{R}^*$
$\mathbf{R} = \mathbf{R}_o$,
while no convergence
do use GABS to find the optimal step size μ^*;
$\mathbf{R} = [\mathbf{R} + \mu^* \nabla U(\mathbf{R})]^+$
return $\mathbf{R}$

TABLE II: Binary search assisted ascent

C. The Rate of Convergence

While the global convergence of both GABS and BSAA is guaranteed by the strict quasiconcavity of $U(\mathbf{R})$ [24], we further study the convergence rate in this section.

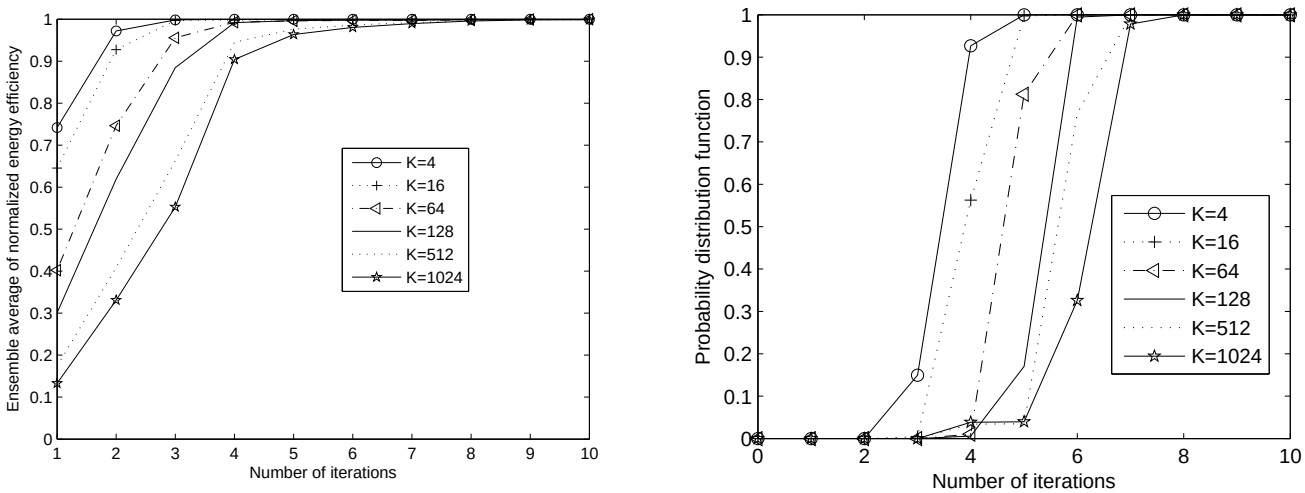
Theorem 3 characterizes the convergence of GABS and is proved in Appendix IV.

Theorem 3. *GABS converges to the globally optimal transmission data rate r^*. A rate r, which satisfies $|r - r^*| \leq \epsilon$, can be found within at most M iterations, where M is the minimum integer such that $M \geq \log_2(\frac{(\alpha-1)r^*}{\epsilon} - 1)$.*

It is difficult to theoretically analyze the global convergence rate of BSAA because of the nonconcavity of $U(\mathbf{R})$. Instead, we run numerical simulations and observe the convergence. Figure 1(a) illustrates the improvement of energy efficiency with iterations. Here we assume the channel gain of each subchannel has Rayleigh distribution with a unit average. The circuit power is 5. The noise power on each subchannel is 0.01. The transmit power is given by Equation (6) with $\Gamma = 0$ dB. The energy efficiency is normalized by the optimal value and the curves are the ensemble averages of 5000 channel instances. Figure 1(b) shows the corresponding probability distribution functions of the numbers of iterations necessary for convergence. In both figures, we vary the number of subchannels to verify its impact on the convergence rate. We can see that BSAA converges very fast to the global optimum, even with 1024 subchannels.

VI. Simulation Results for OFDM

The proposed energy-efficient link adaptation is general and can be applied to different kinds of OFDM, MIMO, and



(a) Converging process: relationship between ensemble average and iterations (b) Probability distribution function of the number of iterations for convergence

Fig. 1: Convergence rate of BSAA.

