Analysis and Optimization of Wireless Transmissions over Fast Fading Channels with Slow Time-Varying Energy Arrival

Jie Gong
School of Data and
Computer Science,
Sun Yat-sen University,
Guangzhou 510006, China
Email: gongj26@mail.sysu.edu.cn

Zhenyu Zhou
State Key Laboratory of
Alternate Electrical Power System
with Renewable Energy Sources,
School of Electrical and Electronic Engineering,
North China Electric Power University,
Beijing 102206, China

Sheng Zhou
Tsinghua National Laboratory for
Information Science and Technology,
Department of Electronic Engineering,
Tsinghua University,
Beijing 100084, China

***Abstract*—In wireless communication systems powered by harvested energy, besides the channel fading, there is another dimension of dynamics induced by energy arrival variations, which makes the design of wireless transmission policies nontrivial. In this paper, we propose a framework for analyzing the energy harvesting powered wireless transmissions where the channel fading and the energy arrival variations are of different timescales. We define the duration between two consecutive changes of energy arrival rate as an *energy harvesting frame*, which consists of N channel fading slots. The power allocation problem can be formulated as an Markov decision process (MDP), and can be decoupled into two sub-problems. The inner problem deals with the power allocation in channel fading timescale in every N slots where the energy arrival rate keeps constant, and the outer problem deals with the energy management in energy harvesting timescale among frames. The two sub-problems can be solved by finite horizon dynamic programming (DP) and infinite horizon DP, respectively. Numerical simulations show that the average rate decreases slightly as N increases, and the rate under i.i.d. channel is higher than that under Markov channel.**

I. Introduction

In wireless sensor networks or body area networks, the sensor nodes or the embedded monitoring devices need to be battery powered to avoid the inconvenience of power line. To prolong the lifetime of these battery powered devices, *energy harvesting* technology has been adopted to charge the battery by gathering energy from ambient environments. However, the resource allocation for wireless data transmissions with energy harvesting confronts great challenges due to the two-dimensional dynamics. On the one hand, wireless links experience dynamic channel fading. On the other hand, the energy arrival process is also random and dynamic. In addition, the channel fading and the energy arrival process in real systems are usually in different timescales, which further increases the difficulty of the problem.

A number of research papers have focused on the optimization of wireless transmissions with energy harvesting in the literature. The energy management problem for stability guarantee is studied in [1] in sensor networks. A discrete-time energy arrival model is adopted in [2] to optimize the data packet transmission. In these works, additive white Gaussian noise (AWGN) channel is assumed, and the focus is mainly on the interaction between energy queue and data queue. There are also research efforts considering fading channels. The optimal offline power allocation policies are proposed in [3] by assuming all the energy arrival and channel states are known in advance. Such an offline analysis has been extended to multiple-input multiple-output (MIMO) channel [4], broadcast channel [5], cooperative relay networks [6], and so on. However, most of the existing works assume that the energy harvesting rate changes at the same timescale of channel fading. The impact of the different timescales on the transmission policies is still unclear.

The timescale issue between channel fading and data packet transmission has been initially considered in [7], [8]. Specifically, in [8], each data packet may be transmitted over several fading blocks. In the energy harvesting systems, the timescale issue is taken into account in refs. [9], [10], [11] by assuming that the energy arrival rate changes much slower than the channel fading, and study the power allocation problems with constant energy arrival rate, i.e., within a single energy harvesting frame. Furthermore, the outage minimization problem over multiple energy harvesting periods has been studied in [12]. Nevertheless, there lacks a general framework for analyzing the interactions between energy harvesting and channel fading when the timescales are different.

In this paper, we consider a general framework to analyze wireless transmissions with energy harvesting, where the energy arrival rate varies in a larger timescale compared with channel fading. Specifically, the channel gain changes in every transmission slot, while the energy harvesting rate changes every $N (N \geq 1)$ slots. In each slot, the transmitter determines its transmit power according to the current channel state and the available battery energy. The problem is formulated as a Markov decision process (MDP) [13], and can be decoupled into two sub-problems: the power allocation problem during

978-1-4673-8999-0/17/$31.00 ©2017 IEEE

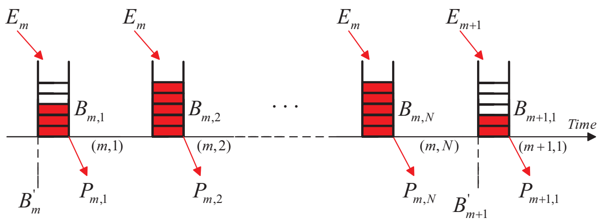

Fig. 1. System model for energy arrival, storage and power allocation.

every N slots with constant energy arrival rate, and the energy management problem with the variations of energy arrival rate. The two sub-problems are solved by the finite horizon dynamic programming (DP) algorithm and the infinite horizon DP algorithm, respectively. Simulations are conducted to illustrate the influence of N on the performance.

