Optimal Energy Supplementary and Data Transmission Schedule for Energy Harvesting Transmitter With Reliable Energy Backup

XIANGCHENG WU\textsuperscript{1}, LIANSHENG TAN\textsuperscript{1,2}, AND SHENGDA TANG\textsuperscript{3}

\textsuperscript{1}Department of Computer Science, Central China Normal University, Wuhan 430079, China
\textsuperscript{2}College of Sciences and Engineering, University of Tasmania, Hobart, TAS 7001, Australia
\textsuperscript{3}School of Mathematics and Statistics, Guangxi Normal University, Guilin 541004, China

Corresponding author: Liansheng Tan (liansheng.tan@utas.edu.au)

This work was supported in part by the National Natural Science Foundation of China under Grant 61761008, in part by the Natural Science Foundation of Guangxi under Grant 2018JJA170024, in part by the Project of Guangxi Colleges and Universities Key Laboratory of Mathematical and Statistical Model under Grant 2017GXKLM002, and in part by the Doctoral Launch Fund Project of Guangxi Normal University under Grant 17A4.

\section*{ABSTRACT}
For the Energy harvesting (EH) transmitter equipped with a Reliable Energy Backup (REB), the following problem is challenging and important: how to minimize the amount of energy supplied by the REB, such that the harvested energy is efficiently utilized for transmitting a given amount of data within a fixed delay constraint. In this paper, we first develop a stochastic model for this problem. We then discuss the optimization issue of delay-constrained data transmission over a fading wireless channel. We transform the delay constraint into the penalty function, and then the energy constrained control problem is modeled as a Markov decision process (MDP) without constraint, by which we obtain the optimal energy supplementary policy and the minimum of the expected energy consumption from REB. In the special case that the energy of the transmitter is supplied by REB alone, we find that the optimal energy supplementary policy is non-decreasing in the elapsed transmission time and for the remaining task of data transmission. This substantially reduces the computational complexity required to implement the optimal energy supplementary policy for a general EH wireless device. Numerical studies validate the theoretical findings, and observations are outlined to demonstrate the characteristics of the optimal energy supplementary policy and the minimum expected energy expenditure from the REB.

\section*{INDEX TERMS}
Energy harvesting transmitter, reliable energy backup (REB), energy supplementary policy, Markov decision process (MDP).

\section*{I. INTRODUCTION}
In order to fulfill the uninterrupted long-term or perpetual communication task, an energy harvesting (EH) transmitter can harvest energy from natural energy sources, for example solar power, wind, magnetic fields, mechanical vibrations, temperature variations. However, in general the energy that is harvested by a typical EH transmitter, is very limited. Moreover, due to the randomness and instability of EH manipulations and the stochastic communication process, the harvested energy may not be able to meet the energy consumption need of the EH transmitter at all times. That is, the EH transmitter if being only powered by the harvested energy may not guarantee the quality-of-service (QoS) of the communications. In this situation, a Reliable Energy Backup (REB) is used to support the communication of the transmitter during any failure appearing in its energy harvesting process. In practice, the REB is some paid-use power supply unit, it refers to any reliable energy source, either traditional or newly developed. For example, the constant battery, the diesel generator, the power grid, the constant energy transfer and so on. The EH transmitter is encouraged to use the free and replenishable harvested energy whenever available, when the residual energy can not meet the required energy consumption of the system, the REB is triggered to input energy to the transmitter to satisfy the energy consumption requirement. In this way, the efficient mixed usage of the harvested energy and reliable backup energy provides a key solution to robust wireless green communications of the EH transmitter.
Although the coexistence of the harvested energy and the backup energy is considered as a promising technology to tackle the problem of energy shortage in the communication system, it gives rise to many new challenges to the management of the energy supplementary from the REB. Since the backup energy is paid source, the EH transmitter should make the best use of the harvested energy and minimize the backup energy consumption while satisfying the QoS requirements. Therefore, energy-efficient mechanisms of the REB should be developed in the protocol stack of the EH transmitter. Obviously, these are theoretically and practically important as well for the design, development and optimization of EH transmitter to achieve uninterrupted long-term or perpetual operations.

There is a significant amount of recent literature on the efficient optimal transmission and energy allocation policies for EH communications based on various communication models [1]–[4], however these works assume that the energies are supplied by a constant energy source. Transmission strategies and power allocation policies for EH nodes in wireless communication systems have been studied in [5]–[9], these works on EH communication systems are based on the assumption that EH is the only source of energy for the nodes. In [10], the authors propose an enhanced date transmission control strategy to minimize the energy delivered from the primary battery while keeping the overflow probability of the data buffer below a desired threshold. In [11], the authors assume that the base stations (BSs) are powered by both on-grid energy and green energy, they develop algorithms to solve the green energy optimization problem. Reference [12] investigates the collaboration between multiple operators to improve the energy utilization, an energy-saving management mechanism is proposed to reduce energy consumption and optimize energy utilization. Reference [13] proposes a model of the flexible configuration of energy, through energy cooperation between base stations and smart grid. Reference [14] presents green energy optimization schemes to maximize energy utilization in cellular networks with hybrid energy supplies, and the algorithms to solve the green energy provisioning problem in BSs while satisfying the QoS requirements are proposed.