MIMO-OFDM systems. To apply it, we only need to find the transmit power relationship $P_T(\mathbf{R})$ of those systems. In this section, we discuss the optimal energy-efficient link adaptation for OFDM with subchannelization as an example.

A. Modeling of OFDM with Subchannelization

In OFDM systems with subchannelization, subcarriers are grouped into subchannels and the subcarriers forming one subchannel may, but not necessarily be adjacent, such as the contiguous and distributed subchannelization schemes in 802.16e [9]. Each subchannel is treated to be flat fading and the *effective channel power gain*, $\overline{g}$, rather than physical channel power gain of each subcarrier, is used as a metric. For simplicity, $\overline{g}$ is the average of channel power gains of all subcarriers within the subchannel. Note that classical OFDM is a special case when each subchannel has one subcarrier. The frame structure is shown in Figure 2. Each transmission slot consists of a data interval, T_s, and a signalling interval, τ. In each data interval, l symbols are transmitted. We use uncoded M-QAM. The transmit power on each subchannel needs to be determined.

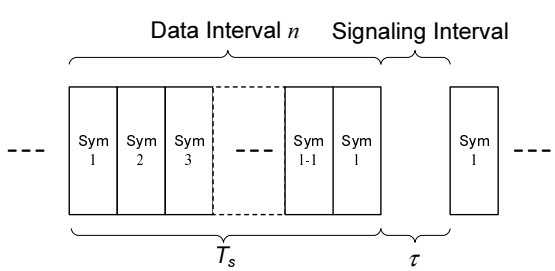


Fig. 2: Frame structure

Consider Subchannel i that consists of c_i subcarriers. The number of bits transmitted per symbols on each subcarrier is $b_i = \frac{\frac{r_i}{c}(T_s+\tau)}{l}$. Hence, the modulation order M is given by $M_i = 2^{b_i} = 2^{B_i r_i}$, where $B_i = \frac{(T_s+\tau)}{c_i l}$. The *bit-error rate* (BER) for coherently detected M-QAM with Gray mapping over an AWGN channel is approximated by [25]

$$P_e(\gamma) \approx 0.2 \exp\left(-\frac{1.5\gamma}{M-1}\right), \tag{29}$$

where γ is the *signal-to-noise ratio* (SNR). For a BER target, P_e, the required SNR on Subchannel i is

$$\gamma_i = \frac{2}{3}(1-M_i)\ln(5P_e) = \frac{2}{3}(1-2^{B_i r_i})\ln(5P_e). \tag{30}$$

Hence, the overall transmit power on Subchannel i is

$$P_{T_i}(r_i) = \frac{\gamma_i c_i N_o W}{\overline{g}_i} = A_i(1-2^{B_i r_i}), \tag{31}$$

where W is the signal bandwidth of each subcarrier and

$$A_i = \frac{2c_i \ln(5P_e) N_o W}{3\overline{g}_i}. \tag{32}$$

Assuming no coupling between transmit powers among subchannels, the overall transmit power will be the cumulative of the transmit powers of all subchannels, that is,

$$P_T(\mathbf{R}) = \sum_{i=1}^{K} P_{Ti}(r_i), \tag{33}$$

which is monotonically increasing and strictly convex in $\mathbf{R}$. The energy-efficient link adaptation immediately follows from Section III.