The rest of the paper is organized as follows. Section II describes the system model and the problem formulation. Section III briefly introduces the DP algorithm. The solution for the problem is presented in Section IV. Simulations are shown in Section V. Finally, Section VI concludes the paper.

II. System Model and Problem Formulation

We consider a wireless communication link where the transmitter is powered by harvesting energy from ambient environments. The system is slotted, and the channel is assumed block fading, i.e., the channel gain remains constant in each slot, but varies among slots. The slot length is normalized for simplicity. The energy harvesting rate also varies over time, and we assume that the energy harvesting rate varies every N slots, and keeps constant during these N slots. In this paper, we assume $N \geq 1$, i.e., energy harvesting rate changes in a slower frequency compared with channel fading. The case that $N < 1$ is left for future work. We view every N slots as an *energy harvesting frame*. Denote $m = 1, 2, \ldots$ as the index of energy harvesting frame, and the n-th transmission slot in frame m is indexed by $(m, n), n = 1, 2, \ldots, N$. At the beginning of the m-th energy harvesting frame, the transmitter detects the energy harvesting rate E_m, which means that the amount of energy E_m arrives in each slot. The harvested energy is stored in the battery with capacity $B_{\max}$. In the slot (m, n), the channel gain is denoted by $\gamma_{m,n}$, the available energy in the battery is denoted by $B_{m,n}$, and the transmit power is denoted by $P_{m,n}$. The model for energy arrival, storage and power allocation is illustrated in Fig. 1.

A. Channel Model

The channel is modeled as a finite state Markov chain (FSMC) with L states. The state transition matrix is denoted by

$$\mathbf{P} = \begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,L} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,L} \\ \vdots & \vdots & \ddots & \vdots \\ p_{L,1} & p_{L,2} & \cdots & p_{L,L} \end{pmatrix}, \tag{1}$$

where $p_{i,j} = \Pr(\gamma_{m,n+1} = j | \gamma_{m,n} = i)$ is the transition probability from state i to j. It satisfies that $p_{i,j} \geq 0, \forall i, j$ and $\sum_j p_{i,j} = 1, \forall i$. And we denote that $\gamma_{m,N+1} = \gamma_{m+1,1}$. Notice that if the rows of matrix $\mathbf{P}$ are identical, i.e., $p_{i,j} = p_{i',j}, \forall i \neq i'$, the channel is i.i.d.

B. Energy Model

The energy amount is quantized, and the battery state is represented by $B_{m,n}$ units of energy, where $B_{m,n} \in \{0, 1, \ldots, B_{\max}\}$. In an energy harvesting frame, the battery energy state is updated as

$$B_{m,n+1} = \min\{B_{m,n} - P_{m,n} + E_m, B_{\max}\} \tag{2}$$

for all $n = 0, 1, \ldots, N-1$. We denote that $B_{m,0} = B_{m-1,N}$ and $P_{m,0} = P_{m-1,N}$. Notice that $P_{m,n}$ also represents the energy amount for transmission as we have normalized the slot length. The transmit power is constrained by the available energy in the battery, i.e.,

$$0 \leq P_{m,n} \leq \min\{B_{m,n}, P_{\max}\}. \tag{3}$$

The energy arrival process is also modeled as an FSMC. Hence, the transition probability can be similarly denoted by $q_{i,j} = \Pr(E_{m+1} = j | E_m = i), \forall i, j \in \{0, 1, \ldots, B_{\max}\}$.

C. Problem Formulation

We aim to adapt the transmit power to maximize the long-term average rate. The objective function can be represented as

$$\max \lim_{M \to +\infty} \frac{1}{MN} \mathbb{E}\left[\sum_{m=1}^{M} \sum_{n=1}^{N} r(P_{m,n}, \gamma_{m,n})\right], \tag{4}$$

where $r(P_{m,n}, \gamma_{m,n})$ is the instantaneous data rate which is a non-decreasing concave function of the channel gain and the transmit power, the expectation is taken over all the possible energy harvesting processes $\{E_m | m = 1, 2, \ldots\}$ and the channel variations $\{\gamma_{m,n} | n = 1, 2, \ldots, N, m = 1, 2, \ldots\}$, and the maximization operation is taken over all the power allocation policies $\{P_{m,n} | n = 1, 2, \ldots, N, m = 1, 2, \ldots\}$. The problem can be modeled as a Markov decision process (MDP) and can be effectively solved by dynamic programming (DP) approach [13], which is briefly introduced in the next section.

