On the Time Scales of Energy Arrival and Channel Fading in Energy Harvesting Communications

Jie Gong, *Member, IEEE*, Zhenyu Zhou, *Senior Member, IEEE*, and Sheng Zhou, *Member, IEEE*

***Abstract*—In wireless communication systems powered by harvested energy, besides the channel fading, there exists another dimension of dynamics, i.e., energy arrival variation. In this paper, we propose a framework for analyzing the energy harvesting powered wireless transmissions where the energy arrival variations and the channel fading are of different timescales. The energy arrival rate changes every $N(N \geq 1)$ time slots, and the channel state changes every $M(M \geq 1)$ slots. We consider a power allocation problem among the time slots, which can be formulated as a Markov decision process and solved by dynamic programming (DP) algorithm. For the special case that $M = 1$, a low-complexity two-stage DP algorithm is proposed, which decouples the original problem into inner and outer 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 when the energy arrival rate changes. Numerical simulations show that the average data rate decreases as N or M increases, and the two-stage DP algorithm can perform close to the DP optimal algorithm.**

***Index Terms*—Energy harvesting, timescale difference, power allocation, two-stage dynamic programming, Markov decision process.**

I. Introduction

By GATHERING energy from ambient environment (solar, wind, radio waves and etc.) to power the wireless devices, *energy harvesting communications* are expected as one of the candidate technologies to support sustainable wireless communications. It also facilitates network deployment without the limitation of power lines. Energy harvesting technology can be widely used in the energy constrained scenarios such as wireless sensor networks [2], green cellular networks [3], [4] and so on. However, due to the randomness of the energy arrival process, energy harvesting communications encounter the dynamics in energy domain in addition to the variations of wireless channels, i.e., channel fading. Therefore, how to deal with the two-dimensional randomness of energy arrival and channel fading becomes a challenging issue.

In the literature, there are plenty of research efforts working on the two-dimensional randomness problem with offline assumption, i.e., both the energy arrivals and the channel states are known in advance. The offline optimal power allocation policy is given in [5] for AWGN channel. In the fading channel case, the optimal power allocation is interpreted as the *directional water-filling* policy [6]. When the harvested energy is insufficient, hybrid energy supply needs to be considered with power grid as supplementaries, and the optimal power allocation is a two-stage water-filling [7]. The offline analysis is further extended to broadcast channel [8], [9], multiple access channel [10], MIMO channel [11], [12], and cooperative relay channel [13]. The advantage of the offline assumption is that some insight on the structure of the optimal power allocation policy can be observed. However, as the dynamics of both energy arrival and channel fading are difficult to be predicted in practice, the offline optimal policy can hardly be applied in real systems.

Under the online assumption that only the current and past energy arrivals and channel states are known, a transmitter needs to make decision based on the current available information and the dependency among time slots due to the existence of battery and channel correlation should be exploited. The packet dropping and blocking probabilities with different sleep and wake-up strategies are analyzed in sensor/mesh networks powered by solar energy [14]. A cross-layer resource allocation problem is considered in [15] to maximize the total system utility. Throughput optimal routing scheme is proposed in [16] for rechargeable sensor networks. Closed-form maximum stable throughput is derived for both cognitive radio networks [17] and cooperative networks [18]. Finite-horizon online transmission scheduling for throughput maximization is studied in [19]. Joint transmitter-receiver optimization with correlated energy arrivals is studied in [20].

Among the works for online policy design, very few of them take the time scale issue into consideration. In fact, the energy harvesting process and the channel variations are usually of different time scales. For instance, the

Manuscript received September 13, 2017; revised November 20, 2017; accepted January 7, 2018. Date of publication January 10, 2018; date of current version May 17, 2018. This work was supported in part by NSFC under Grant 61771495, Grant 61601181, and Grant 61571265, in part by the Fundamental Research Funds for the Central Universities under Grant 161gpy37 and Grant 17lgjc40, in part by the EU's Horizon 2020 Research and Innovation Staff Exchange programm (TESTBED Project) under Grant 734325, and in part by the National Key Research and Development Program of China under Grant 2017YFB1001703. This work was presented in part at the IEEE International Conference on Communications, Paris, France, May 2017 [1]. The associate editor coordinating the review of this paper and approving it for publication was E. Ayanoglu. *(Corresponding author: Zhenyu Zhou.)*

J. Gong is with the Guangdong Key Laboratory of Information Security Technology, School of Data and Computer Science, Sun Yat-sen University, Guangzhou 510006, China (e-mail: gongj26@mail.sysu.edu.cn).

Z. Zhou is with the 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 (e-mail: zhenyu_zhou@ncepu.edu.cn).

S. Zhou is with the Tsinghua National Laboratory for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: sheng.zhou@tsinghua.edu.cn).

Digital Object Identifier 10.1109/TGCN.2018.2791629

solar power in a sunny day is constant over several seconds or minutes compared with the mobile terminal where the channel state changes in mini-seconds due to multi-path fading. While in a cloudy day, the solar power may change every second due to the movement of clouds. But if line-of-sight link is available and the terminal moves slowly, the wireless channel state may change very slow. The difference on time scales between energy arrivals and channel fading can be utilized as a pre-knowledge to simplify the problem. For instance, if the energy arrival rate changes much slower than the channel fading, the power allocation problem is constrained by a constant energy arrival rate. Thus, the randomness in energy domain disappears, and there is only an energy causality constraint. Under this assumption, power allocation and scheduling policies have been proposed for relay networks [21], distributed networks [22], and network MIMO transmissions [23]. Furthermore, the outage minimization problem over multiple energy harvesting periods has been studied in [24]. On the other hand, the channel coherence time is also shown to have significant impact on the performance [25], [26]. Nevertheless, there lacks a general framework for the time scale issue in energy harvesting communications.

In this paper, we build up a unified framework to analyze wireless transmissions with energy harvesting, where the energy arrival and channel fading vary in different time scales. Specifically, the channel state changes in every $M(M \geq 1)$ time slots, while the energy arrival rate changes every $N(N \geq 1)$ slots. In each time slot, the transmitter determines its transmit power according to the current channel state, the available battery energy, as well as how long the current state (including energy arrival rate and channel state) has lasted. The problem is formulated as a Markov decision process (MDP) [27], and can be solved by the dynamic programming (DP) algorithm. For the case that $M = 1$, i.e., the energy arrival rate changes slower than the channel fading, the benefit of slow change in energy arrival rate can be exploited by decoupling the original problem into two sub-problems: the power allocation problem during every N slots with constant energy arrival rate, and the energy management problem with the variations of energy arrival rate. A two-stage DP algorithm is proposed to solve the two sub-problems by the finite horizon DP algorithm and the infinite horizon DP algorithm, respectively. Numerical results are given to show the impact of parameters such as M and N on the average data rate, and to evaluate the performance of the proposed two-stage DP algorithm. The main contributions of this paper are summarized as follows.

- A unified framework is developed, which can be used to analyze the energy harvesting communication system with any time scales for energy arrivals and channel fading. In this framework, an average rate maximization problem is formulated and a DP optimal algorithm is proposed to solve the problem.
- For the case that the energy arrival rate changes slower than the channel fading ($M = 1$), a two-stage DP algorithm is proposed, and the computational complexity is greatly reduced compared with the DP optimal algorithm. Numerical results show that the two-stage DP algorithm can achieve close-to-optimal performance.
- The impacts of parameters including M, N, battery capacity on the performance are extensively analyzed. The results show that: (1) Larger M or N leads to smaller average data rate; (2) As M or N becomes large, the average data rate quickly converges to a lower bound which can be calculated based on the distributions of channel fading and energy arrival; (3) As the battery capacity becomes large, the average data rate converges to an upper bound that can be calculated based on the average energy arrival rate.