B. Performance Comparison

In this section, we compare the performance of energy-efficient OFDM transmission with that of traditional transmission schemes. The system parameters are listed in Table III. The *International Telecommunication Union* (ITU) pedestrian channel model B [26] is used to implement the multipath frequency-selective fading. We implement two subchannelization schemes as in Figure 3, fixed-interval and contiguous, both of which group 10 subcarriers into a subchannel. In the

Carrier frequency	1.5 GHz
Subcarrier number	256
Subcarrier bandwidth	10 kHz
BER requirement	10^{-3}
Symbol number of data interval, l	100
Time duration of data interval, T_s	0.01s
Time duration of signalling interval, τ	0.001s
Thermal noise power, N_o	-141 dBW/MHz
User antenna height	1.6 m
BS antenna height	40 m
Environment	Macro cell in urban area
Circuit power, P_C	100 mW
Modulation	Uncoded M-QAM
Subchannelization	Fixed-interval and contiguous
Propagation Model	Okumura-Hata model
Shadowing	Log-normal with standard deviation of 10 dB
Frequency-selective fading	ITU pedestrian channel B
User speed	3 km/h

TABLE III: System parameters

fixed-interval subchannelization, one draws subcarriers out of all subcarriers with a fixed interval to form a subchannel, while in the contiguous one, each subchannel consists of a block of contiguous subcarriers.

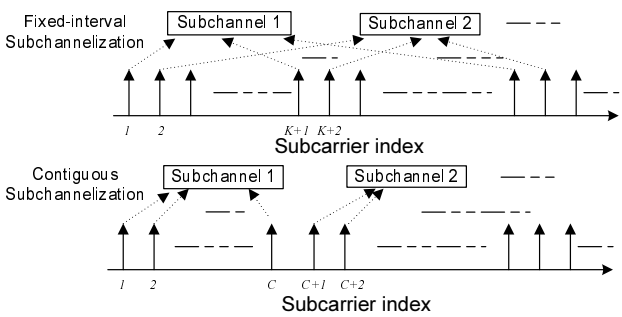
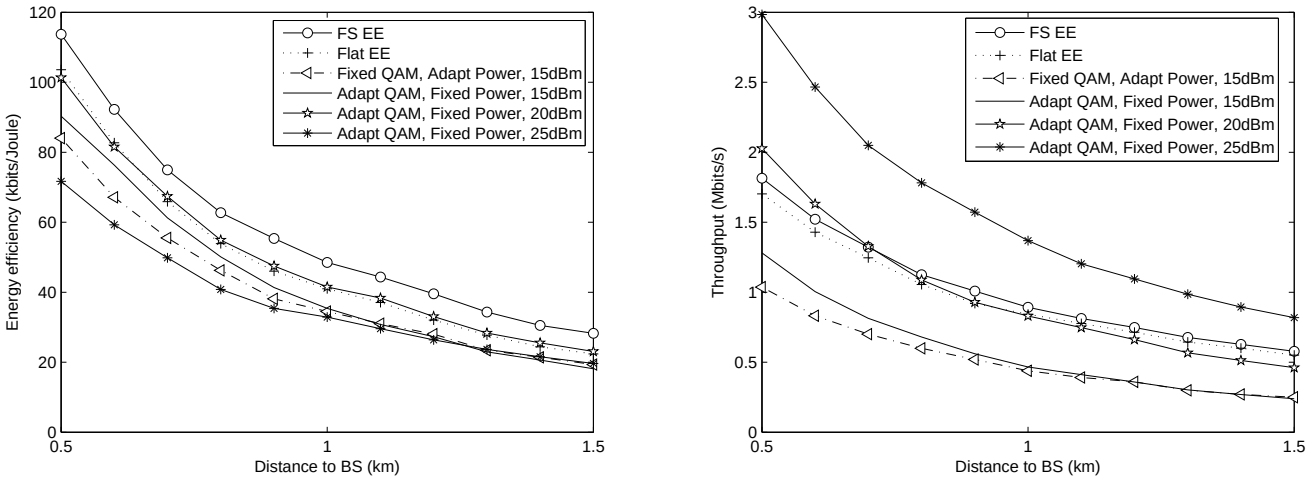


Fig. 3: OFDM subchannelization (K subchannels, each with c subcarriers)