III. Dynamic Programming Basics

The DP algorithm deals with a set of MDP problems which can be divided into *stages*. The problems have two principal features: (1) There is an underlying discrete time dynamic system. (2) The *reward*[1] function is additive over time. The dynamic system expresses the evolution of the system *states*, under the influence of *decisions* taken at discrete instances of time (stage). The system has the form

$$x_{k+1} = f(x_k, u_k, w_k), \quad k = 0, 1, 2, \ldots, K-1 \tag{5}$$

where k is the index of stage, $x_k \in \mathcal{S}$ is the system state, u_k is the control variable, w_k is a random parameter, K is the

[1] In this paper, we use the term "reward" instead of *cost* as in [13], since we consider the rate maximization problem, rather than minimization problems.

number of concerned stages, and f is a function that describes the mechanism by which the state is updated.

The reward function in stage k, denoted by $g(x_k, u_k, w_k)$, is additive over time. For the *finite horizon problems* where K is finite, the maximum reward with the initial state x_0 is

$$J^*(x_0) = \max_{\pi} \mathbb{E}\Bigg[\sum_{k=0}^{K-1} g(x_k, u_k, w_k) + g(x_K)\Bigg], \quad (6)$$

where the expectation is taken over all the random parameters, and the maximization is taken over all the possible policies $\pi = \{\mu_0, \mu_1, \dots, \mu_{K-1}\}$ which define the mappings from the states to the controls, i.e., $u_k = \mu_k(x_k)$. The problem can be solved by DP algorithm [13, Vol. I, Prop. 1.3.1], which recursively proceeds the following optimizations

$$J_K(x_K) = g_K(x_K), \quad (7)$$

$$J_n(x_k) = \max_{u_k \in \mathcal{U}(x_k)} \mathbb{E}\{g(x_k,u_k,w_k)+J_{k+1}(f(x_k,u_k,w_k))\},$$
$$k = 0, 1, \dots, K-1, \quad (8)$$

where $\mathcal{U}(x_k)$ is the feasible control set in state x_k.

For the *infinite horizon problems* where K tends to infinity, the objective is to maximize the average per-stage reward, which can be formulated as

$$J^*(x_0) = \max_{\pi} \lim_{K\to\infty} \frac{1}{K}\mathbb{E}\Bigg[\sum_{k=0}^{K-1} g(x_k, u_k, w_k)\Bigg]. \quad (9)$$

In general, the optimal average reward is independent of x_0 [13, Vol. II, Sec. 4.2]. In this case, suppose a scalar λ and a vector $\boldsymbol{h} = \{h(i)|i \in \mathcal{S}\}$ satisfy Bellman's equation

$$\lambda + h(i) = \max_{u\in\mathcal{U}(i)} \Big[g(i, u) + \sum_{j\in\mathcal{S}} p_{ij}(u)h(j)\Big], \quad (10)$$

where $g(i, u) = \mathbb{E}_w[g(i, u, w)]$, and $p_{ij}(u) = \Pr(x_{k+1} = j|x_k = i, u_k = u)$. Then λ is the optimal average reward, and if $u = \mu^*(i)$ attains the maximum in (10) for all i, the stationary policy $\pi^* = \{\mu^*, \mu^*, \cdots\}$ is optimal.

The problem (9) can be solved by the *value iteration algorithm* [13, Vol. II, Sec. 4.4] based on (10). Specifically, we initialize $h^{(0)}(i) = 0, \forall i \in \mathcal{S}$. Then we fix a state s, and denote the output of the k-th iteration as $\boldsymbol{h}^{(k)} = \{h^{(k)}(i)|i \in \mathcal{S}\}$. For the $(k+1)$-th iteration, we update the vector $\boldsymbol{h}$ as

$$h^{(k+1)}(i) = \max_{u\in\mathcal{U}(i)} \Big[g(i, u) + \sum_{j\in\mathcal{S}} p_{ij}(u)h^{(k)}(j)\Big] - \max_{u\in\mathcal{U}(s)} \Big[g(s, u) + \sum_{j\in\mathcal{S}} p_{sj}(u)h^{(k)}(j)\Big]. \quad (11)$$

