Optimal Energy Management for Energy Harvesting Transmitters under Battery Usage Constraint

Xianwen Wu, Jing Yang, and Jingxian Wu
Department of Electrical Engineering,
University of Arkansas, Fayetteville, AR, 72701, U.S.A.
xw002@email.uark.edu *jingyang@uark.edu* *wuj@uark.edu*

***Abstract*—This paper takes the impact of charging and discharging operations on battery degradation into consideration, and studies the optimal energy management policy for an energy harvesting communication system under a battery usage constraint. Specifically, in each time slot, we assume the harvested energy can be used to power the transmitter immediately without entering into the battery, or stored into the battery for now and retrieved later for transmission. Whenever the battery is charged or discharged, a cost will be incurred to account for its impact on battery degradation. We impose an long-term average cost constraint on the battery, which is translated to the average number of charge/discharge operations per unit time. The objective is to develop an online policy to maximize the long-term average throughput of the transmitter under energy causality constraint and the battery usage constraint.**

We first relax the energy causality constraint on the system, and impose an energy flow conservation constraint instead. We show that the optimal energy management policy has a double-threshold structure: if the amount of energy arrives in each time slot lies in between the two thresholds, it will be used immediately without involving the battery; otherwise, the battery will be charged or discharged accordingly to maintain a constant transmit power. We then modify the double-threshold policy slightly to accommodate the energy causality constraint, and analyze its long-term performance. We show that the system achieves the same long-term average performance, thus it is optimal.

I. INTRODUCTION

The random and intermittent nature of harvested energy imposes critical challenges on the design of sustainable and reliable energy harvesting wireless sensor networks. Rechargeable batteries are usually employed as an energy buffer to filter out the fluctuations in the energy harvesting process and maintain a continuous and stable energy output. A large number of energy management schemes have been proposed to optimize the performances of such systems.

Modeling the battery as an ideal energy buffer for energy storage and retrieval, researchers have developed various energy management schemes to optimize different performance metrics under infinite battery setting [1]–[3] and finite battery setting [4]–[9]. The performance metrics include channel capacity [3], transmission delay [1], throughput [4]–[6], etc.

However, modeling batteries as perfect energy buffers may not be realistic, since battery operations involve very complicated mechanisms, which lead to inevitable energy storage imperfections and battery degradation. In this context, some works aim to take more practical battery characteristics into the optimization framework, and investigate their impacts on the optimal energy management policies and system performances. In [10], the authors consider battery storage imperfections where stored energy leaks in time, and the battery degrades at the same time. An optimal throughput maximization policy is proposed under an offline setting. Reference [11] proposes a battery health model to capture the dependency of battery degradation on its discharge depth, and investigates degradation-aware policy to improve the lifetime of the battery while guaranteeing the minimum QoS requirement. The problem is casted into the framework of Markov Decision Processes, and solved independently for each health state by exploiting the timescale separation between the communication time-slot and the battery degradation process. [12] investigates the scenario where a portion of energy is lost instantaneously when it enters the battery, and proposes optimal offline transmission policies under various settings. The optimal policy has a double-threshold structure, where the battery charges/discharges when the harvested energy is above/below the thresholds and transmits with the corresponding threshold.

It has been shown that the battery lifetime is closely related to its charge/discharge cycles. Frequent battery charge/discharge operations result in irreversible battery capacity degradation and jeopardize its battery lifetime. In this paper, we take the impact of charge/discharge operations on battery lifetime into consideration, and study the optimal energy management policy for an energy harvesting communication system under a battery usage constraint. Specifically, in each time slot, we assume the harvested energy can be used to power the transmitter immediately without entering into the battery, or stored into the battery for now and retrieved later for transmission. Besides the energy causality constraints, we impose a battery cost constraint, which is translated into the average number of charge/discharge operations per unit time. The objective is to maximize the long-term average throughput of the transmitter under energy causality constraint and the battery usage constraint. We do not consider battery degradation explicitly in this setup, as we assume that the aging process happens over a time scale that is much longer than the communication period we consider about, and the battery storage capacity is always sufficiently

This work was supported in part by the U. S. National Science Foundation (NSF) under Grants ECCS-1405403, ECCS-1454471 and ECCS-1202075.

large to prevent any energy overflow in our setting.