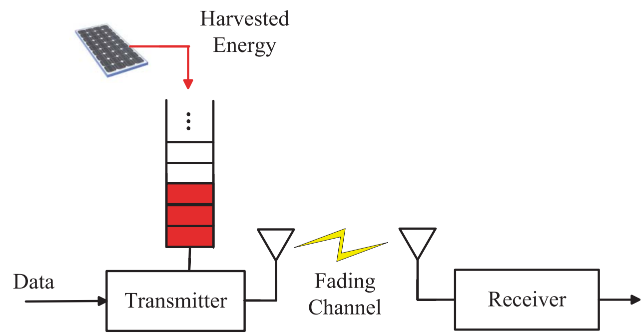

Fig. 1. Wireless communication link with energy harvesting transmitter.

The rest of the paper is organized as follows. Section II describes the system model and the problem formulation. Section III introduces the DP optimal algorithm. The two-stage DP algorithm for $M = 1$ is presented in Section IV. Simulations are shown in Section V. Finally, Section VI concludes the paper.

II. System Model and Problem Formulation

Consider a wireless communication link where the transmitter is powered by the energy harvested from ambient environments as shown in Fig. 1. The system is slotted with normalized slot length. In each slot, the harvested energy is stored in an energy battery with capacity $B_{\max}$. The transmitter determines the transmission mode depending on the channel and energy state. In this paper, we extend the model in [24] to a generalized one where the channel states can change faster, equal or slower than the energy arrival rate. The channel gain remains constant during every M slots, named as a *channel fading frame*, but varies among frames. The energy arrival rate varies every N slots, and keeps constant during these N slots. Every N slots with constant energy arrival rate is named as an *energy harvesting frame*. The values of M and N imply the relative dynamics between channel fading and energy arrival. For instance, $M < N$ refers to fast channel fading and slow change of energy arrival rate, and vice versa. In slot k, the channel gain is denoted by γ_k, the energy arrival rate is E_k, the battery state is B_k, and the transmit power is denoted by P_k.

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 $p_{i,j} = \Pr(\gamma_{k+1} = j|\gamma_k = i)$ is the transition probability from state i to j. It satisfies that $p_{i,j} \geq 0, \forall i, j$ and $\sum_j p_{i,j} = 1, \forall i$. The independently and identically distributed (i.i.d.) channel can also be modeled in this framework. If the channel is i.i.d., the rows of matrix $\mathbf{P}$ are identical, i.e., $p_{i,j} = p_{i',j}, \forall i \neq i'$.

B. Energy Model

Assume the amount of energy is quantized. Since the slot length is normalized, energy and power can be interchangeably used in the rest of the paper. The value of the battery state is taken in a finite set $B_k \in \{0, 1, \ldots, B_{\max}\}$. In slot k, the energy P_k is used for data transmission, and the transmitter harvests the energy E_k and stores it in the battery. By the end of slot k, the battery energy state is updated as

$$B_{k+1} = \min\{B_k - P_k + E_k, B_{\max}\}, \tag{2}$$

where the transmit power is constrained by both the available energy in the battery and the maximum output power of the power amplifier $P_{\max}$. We have

$$0 \leq P_k \leq \min\{B_k, P_{\max}\} \tag{3}$$

which satisfies $P_k \in \{0, 1, \ldots, \min\{B_k, P_{\max}\}\}$. The energy arrival process is also modeled as an FSMC in general. Hence, the transition probability can be similarly denoted by $q_{i,j} = \Pr(E_{k+1} = j|E_k = i), \forall i, j \in \{0, 1, \ldots, E_{\max}\}$, where $E_{\max}$ is the maximum energy arrival rate.

C. Problem Formulation

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

$$\lim_{K\to+\infty} \max_{P_1,P_2,\ldots,P_K} \frac{1}{K}\mathbb{E}\left[\sum_{k=1}^{K} r(P_k, \gamma_k)\right], \tag{4}$$

where $r(P_k, \gamma_k)$ is the per-slot data rate, the expectation is taken over all the possible energy harvesting processes $\{E_k|k = 1, 2, \ldots\}$ and the channel variations $\{\gamma_k|k = 1, 2, \ldots\}$, and the maximization operation is taken over all the power allocation policies $\{P_k|k = 1, 2, \ldots\}$. Due to the nature of Markov chain as well as the capability of energy storage, the power allocation among slots correlates with one another. As a result, the optimization over time should be jointly considered. MDP [27] is a general framework to deal with such a time dependent decision making problem, and DP is an effective algorithm to solve this problem. In the next section, we will show how to apply the DP algorithm into the average rate maximization problem.

III. Optimal Control for Average Rate Maximization

In this section, the basic idea of DP algorithm for MDP problems is firstly introduced, based on which the optimal DP algorithm to solve the problem (4) is proposed.

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

$$J^*(x_0) = \max_{\pi} \mathbb{E}\left[\sum_{k=0}^{K-1} g(x_k, u_k, w_k) + g(x_K)\right], \tag{5}$$

where the expectation is taken over all the random parameters, and the maximization is taken over all the possible policies $\pi = \{\mu_0, \mu_1, \ldots, \mu_{K-1}\}$ 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 [27, Vol. I, Prop. 1.3.1], which recursively proceeds the following optimizations

$$J_K(x_K) = g_K(x_K), \tag{6}$$

$$J_n(x_k) = \max_{u_k\in\mathcal{U}(x_k)} \mathbb{E}\{g(x_k,u_k,w_k)+J_{k+1}(f(x_k,u_k,w_k))\},$$
$$k = K-1, \ldots, 1, 0, \tag{7}$$

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

$$J^*(x_0) = \max_{\pi} \lim_{K\to\infty} \frac{1}{K}\mathbb{E}\left[\sum_{k=0}^{K-1} g(x_k, u_k, w_k)\right]. \tag{8}$$

In general, the optimal average reward is independent of x_0 [27, 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

$$\lambda + h(i) = \max_{u\in\mathcal{U}(i)} \left[g(i, u) + \sum_{j\in\mathcal{S}} p_{ij}(u)h(j)\right], \tag{9}$$

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 (9) for all i, the stationary policy

[1] In this paper, we use the term "reward" instead of *cost* as in [27], since we consider the rate maximization problem, rather than minimization problems.

$\pi^* = \{\mu^*, \mu^*, \ldots\}$ is optimal, where μ^* is a mapping from state space to action space.

The problem (8) can be solved by the *value iteration algorithm* [27, Vol. II, Sec. 4.4] based on (9). 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

$$h^{(k+1)}(i) = \max_{u \in \mathcal{U}(i)} \left[g(i,u) + \sum_{j \in \mathcal{S}} p_{ij}(u) h^{(k)}(j) \right] - \max_{u \in \mathcal{U}(s)} \left[g(s,u) + \sum_{j \in \mathcal{S}} p_{sj}(u) h^{(k)}(j) \right]. \quad (10)$$