In this paper, we consider a point-to-point communication of an EH transmitter with REB. Our goal is to minimize the energy consumed supplied by the REB for transmission of a given amount of data within a time deadline over a fading wireless channel. Compared with the literature mentioned above, the main differences of our work are given as follows.

- Firstly, we propose a new stochastic system model to characterize a point-to-point data transmission of EH transmitter with REB. In this model, we consider a delay-constrained data transmission, that is, the data task should be transmitted completed in a given deadline. We use two stochastic processes to model the energy harvesting process and the channel fading process respectively, based on which the dynamic of the energy level of the transmitter and the evolution of the remaining data task process are characterized.
- Secondly, the attacked control problem is modelled as a Markov decision process (MDP), which enables us to obtain the desired optimal energy supplementary policy of the REB and the minimum expected energy consumption supplied by the REB. In the case that the transmitter is supplied only by REB, using the concept of super-modularity, we prove that the optimal energy supplementary policy is monotonous increasing function of the residual transmission time and the remaining amount of the data task. These properties substantially reduce the computational complexity required to implement the optimal transmission scheduling policy.

The rest of the paper is organized as follows. Section II gives the analytical model for the EH transmitter with REB. Section III outlines the main results. Section IV provides numerical studies to support the theoretical results. Finally conclusions are given in Section V.

II. SYSTEM MODEL AND CONTROL PROBLEM

It is assumed that the time duration is partitioned into consecutive slots and the length of each slot is denoted by $\tau$. They are indexed by a set $T = \{0, 1, \ldots\}$, where the $n$-th time slot is denoted by $[n, n+1)$.

We consider a point-to-point delay-sensitive data transmission of an EH transmitter with REB. More specially, there is a given amount of data task denoted by $D[\zeta, N_d]$, where $\zeta$ is the amount of the data task, and $N_d$ is the execution deadline, the transmitter should transmit the data task $D[\zeta, N_d]$ to the receiver over a time-varying channel within $N_d$ time slots strictly. We assume that transmission consumes most of the energy in the transmitter and ignore other causes of energy consumption.

As shown in Fig. 1, the transmitter under consideration has two types of energy source to meet its energy consumption: the harvested energy extracted from the ambient, and the backup energy supplied by the REB. The harvested energy is clean green energy and it can be free to use, while the backup energy supplied by the REB is un-renewable and paid-use resource. The transmitter should make the best use of the

![Figure 1. The architecture and working mechanism of EH-transmitter with an REB.](image-url)
harvested energy and minimize the supplementary energy from the REB while satisfying the delay constraint.

A. ENERGY HARVESTING MODEL

Similar to [15], [16], we use the Markov energy harvesting model denoted by $X^h = \{X^h_n, n \in T\}$ to characterize the studied energy harvesting process. Specially, we assume the energy harvesting process has $h$ states which is denoted by $\Omega^h = \{h_1, h_2, \ldots, h_h\}$, the state of the energy harvesting process remains static within each time slot, but may vary at the end of each slot. We let $p^h_{h_{\ell} h_{\ell}'}, h, h' \in \Omega^h$, be the transition probability of $X^h$ from state $h$ to $h'$, and denote the transition probability matrix by $P^h = [p^h_{h_{\ell} h_{\ell}'}]_{h, h'}$. When the state is $h \in \Omega^h$, there are $\beta(h)$ Joules of energy can be drawn from the ambient in each slot. The harvested energy is stored in an energy buffer with capacity $B$. When the accumulated energy reaches $B$, the excess energy will be lost. Due to the delay of the energy transformation, we assume that the energy harvested in the $n$-th slot can only be used for transmission in the $(n+1)$-th slot onward.

B. CHANNEL FADING MODEL

The channel fading process is reflected by $X^c = \{X^c_n, n \in T\}$, and is then modelled by a finite state Markov chain [17], [18]. Specifically, the channel state space can be partitioned into $n^c$ non-overlapping ranges. Note that there is a map, in which each range is mapped into one channel state. Let the channel state space be $\Omega^c = \{c_1, c_2, \ldots, c_{n^c}\}$, and in channel state $c \in \Omega^c$, we assume the corresponding channel state is $\alpha(c)$. The channel state is assumed to remain static within each time slot, but may change at the slot boundary. We let the transition probability matrix of $X^c$ be $P^c = [p^c_{c_{\ell} c_{\ell}'}]_{c, c'}$, where $p^c_{c_{\ell} c_{\ell}'}$, $c, c' \in \Omega^c$, is the transition probability of $X^c$ from state $c$ to $c'$.