Figures 4(a) and 4(b) and Figures 5(a) and 5(b) compare energy efficiency and throughput of different transmission schemes with contiguous subchannelization and with fixed-interval subchannelization respectively. Two energy-efficient OFDM transmission schemes are implemented: FS EE, that is the optimal energy-efficient transmission developed in this paper, and flat EE, that treats the channel as flat fading. Transmissions with both fixed and adaptive QAM modulations are also implemented for comparison. For fixed modulation, the transmit power is adapted to meet BER requirement while not exceeding 15 dBm maximum power constraint. For adaptive modulation, transmit power is equally distributed over all subchannels and the modulation is adapted to meet BER requirement. From Figures 4(a) and 4(b), fixed and adaptive modulations perform closely to each other, especially when far away from BS, for both energy efficiency and throughput, when the maximum transmit power is 15 dBm. By increasing the transmit power from 15 dBm to 25 dBm, the throughput of adaptive modulation increases, however, the energy efficiency first increased and then decreases. Due to the global optimality, the proposed energy-efficient transmission for frequency-selective channels always achieves the highest energy efficiency, and outperforms the others by at least 15%. However, the throughput is not necessarily maximum; the other schemes, especially the adaptive QAM modulation with 25 dBm transmit power, sacrifice power to obtain higher throughput. Similar results can also be observed in Figures 5(a) and 5(b). Furthermore, we note that when fixed-interval subchannelization is used, different subchannels have trivial differences in average channel gain and the energy-efficient transmission treating the channel to be flat fading performs the same as the one considering the difference of different subchannels. This indicates energy-efficient link adaptation treating channels to be flat fading is sufficient for performance optimization.

VII. Conclusion

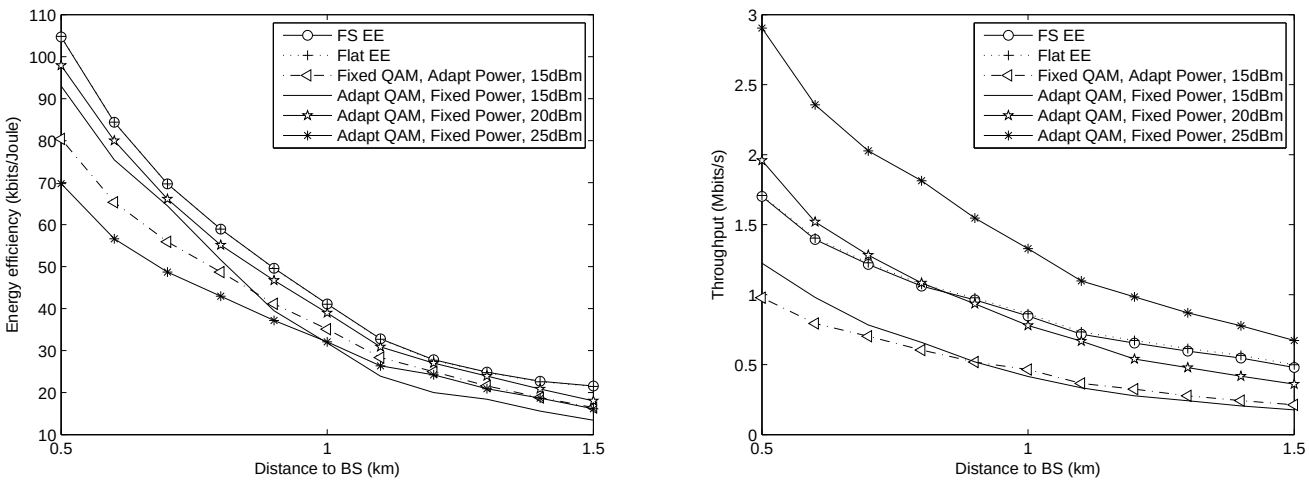
In this paper, we have investigated the energy-efficient link adaptation. While the usefulness of the proposed technique is illustrated using frequency-selective OFDM as an example, the solution developed is applicable to more general transmission scenarios where transmission occurs over resources experiencing different channel conditions. Joint circuit and transmit power consumptions are taken into account to maximize energy efficiency rather than throughput. We demonstrate the existence of a unique globally optimal link adaptation solution and provide iterative algorithms to obtain this optimum. The optimal power allocation is shown to be a dynamic water-filling where the water level is determined by the maximum energy efficiency. We further consider a special case when the channel is experiencing flat fading and show the upper bound of energy efficiency as well as two ways to achieve this bound. We explicitly demonstrate that energy efficiency



(a) Energy efficiency

(b) Throughput

Fig. 4: Performance comparison for contiguous subchannelization.