Under the aperiodic-type conditions [27, 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

$$\underline{c}_k = \min_i \left[h^{(k+1)}(i) + \lambda^{(k+1)} - h^{(k)}(i) \right], \quad (11)$$

$$\overline{c}_k = \max_i \left[h^{(k+1)}(i) + \lambda^{(k+1)} - h^{(k)}(i) \right], \quad (12)$$

we have $\underline{c}_k \leq \underline{c}_{k+1} \leq \lambda \leq \overline{c}_{k+1} \leq \overline{c}_k$ [27, Vol. II, Prop. 4.3.3], which is used as the stopping criterion. For the general case, a variant iteration scheme is given as

$$\begin{aligned} h^{(k+1)}(i) &= (1-\tau) h^{(k)}(i) + \max_{u \in \mathcal{U}(i)} \left[g(i,u) + \tau \sum_{j \in \mathcal{S}} p_{ij}(u) h^{(k)}(j) \right] \\ &- \max_{u \in \mathcal{U}(s)} \left[g(s,u) + \tau \sum_{j \in \mathcal{S}} p_{sj}(u) h^{(k)}(j) \right], \end{aligned} \quad (13)$$

where τ is a scalar such that $0 < \tau < 1$. The above equation is guaranteed to converge [27, Vol. II, Prop. 4.3.4], and $\max_{u \in \mathcal{U}(s)} [g(s,u) + \tau \sum_{j \in \mathcal{S}} p_{sj}(u) h^{(k)}(j)]$ converges to the optimal reward λ. The convergence speed can be controlled by tuning the parameter τ. In particular, small value of τ may result in slow convergence as the update difference between $h^{(k+1)}(i)$ and $h^{(k)}(i)$ is small. On the other hand, if the value of τ is close to 1, the convergence may be also slow as the aperiodic-type condition may not hold [27, Vol. II, Sec. 4.3]. Therefore, with the increase of τ, the convergence rate firstly increases and then decreases. We will show the behavior in the numerical results.

B. DP Algorithm for Average Rate Maximization

To formulate the average rate maximization problem as an MDP and solve it by DP algorithm, the system state, the control action, the per-stage reward function and the state transition need to be determined. In the conventional energy harvesting communication problem, channel state and energy state are sufficient for decision making. However, when different time scales are considered, how long the channel state and energy state has lasted needs to taken into account. Specifically, the system state is a five-dimensional parameter

$$x_k = (B_k, \gamma_k, E_k, m_k, n_k),$$

where B_k, γ_k and E_k are defined in Section II, $m_k \in \{1, 2, \ldots, M\}$ refers to the number of slots the current channel state γ_k has lasted, and $n_k \in \{1, 2, \ldots, N\}$ refers to the number of slots the current energy arrival rate E_k has lasted. The control action is the transmit power P_k in each slot, and the per-stage reward is $r(P_k, \gamma_k)$. According to the conditional probability, the state transition probability can be written as

$$\begin{aligned} &\Pr(x_{k+1} | x_k, P_k) \\ &\quad = \Pr(B_{k+1} | B_k, E_k, P_k) \Pr(\gamma_{k+1} | \gamma_k, m_k) \\ &\quad \times \Pr(E_{k+1} | E_k, n_k) \Pr(m_{k+1} | m_k) \Pr(n_{k+1} | n_k) \end{aligned} \quad (14)$$

where B_k is updated according to (2), i.e.,

$$\Pr(B_{k+1} | B_k, E_k, P_k) = \begin{cases} 1, & \text{if } B_{k+1} = \min\{B_k - P_k + E_k, B_{\max}\} \\ 0. & \text{otherwise} \end{cases} \quad (15)$$

m_k and n_k increase by one in each stage if they are not equal to M and N, respectively. We have

$$\Pr(m_{k+1} | m_k) = \begin{cases} 1, & \text{if } m_{k+1} = (m_k \mod M) + 1 \\ 0, & \text{otherwise} \end{cases} \quad (16)$$

$$\Pr(n_{k+1} | n_k) = \begin{cases} 1, & \text{if } n_{k+1} = (n_k \mod N) + 1 \\ 0, & \text{otherwise} \end{cases} \quad (17)$$

where $(x \mod y)$ is modulus operation which returns the remainder after division of x by y. Whether γ_k changes or not depends on the value of m_k. We have

$$\Pr(\gamma_{k+1} | \gamma_k, m_k) = \begin{cases} 1, & \text{if } m_k < M \text{ and } \gamma_{k+1} = \gamma_k \\ p_{i,j}, & \text{if } m_k = M \text{ and } \gamma_{k+1} = j, \gamma_k = i \\ 0, & \text{otherwise} \end{cases} \quad (18)$$

where $p_{i,j}$ is defined in (1) and $i, j \in \{1, 2, \ldots, L\}$. Similarly, if $n_k < N$, we have $E_{k+1} = E_k$, if $n_k = N$, E_k changes according to $q_{i,j}$, i.e.,

$$\Pr(E_{k+1} | E_k, n_k) = \begin{cases} 1, & \text{if } n_k < N \text{ and } E_{k+1} = E_k \\ q_{i,j}, & \text{if } n_k = N \text{ and } E_{k+1} = j, E_k = i \\ 0, & \text{otherwise} \end{cases} \quad (19)$$

To this end, the average rate maximization problem is formulated as an infinite horizon MDP, and can be solved by the value iteration as detailed in Algorithm 1.

In this algorithm, τ is a predefined scalar that satisfies $0 < \tau \leq 1$. If $\tau = 1$, we obtain the standard value iteration. If the standard one does not converge, $\tau < 1$ should be adopted. ϵ is the target accuracy. The convergence speed of the algorithm is determined jointly by the values of τ and ϵ.

Notice that it is hard to get some insight on the structure of the optimal solution via the DP algorithm. However, some intuition can be obtained based on asymptotic analysis. It is obvious that if the battery size is sufficiently large, the original problem degrades to the power allocation problem over fading channels under average power constraint, whose optimal solution follows water-filling structure. Thus, with finite battery size, the optimal solution is expected as follows. In general, it also follows the water-filling structure. But when the battery is full, the energy needs to be used greedily to avoid

Algorithm 1 Value Iteration Algorithm for Average Rate Maximization

Initialize $h^{(0)}(x) = 0, \forall x = (B, \gamma, E, m, n)$, $\lambda^{(0)} = 0, k = 0$.
repeat
1. Update the scalar λ as
$$\lambda^{(k+1)} = \max_P \left[r(x_0, P) + \tau \sum_{x'} \Pr(x'|x_0) h^{(k)}(x') \right],$$
where $x_0 = (B_0, \gamma_0, E_0, m_0, n_0)$ is a fixed state, and $0 \leq P \leq \min\{B_0, P_{\max}\}$.
2. Update the vector $\boldsymbol{h}$ as
$$h^{(k+1)}(x) = (1-\tau)h^{(k)} + \max_P \left[r(x, P) + \tau \sum_{x'} \Pr(x'|x) h^{(k)}(x') \right] - \lambda^{(k+1)}.$$
3. Update $k = k + 1$.

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

energy waste. As the battery size gets larger, full battery is less probably, and hence, the optimal solution is more like the water-filling structure.

C. Complexity Analysis

It can be easily found that the number of states is $(B_{\max} + 1) \times L \times (E_{\max} + 1) \times M \times N$, and the number of actions is $P_{\max} + 1$. As the system state is a five-dimensional vector, the state space may be very large if some of the elements are of large size. As a result, the DP optimal algorithm may encounter the *curse of dimensionality* [27]. In this case, low-complexity algorithms are urgently required. Actually, the time scale difference between channel fading and energy arrival can be exploited to design a low-complexity algorithm. Specifically, when the energy arrival rate changes slower than channel fading, one can firstly solve the power allocation problem for a fixed energy arrival rate, and then optimize energy scheduling among energy harvesting frames. The low-complexity algorithm design for this special case is detailed in the next section.