When $n^c = 2$, this model is reduced to be Gilbert-Elliot channel model.

C. COMMUNICATION MODEL

We assume that the transmitter has finite different discrete transmit power rate levels, which is denoted by

$$\mathcal{A} = \{a_1, a_2, \ldots, a_{n^a}\},$$

where $0 = a_1 < a_2 < \ldots < a_{n^a} < \infty$, we also refer to $\mathcal{A}$ as an action space, and $a_i \in \mathcal{A}$ as an action. At the beginning of slot $n$, $n \in T$, the transmitter needs to choose a power rate $\pi_n \in \mathcal{A}$ to transmit the data task. Let $\pi = \{\pi_0, \pi_1, \ldots, \}$, and we refer to $\pi$ as the transmit power control policy of the transmitter. Given the transmit power control policy $\pi_n$ and the channel state $c \in \Omega^c$ in the $n$-th slot, then the amount of data transmitted in this slot is given as follows:

$$\mu(\alpha[c], \pi_n) = W \log_2 \left(1 + \frac{\alpha[c]\pi_n}{W N_0}\right),$$

(1)

where the component $N_0$ is the power spectral density of the Gaussian noise, the component $W$ is the bandwidth of the channel. One sees that $\mu(\alpha[c], \pi_n)$ is a monotonically nondecreasing and in the meantime a concave function in $\pi_n$ for the given value $\alpha[c]$.

Thus, at the beginning of slot $n$ the remaining amount of the (uncompleted) data task can be defined as

$$X^d_n = \left[\zeta - \sum_{k=0}^{n-1} \mu(\alpha[X^c_k], \pi_k)\right]^+,$$

(2)

where the notation $[x]^+ = \max(0, x)$, and let $\sum_{k=0}^{n-1}(\cdots) = 0$. We refer to $X^d = \{X^d_n, n = 0, 1, \ldots\}$ as a remaining data task process of the transmitter, and denote the state space by $\Omega^d = [0, \zeta]$.

Let $X^e_n, n \in T$, be the energy level in the energy buffer at the beginning of slot $n$, then we have

$$X^e_n = \min\{B, [X^e_{n-1} - \pi_n]^+ + \beta[X^h_n]\},$$

(3)

and we refer to $X^e = \{X^e_n, n \in T\}$ as the energy level process of the transmitter with state space $\Omega^e = [0, B]$.

Now, we combine the above four stochastic processes $X^c$, $X^h$, $X^d$, and $X^e$ to build our system model denoted by

$$X = \{X_n|X^c_n = X^c_n, X^h_n = X^h_n, X^d_n = X^d_n, n \in T\}.$$

We call $X_n$ the state of the system in the $n$-th slot and let the state space of $X$ be $\Omega = \Omega^c \times \Omega^h \times \Omega^d \times \Omega^e$.

D. ENERGY CONSUMPTION MINIMIZATION PROBLEM

In our energy supply mechanism, the transmitter exploits the replenishable harvested energy firstly, when the harvested energy can not meet the required energy consumption, then the REB is triggered to supply energy to satisfy the energy consumption of the system.

Given the energy level $x^e_n$ and the transmit power rate $\pi_n$ in the $n$-th slot, let $s_n(x^e_n, \pi_n)$ be the supplementary energy rate supplied by REB in this slot, then $s_n$ can be defined as

$$s_n(x^e_n, \pi_n) = x^e_n - x^e_{n-1}^+,$$

(4)

Let $s = \{s_0, s_1, \ldots\}$, and we refer to $s$ as the energy supplementary policy of the REB.

In practical application, the determining of the amount for backup energy $s_n$ from the REB is very important. If $s_n$ is too small, the system may be always in a status of energy shortage, the data task will be transmitted with a low rate, and the violation of the delay constraint will be occurred, which results in the lower level of QoS. On the other hand, if $s_n$ is too large, more backup energy is supplied, and the cost of the communication increases too. Furthermore, the power dissipation may be occurred since the finite capacity of the energy buffer, which lead to an inefficient use of harvested energy. Thus, developing an optimal energy supplementary policy $s$ of the REB plays a critical role for the efficient use of energy of the EH transmitter with REB.

Let $N_d$ be an actual transmission time of the task $D[\zeta, N_d]$. At the beginning of slot $n$, given the system state $X_n = (c, h, d, e)$, and the transmit power policy $\pi$, the expected
supplementary energy from the REB until the data task is transmitted completely is defined by

$$V^\pi_n(c, h, d, e) = \mathbb{E}^T \left[ \sum_{k=0}^{N_n} s_n(X^c_k, \pi_k) \mid c, h, d, e \right],$$

(5)

where $\mathbb{E}^T [\cdot | c, h, d, e]$ represents the conditional expectation given the initial state $(c, h, d, e)$ under the policy $\pi$. We call $V^\pi_n$ the cost function at the beginning slot $n$.