We first relax the energy causality constraint on the system, and impose a long-term energy flow conservation constraint instead. We show that the optimal energy management policy has a double-threshold structure: if the amount of energy arrives in each time slot lies in between the two thresholds, it will be used immediately without involving the battery; otherwise, the battery will be charged or discharged accordingly to maintain a constant transmit power. We then modify the two-threshold policy slightly to accommodate the energy causality constraint, and analyze its long-term performance. We show that the system achieves the same long-term average performance, thus it is optimal.

Despite a similar double-threshold structure, our policy is fundamentally different from that studied in [12] due to different constraints we impose on the system. Essentially, under the battery inefficiency assumption that a ratio of the saved energy will be lost in [12], the *amount* of energy to be saved in the battery is the key factor, which can be identified by solving the standard convex optimization problem. While under the battery usage constraint, the *number* of charge/discharge operations matters. Thus, our optimization problem has a combinatorial flavor, which cannot be solved straightforwardly via convex optimization. As a result, under our policy, the transmitter always tries to equalize the transmit power whenever it charges or discharges, while in [12], the transmitter transmits with the corresponding thresholds.

II. System Model and Problem Formulation

Consider a time slotted energy harvesting communication system. Let A_t be the energy harvested from the ambient environment in time slot t, $t = 1, 2, \ldots, T$. A_ts are i.i.d random variables with known probability density function (pdf) $p_A(\cdot)$. Energy can be used to transmit data from a backlogged buffer, or stored in a battery for later use, as shown in Fig. 1. Let B_t be the amount of energy that enters the battery in time t, and C_t be the remaining amount from A_t. Then,

$$A_t = B_t + C_t \tag{1}$$

Let D_t be the energy drawn from the battery in time t. The total amount of energy used for transmission in time slot t is then equal to $P_t := D_t + C_t$. Then, the battery level evolves according to

$$E_{t+1} = E_t - D_t + B_t, \quad D_t \leq E_t \tag{2}$$

with $E_0 = 0$.

Assume the transmission rate is a concave function of P_t, denoted as $R(P_t)$. Our objective is to optimize the long-term average transmission rate under the energy causality constraint and the battery usage constraint, which is denoted as the expected number of charge/discharge operations per time slot. Then, the optimization problem is formulated as

$$\max_{\{C_t, D_t\}} \quad \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}[R(P_t)] \tag{3}$$

$$\text{s.t.} \quad (1)-(2), \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}\left(\mathbf{1}_{D_t} + \mathbf{1}_{B_t}\right) \leq \rho \tag{4}$$

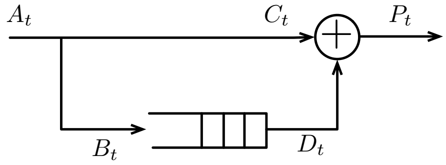

Fig. 1: System model

The expectations in the objective function and the constraint are taken over all possible energy harvesting sample paths. The optimization problem has a combinatorial flavor, as we need to decide in which time slots the system should charge or discharge the battery. Thus, to make the problem tractable, in the following, we will first relax the energy causality constraint and study the problem with a relaxed long-term energy flow conservation constraint for the battery. With the structured optimal energy management policy obtained for this case, we will propose a best-effort transmission policy which obeys the energy causality constraint and prove that it achieves the same performance as time T goes to infinity. Therefore, it is optimal.