(a) Energy efficiency

(b) Throughput

Fig. 5: Performance comparison for fixed-interval subchannelization.

is improved by increasing channel power gain, bandwidth, and by reducing circuit power consumption. From the simulation results, we observed at least 15% improvement in energy utilization when frequency selectivity is exploited and the improvement depends on how much frequency diversity exists within the channels.

Appendix I
Proof of Lemma 1

Proof: Denote the upper contour sets of $U(\mathbf{R})$ as

$$S_\alpha = \{\mathbf{R} \succeq \mathbf{0} | U(\mathbf{R}) \geq \alpha\}, \tag{I.34}$$

where symbol $\succeq$ denotes vector inequality and $\mathbf{R} \succeq \mathbf{0}$ means each element of $\mathbf{R}$ is nonnegative. According to Proposition C.9 of [21], $U(\mathbf{R})$ is strictly quasiconcave if and only if S_α is strictly convex for any real number α. When $\alpha < 0$, no points exist on the contour $U(\mathbf{R}) = \alpha$. When $\alpha = 0$, only $\mathbf{0}$ is on the contour $U(\mathbf{0}) = \alpha$. Hence, S_α is strictly convex when $\alpha \leq 0$. Now we investigate the case when $\alpha > 0$. S_α is equivalent to $S_\alpha = \{\mathbf{R} \succeq \mathbf{0} | \alpha P_C + \alpha P_T(\mathbf{R}) - R \leq 0\}$. Since $P_T(\mathbf{R})$ is strictly convex in $\mathbf{R}$, S_α is also strictly convex. Hence, we have the strict quasiconcavity of $U(\mathbf{R})$.

The partial derivative of $U(\mathbf{R})$ with r_i is

$$\frac{\partial U(\mathbf{R})}{\partial r_i} = \frac{P_C + P_T(\mathbf{R}) - R P_T'(\mathbf{R})}{(P_C + P_T(\mathbf{R}))^2} \triangleq \frac{\beta(r_i)}{(P_C + P_T(\mathbf{R}))^2}, \tag{I.35}$$

where $P_T'(\mathbf{R})$ is the first partial derivative of $P_T(\mathbf{R})$ with respect to r_i. According to Lemma 1, if r_i^* exists such that $\frac{\partial U(\mathbf{R})}{\partial r_i}\Big|_{r_i = r_i^*} = 0$, it is unique, i.e. if there is a r_i^* such that $\beta(r_i^*) = 0$, it is unique. In the following, we investigate the conditions when r_i^* exists.

The derivative of $\beta(r_i)$ is

$$\beta^{'}(r_i) = -RP_T^{''}(\mathbf{R}) < 0, \tag{I.36}$$

where $P_T^{''}(\mathbf{R})$ is the second partial derivative of $P_T(\mathbf{R})$ with respect to r_i. Hence, $\beta(r_i)$ is strictly decreasing. According to the L'Hopital's rule, it is easy to show that

$$\begin{aligned}\lim_{r_i->\infty}\beta(r_i) &= \lim_{r_i->\infty}(P_C + P_T(\mathbf{R}) - RP_T^{'}(\mathbf{R})) \\ &= \lim_{r_i->\infty}\left(\frac{P_C + P_T(\mathbf{R}) - RP_T^{'}(\mathbf{R})}{r_i}r_i\right) \\ &= \lim_{r_i->\infty}\left(\frac{P_T^{'}(\mathbf{R}) - P_T^{'}(\mathbf{R}) - RP_T^{''}(\mathbf{R})}{1}r_i\right) \\ &= \lim_{r_i->\infty} -P_T^{''}(\mathbf{R})Rr_i < 0.\end{aligned} \tag{I.37}$$