In this paper, the task is to find an energy supplementary policy $s$ in order to minimize the total expected supplementary energy supplied by the REB and also to ensure that the data task can be transmitted within a strict delay constraint. From (4), when we determine the optimal transmit power $\pi$, we can obtain the optimal energy supplementary policy. Thus, we formulate the at hand constrained control problem in the slotted system as follows: given a data task $D[\xi, N_d]$, the initial channel state $c$, the EH state $h$, and the energy level $e$, find a transmit power control policy $\pi^*$ such that,

$$V_0(c, h, \xi, e) = V^\pi_0(c, h, \xi, e) = \min_{\pi} \left[ V^\pi_n(c, h, \xi, e) \right]$$

s.t. $C1: \sum_{k=0}^{N_n} \mu(\alpha(x^c_k), \pi_k) = \xi$;

$$C2: N_e \leq N_d \text{ P-a.s.}.$$  

We call $\pi^* = \{\pi^*_0, \pi^*_1, \ldots\}$ the optimal transmit power control policy, and then we obtain the optimal energy replenishment policy $s^\pi_n$ by

$$s^\pi_n(x^c_n, \pi_n) = [\pi^*_0 - x^c_0]^+, \quad n \in \mathcal{T}.$$  

(7)

In the following, we will reformulate the control problem (6) into a mathematical framework for analytical tractability.

To simplify the notation, we define $x = (c, h, d, e) \in \Omega^4$ and $x' = (c', h', d', e') \in \Omega^4$, respectively. Herein no distinction between these two symbols is made, the symbols are used alternatively according to the convenience of the illustrating purpose.

Remark 1: Different from the expectation delay constraint in traditional control problem, we use an $P$-a.s., i.e., “almost sure”, delay constraint in (6) to characterize the hard deadline, and the convex analytic approach is problematic in our system model, we need to transform this constraint condition into some other mathematical form easy to be processed. To address this challenge, similar to [19], [20] and [21], we transform the $P$-a.s. delay constraint into the penalty function, and ultimately we reformulate the control problem (6) into an finite horizon MDP with no constraint, the transmit power control policy and the minimum expected supplementary energy of the REB are obtained by solving the MDP.

III. MDP REFORMULATION

A. NON-STATIONARY MDP

Based on the system model, a discrete-time finite-horizon (non-stationary) MDP with a terminal data loss cost is built up as follows.

1) Decision time epochs: $T_d = \{0, 1, \ldots, N_d - 1\}$.
2) State space: let the state be $x_n = (x^c_n, x^h_n, x^n_d, x^n_e)$ at time epoch $n$, i.e., the state space is $\Omega^4$.
3) Action space: $\mathcal{A}$.
4) Transition probabilities: Denote $P^{[a:a]}_{xx'}$ be the probability that the state at time $(n+1)$ is $x'$ if the state at time $n$ is $x$ and the action $a \in \mathcal{A}$ is taken. Then we have

$$P^{[a:a]}_{xx'} = P(X_{n+1} = x' | x_n = x, \pi_n = a) = P(X_{n+1} = (c', h', d', e') | X_n = (c, h, d, e), \pi_n = a) = P(X_{n+1} = c' | X_n = c)P(X_{n+1}^h = h' | X_n^h = h)\times I(d' = d - e + \beta(h)),$$

where $I(x)$ is an index function, $\mu(a[c],\pi_n)$ can be obtained by (1). We then define the transition matrix under action $a$ at time $n$ by $P^{[a:a]}_{xx'} = [P^{[a:a]}_{xx'}]_{x,x'}$.

5) One-step cost: Let $g^{[a]}_n(e)$ be the energy consumption from REB in slot $n$, when the current energy level is $e$, and the action $a \in \mathcal{A}$ is taken, then we have

$$g^{[a]}_n(e) = s_n(a, e) = [a - e]_+, \quad n \in T_d.$$  

6) Terminal cost: To exclude these infeasible control policies that result in the violation of the delay constraint, as mentioned in remark 1, a terminal penalty cost is added at the end of the horizon as

$$g_{N_d} = \rho(x^c_{N_d}, h_{N_d} > 0),$$

where we let $\rho(\cdot)$ be a convex increasing function in the uncompleted data task $x^c_{N_d}$ at the end of the horizon.