Under the aperiodic-type conditions [13, Vol. II, Prop. 4.3.2], $\lambda^{(k)} = \max_{u\in\mathcal{U}(s)}[g(s, u) + \sum_{j\in\mathcal{S}} p_{sj}(u)h^{(k)}(j)]$ converges to the optimal reward λ. In addition, by defining

$$\underline{c}_k = \min_i \big[h^{(k+1)}(i) + \lambda^{(k+1)} - h^{(k)}(i)\big], \quad (12)$$

$$\overline{c}_k = \max_i \big[h^{(k+1)}(i) + \lambda^{(k+1)} - h^{(k)}(i)\big], \quad (13)$$

we have $\underline{c}_k \le \underline{c}_{k+1} \le \lambda \le \overline{c}_{k+1} \le \overline{c}_k$ [13, Vol. II, Prop. 4.3.3], which is used as the stopping criterion.

IV. Solution for Average Rate Maximization

In this section, we solve the average rate maximization problem via DP algorithm. The objective function (4) can be re-written as

$$\max \lim_{M\to+\infty} \frac{1}{M}\mathbb{E}_{E_m}\Bigg[\sum_{m=1}^{M} \mathbb{E}_{\gamma_{m,n}}\Bigg[\sum_{n=1}^{N} \frac{1}{N} r(P_{m,n}, \gamma_{m,n})\Bigg]\Bigg]. \quad (14)$$

Based on this representation, the original problem can be decoupled into two sub-problems. For given initial battery energy state and initial channel state of an energy harvesting frame, if we fix the terminal battery energy state, the optimal power allocation in this frame is irrelevant to the other frames since the energy arrival rate is known and fixed. In this sense, we can formulate an *inner problem* of finite horizon power allocation optimization within a frame. On the other hand, given the expected per-frame reward obtained by solving the inner problem, the optimization among frames can be formulated as an *outer problem* that determines the initial/terminal battery energy of each frame, which is equivalent to determining the optimal amount of energy to be used in each frame. The inner and outer problems can be solved by finite horizon DP algorithm and infinite horizon DP algorithm, respectively. We discuss the two sub-problems in detail as follows.

A. Inner Problem for Per-frame Power Allocation

To describe the finite horizon inner problem in terms of MDP, we need to clarify the following items: state, action, per-stage reward and state transition. Since the energy arrival rate E_m is known and constant over N slots, the system state can be denoted by $(B_{m,n}, \gamma_{m,n})$. The action is the transmit power $P_{m,n}$, and the per-stage the reward is given by

$$g(B_{m,n}, \gamma_{m,n}, P_{m,n}) = \frac{1}{N} r(P_{m,n}, \gamma_{m,n}). \quad (15)$$

The state transition of $B_{m,n}$ is determined by (2), and that of $\gamma_{m,n}$ is independed with $B_{m,n}$ and determined by the state transition matrix $\mathbf{P}$.

The DP algorithm recursively solves the following per-slot optimization sub-problems

$$J_n(B_{m,n}, \gamma_{m,n}) = \max_{P_{m,n}} \Big[\frac{1}{N} r(P_{m,n}, \gamma_{m,n}) + \sum_{\gamma} p_{\gamma_{m,n},\gamma} J_{n+1}(B_{m,n+1}, \gamma)\Big] \quad (16)$$

under the constraint (3) for all $1 \le n \le N$. Recall that $p_{\gamma_{m,n},\gamma} = \Pr(\gamma|\gamma_{m,n})$ is the state transition probability. To determine the boundary value J_{N+1}, as shown in Fig. 1, we insert an *artificial* battery state $B_{m,N+1}$ before the start of frame $m + 1$ with taking the new arrival E_{m+1} into account, i.e., $B_{m,N+1} = B_{m,N} - P_{m,N}$. Thus, the boundary condition is defined as

$$J_{N+1}(B_{m,N+1}, \gamma_{m,N+1}) = \begin{cases} 0, & \text{if } B_{m,N+1} = B'_{m+1} \\ -\infty, & \text{else} \end{cases} \quad (17)$$

where B'_{m+1} denotes the amount of energy reserved for the future usage, which is determined by the infinite horizon outer problem as detailed in the next subsection.