III. Optimal Policy Without Causality Constraints

In the following, we will first consider a relaxed optimization problem, where we replace the energy causality constraint in (2) with the following long-term energy flow conservation constraint for the battery:

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} D_t \leq \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} B_t \tag{5}$$

Assume $\mathcal{Q}$ is the optimal policy satisfying the battery usage constraint in (4) and the energy flow conservation constraint in (5). In general, under $\mathcal{Q}$, the transmit power P_t may depend on the current energy arrival A_t, as well as the energy arrival and departure history up to $t-1$, denoted as H^{t-1}. With a little abuse of notation, in this section, we use P_t to denote the transmit power in time slot t under policy $\mathcal{Q}$. We assume P_t is a deterministic function of A_t and H^{t-1}, denoted as $P_t = Q(A_t, H^{t-1})$. In the following, we will identify the structural properties of $\mathcal{Q}$, and show that it can be explicitly obtained using a simple approach. Our analysis can be directly extended to handle any randomized policy as well.

Define $\mathcal{A}_t := \{(A_t, H^{t-1}) | A_t \neq Q(A_t, H^{t-1})\}$, $t = 1, 2, \ldots$, i.e., the set of states in which the battery charges or discharges in time slot t under $\mathcal{Q}$. Define

$$P_0 = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}[A_t | (A_t, H^{t-1}) \in \mathcal{A}_t], \tag{6}$$

i.e., the average amount of energy harvested during the states included in $\cup_{t=1}^{\infty}\mathcal{A}_t$. We assume the limit exists. Then, we have the following observations.

Lemma 1 *Under the optimal policy $\mathcal{Q}$, B_t and D_t cannot be positive in the same slot t.*

This is obvious due to the fact that if B_t and D_t are both positive, we can always adjust the values of B_t and D_t to make one of them to be zero, and achieve the same transmit power P_t with a reduced battery usage cost.

Lemma 2 *Under the optimal policy $\mathcal{Q}$, whenever the battery charges or discharges, the transmit power P_t should be a constant and equal to P_0.*

Lemma 2 can be proved by Jensen's inequality. Due to space limitations, all of the proofs except that of Theorem 1 will be omitted. Based on Lemmas 1 and 2, we have the following theorem.

Theorem 1 *The optimal policy under the relaxed long-term energy flow conservation constraint depends on the instantaneous energy arrival only, and has a double-threshold structure, i.e., if $A_t < \tau_1$, we must have $D_t = P_0 - A_t$, $P_t = P_0$; if $A_t > \tau_2$, we must have $B_t = A_t - P_0$, $P_t = P_0$, where P_0, τ_1 and τ_2 are the solution to the following optimization problem*

$$\max_{P_0,\tau_1,\tau_2} \quad R(P_0)\rho + \int_{\tau_1}^{\tau_2} R(x)p_A(x)dx \tag{7}$$

$$\text{s.t.} \quad \mathbb{P}[A_t > \tau_2] + \mathbb{P}[A_t < \tau_1] = \rho \tag{8}$$

$$\mathbb{E}[A_t - P_0 | A_t > \tau_2] = \mathbb{E}[P_0 - A_t | A_t < \tau_1] \tag{9}$$

$$\tau_1 \leq P_0 \leq \tau_2 \tag{10}$$

Theorem 1 can be proved through contradiction. Assume that $\mathcal{Q}$ does not have such double-threshold structure. Then, we can always construct another policy to outperform it without violating the constraints in (4) and (5). The detailed proof is provided in the Appendix.

Theorem 1 provides an upper bound on any energy management policy satisfying the energy causality constraint and the battery usage constraint.

Theorem 2 *The objective function (7) can be reduced to a function with a single variable τ_1. Moreover, it first increases then decrease in τ_1, and the maximum point corresponds to the optimal solution satisfying (8)-(10).*

Theorem 2 suggest a computationally efficient way to solve the optimization problem described in Theorem 1. Starting with $\tau_1 = 0$, we first solve (8)(9) to get τ_2 and P_0 and evaluate the objective function. We gradually increase τ_1, repeat the process, and keep track of the objective function value until we observe a decrease. The turning point corresponds the optimal solution.