We denote the above MDP by $[T_d, \Omega^4, \mathcal{A}, P^{[a:a]}_{xx'}, g^{[a]}_n, g_{N_d}]$. In the reformulation, we replace the constraint with a terminal data loss cost. The terminal cost means that if the deadline has not been met at the end, the residual data task is abandoned, and the penalty value is, and the increment of the penalty increases with respect to the amount of the residual data task. In practical application, $\rho$ can just be set as an arbitrary sufficient big number. For example, let $\rho = +\infty$ for the control policies that lead to the violation of time constraint. Thus, by our assumption, at the end of the execution deadline, the data task should be completed, otherwise, a big cost of violating the constraint will incur.

Since the MDP $[T_d, \Omega^4, \mathcal{A}, P^{[a:a]}_{xx'}, g^{[a]}_n, g_{N_d}]$ is a finite-horizon MDP with finite action set, the minimum cost function $V$ of the control problem (6) exists and it is the unique solution of the following dynamic programming equation.

$$V^\pi_n(x) = \min_{a \in \mathcal{A}} \left\{ g^{[a]}_n(e) + \sum_{x' \in \Omega^4} P^{[a:a]}_{xx'} V^\pi_{n+1}(x') \right\}.$$  

At decision epoch $n$, $n \in T_d$, given system state $x = (c, h, d, e) \in \Omega^4$, we choose $\pi^*_n$ to be the smallest action that realizes the minimum on the right of the dynamic programming equation, then $\pi^* = \{\pi^*_0, \pi^*_1, \ldots, \pi^*_{N_d-1}\}$ is the unique optimal policy [22]. By (7), we obtain the decision
consequence \( s^* = \{ s_0, s_1, \ldots, s_{N_d-1} \} \) is the optimal energy supplementary policy of the REB.

For \( \forall x = (c, h, d, e) \in \Omega^x \), the minimum cost \( V_n(x) \) and the optimal transmit power policy \( \pi^* \) can be obtained by the following iteration algorithm.

\[
V_{N_d}(x) = \phi(d)M_{d>0}, \tag{8}
\]

for \( n = N_d - 1, \ldots, 1, 0 \), we have

\[
V_n(x) = \min_{a \in A} [W_n^{[a]}(x)], \tag{9}
\]

\[
\pi^*_n(x) = \arg \min_{a \in A} [W_n^{[a]}(x)], \tag{10}
\]

\[
s^*_n(x) = [\pi^*_n(x) - e]^+, \tag{11}
\]

with the state action cost function

\[
W_n^{[a]}(x) = g_n^{[a]}(e) + \sum_{x' \in \Omega^x} p_n^{[a]a}(x') V_{n+1}(x'). \tag{12}
\]

We can see that the optimal energy supplementary policy of the REB are offline power control policies. That is, the policies can be computed offline first and then be compiled into simple look-up tables in the transmitter. At each decision epoch, what the transmitter should do is only to look up the table and choose the corresponding optimal energy supplementary policy to match the current state. Therefore, the energy supplementary policy is suitable for the transmitter with limited computational capability.

The following property is intuitive, it means that if more amount of the residual data task is left, i.e., a larger \( d \), we should spend more backup energy from REB on the transmission, then \( V \) increases. Similarly, if more transmission time is passed, i.e., a larger \( n \), more energy for transmission should be allocated, which will subsequently lead to a larger total energy consumption \( V \).

**Property 1:** For the MDP \( \{T_d, \Omega^x, A, p_n^{[a]}, g_n^{[a]}, g_{N_d} \} \) described above, the following two statements hold for \( n = 0, 1, \ldots, N_d \):

(a) \( V_n(c, h, d, e) \) is a non-decreasing function in \( d \) when other parameters are fixed;

(b) \( V_n(c, h, d, e) \) is a non-decreasing function in \( n \) when other parameters are fixed.

The proof can be found in Appendix A.

### B. SPECIAL CASE: TRANSMITTER WITH REB

In the following, we consider a special case where the energy consumption of the transmitter is supplied by the REB. In this case, we investigate some structural results of the optimal energy supplementary policy, which can be used to reduce the computational complexity of the value iteration algorithm.

When the energy of the transmitter is supplied by the REB, as a result, the system state is left with two elements: the channel state and the residual data task, then the system model is denoted by

\[
Y = \left\{ Y_n | Y_n = (X_n^c, X_n^d), n \in T \right\}
\]

with state space \( \Omega^y = \Omega^c \times \Omega^d \). We define the notation \( y = (c, d) \in \Omega^y \) and \( y' = (c', d') \in \Omega^y \), respectively. The transition probability of the system state reduces to be

\[
p_n^{[a]}(c', d'|c, d) = P(Y_{n+1} = (c', d') | Y_n = (c, d), \pi_n = a) = p_n^{[a]}(d' = |d - \mu(c, a)|^+). \tag{13}
\]

It should be noted that, for the transmitter with REB, there is no difference between the transmit power policy and the energy supplementary policy. Thus, we also refer to \( \pi \) as the energy supplementary policy.