Besides,

$$\begin{aligned}\lim_{r_i->0}\beta(r_i) &= \lim_{r_i->0}(P_C + P_T(\mathbf{R}) - RP_T^{'}(\mathbf{R})) \\ &= P_C + P_T(\mathbf{R}_i^{(0)}) - R_i^{(0)}P_T^{'}(\mathbf{R}_i^{(0)}),\end{aligned} \tag{I.38}$$

where $\mathbf{R}_i^{(0)} = [r_1, r_2, \cdots, r_{i-1}, 0, r_{i+1}, \cdots, r_K]^T$ and $R_i^{(0)} = \sum_{j\neq i} r_j$.

(1°)When $P_C + P_T(\mathbf{R}_i^{(0)}) - R_i^{(0)}P_T^{'}(\mathbf{R}_i^{(0)}) \geq 0$, $\lim_{r_i->0}\beta(r_i) \geq 0$. Together with (I.37), we see that t_i^* exists and $U(\mathbf{R})$ is first strictly increasing and then strictly decreasing in r_i.

(2°)When $P_C + P_T(\mathbf{R}_i^{(0)}) - R_i^{(0)}P_T^{'}(\mathbf{R}_i^{(0)}) < 0$, $\lim_{r_i->0}\beta(r_i) < 0$. Together with (I.36) and (I.37), t_i^* does not exist. However, $U(\mathbf{R})$ is always strictly decreasing in r_i. Hence, $U(\mathbf{R})$ is maximized at $r_i = 0$.

Lemma 1 is readily obtained. ■

Appendix II
Proof of the Upperbound in Theorem 2

Proof: $U(R) = \frac{R}{P_C+P_T(R)} \leq \frac{R}{P_T(R)}$. Denote $\widehat{U}(R) = \frac{R}{P_T(R)}$. $\widehat{U}^{'}(R) = \frac{d\widehat{U}(R)}{dR} = \frac{P_T(R)-RP_T^{'}(R)}{P_T^2(R)}$. According to the L'Hopital's rule, $\lim_{R\to 0}\widehat{U}^{'}(R) = \lim_{R\to 0}\frac{P_T(R)-RP_T^{'}(R)}{P_T^2(R)} = \lim_{R\to 0}\frac{-RP_T^{''}(R)}{2P_T(R)P_T^{'}(R)} = \lim_{R\to 0}\frac{-P_T^{''}(R)}{2(P_T^{'}(R))^2} \leq 0$. Besides $P_T(R)-RP_T^{'}(R)$ is 0 when $R = 0$ and has negative derivative when $R > 0$. Hence, $P_T(R) - RP_T^{'}(R) < 0$ when $R > 0$. Thus, $\widehat{U}^{'}(R)$ is negative when $R > 0$ and $\widehat{U}(R)$ is maximized when R approaches zero, i.e. $U(R) \leq \lim_{R\to 0}\frac{R}{P_T(R)} = \frac{1}{P_T^{'}(0)}$. ■

Appendix III
Proof of Propositions 1, 2, and 3

Proof: Denote $P_R(r)$ to be the received power on a subchannel for reliable detection when the data rate on the subchannel is r. We have $P_T(R) = \frac{KP_R(r)}{g} = \frac{KP_R(\frac{R}{K})}{g}$, where g is the channel power gain. It is easy to see that $P_R(r)$ is monotonically increasing and strictly convex, and $P_T(0) = P_R(0) = 0$. According to Theorem 2, we have $R^*P_T^{'}(R^*) = P_C + P_T(R^*)$, which is equivalent to $R^*P_R^{'}(\frac{R^*}{K}) - KP_R(\frac{R^*}{K}) = P_Cg$. By differentiating the left hand side with respect to R^*, $\frac{\partial\left(R^*P_R^{'}(\frac{R^*}{K})-cP_R(\frac{R^*}{K})\right)}{\partial R^*} = \frac{R^*}{K}P_R^{''}(\frac{R^*}{K}) > 0$. Hence, the left hand side is strictly increasing in R^*. Therefore, higher data rate should be used when the channel has higher power gain. Suppose $g_1 > g_2$, and the corresponding optimal modulation and coding result in data rates R_1^* and R_2^* respectively. Hence, $U_1(R_1^*) > U_1(R_2^*)$. Besides, $U_1(R_2^*) = \frac{R_2^*}{P_C+\frac{KP_R(R_2^*/K)}{g_1}} > \frac{R_2^*}{P_C+\frac{KP_R(R_2^*/K)}{g_2}} = U_2(R_2^*)$. Hence, the energy efficiency increase with channel gain.