IV. Optimal Policy under Causality Constraints

Let τ_1, τ_2, P_0 be the optimal solution to the optimization problem described in Theorem 1. Let $\mathcal{B} = [0, \tau_1] \cup [\tau_2, \infty]$. Then, we define a best-effort transmission policy as follows.

Definition 1 (Best-effort transmission policy) *In each time slot t, if $A_t \notin \mathcal{B}$, the transmitter transmits with the harvested energy A_t. Otherwise, if $A_t > \tau_2$, the battery is charged with amount $A_t - P_0$, and transmitter transmits with P_0; if $A_t < \tau_1$ and $E_t \neq 0$, the battery is discharged with amount $\min\{E_t, P_0 - A_t\}$, and the transmitter transmits with $\min\{E_t, P_0 - A_t\} + A_t$.*

We note that the energy causality constraint is always satisfied under the proposed best-effort transmission policy. Besides, the battery usage constraint is satisfied as well. Due to the energy causality constraint, the transmitter may not be able to transmit with power P_0 if $A_t < \tau_1$ and E_t is not sufficiently large. This may result in some performance degradation. However, as we will show in the following theorem, the probability of such scenario will decrease exponentially fast as T increases. Thus, the long-term average throughput will converge to that upper bound exponentially, which indicates the optimality of the proposed best-effort policy.

Define the planned charge/discharge process as

$$A_t^* = \begin{cases} A_t - P_0 & A_t \in \mathcal{B} \\ 0 & A_t \notin \mathcal{B} \end{cases} \tag{11}$$

Then, under the proposed best effort policy, we have $E_{t+1} = \max\{E_t + A_t^*, 0\}$, and the energy spent at t is

$$P_t = A_t + E_t - E_{t+1} \tag{12}$$

We define

$$Q_t = \begin{cases} P_0 & A_t \in \mathcal{B} \\ A_t & A_t \notin \mathcal{B} \end{cases} \tag{13}$$

Thus, $P_t \neq Q_t$ if and only if $E_t + A_t < P_0$, and $P_t \leq Q_t$, $\forall t$. Note that Q_t is exactly the optimal policy defined in Theorem 1.

Theorem 3 *Assume $|A_t| \leq M$ and $R(\cdot)$ is Lipschitz. Under the best-effort transmission policy,*

$$\lim_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} (Q_t - P_t) = 0, \quad a.s. \tag{14}$$

Theorem 3 indicates that the best-effort transmission policy converges to the optimal policy described in Theorem 1 almost surely.

Theorem 4 *The best-effort transmission policy achieves the upper bound on the long-term expected throughput characterized in Theorem 1 almost surely. Therefore, it is optimal.*

V. Numerical Results

In this section, we use numerical results to illustrate the proposed best-effort transmission policy and evaluate its performance.

We assume the energy arrivals are i.i.d. random variables uniformly distributed over $[0, 6]$. We let $\rho = 0.3$, i.e., the battery can only be charged or discharged for 30% of the time, and the rate function $R(x) = \frac{1}{2}\log(1 + x)$. We first numerical solve the equations in Theorem 1, and identify the corresponding thresholds $\tau_1 = 1.0158$, $\tau_2 = 5.2158$, and $P_0 = 2.7298$. The corresponding time-average transmission rate is 0.6761, which is the upper bound for any online policy.