We solve the inner problem starting from slot (m,N). In this slot, we have

$$\begin{aligned} & J_N(B_{m,N},\gamma_{m,N}) \\ &= \max_{P_{m,N}} \left[\frac{1}{N} r(P_{m,N},\gamma_{m,N}) + \sum_{\gamma} p_{\gamma_{m,N},\gamma} J_{N+1}(B_{m,N+1},\gamma)\right] \\ &= \begin{cases} \frac{1}{N} r(B_{m,N}-B'_{m+1},\gamma_{m,N}), \\ \qquad\qquad \text{if } \max\{0,B_{m,N}-P_{\max}\} \le B'_{m+1} \le B_{m,N} \\ -\infty, \qquad \text{else} \end{cases} \end{aligned} \tag{18}$$

where the maximum is achieved when the optimal power is $P^*_{m,N} = B_{m,N} - B'_{m+1}$. However, due to the transmit power constraint (3), if $B_{m,N} > P_{\max}$ and $B'_{m+1} < B_{m,N} - P_{\max}$, $B_{m,N+1} = B'_{m+1}$ is not achievable, and J_N tends to infinity. Hence, finite reward is obtained only when $\max\{0, B_{m,N} - P_{\max}\} \le B'_{m+1} \le B_{m,N}$.

Similarly, the rest $J_n(B_{m,n},\gamma_{m,n}), n < N$ can be calculated recursively according to (16). Finally, when $n = 1$, the expected optimal reward $J_1(B_{m,1},\gamma_{m,1})$ in frame m starting from initial state $(B_{m,1},\gamma_{m,1})$ is obtained for given energy arrival rate E_m and remained energy B'_{m+1}.

B. Outer Problem for Inter-frame Energy Management

The infinite horizon outer problem can also be solved by DP based on Bellman's equation. For simplicity, we ignore the index of slot n, and denote that $B_m = B_{m,1}, \gamma_m = \gamma_{m,1}$. At the beginning of frame m, the transmitter is aware of the artificial battery energy B'_m reserved in the last frame, the new energy arrival rate E_m, and the current channel gain γ_m, which altogether can be viewed as the system state for the outer problem. Based on the state, the transmitter determines how much energy should be reserved for the next frame. Hence, the action is denoted by B'_{m+1}. The reward function associated with the given state and action is the optimal per-frame reward J_1 obtained by solving the finite horizon problem, i.e.,

$$G(B'_m, E_m, \gamma_m, B'_{m+1}) = J_1(\min\{B'_m + E_m, B_{\max}\}, \gamma_m), \tag{19}$$

where J_1 is calculated by the last subsection giving that E_m is constant and B'_{m+1} is known.

The channel state transition for the outer problem is the N-step probability transition matrix calculated by $\mathbf{P}^N$, where the item at the i-th row and the j-th column is denoted as $p^{(N)}_{i,j}$. Recall that the state transition of energy arrival rate is $q_{i,j}$. Based on the Bellman's equation, if there exists a scalar λ and a vector $\boldsymbol{h}$ such that

$$\lambda + h(b,e,\gamma) = \max_{b'} \Bigg[G(b,e,\gamma,b') + \sum_{\gamma'} p^{(N)}_{\gamma,\gamma'} \sum_{e'} q_{e,e'} h(b',e',\gamma') \Bigg], \tag{20}$$

the optimal average rate is λ. The value iteration algorithm to calculate the optimal rate is summarized in Algorithm 1.

Algorithm 1 Value Iteration Algorithm

Initialize $h^{(0)}(b,e,\gamma) = 0, \forall b,e,\gamma,\ \lambda^{(0)} = 0, k = 0$.
repeat
1. Update the scalar λ as
$$\lambda^{(k+1)} = \max_{b'} \Bigg[G(b_0,e_0,\gamma_0,b') + \sum_{\gamma'} p^{(N)}_{\gamma_0,\gamma'} \sum_{e'} q_{e_0,e'} h^{(k)}(b',e',\gamma') \Bigg],$$
where (b_0,e_0,γ_0) is a fixed state.
2. Update the vector $\boldsymbol{h}$ as
$$h^{(k+1)}(b,e,\gamma) = \max_{b'} \Bigg[G(b,e,\gamma,b') + \sum_{\gamma'} p^{(N)}_{\gamma,\gamma'} \sum_{e'} q_{e,e'} h^{(k)}(b',e',\gamma') \Bigg] - \lambda^{(k+1)}.$$
3. Update $k = k+1$.

until $\bar{c}_k - \underline{c}_k < \epsilon$

C. Complexity Analysis

As DP algorithm usually encounters the *curse of dimensionality* [13] as it is actually an exhaustive search algorithm, it is necessary to discuss the complexity issue of the proposed algorithm. In the inner problem, the number of states is $L \times (B_{\max}+1)$, and the number of actions is $P_{\max}+1$. In the outer problem, the numbers are $L \times (B_{\max}+1)^2$ and $B_{\max}+1$, respectively. The value of $B_{\max}$ and $P_{\max}$ depends on the volume of a unit energy. In real systems, transmission modes are discrete, including discrete modulation (QAM) and coding scheme, as well as quantized transmit power levels. Thus, with appropriate quantization of transmit power levels, the number of states and actions can be maintained small enough so that DP algorithm is applicable.