According to Theorem 2, $R^*P_T^{'}(R^*) - P_T(R^*) = P_C$. The derivative of the left hand side is $R^*P_T^{''}(R^*) > 0$. Hence, R^* increases with P_C. The proof that the energy efficiency decreases with circuit power is similar to the proof that energy efficiency increases with channel gain. When $P_C = 0$, according to proof in II, $U(R)$ is maximized when R approaches zero, i.e. $U_{max} = \lim_{R\to 0}\frac{R}{P_T(R)} = \frac{1}{P_T^{'}(0)}$.

$R = Kr$ and $P_T(R) = K\overline{P}_T(\frac{R}{K})$, where $\overline{P}_T(r)$ is the transmit power on each subchannel, and is monotonically increasing and strictly convex in r. According to Theorem 2, we have $R^*\overline{P}_T^{'}(\frac{R^*}{K}) = P_C + K\overline{P}_T(\frac{R^*}{K})$, which is equivalent to $r^*\overline{P}_T^{'}(r^*) - P_T(r^*) = \frac{P_C}{K}$. The left hand side is increasing in r^* while the right hand side is decreasing in K. Hence, the data rate on each subchannel should decrease with increasing number of subchannels assigned. The proof that the energy efficiency increases with the number of subchannels assigned is also similar to the proof in III and is omitted. The highest energy efficiency is obtained with infinite number of subchannels, i.e. $U(R) = \lim_{K\to\infty}\frac{R}{P_C+P_T(R)} = \frac{r}{\overline{P}_T(r)}$. Similar to the proof in II, $U(R)$ is maximized when r approaches zero. We have $U_{max} = \lim_{r\to 0}\frac{r}{\overline{P}_T(r)} = \frac{1}{P_T^{'}(0)}$. ■

Appendix IV
Proof of Theorem 3

Proof: The global convergence is straightforward from Lemmas 1. Since $r_2^{[0]} = \alpha r_1^{[0]}$ and $r_1^{[i]} \leq r^* \leq r_2^{[i]}$, with induction, we have $r_2^{[i]} - r_1^{[i]} = \frac{r_2^{[0]}-r_1^{[0]}}{2^i} \leq \frac{(\alpha-1)r^*}{2^i}$. Hence, $\widehat{r}^{[i]} = \frac{r_1^{[i]}+r_2^{[i]}}{2} \geq (2r_2^{[i]} - \frac{(\alpha-1)r^*}{2^i})/2 \geq r^* - \frac{(\alpha-1)r^*}{2^{i+1}}$ and $\widehat{r}^{[i]} \leq r^* + \frac{(\alpha-1)r^*}{2^{i+1}}$. Then $|\widehat{r}^{[i]} - r^*| \leq \frac{(\alpha-1)r^*}{2^{i+1}}$. Let $\frac{(\alpha-1)r^*}{2^{i+1}} \leq \epsilon$. We have $i \geq \log_2(\frac{(\alpha-1)r^*}{\epsilon} - 1)$. Theorem 3 follows immediately. ■

References

[1] R. G. Gallager, *Information Theory and Relaible Communication*. John Wiley & Sons, Inc., 1968.

[2] S. Sampei, S. Komaki, S., and N. Morinaga, "Adaptive modulation/tdma scheme for personal multimedia communication systems," in *Proc. of 1994 IEEE Global Telecommun. Conf.*, Nov. 1994, pp. 989–993.