We then plot one sample path of the energy arrivals for the first 20 time slots in Fig. 2, and indicate the corresponding transmit power under the proposed policy. As expected, the transmit power equals A_t if A_t falls between those two thresholds, and equals P_0 otherwise, except when $t = 9, 14$. In those time slots, the battery does not have sufficient energy

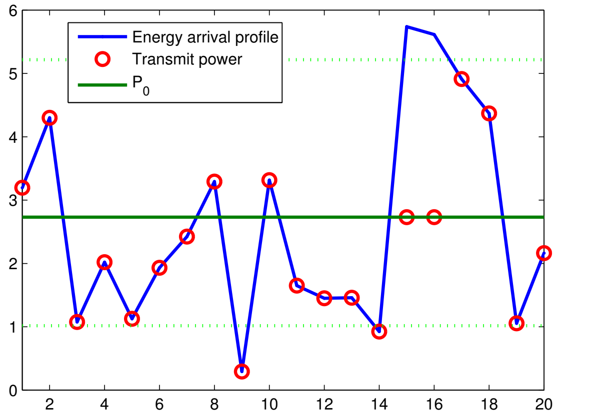

Fig. 2: A sample path of the energy arrivals and the transmit policy.

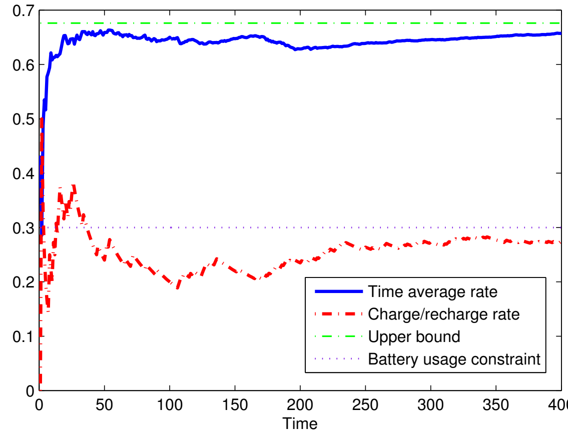

Fig. 3: A sample path of the time-average transmit rate and the battery charge/discharge rate.

to meet the power demand P_0, and the transmitter transmits with all the power the system has at that time.

We then evaluate the time-average transmission rate and the average number of battery charge/discharge operations per time slot. We plot a sample path in Fig. 3. We observe that both curves fluctuate at the beginning, and become stable after about 250 time slots. This corroborates with our theoretical results that the performance of the best-effort transmission policy converges to the upper bound almost surely. Finally, we run the simulation 1000 times, and plot the sample average of $\frac{1}{T}\sum_{t=1}^{T} R(P_t)$ as a function of T in Fig. 4. The sample average of battery charge/discharge rate is also plotted in the same figure. We observe that the sample average of $\frac{1}{T}\sum_{t=1}^{T} R(P_t)$ converges to the upper bound as expected. The sample average of battery charge/discharge rate is very close to the battery usage constraint after a short time period. This implies that the desired battery usage constraint is satisfied under the proposed policy.

Appendix

Assume under the optimal policy $\mathcal{Q}$, the transmit power does not obey the double-threshold structure. Define

$$\mathcal{A}_t^- = \{(A_t, H^{t-1}) | (A_t, H^{t-1}) \in \mathcal{A}_t, A_t < P_0, \}, t = 1, 2, \ldots$$
$$\mathcal{A}_t^+ = \{(A_t, H^{t-1}) | (A_t, H^{t-1}) \in \mathcal{A}_t, A_t \geq P_0, \}, t = 1, 2, \ldots$$

and

$$\mathbb{P}[\mathcal{A}^-] = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{P}[\mathcal{A}_t^-], \quad \mathbb{P}[\mathcal{A}^+] = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{P}[\mathcal{A}_t^+].$$

Then, we define

$$\underline{\mathcal{A}}_t^- = \{(x, H^{t-1}) | \mathbb{P}[0 \leq A_t \leq x] \leq \mathbb{P}[\mathcal{A}^-]\}, t = 1, 2, \ldots$$
$$\bar{\mathcal{A}}_t^+ = \{(x, H^{t-1}) | \mathbb{P}[A_t \geq x] \leq \mathbb{P}[\mathcal{A}^+]\}, t = 1, 2, \ldots$$