IV. Sub-Optimal Algorithm Design for the Special Case With $M = 1$

In the case that $M = 1$ and $N \geq 1$, i.e., the change of energy harvesting rate is no faster than the channel fading, the original problem can be decoupled into two sub-problems: finite horizon power allocation in each energy harvesting frame, and infinite horizon energy management among frames. For notation clarity, each energy harvesting frame is re-indexed as k, each slot in the energy harvesting frame is indexed as n, and the energy harvesting rate is simplified as $E_{k,n} = E_k, \forall n = 1, 2, \ldots, N$. Thus, the objective function (4) can be re-written as
$$\max \lim_{K \to +\infty} \frac{1}{K} \mathbb{E}_{E_k} \left[\sum_{k=1}^{K} \mathbb{E}_{\gamma_{k,n}} \left[\sum_{n=1}^{N} \frac{1}{N} r\big(P_{k,n}, \gamma_{k,n}\big) \right] \right]. \quad (20)$$

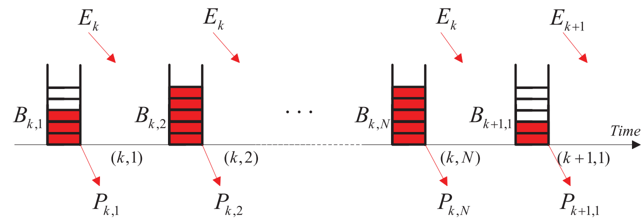

Fig. 2. System model of energy arrival, storage and power allocation for $M = 1$.

The energy arrival, storage and power allocation process is depicted in Fig. 2. Denote $B_k = B_{k,1}$ as the initial battery energy of the energy harvesting frame k. If the initial battery state B_k and the terminal battery state B_{k+1} are given, the optimal power allocation in frame k 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.

It is remarkable that the case with $M = 1$ can be extended to the case where N is any integer multiple of M. In particular, suppose $N = TM$, where T is an positive integer. As both energy arrival and channel state do not change during every M slots, they can be viewed as a single *super-slot* indexed by t. In each super-slot (k, t), the energy arrival becomes ME_k, the energy consumption is $MP_{k,t}, t = 1, 2, \ldots, T$ with constant transmit power $P_{k,t}$ for all M slots, the channel gain $\gamma_{k,t}$ is also constant for all M slots, and the per-super-slot reward is $\frac{1}{T} r(P_{k,t}, \gamma_{k,t})$. With super-slot as stage, the problem with $N = TM$ is the same as the special case with $M = 1$, and can be solved in the same way.

A. Inner Problem for Per-Frame Power Allocation

In the finite horizon inner problem, since the energy arrival rate E_k is known and constant over N slots, the system state can be denoted by $(B_{k,n}, \gamma_{k,n})$. The action is the transmit power $P_{k,n}$, and the per-stage reward is $\frac{1}{N} r(P_{k,n}, \gamma_{k,n})$. The state transition of $B_{k,n}$ is determined by (2), and that of $\gamma_{k,n}$ in each frame is determined by the state transition matrix $\mathbf{P}$.

The DP algorithm recursively solves the following per-slot optimization sub-problems
$$J_n\big(B_{k,n}, \gamma_{k,n}\big) = \max_{P_{k,n}} \left[\frac{1}{N} r\big(P_{k,n}, \gamma_{k,n}\big) + \sum_{\gamma} p_{\gamma_{k,n},\gamma} J_{n+1}\big(B_{k,n+1}, \gamma\big) \right] \quad (21)$$

under the constraint (3) for all $1 \leq n \leq N$. Recall that $p_{\gamma_{k,n},\gamma} = \Pr(\gamma|\gamma_{k,n})$ is the state transition probability. Denote $B_{k,N+1} = B_{k+1,1}$, and define the boundary value J_{N+1} as

$$J_{N+1}\big(B_{k,N+1}, \gamma_{k,N+1}\big) = \begin{cases} 0, & \text{if } B_{k,N+1} = B_{k+1} \\ -\infty. & \text{else} \end{cases} \tag{22}$$

Then the inner problem is solved recursively starting from slot (k, N). In this slot, we have

$$\begin{aligned} &J_N\big(B_{k,N}, \gamma_{k,N}\big) \\ &= \max_{P_{k,N}} \left[\frac{1}{N} r\big(P_{k,N}, \gamma_{k,N}\big) + \sum_{\gamma} p_{\gamma_{k,N},\gamma} J_{N+1}(B_{k,N+1}, \gamma) \right] \\ &= \begin{cases} \frac{1}{N} r\big(B_{k,N} + E_k - B_{k+1}, \gamma_{k,N}\big), \\ \qquad \text{if } B_{k+1} - E_k \leq B_{k,N} \leq B_{k+1} - E_k + P_{\max} \\ -\infty, \qquad \text{else} \end{cases} \end{aligned} \tag{23}$$

where the condition is obtained based on $B_{k,N+1} = B_{k+1}$ and $0 \leq P_{k,N} \leq P_{\max}$.

Similarly, the rest $J_n(B_{k,n}, \gamma_{k,n}), n < N$ can be calculated recursively according to (21). Finally, when $n = 1$, the expected optimal reward $J_1(B_{k,1}, \gamma_{k,1})$ in frame k starting from initial state $(B_{k,1}, \gamma_{k,1})$ is obtained for given energy arrival rate E_k and remained energy B_{k+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 $\gamma_k = \gamma_{k,1}$. At the beginning of frame k, the transmitter is aware of the initial battery energy B_k reserved in the last frame, the new energy arrival rate E_k, and the current channel gain γ_k, 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_{k+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.,

$$G(B_k, E_k, \gamma_k, B_{k+1}) = J_1(B_k, \gamma_k), \tag{24}$$

where J_1 is calculated by the last subsection giving that E_k is constant and B_{k+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 $p_{i,j}^{(N)}$. 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

$$\begin{aligned} &\lambda + h(b, e, \gamma) \\ &= \max_{b'} \left[G\big(b, e, \gamma, b'\big) + \sum_{\gamma'} p_{\gamma,\gamma'}^{(N)} \sum_{e'} q_{e,e'} h\big(b', e', \gamma'\big) \right], \end{aligned} \tag{25}$$

the optimal average rate is λ. The value iteration algorithm to solve the optimal average rate is similar to Algorithm 1.

C. Complexity Analysis

We now compare the complexity of the suboptimal algorithm with that of the DP optimal algorithm. In the inner problem, the number of states is $(B_{\max} + 1) \times L$ per-stage. As there are N stages, the total number of states is $(B_{\max} + 1) \times L \times N$. In the outer problem, the number of states is $(B_{\max} + 1) \times L \times (E_{\max} + 1)$. As the two subproblems are solved successively, compared with the one for DP optimal algorithm in Section III-C where the number of states is $(B_{\max}+1)\times L\times(E_{\max}+1)\times N$, the complexity of the proposed two-stage DP algorithm is much lower than the DP optimal algorithm. Numerical results shown later also show that the performance of two-stage DP algorithm can be very close to the DP optimal algorithm.

D. Case Study: Binary Energy Arrival and Channel State

For the special case that both energy arrival and channel fading are binary, some analytical results can be derived. Specifically, both the channel gain and the energy/power take only two values, i.e., the channel gain $\gamma_{k,n} \in \{0, 1\}$, the energy arrival rate $E_k \in \{0, 1\}$, and the transmit power $P_{k,n} \in \{0, 1\}$. The channel fading is assumed i.i.d., and $\Pr(\gamma_{k,n} = 1) = p$, where $\gamma_{k,n} = 1$ refers to a good channel, and the corresponding data rate is high, denoted as $r(1, 1) = r_1$. On the other hand, $\Pr(\gamma_{k,n} = 0) = 1 - p$, where $\gamma_{k,n} = 0$ refers to a bad channel, which results in a much lower data rate denoted as $r(1, 0) = r_0$. We have $r_1 > r_0$, and the data rate with $P_{k,n} = 0$ is zero. The binary energy arrival means that there is either a unit energy ($E_k = 1$) or nothing ($E_k = 0$) that can be harvested in each slot of energy harvesting frame k. The energy arrival process is also assumed i.i.d. with probability $\Pr(E_k = 1) = q$ and $\Pr(E_k = 0) = 1 - q$.