For \( \forall (c, d) \in \Omega^y \), we denote the minimum expected cost function at decision epoch \( n \) by \( \tilde{V}_n(c, d) \), which satisfies the following Bellman’s equation of optimality:

\[
\tilde{V}_n(c, d) = \min_{a \in A} [\tilde{W}_n^{[a]}(c, d)]. \tag{14}
\]

with

\[
\tilde{W}_n^{[a]}(c, d) = g_n^{[a]}(a) + \sum_{y' \in \Omega^y} p_n^{[a]}(c', d'|c, d) \tilde{V}_{n+1}(c', d'). \tag{15}
\]

One is able to solve Equation (14) using the value iteration algorithm as described in Section III-A. However, before that let us discuss some properties related to \( \tilde{V}_n(c, d) \) as outlined in Lemmas 1 and 2. These results are the pathways to the monotone policy in Theorem 1.

**Lemma 1:** The following two statements hold for \( n = 0, 1, \ldots, N_d \):

(a) Given \( n, c, \tilde{V}_n(c, d) \) is non-decreasing in \( d \);

(b) Given \( c, d, \tilde{V}_n(c, d) \) is a non-decreasing in \( n \).

The proof is similar to that for Property 1, we thus omit the details.

**Lemma 2:** The cost function \( \tilde{V}_n(c, d) \) and \( \tilde{W}_n^{[a]}(c, d) \) are concave functions in \( d \) for \( n = 0, 1, \ldots, N_d \).

The proof can be found in Appendix A.

The following result states that the optimal energy supplementary policy is non-decreasing in the transmission time that is passed and the remaining amount of the data task, respectively.

**Theorem 1:** For the transmitter with REB, the optimal energy supplementary policy \( \pi^*_n(c, d) \) is non-decreasing in \( d \) and \( n \), respectively. That is, for any \( d' \geq d \), where \( d', d \in [0, \xi] \), we have

\[
\pi^*_n(c, d') \geq \pi^*_n(c, d), \tag{16}
\]

\[
\pi^*_{n+1}(c, d) \geq \pi^*_n(c, d). \tag{17}
\]

With this monotonic structure obtained in theorem 1, we can significantly reduce the computational complexity of the value iteration algorithm. We summary the algorithm into Algorithm 1.

The main difference between Algorithm 1 and the traditional iterative algorithm is that the procedure in computing \( W_n \) in (20), which inherits lower complexity than the traditional iterative algorithm. Particularly, when we compute \( \tilde{W}_n \), by the monotonic structure of the optimal policy, we just need
Algorithm 1 Computing the Minimum Energy Consumption and the Optimal Transmission Power Policy When the Energy Is Supported by the REB Only

1. Set \( n = N_d \) and for all \((c, d) \in \Omega^d:\)
\[
\tilde{V}_{N_d}(c, d) = \begin{cases} 
+\infty & \text{if } d > 0, \\
0 & \text{if } d = 0.
\end{cases}
\] (18)

2. Set \( n = n - 1 \) and compute for all \((c, d) \in \Omega^d:\)
\[
\tilde{V}_{n}(c, d) = \min_{a \in A_n(c, d)} \{ \tilde{W}_n(c, d, a) \};
\]
\[
\pi^*_n(c, d) = \arg \min_{a \in A_n(c, d)} \{ \tilde{W}_n(c, d, a) \},
\]
with
\[
A_n(c, d) = \{ a_1, a_2, \ldots, \pi^*_{n+1}(c, d) \};
\]
\[
\tilde{W}_n(c, d, a) = g_n[a](c, d) + \sum_{(c', d') \in \Omega^d} p_{(a,c,d)}(c', d') \tilde{V}_{n+1}(c', d').
\] (20)

3. If \( n = 0 \), output the cost function \( \tilde{V}_0(c, d) \), and the optimal policy \( \pi = \{ \pi^*_0, \pi^*_1, \ldots, \pi^*_N_d \} \). Otherwise, go to step 2.

To find the optimal solution in action set \( \{ a_1, a_2, \ldots, \pi^*_{n+1} \} \) instead of the larger set \( A = \{ a_1, a_2, \ldots, a_{n+1} \} \), the computational complexity of the algorithm 1 can be further reduced.

IV. NUMERICAL RESULTS

In this section, our focus will in giving a numerical example to illustrate the computation of the optimal control policy, and we will also investigate the characteristics of the energy supplementary policy and the minimum expected total consumed energy from REB.

A. NUMERICAL SETTING

Let the length of each time slot be \( \tau = 0.1 \) second, and we assume that the amount of the data task that the transmitter will be transmitted is \( \zeta = 5 \) M, and the execution deadline is \( \tau_d = 2 \) seconds.