[3] M. Ouchi, H. J. Lee, S. Komaki, and N. Morinaga, "Proposal for modulation level controlled radio system applied to atm networks," in *Proc. of Fourth European Conference on Radio Relay Systems*, Oct. 1993, pp. 322–327.

[4] K. Lahiri, A. Raghunathan, S. Dey, and D. Panigrahi, "Battery-driven system design: A new frontier in low power design," in *Proc. Intl. Conf. on VLSI Design*, Bangalore, India, Jan. 2002, pp. 261–267.

[5] S. Verdu, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory.*, vol. 48, no. 6, pp. 1319–1343, June 2002.

[6] F. Meshkati, H. V. Poor, S. C. Schwartz, and N. B. Mandayam, "An energy-efficient approach to power control and receiver design in wireless networks," *IEEE Trans. Commun.*, vol. 5, no. 1, pp. 3306–3315, Nov. 2006.

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

[8] A. Y. Wang, S. Chao, C. G. Sodini, and A. P. Chandrakasan, "Energy efficient modulation and MAC for asymmetric RF microsensor system," in *Int. Symp. Low Power Electronics and Design*, Huntington Beach, CA, 2001, pp. 106–111.

[9] Draft IEEE std 802.16e/D9, "IEEE standard for local and metropolitan area networkspart 16: air interface for fixed and mobile broadband wireless access systems," 2005.

[10] 3GPP, TR 25.913, "Requirements for evolved utra and evolved utran," www.3gpp.org.

[11] Y. (G.) Li and G. L. Stuber, *OFDM for Wireless Communications*. Springer, 2006.

[12] G. Miao, and Z. Niu, "Practical feedback design based ofdm link adaptive communications over frequency selective channels," in *Proc. IEEE Conf. Commun. (ICC' 2006)*, Istanbul, Turkey, 2006, pp. 4624–4629.

[13] R. J. McEliece and W. E. Stark, "Channels with block interference," *IEEE Trans. Inf. Theory.*, vol. 30, no. 1, pp. 44–53, Jan. 1984.

[14] D Tse and P Viswanath , *Fundamentals of wireless communication*. Cambridge University Press, 2005.

[15] J. M. Cioffi, *A Multicarrier Primer*. ANSI T1E1, 1999.

[16] C. E. Shannon, "Communication in the presence of noise," in *Proc. IRE*, vol. 37, Jan. 1949, pp. 10–21.

[17] S. Cui A. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," in *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, Aug. 2004, pp. 1089–1098.

[18] R. G. Gallager, "Power limited channels: Coding, multiaccess, and spread spectrum," in *Proc. Conf. Inform. Sci. and Syst.*, vol. 1, Mar. 1988.

[19] D. Goodman and N. Mandayam, "Power control for wireless data," vol. 7, no. 2, pp. 48–54, Apr. 2000.

[20] N. Feng, S. C. Mau, and N. B. Mandayam, "Pricing and power control for joint network-centric and user-centric radio resource management," *IEEE Trans. Commun.*, vol. 52, no. 9, pp. 1547–1557, Sept. 2004.

[21] E. Wolfstetter, *Topics in Microeconomics: Industrial Organization, Auctions, and Incentives*. Cambridge University Press, 1999.

[22] B. Prabhakar, E. U. Biyikoglu, and A. E. Gamal, "Energy-efficient transmission over a wireless link via lazy packet scheduling," in *Proc. IEEE Infocom 2001*, vol. 1, 2001, pp. 386–394.

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

[24] K. C. Kiwiel and K. Murty, "Convergence of the steepest descent method for minimizing quasiconvex functions," in *J. of Optimization Theory and Applications*, vol. 89, no. 1, Sept. 2005, pp. 221–226.

[25] A.J.Goldsmith and S. G. Chua, "Variable-rate variable-power mqam for fading channels," *IEEE Trans. Commun.*, vol. 45, no. 10, pp. 1218–1230, Oct. 1997.

[26] ITU-R Recommendation M.1225, "Guidelines for evaluation of radio transmission technologies for imt-2000," 1997.