Denote $\underline{\mathcal{A}}_t = \underline{\mathcal{A}}_t^- \cup \mathcal{A}_t^+$, $\bar{\mathcal{A}}_t = \mathcal{A}_t^- \cup \bar{\mathcal{A}}_t^+$. Define

$$\underline{P}_0 := \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}[A_t | (A_t, H^{t-1}) \in \underline{\mathcal{A}}_t] \quad (15)$$

$$\bar{P}_0 := \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}[A_t | (A_t, H^{t-1}) \in \bar{\mathcal{A}}_t] \quad (16)$$

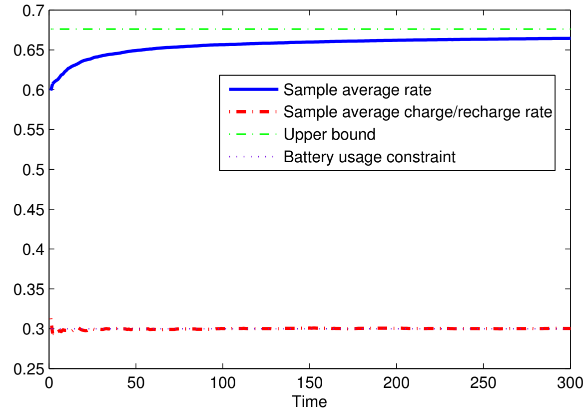

Fig. 4: Sample path of the energy arrivals and the transmit policy.

We define two policies $\underline{\mathcal{Q}}$ and $\bar{\mathcal{Q}}$ under which in each time slot t, the transmitter power is defined as follows respectively.

$$\underline{P}_t = \begin{cases} A_t, & (A_t, H^{t-1}) \notin \underline{\mathcal{A}}_t \\ \underline{P}_0, & (A_t, H^{t-1}) \in \underline{\mathcal{A}}_t \end{cases} \quad (17)$$

$$\bar{P}_t = \begin{cases} A_t, & (A_t, H^{t-1}) \notin \bar{\mathcal{A}}_t \\ \bar{P}_0, & (A_t, H^{t-1}) \in \bar{\mathcal{A}}_t \end{cases} \quad (18)$$

Denote

$$R(\mathcal{Q}) := \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} R(P_t)$$

$$R(\underline{\mathcal{Q}}) := \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} R(\underline{P}_t)$$

$$R(\bar{\mathcal{Q}}) := \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} R(\bar{P}_t).$$

We aim to show that

$$\mathbb{E}[R(\mathcal{Q})] \leq \mathbb{E}[R(\underline{\mathcal{Q}})], \quad \mathbb{E}[R(\mathcal{Q})] \leq \mathbb{E}[R(\bar{\mathcal{Q}})],$$

based on which we can claim that a necessary condition for $\mathcal{Q}$ to be optimal is, in each time slot t,

$$\underline{\mathcal{A}}_t^- = \mathcal{A}_t^-, \quad \bar{\mathcal{A}}_t^+ = \mathcal{A}_t^+,$$
$$P_t = P_0, \quad \forall (A_t, H^{t-1}) \in \mathcal{A}_t,$$

i.e., a double-threshold structure.

Definition 2 *Let f, g be two increasing functions defined over $I_c := [0, c]$. We say $f \prec g$ if*

1) $\forall t \in I_c$, $\int_0^t f(s)ds \geq \int_0^t g(s)ds$.
2) $\int_0^c f(s)ds = \int_0^c g(s)ds$.