Firstly, consider the inner problem for per-frame power allocation to calculate the function $G(B_k, E_k, \gamma_k, B_{k+1})$. We have the following observation.

Lemma 1: With binary model, if $E_k = 0$, we have $B_{k,n} \leq B_k, \forall n = 1, 2, \ldots, N$. If $E_k = 1$ on the other hand, we have $B_{k,n} \leq B_{k+1}, \forall n = 1, 2, \ldots, N$.

Proof: The case that $E_k = 0$ is trivial as the battery energy will not increase. In the case that $E_k = 1$, the battery energy increases if $P_{k,n} = 0$. It continuously increases if the transmitter keeps silent in consecutive slots. However, to guarantee the battery state by the end of an energy harvesting frame equals to B_{k+1}, the transmitter has to transmit in the slots $n \geq n_0$ once $B_{k,n_0} = B_{k+1}$ for some $n_0 \leq N$. Hence, the lemma is proved. ■

Based on Lemma 1, the battery capacity is not violated through a whole energy harvesting frame if it is not violated by the end of the frame. Thus, the action B_{k+1} can be equivalently replaced by the number of transmissions in each frame, and G can be calculated based on the number of transmissions. Denote $f(T, N, \gamma_0)$ as the expected maximum sum rate by transmitting T times in N slots for a given initial channel state γ_0, where $T \leq N$. We have

$$\begin{aligned} &G(B_k, E_k, \gamma_k, B_{k+1}) \\ &= \begin{cases} \frac{1}{N} f(T, N, \gamma_k), & \text{if } B_k \geq 1 \text{ and } 0 \leq T \leq N, \\ \frac{1}{N}(pf(T, N-1, 1) + (1-p)f(T, N-1, 0)), & \\ & \text{if } B_k = 0 \text{ and } 0 \leq T \leq N-1, \end{cases} \end{aligned} \tag{26}$$

where $T = B_k + NE_k - B_{k+1}$, and the function f can be calculated according to the following theorem.

Theorem 1: With binary model, the function f can be calculated recursively as

$$\begin{aligned} f(T, N, 0) &= Tr_1 - (r_1 - r_0)(1-p)^{N-T} \\ &\quad \times \left(T + \sum_{i=1}^{T-1} (T-i) p^i \sum_{j=1}^{N-T-1} a_{i,j} \right), \end{aligned} \tag{27}$$

$$f(T, N, 1) = f(T-1, N, 0) + r_1 \tag{28}$$

for all $1 \leq T < N$, where $a_{1,j} = 1, \forall j$, $a_{i,j} = \sum_{s=1}^{j} a_{i-1,s}, \forall i \geq 2$, and the boundary conditions are

$$f(0, N, 0) = f(0, N, 1) = 0, \tag{29}$$

$$f(N, N, 1) = r_1 + (N-1)(pr_1 + (1-p)r_0), \tag{30}$$

$$f(N, N, 0) = r_0 + (N-1)(pr_1 + (1-p)r_0). \tag{31}$$

Proof: See the Appendix. ∎

In this special case with binary energy arrival and channel fading, the sub-optimality of the proposed two-stage DP algorithm can be easily explained. Suppose $p = q$ and the battery capacity is sufficiently large, obviously, the optimal policy is only to transmit when the channel is good. While in the two-stage DP algorithm, as the number of transmissions in each energy harvesting frame is pre-determined at the beginning of the frame, it is inevitable that some data will be transmitted in bad channel. Nevertheless, with a larger N, the gap between the pre-determined number of transmissions and the good channel realizations will diminish.

V. Numerical Simulations

We run some numerical simulations to analyze the influence of different timescales. Two types of parameter settings are considered: (1) The channel states are finite and i.i.d., and the energy arrival rate follows a Poisson distribution; (2) Both the channel and energy arrival states are binary. The detailed numerical results are shown as follows.

A. Poisson Energy Arrival and i.i.d. Channel

In this settings, Rayleigh channel distribution with average $\gamma_{\text{avg}} = 10$ dB is adopted, i.e., the density function of channel gain is

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

The channel gain is measured with reference transmit power $P_{\text{ref}} = 5$. The whole region of γ is partitioned into $K = 6$ non-overlapping consecutive intervals according to [28] so that the probability of the channel gain lying in each interval is equal. Each interval is represented by its average value. The channel

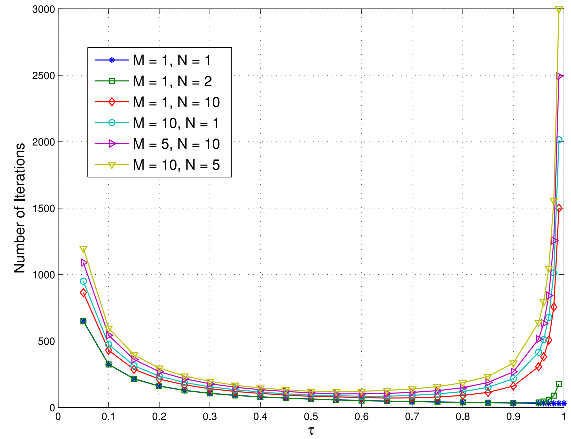

Fig. 3. Number of iterations versus τ.

state changes among these K average values with independent and identical state transition probability. Hence, we have

$$p_{i,j} = \frac{1}{K}. \quad \forall i, j$$

The energy arrival is ax Poisson process with average rate E_{avg}. Hence, we have

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

which is irrelevant to the previous state i. We set $E_{\text{avg}} = 5$, $P_{\max} = 10$, and adopt Shannon's equation to calculate the rate function, i.e.,

$$r(P_{k,n}, \gamma_{k,n}) = \log_2\left(1 + \frac{P_{k,n}}{P_{\text{ref}}} \gamma_{k,n}\right).$$

Firstly, the convergence speed for variable length of energy harvesting frame and channel fading frame is shown in Fig. 3. The battery capacity is set to $B_{\max} = 20$. It can be seen that when $M = N = 1$, the standard value iteration algorithm ($\tau = 1$) converges, and the convergence speed is fastest compared with $\tau < 1$. However, when $M \neq 1$ or $N \neq 1$, the standard value iteration algorithm does not converge. The reason is probably that there are periodic states caused by the repetitions of energy arrival or channel fading. Thus, the generalized algorithm with $\tau < 1$ is necessary. In addition, it can be found that the fastest convergence speed is achieved with $\tau \approx 0.6$, with which the iteration step length is well controlled to speed up while guaranteeing the convergence. Thus, we adopt $\tau = 0.6$ in the rest of the simulations.

Secondly, the average rates with different values of N and M are depicted in Figs. 4 and 5. It is seen that with fixed M, the average rates decrease as N increases. 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 may not be fully utilized. For instance, when the energy arrival rate changes slower than the channel fading and

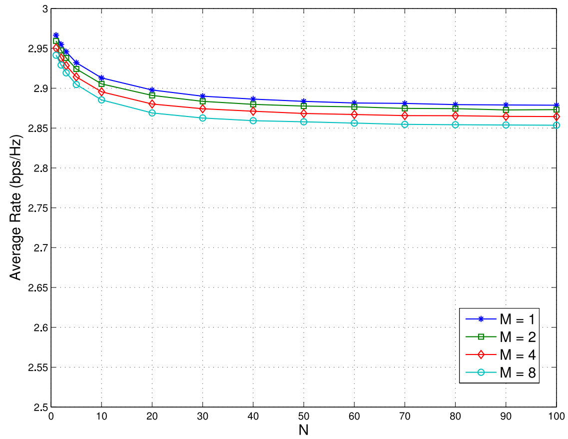

Fig. 4. Average rate versus N for Poisson energy arrival and i.i.d. channel.