V. Numerical Simulations

We run some simulations to analyze the influence of different timescales. The energy arrival process follows Poisson distribution with average rate E_{avg}. Hence, we have

$$q_{i,j} = \frac{E^j_{\text{avg}}}{j!} e^{-E_{\text{avg}}}, \quad j = 0, 1, \cdots$$

which is irrelevant to the previous state i. We set $P_{\max} = 10$, and the channel gain $\gamma_{m,n}$ is measured based on the reference transmit power $P_{\text{ref}} = 5$. We adopt Shannon's equation to calculate the rate function, i.e.,

$$r(P_{m,n},\gamma_{m,n}) = \log_2\Big(1 + \frac{P_{m,n}}{P_{\text{ref}}}\gamma_{m,n}\Big).$$

The channel gain follows Rayleigh distribution with average γ_{avg}, i.e.,

$$p_\gamma = \frac{1}{\gamma_{\text{avg}}} \exp\Big(-\frac{\gamma}{\gamma_{\text{avg}}}\Big).$$

And the whole region is divided into $K = 6$ non-overlapping consecutive intervals with boundary values 2.61dB, 6.08dB,

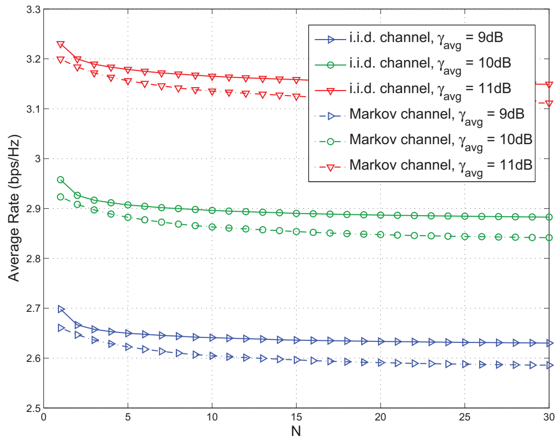

Fig. 2. Average rate versus number of slots N in an energy harvesting frame under different channel gains. $E_{\text{avg}} = 5, B_{\max} = 20$.

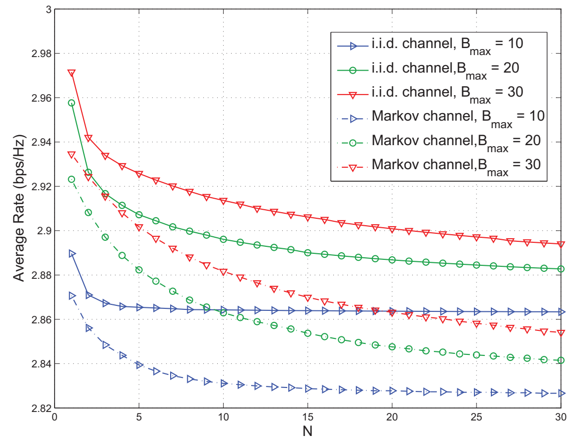

Fig. 4. Average rate versus number of slots N in an energy harvesting frame under different battery capacities. $E_{\text{avg}} = 5, \gamma_{\text{avg}} = 10$dB.

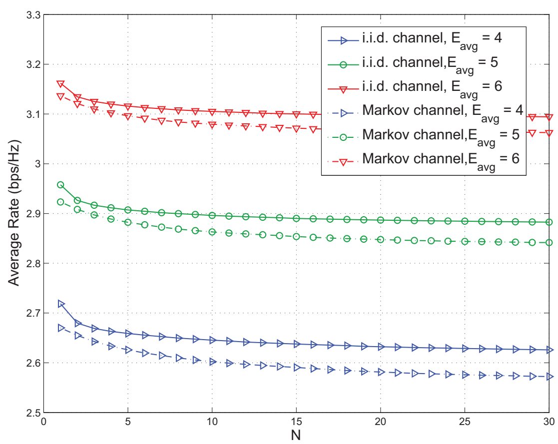

Fig. 3. Average rate versus number of slots N in an energy harvesting frame under different average energy arrival rates. $\gamma_{\text{avg}} = 10$dB, $B_{\max} = 20$.