Lemma 3 *If $f \prec g$, then for any concave function $r(\cdot)$,*

$$\int_0^c r(f(s))ds \geq \int_0^c r(g(s))ds$$

Given $\mathcal{A}_t$, $t = 1, 2, \ldots$ define the sub-level function

$$\phi_{\mathcal{A}}(x) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{P}[A_t \leq x, (A_t, H^{t-1}) \in \mathcal{A}_t]$$

Take one of its quasi-inverse, denote as $x_{\mathcal{A}}(\phi)$. Note $\phi_{\mathcal{A}}(x_{\mathcal{A}}(\phi_{\mathcal{A}}(x))) = \phi_{\mathcal{A}}(x)$. Assume both are increasing. Then,

$$\phi_{\mathcal{A}}(\infty) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{P}[\mathcal{A}_t] := \mathbb{P}[\mathcal{A}]$$

We note that

$$\begin{aligned} \phi_{\underline{\mathcal{A}}}(x) &\geq \phi_{\mathcal{A}}(x), \quad \text{if } x \in [0, P_0) \\ \phi_{\underline{\mathcal{A}}}(x) &= \phi_{\mathcal{A}}(x), \quad \text{if } x \in [P_0, \infty) \end{aligned}$$

Thus,

$$\begin{aligned} x_{\underline{\mathcal{A}}}(\phi) &\leq x_{\mathcal{A}}(\phi), \quad \text{if } \phi \in [0, \phi_{\mathcal{A}}(P_0)) \\ x_{\underline{\mathcal{A}}}(\phi) &= x_{\mathcal{A}}(\phi), \quad \text{if } \phi \in [\phi_{\mathcal{A}}(P_0), \phi_{\mathcal{A}}(\infty)) \end{aligned}$$

Define

$$f(\phi) = \begin{cases} x_{\mathcal{A}}(\phi) & \phi \in [0, \phi_{\mathcal{A}}(\underline{P}_0)) \\ \underline{P}_0 & \phi \in [\phi_{\mathcal{A}}(\underline{P}_0), \phi_{\mathcal{A}}(\underline{P}_0) + \phi_{\mathcal{A}}(\infty)] \\ x_{\mathcal{A}}(\phi - \phi_{\mathcal{A}}(\infty)) & \phi \in (\phi_{\mathcal{A}}(\underline{P}_0) + \phi_{\mathcal{A}}(\infty), 2\phi_{\mathcal{A}}(\infty)] \end{cases}$$

$$g(\phi) = \begin{cases} x_{\underline{\mathcal{A}}}(\phi) & \phi \in [0, \phi_{\underline{\mathcal{A}}}(P_0)) \\ P_0 & \phi \in [\phi_{\underline{\mathcal{A}}}(P_0), \phi_{\underline{\mathcal{A}}}(P_0) + \phi_{\underline{\mathcal{A}}}(\infty)] \\ x_{\underline{\mathcal{A}}}(\phi - \phi_{\underline{\mathcal{A}}}(\infty)) & \phi \in (\phi_{\underline{\mathcal{A}}}(P_0) + \phi_{\underline{\mathcal{A}}}(\infty), 2\phi_{\underline{\mathcal{A}}}(\infty)] \end{cases}$$

Then, we have the following Lemmas.

Lemma 4 *$f \prec g$ on $[0, 2\phi_{\mathcal{A}}(\infty)]$.*

Lemma 5 *Denote $\int_{\mathcal{A}} R(\cdot) = \mathbb{E}[R(\cdot)|\mathcal{A}] \cdot \mathbb{P}[\mathcal{A}]$. We have*

$$\int_0^{2\phi_{\mathcal{A}}(\infty)} R(f(t))dt = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \left(\int_{\underline{\mathcal{A}}_t} R(\underline{P}_0) + \int_{\mathcal{A}_t} R(A_t) \right)$$

$$\int_0^{2\phi_{\mathcal{A}}(\infty)} R(g(t))dt = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \left(\int_{\mathcal{A}_t} R(P_0) + \int_{\underline{\mathcal{A}}_t} R(A_t) \right)$$

In order to show $\mathbb{E}(R(\mathcal{Q})) \leq \mathbb{E}(R(\underline{\mathcal{Q}}))$, it suffices to show that

$$\begin{aligned} &\mathbb{E}[R(\mathcal{Q})] - \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}[R(A_t)] \\ &\leq \mathbb{E}[R(\underline{\mathcal{Q}})] - \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}[R(A_t)] \end{aligned} \tag{19}$$