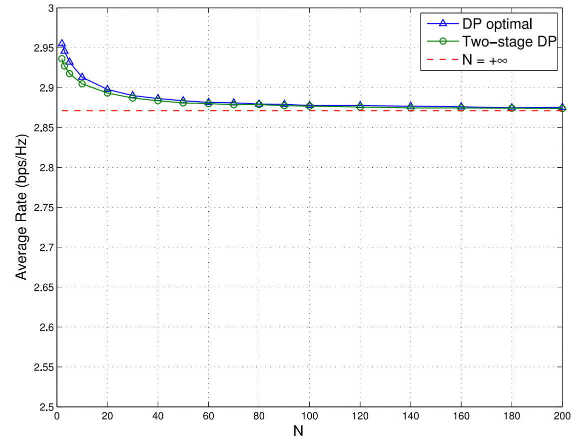

Fig. 6. Comparison of two-stage DP algorithm with DP optimal algorithm for Poisson energy arrival and i.i.d. channel with $M = 1$.

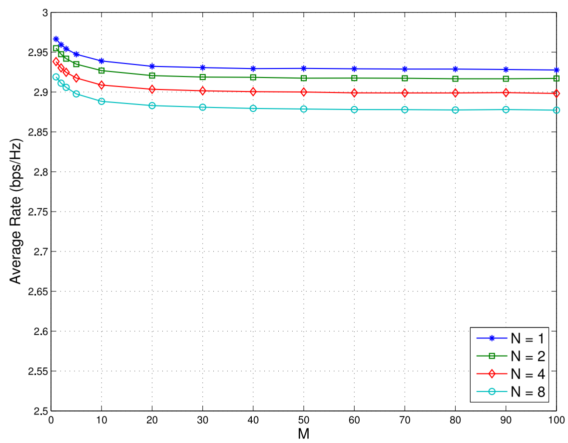

Fig. 5. Average rate versus M for Poisson energy arrival and i.i.d. channel.

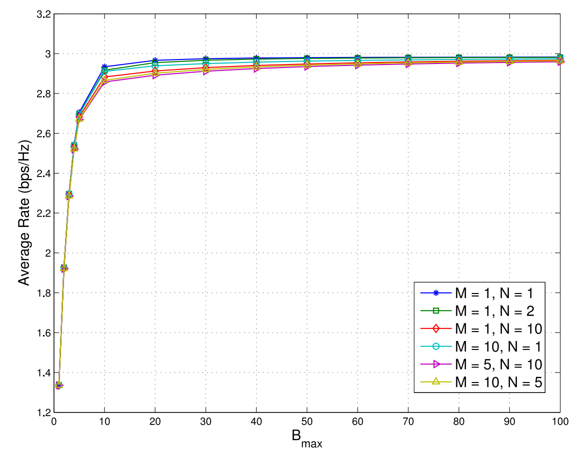

Fig. 7. Average rate versus $B_{\max}$ for Poisson energy arrival and i.i.d. channel.

it equals to 0 in an energy harvesting frame, the data cannot be transmitted even though the channel is good. Similarly, with fixed N, the average rate decreases as M increases. As N or M is sufficiently large, the average rate converges to a fixed value, which can be calculated as the weighted sum of ergodic rate of each energy harvesting frame or channel fading frame, where the weight is the steady state distribution of energy state transition and channel state transition, respectively.

Then for $M = 1$, the low-complex two-stage DP algorithm is compared with the DP optimal algorithm in Fig. 6. It is shown that when $N = 2$, the performance gap is the largest. As N increases, the gap diminishes. The two curves gradually converge to a fixed value given by the red dashed line. Since the outer stage of the two-stage DP algorithm is optimized based on the average performance of the inner stage, with larger value of N, the average performance will approximate each realization much better. This is the reason why the gap diminishes as N increases. In addition, as $N \to \infty$, the average performance of the inner stage equals to the performance of a realization with a long trajectory.

Finally, the performance versus the battery capacity is illustrated in Fig. 7. It can be seen that the average rate increases as the battery capacity becomes larger. It is greatly reduced when the battery capacity is lower than the average energy arrival rate, i.e., $B_{\max} < E_{\text{avg}} = 5$, as a large amount of energy is wasted due to the battery overflow. When $B_{\max} > E_{\text{avg}}$, the change of rate is marginal. When the battery capacity is sufficiently large, the curves with different values of M and N converges to the same value. This is because larger battery capacity can eliminate the dynamics in energy domain, and hence, the maximum average rate for any M and N is equivalent to the capacity of fading channel.

B. Binary Channel State and Energy Arrival

We further study the performance with binary setup. The channel state is good with probability $p = 0.7$, and is bad with probability $1-p = 0.3$. There is a constant energy arrival with probability $q = 0.7$, and is no energy arrival with probability $1 - q = 0.3$. In the good channel state, the received SNR is 20 dB, and in the bad channel state, the received SNR is 0 dB. The battery capacity is 4.

The performance versus M and N are depicted in Figs. 8 and 9 with $B_{\max} = 4$. It can be found that similar to the previous section, the average rate decreases as M or N increases. However, the difference is that all the curves converge to the same value as M or N is sufficiently large. It

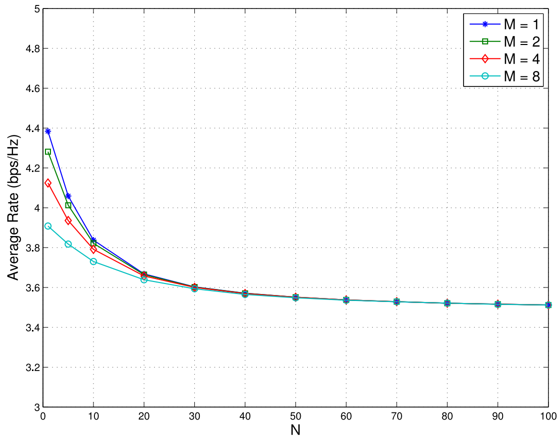

Fig. 8. Average rate versus N for binary energy arrival binary channel.

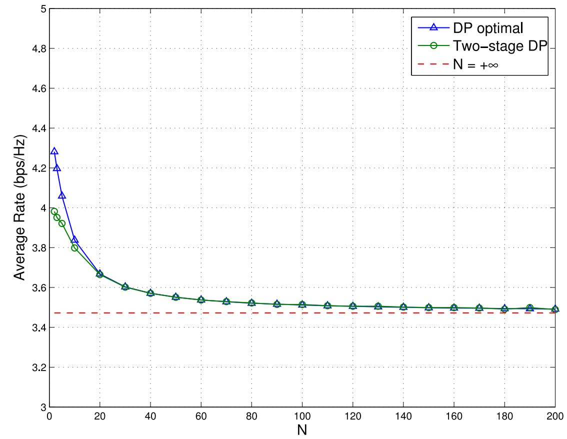

Fig. 10. Comparison of the two-stage DP algorithm with the DP optimal algorithm for binary energy arrival binary channel with $M = 1$.