We consider a bandlimited Gilbert-Elliot channel, the channel has two states \([0, 1]\), representing the “bad” and “good” states of the channel, with the corresponding channel gains as 0 and \( 1 \times 10^{-12} \). The transition matrix of the channel fading process is assumed to be
\[
P^c = \begin{bmatrix} 0.5 & 0.5 \\ 0.4 & 0.6 \end{bmatrix}.
\]

We assume the channel bandwidth is \( W = 20 \) MHz, and the noise power spectral density \( N_0 = 2.0 \times 10^{-18} \) W/Hz.

We use \([0, 1, 2]\) to represent the “low”, “normal” and “high” states of the energy harvesting process with the energy harvesting rates as 0 mW, 0.10 mW and 0.30 mW respectively. The transition matrix is assumed to be
\[
P^h = \begin{bmatrix} 0.2 & 0.4 & 0.4 \\ 0.0 & 0.5 & 0.5 \\ 0.6 & 0.0 & 0.4 \end{bmatrix}.
\]

The capacity of the energy buffer is set to be \( B = 2 \) mJ. That is, in each slot the maximum transmit power rate the buffer can supplied is \( 2/0.1 \) mJ/s = 20 mW.

The admissible transit power rate set of the transmitter is denoted by \( A = \{ a_k = k\delta^a, k = 0, 1, \ldots, n^a \} \), where \( \delta^a = 0.01 \) mW is the TPR grid, and we let \( n^a = 400 \).

Noting that the state space \( \Omega^c \) is a hybrid-state space, in which \( \Omega^c, \Omega^h \) are discrete and \( \Omega^d \) are continuous, we use a state-space discretization to yield recursive approximations to the MDP we obtained, and the analysis of the convergence and the complexity can be found in [24]. We let the step-size of the power rate be 0.01 mW, we then obtain a finite horizon MDP with 756 states and 400 actions based on the method developed in this paper. Then the optimal energy supplementary policy and the minimum expected energy consumption from REB can be obtained. We list some parts of look-up table in Table 1, which shows the availability of our algorithm.

### Table 1. The optimal energy replenishment policies under some states, where * represents arbitrary state.

| \((c, h, d, a)\) | \(t = 0.1\) | \(t = 0.3\) | \(t = 0.5\) | \(t = 1.0\) |
|---------------|----------|----------|----------|----------|
| \((0, *, *, *)\) | 0.00     | 0.00     | 0.00     | 0.00     |
| \((1, 0, 1, 0)\) | 0.07     | 0.07     | 0.07     | 0.07     |
| \((1, 1, 1, 0)\) | 0.00     | 0.00     | 0.00     | 0.00     |
| \((1, 2, 0, 0)\) | 0.07     | 0.21     | 0.21     | 0.21     |
| \((1, 0, 2, 0)\) | 0.00     | 0.11     | 0.11     | 0.11     |
| \((1, 2, 0, 1)\) | 0.00     | 0.11     | 0.11     | 0.11     |
| \((1, 0, 3, 0)\) | 0.07     | 0.35     | 0.35     | 0.35     |
| \((1, 2, 3, 0)\) | 0.07     | 0.35     | 0.35     | 0.35     |
| \((1, 2, 3, 1)\) | 0.00     | 0.00     | 0.00     | 0.00     |
| \((1, 2, 3, 2)\) | 0.00     | 0.15     | 0.15     | 0.15     |

**B. OBSERVATIONS**

From Table 1, some observations are given as follows. We can see that when the state is \((0, *, *, *)\), where * means arbitrary state, the energy supplementary is zero, this because when the channel state is 0, the channel gain becomes 0, in this case no data can be transmitted successfully in this channel state. Thus, no data transmission is the best decision. We can see that the energy supplementary under state \((1, 1, 1, 0)\) is zero, this because that the residual energy in the buffer can meet the optimal transmit power rate, then the REB need not supply energy for the transmission. In Table 1, the optimal energy supplementary under the state \((1, 0, 2, 0)\) is same as energy supplementary under the state \((1, 2, 0, 1)\), this observation can be explained by the fact that the energy harvested in the current can not be used due to the delay of the energy transformation.

We also plot some of our numerical results in the following Fig. 2-5 to highlight some characteristics of the optimal energy supplementary policy and the minimum expected
energy supplied by REB when the energy of the transmitter is supplied only by the REB.

Fig. 2 - 5 show that the optimal energy supplementary policy and minimum expected energy expenditure from REB with respect to remaining amount of data task and the transmission time that passed. From Fig. 2, we can see that when the channel state is 0, the optimal energy supplementary is zero no matter than what the residual data task and remaining time are. The explanation of this observation is same to the case of the state (0, 0, *, *, *) in Table 1.