We note that under $\mathcal{Q}$, we have $P_t = A_t$ if $(A_t, H^{t-1}) \notin \mathcal{A}_t$; Similarly, under $\underline{\mathcal{Q}}$, we have $\underline{P}_t = A_t$ if $(A_t, H^{t-1}) \notin \underline{\mathcal{A}}_t$. Thus, (19) is equivalent to

$$\begin{aligned} &\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \int_{\mathcal{A}_t} R(P_t) - R(A_t) \\ &\leq \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \int_{\underline{\mathcal{A}}_t} R(\underline{P}_t) - R(A_t) \end{aligned}$$

i.e.,

$$\begin{aligned} &\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \left(\int_{\mathcal{A}_t} R(P_t) + \int_{\underline{\mathcal{A}}_t} R(A_t) \right) \\ &\leq \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \left(\int_{\underline{\mathcal{A}}_t} R(\underline{P}_t) + \int_{\mathcal{A}_t} R(A_t) \right) \end{aligned}$$

which is then true due to Lemma 2, the definition of $\underline{P}_t$, Lemma 4 and Lemma 5.

Similarly, we can show that $\mathbb{E}(R(\mathcal{Q})) \leq \mathbb{E}(R(\bar{\mathcal{Q}}))$. Therefore, the optimal policy must have the double-threshold structure specified in Theorem 1.

REFERENCES

[1] 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, January 2012.

[2] 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, September 2011.

[3] O. Ozel and S. Ulukus, "Achieving AWGN capacity under stochastic energy harvesting," *IEEE Transactions on Information Theory*, vol. 58, no. 10, pp. 6471–6483, 2012.

[4] K. Tutuncuoglu and A. Yener, "Optimum transmission policies for battery limited energy harvesting nodes," *IEEE Transactions on Wireless Communications*, vol. 11, no. 3, pp. 1180–1189, March 2012.

[5] O. Ozel, J. Yang, and S. Ulukus, "Optimal broadcast scheduling for an energy harvesting rechargeable transmitter with a finite capacity battery," *IEEE Transactions on Wireless Commmunications*, vol. 11, no. 6, pp. 2193–2203, June 2012.

[6] ——, "Optimal transmission schemes for parallel and fading Gaussian broadcast channels with an energy harvesting rechargeable transmitter," *Elsevier Computer Communications, special issue for selected papers from WiOpt 2011*, vol. 36, no. 12, pp. 1360–1372, 2013.

[7] J. Lei, R. Yates, and L. Greenstein, "A generic model for optimizing single-hop transmission policy of replenishable sensors," *IEEE Transactions on Wireless Communications*, vol. 8, no. 2, pp. 547–551, Feburary 2009.

[8] R. Srivastava and C. E. Koksal, "Basic performance limits and tradeoffs in energy-harvesting sensor nodes with finite data and energy storage," *IEEE/ACM Transactions on Networking*, vol. 21, no. 4, pp. 1049–1062, August 2013.

[9] K. Kar, A. Krishnamurthy, and N. Jaggi, "Dynamic node activation in networks of rechargeable sensors," *IEEE/ACM Transactions on Networking*, vol. 14, no. 1, pp. 15–26, Feburary 2006.

[10] B. Devillers and D. Gunduz, "A general framework for the optimization of energy harvesting communication systems with battery imperfections," *Journal of Communications and Networks*, vol. 14, no. 2, pp. 130–139, April 2012.

[11] N. Michelusi, L. Badia, R. Carli, L. Corradini, and M. Zorzi, "Energy management policies for harvesting-based wireless sensor devices with battery degradation," *IEEE Transactions on Communications*, vol. 61, no. 12, pp. 4934–4947, 2013.

[12] K. Tutuncuoglu, A. Yener, and S. Ulukus, "Optimum policies for an energy harvesting transmitter under energy storage losses," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 3, pp. 467–481, 2015.