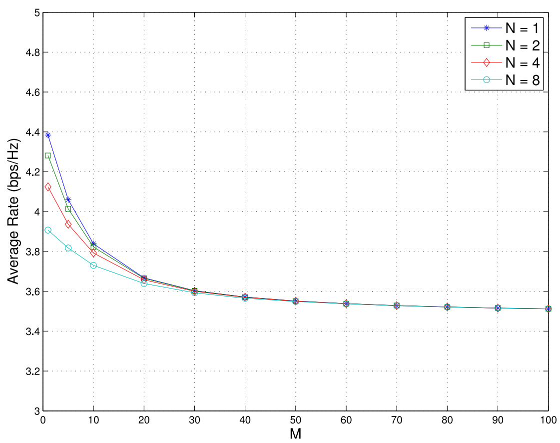

Fig. 9. Average rate versus M for binary energy arrival binary channel.

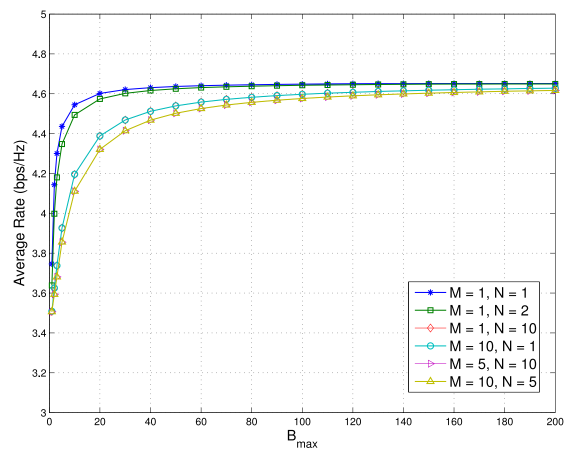

Fig. 11. Average rate versus $B_{\max}$ for binary energy arrival binary channel.

means that the change in M or N has little impact on the average rate when the other parameter is large. This is due to the nature of binary settings. For example, when N is sufficiently large, the energy arrival rate is either 0 or 1 for a long time. If it is 0, the transmit power is 0 for almost all the slots. On the other hand, if it is 1, the transmit power is almost always 1. Thus, the average rate is irrelevant to the time scale of channel fading, but is determined by the channel distribution.

For $M = 1$, the two-stage DP algorithm is compared with the DP optimal algorithm in Fig. 10. It can be found that the proposed low-complex algorithm converges to the optimal one very quickly. When $N > 10$, they perform almost the same. As a result, the proposed two-stage DP algorithm can achieve the close-to-optimal performance in most cases. Besides, both the curves converge to a steady value (shown by the red dashed line) when N is sufficiently large.

At last, the impact of battery capacity on the performance is shown in Fig. 11. For all the simulated parameters, the curves converge when $B_{\max} > 100$, and the convergence speed is faster when the values of M and N are smaller. It can also be seen that the curves with $M = 1, N = 10$ and $M = 10, N = 1$ overlap with each other, and those with $M = 5, N = 10$ and $M = 10, N = 5$ overlap as well. That is, the influence of M and N is symmetric. The reason is that with binary settings, the behaviors in a channel fading frame or an energy harvesting frame are similar. If $M < N$, the transmitter in an energy harvesting frame with $E_k = 1$ greedily uses energy to avoid battery overflow. Similarly, if $M > N$, the transmitter in a channel fading frame with $\gamma_k = 1$ also greedily uses energy to fully utilize the good channel. Consequently, the influence on the average data rate is the same.

VI. Conclusion

For the power allocation problem with two-dimensional dynamics including channel fading and energy arrival variation of different timescales, the lengths of energy harvesting frame and channel fading frame have significant impact on the average rate performance. The longer the lengths are, the less the chance for opportunistic scheduling, i.e., the energy has to be used in bad channel to avoid battery overflow or the good channel is not used for transmission due to energy deficiency. Hence, the average rate is lower. Full opportunistic power allocation can be explored if both the energy arrival rate and the channel state change every slot. It tells us that the dynamics of energy arrival actually benefit the energy harvesting communications. Besides, asymptotic behaviors are also observed. In particular, the average rate converges to a lower bound as

M or N tends to infinity, and converges to an upper bound as the battery capacity $B_{\max}$ tends to infinity. The performance bound can be used to guide the system design and performance evaluation.

APPENDIX
PROOF OF THEOREM 1

To achieve the maximum data rate, the slots with good channel should be chosen for transmission. Thus, the optimal transmission policy is obvious. That is, if the remaining slots are strictly larger than the remaining transmission times, the data is transmitted if the channel is good, and is not transmitted otherwise. Once the remaining slots are equal to the remaining transmission times, the data is transmitted in all remaining slots.

Firstly, the boundary values can be calculated straightforward. Then for the case that $1 \le T < N$, the conclusion can be proved by induction as follows.

For $T = 1$ and any $N > 1$, it is obvious that $f(1, N, 1) = r_1 = f(0, N, 0) + r_1$, where the high data rate r_1 is achieved in the first slot with $\gamma_0 = 1$. Thus, (28) is validated for $T = 1$ and $N \ge 1$. In addition, we have for $N > 1$,

$$\begin{aligned} f(1,N,0) &= pf(1,N-1,1) + (1-p)f(1,N-1,0) \\ &= pr_1 + (1-p)(pr_1 + (1-p)f(1,N-2,0)) \\ &= pr_1(1+(1-p)) + (1-p)^2 f(1,N-2,0) \\ &= pr_1\Big(1 + (1-p) + \ldots + (1-p)^{N-2}\Big) \\ &\quad + (1-p)^{N-1} f(1,1,0) \\ &= r_1\Big(1-(1-p)^{N-1}\Big) + (1-p)^{N-1} r_0 \\ &= r_1 - (r_1 - r_0)(1-p)^{N-1}. \end{aligned} \tag{32}$$

As a result, (27) is also validated for $T = 1$ and $N > 1$.

Suppose (27) and (28) hold for some T and any $N > T$, we now prove that they also hold for $T + 1$ and any $N > T + 1$. We have

$$\begin{aligned} &f(T+1,N,1) \\ &= r_1 + pf(T,N-1,1) + (1-p)f(T,N-1,0) \\ &= r_1 + p(r_1 + f(T-1,N-1,0)) + (1-p)f(T,N-1,0) \\ &= (T+1)r_1 - (r_1-r_0)(1-p)^{N-T} \\ &\quad \times \left(T + \sum_{i=1}^{T-1}(T-i)p^i \sum_{j=1}^{N-T-1} a_{i,j}\right) \\ &= r_1 + f(T,N,0), \end{aligned} \tag{33}$$

and similar to the derivation of (32),

$$\begin{aligned} &f(T+1,N,0) \\ &\quad = pf(T+1,N-1,1) + (1-p)f(T+1,N-1,0) \\ &\quad = pf(T,N-1,0) + pr_1 + (1-p)f(T+1,N-1,0) \\ &\quad = p\Big(f(T,N-1,0)+\cdots+(1-p)^{N-K-3}f(T,T+2,0)\Big) \\ &\qquad + pr_1\Big(1+\ldots+(1-p)^{N-T-3}\Big) \\ &\qquad + (1-p)^{N-T-2} f(T+1,T+2,0) \\ &\quad = p\Big(f(T,N-1,0)+\cdots+(1-p)^{N-K-3}f(T,T+2,0)\Big) \\ &\qquad + r_1\Big(1-(1-p)^{N-T-2}\Big) \\ &\qquad + (1-p)^{N-T-2}(T+1)(pr_1 + (1-p)r_2) \\ &\quad = (T+1)r_1 - (r_1-r_0)(1-p)^{N-T-1} \\ &\qquad \times \left((T+1) + \sum_{i=1}^{T}(T+1-i)p^i \sum_{j=1}^{N-T-2} a_{i,j}\right). \end{aligned} \tag{34}$$

That is, (27) and (28) hold for $T + 1$ and any $N > T + 1$. Therefore, (27) and (28) hold for all $1 \le T < N$. The proof is completed.