8.41dB, 10.41dB, and 12.53dB which are obtained by keeping uniform probability distribution for the reference gain $\gamma_{\text{avg}} = 10$dB, i.e., the probability that the channel gain lies in each interval is equal. Each interval is represented by its average value. The channel state changes among these K average values. We study two types of channel transitions: (1) the channel gain is i.i.d. among slots and is generated according to the probability with which the value lies in each interval; (2) the channel is Markov, i.e, the current channel state depends on the previous channel state, and the channel state transition happens only between adjacent states, and the transition probability is determined according to [14]. The former refers to the i.i.d. fast fading channel, and the latter refers to the Markov slow fading channel.

Figs. 2, 3, and 4 show the average rate performance versus N for different values of average channel gain, average energy arrival rate, and battery capacity, respectively. It can be observed that with the increase of N, the average rate decreases for all the curves. It can be explained as follows. Intuitively, high transmit power should be allocated when the channel state is good, which can be viewed as *opportunistic* power allocation. If the energy arrival rate changes in the same timescale of channel fading, full opportunistic power allocation can be explored. However, if the energy arrival and the channel fading are of different timescales, the good channel states under the low energy arrival rate may not be fully utilized, and the bad channel states under the high energy arrival rate have to be utilized to avoid the battery energy overflow. When N is sufficiently large, the average rate converges as the per-frame average rate over N slots becomes independent with the adjacent frames.

In addition, it is also observed that the average rate of Markov channel is lower than that of i.i.d. channel. The reason is that in Markov channel, the channel can only transit between adjacent states. If the channel is in deep fading, it cannot become good enough in a short time. Consequently, the harvested energy is either used in bad channel states or wasted due to battery overflow. It can be concluded that the ability of exploring opportunistic power allocation of Markov channel is limited compared with that of i.i.d. channel.

The asymptotic performance of the average rate versus battery capacity is depicted in Fig. 5. It can be seen that for i.i.d. channel with $N = 1$, the average rate converges when $B_{\max} > 80$. While for larger value of N or for Markov channel, the convergence speed is much slower. This is because in these cases, larger battery capacity is required to combat miss match between channel fading and energy arrival.

In Fig. 6, the number of iterations in value iteration algorithm for the outer problem is depicted for both channels. The terminate criterion is that $\overline{c}_k - \underline{c}_k < 10^{-3}$. It can be seen that more iterations are required for Markov channel model compared with i.i.d. channel model. In i.i.d. channel, the expected per-frame rate is relevant only to the initial energy

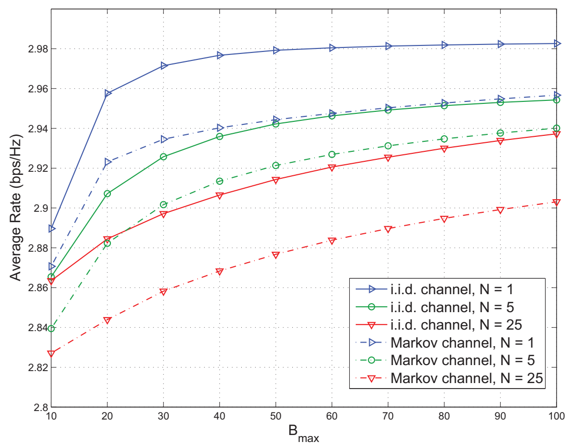

Fig. 5. Average rate versus battery capacity $B_{\max}$ in an energy harvesting frame under different energy harvesting frame length N. $E_{\text{avg}} = 5, \gamma_{\text{avg}} = 10\text{dB}$.

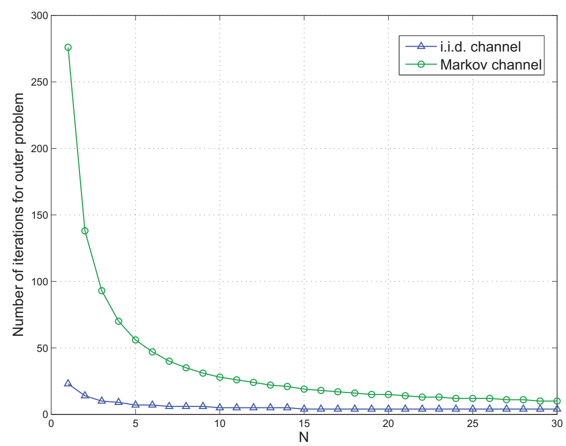

Fig. 6. The number of iterations in value iteration algorithm for the outer problem. $E_{\text{avg}} = 5, \gamma_{\text{avg}} = 10\text{dB}, B_{\max} = 20$.

condition. On the contrary, in Markov channel, it is influenced by both the initial energy condition and the initial channel condition. Hence, the latter has more h values to be calculated, which results in more iterations before convergence.