REFERENCES

[1] J. Gong, Z. Zhou, and S. Zhou, "Analysis and optimization of wireless transmissions over fast fading channels with slow time-varying energy arrival," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, May 2017, pp. 1–6.

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

[3] T. Han and N. Ansari, "On optimizing green energy utilization for cellular networks with hybrid energy supplies," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 3872–3882, Aug. 2013.

[4] J. Gong, J. S. Thompson, S. Zhou, and Z. Niu, "Base station sleeping and resource allocation in renewable energy powered cellular networks," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 3801–3813, Nov. 2014.

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

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

[7] J. Gong, S. Zhou, and Z. Niu, "Optimal power allocation for energy harvesting and power grid coexisting wireless communication systems," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 3040–3049, Jul. 2013.

[8] M. A. Antepli, E. Uysal-Biyikoglu, and H. Erkal, "Optimal packet scheduling on an energy harvesting broadcast link," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1721–1731, Sep. 2011.

[9] J. Yang, O. Ozel, and S. Ulukus, "Broadcasting with an energy harvesting rechargeable transmitter," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 571–583, Feb. 2012.

[10] J. Yang and S. Ulukus, "Optimal packet scheduling in a multiple access channel with energy harvesting transmitters," *J. Commun. Netw.*, vol. 14, no. 2, pp. 140–150, 2012.

[11] M. Gregori and M. Payaró, "On the precoder design of a wireless energy harvesting node in linear vector Gaussian channels with arbitrary input distribution," *IEEE Trans. Commun.*, vol. 61, no. 5, pp. 1868–1879, May 2013.

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

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

[14] D. Niyato, E. Hossain, and A. Fallahi, "Sleep and wakeup strategies in solar-powered wireless sensor/mesh networks: Performance analysis and optimization," *IEEE Trans. Mobile Comput.*, vol. 6, no. 2, pp. 221–236, Feb. 2007.

[15] M. Gatzianas, L. Georgiadis, and L. Tassiulas, "Control of wireless networks with rechargeable batteries [transactions papers]," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 581–593, Feb. 2010.

[16] S. Chen, P. Sinha, N. B. Shroff, and C. Joo, "Finite-horizon energy allocation and routing scheme in rechargeable sensor networks," in *Proc. IEEE INFOCOM*, Shanghai, China, 2011, pp. 2273–2281.

[17] N. Pappas, J. Jeon, A. Ephremides, and A. Traganitis, "Optimal utilization of a cognitive shared channel with a rechargeable primary source node," *J. Commun. Netw.*, vol. 14, no. 2, pp. 162–168, Apr. 2012.

[18] I. Krikidis, T. Charalambous, and J. S. Thompson, "Stability analysis and power optimization for energy harvesting cooperative networks," *IEEE Signal Process. Lett.*, vol. 19, no. 1, pp. 20–23, Jan. 2012.

[19] B. T. Bacinoglu and E. Uysal-Biyikoglu, "Finite-horizon online transmission scheduling on an energy harvesting communication link with a discrete set of rates," *J. Commun. Netw.*, vol. 16, no. 3, pp. 293–300, Jun. 2014.

[20] S. Zhou, T. Chen, W. Chen, and Z. Niu, "Outage minimization for a fading wireless link with energy harvesting transmitter and receiver," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 3, pp. 496–511, Mar. 2015.

[21] C. Huang, J. Zhang, P. Zhang, and S. Cui, "Threshold-based transmissions for large relay networks powered by renewable energy," in *Proc. IEEE Glob. Commun. Conf. (Globecom)*, Atlanta, GA, USA, Dec. 2013, pp. 1921–1926.

[22] H. Li, C. Huang, S. Cui, and J. Zhang, "Distributed opportunistic scheduling for wireless networks powered by renewable energy sources," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Toronto, ON, Canada, 2014, pp. 898–906.

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

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

[25] 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 *Proc. IEEE Glob. Conf. Signal Inf. Process. (GlobalSIP)*, Atlanta, GA, USA, Dec. 2014, pp. 103–107.

[26] X. Chen and W. Chen, "Joint channel-buffer aware energy-efficient scheduling over fading channels with short coherent time," in *Proc. IEEE Int. Conf. Commun. (ICC)*, London, U.K., Jun. 2015, pp. 2810–2815.

[27] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, MA, USA: Athena Sci., 2005.

[28] Q. Liu, S. Zhou, and G. B. Giannakis, "Queuing with adaptive modulation and coding over wireless links: Cross-layer analysis and design," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 1142–1153, May 2005.

Jie Gong (S'09–M'13) received the B.S. and Ph.D. degrees from the Department of Electronic Engineering, Tsinghua University, Beijing, China, in 2008 and 2013, respectively. From 2012 to 2013, he visited the Institute of Digital Communications, University of Edinburgh, Edinburgh, U.K. From 2013 to 2015, he was a Post-Doctoral Scholar with the Department of Electronic Engineering, Tsinghua University, Beijing, China. He is currently an Associate Research Fellow with the School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China. His research interests include Cloud RAN, energy harvesting technology and green wireless communications. He was a co-recipient of the Best Paper Award from IEEE Communications Society Asia–Pacific Board in 2013. He served as the Workshop Publicity Co-Chair for the IEEE WCNC 2018 Workshop on Intelligent Computing and Caching at the Network Edge, and a TPC member for the IEEE/CIC ICCC 2016 and 2017, the IEEE WCNC 2016, 2017, and 2018, the IEEE Globecom 2017, the APCC 2017, and the ICC 2018. He was selected as an IEEE Wireless Communications Letters Exemplary Reviewer in 2016.

Zhenyu Zhou (M'11–SM'17) received the M.E. and Ph.D. degrees from Waseda University, Tokyo, Japan, in 2008 and 2011, respectively. From 2012 to 2013, he was the Chief Researcher with the Department of Technology, KDDI, Tokyo. Since 2013, he has been an Associate Professor with the School of Electrical and Electronic Engineering, North China Electric Power University, China. Since 2014, he has also been a Visiting Scholar with Tsinghua-Hitachi Joint Laboratory on Environment-Harmonious ICT, University of Tsinghua, Beijing. His research interests include green communications, vehicular communications, and smart grid communications. He was a recipient of the IEEE Vehicular Technology Society Young Researcher Encouragement Award in 2009, the Beijing Outstanding Young Talent Award in 2016, the IET Premium Award in 2017, and the IEEE ComSoc Green Communications and Computing Technical Committee 2017 Best Paper Award. He served as an Associate Editor for IEEE ACCESS, and a Guest Editor for the *IEEE Communications Magazine* and TRANSACTIONS ON EMERGING TELECOMMUNICATIONS TECHNOLOGIES. He also served as the Workshop Co-Chair for IEEE ISADS 2015, and a TPC member for IEEE Globecom, IEEE CCNC, IEEE ICC, IEEE APCC, IEEE VTC, and IEEE Africon. He is a voting member of the P1932.1 Working Group.

Sheng Zhou (S'06–M'12) received the B.E. and Ph.D. degrees in electronic engineering from Tsinghua University, Beijing, China, in 2005 and 2011, respectively. In 2010, he was a visiting student with the Wireless System Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA, USA, for five months. From 2014 to 2015, he was a Visiting Researcher with the Central Research Laboratory, Hitachi Ltd., Japan. He is currently an Associate Professor with the Department of Electronic Engineering, Tsinghua University. His research interests include cross-layer design for multiple antenna systems, mobile edge computing, and green wireless communications.