VI. Conclusion

For the power allocation problem with two-dimensional dynamics including channel fading and energy arrival variation of different timescales, we adopted the finite horizon DP algorithm to solve the inner sub-problem and the infinite horizon DP algorithm to solve the outer sub-problem. Some observations were obtained based on the numerical simulations. Specifically, the ability of exploring opportunistic power allocation is influenced by two factors: the timescale of energy arrival variation and that of channel fading. Full opportunistic power allocation can be explored if both the energy arrival rate and the channel state changes in the same timescale. On the other hand, if the energy arrival rate or the channel state changes slower than the other, the ability of exploring opportunistic power allocation is limited.

Future work includes an extension to the case that $N < 1$, i.e., the energy arrival rate changes faster than the channel state.

Acknowledgment

This work is sponsored in part by the Fundamental Research Funds for the Central Universities, the Nature Science Foundation of China (No. 61601180, No. 61601181, No. 61571265, No. 91638204, No. 61461136004), and Hitachi R&D Headquarter.

References

[1] V. Sharma, U. Mukherji, V. Joseph, and S. Gupta, "Optimal energy management policies for energy harvesting sensor nodes," *IEEE Trans. on Wireless Communications*, vol. 9, no. 4, pp. 1326–1336, Apr. 2010.

[2] J. Yang and S. Ulukus, "Optimal packet scheduling in an energy harvesting communication system," *IEEE Transactions on Communications*, vol. 60, no. 1, pp. 220–230, Jan. 2012.

[3] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, "Transmission with energy harvesting nodes in fading wireless channels: Optimal policies," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 8, pp. 1732–1743, Aug. 2011.

[4] C. Hu, J. Gong, X. Wang, S. Zhou, and Z. Niu, "Optimal green energy utilization in MIMO systems with hybrid energy supplies," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 8, pp. 3675–3688, Aug 2015.

[5] J. Yang, O. Ozel, and S. Ulukus, "Broadcasting with an energy harvesting rechargeable transmitter," *IEEE Transactions on Wireless Communications*, vol. 11, no. 2, pp. 571–583, 2012.

[6] A. Minasian, S. ShahbazPanahi, and R. S. Adve, "Energy harvesting cooperative communication systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 11, pp. 6118–6131, Nov. 2014.

[7] X. Chen and W. Chen, "A joint channel-aware and buffer-aware scheduling for energy-efficient transmission over fading channels with long coherent time," in *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Dec 2014, pp. 103–107.

[8] ——, "Joint channel-buffer aware energy-efficient scheduling over fading channels with short coherent time," in *IEEE International Conference on Communications (ICC)*, June 2015, pp. 2810–2815.

[9] C. Huang, J. Zhang, P. Zhang, and S. Cui, "Threshold-based transmissions for large relay networks powered by renewable energy," in *IEEE Globel Communications Conference (Globecom)*, Dec. 2013.

[10] H. Li, C. Huang, S. Cui, and J. Zhang, "Distributed opportunistic scheduling for wireless networks powered by renewable energy sources," in *IEEE Conference on Computer Communications(Infocom'14)*. IEEE, 2014, pp. 898–906.

[11] J. Gong, S. Zhou, and Z. Zhou, "Networked MIMO with fractional joint transmission in energy harvesting systems," *IEEE Transactions on Communications*, vol. 64, no. 8, pp. 3323–3336, Aug. 2016.

[12] C. Huang, R. Zhang, and S. Cui, "Optimal power allocation for outage probability minimization in fading channels with energy harvesting constraints," *IEEE Transactions on Wireless Communications*, vol. 13, no. 2, pp. 1074–1087, Feb. 2014.

[13] D. P. Bertsekas, *Dynamic programming and optimal control*. Athena Scientific Belmont, MA, 2005.

[14] J. Razavilar, K. R. Liu, and S. I. Marcus, "Jointly optimized bit-rate/delay control policy for wireless packet networks with fading channels," *IEEE Transactions on Communications*, vol. 50, no. 3, pp. 484–494, Mar. 2002.