From Fig. 4 and 5, we can see that given the residual data task, the energy supplementary from REB is increasing in the transmission time that has been used; and given the residual transmission time, it is increasing in the remaining amount of the data task, which are consistent with the results in Theorem 1. This observation means that an aggressive transmit power strategy must be implemented when the residual transmission time is less or the uncompleted data task is larger. Such characteristic can also be seen from Fig. 3.

We should note that when the residual transmission time is too small and the remaining data task is too large, in this case the transmitter can not transmit the remaining data task under the deadline even with the largest transmit power rate, then a high penalty value will be incurred due to the violation of the delay constraint. This fact explains why the energy consumption shown in Fig. 4 is much high when the elapsed transmission time is large.

V. CONCLUSION

In this paper, we consider an wireless communication of an EH transmitter with REB over a wireless fading channel. We consider optimization of energy supplementary policy to minimizing the total expected energy consumption supplied by REB while the data task should be transmitted in a deadline strictly. We reformulate our control problem into a finite horizon MDP, and obtain the optimal energy supplementary and the minimum total expected energy supplementary from the REB by solving the Bellman equation. When the energy of the transmitter supplied only by the REB, we prove that the optimal energy supplementary policy is non-decreasing in the transmission time that passed and the remaining amount of the data task. Finally, the numerical experiments are used to illustrate the theoretical findings numerically. Our model can provide guidance on determining the optimal system parameters. One of the future research of this model is the performance metrics evaluation of the EH transmitter with REB under the optimal energy management strategy, such as the average energy consumption rate of the EH transmitter, the residual energy distribution, the data queue length in data buffer, and the packet blocking probability due to the delay constraint, and so on.
APPENDIX A
PROOFS OF PROPERTY 1 LEMMA 2 AND THEOREM 1

A. PROOF OF PROPERTY 1

Proof: We first prove (a) by mathematical induction. It is easy to see that for $V_{N_{d}}(c, h, d, e) = g_{N_{d}}(d)$ the statement holds. We assume that $V_{n+1}(c, h, d, e)$ is a non-decreasing function in $d$ when $n, c, h, e$ are fixed. In the following, we should prove the statement holds for $V_{n}(c, h, d, e)$. To do that, we first show the state action cost function $W^{[a]}_{n}(c, h, d, e)$ is a non-decreasing function in $d$. We rewrite (12) equivalently as follows

$$W^{[a]}_{n}(c, h, d, e) = s^{[a]}_{n}(e) + \sum_{c' \in \Omega_{c}, h' \in \Omega_{h}} p^{c}_{cc} p_{hh}^{b} \times V_{n+1}(c', h', [d - \mu(a[c], a)]^{+}, \min[B, [e - a]^{+} + \beta(h)]) \quad \text{(21)}$$

The first term of the right hand side is $g^{[a]}_{n}(e)$, which is independent of $d$. By the induction hypothesis, the second term of the right hand side of (21), being a linear combination of non-decreasing functions, is non-decreasing in $d$. Thus, from (21), $W^{[a]}_{n}(c, h, d, e)$ is a non-decreasing function in $d$. By (9), we have that the cost function $V_{n}(c, h, d, e)$ is non-decreasing in $d$.

To prove (b), we let $a = 0$, and from (9) we have

$$V_{n+1}(c, h, d, e) \leq V_{n+1}(c, h, d, e), \text{ and from (12), we get } W^{[a]}_{n}(c, h, d, e) \leq W^{[a]}_{n+1}(c, h, d, e). \text{ Then we have}$$

$$V_{n}(c, h, d, e) = \min_{a \in \mathcal{A}} \{W^{[a]}_{n}(c, h, d, e)\} \leq \min_{a \in \mathcal{A}} \{W^{[a]}_{n+1}(c, h, d, e)\} = V_{n+1}(c, h, d, e).$$

By mathematical induction, we have shown that $V_{n}(c, h, d, e)$ is a non-decreasing function in $n$ when $c, h, d, e$ are fixed.

B. PROOF OF LEMMA 2

Proof: We prove this statement by mathematical induction.

Since $V_{N_{d}}(c, d) = V_{N_{d}}(c, d) = g_{N_{d}}(d)$, and the statement holds.

Now suppose that $V_{n+1}(c, d)$ and $W^{[a]}_{n+1}(c, d)$ are a concave function in $d$.

In the following, we first prove that $W^{[a]}_{n}(c, d)$ is a concave function in $d$. Similar to (21), we can write $W^{[a]}_{n}(c, d)$ as

$$W^{[a]}_{n}(c, d) = g^{[a]}_{n} + \sum_{c' \in \Omega_{c}} p^{c'}_{cc} \tilde{V}_{n+1}(c', [d - \mu(a[c], a)]^{+}).$$

The first term $g^{[a]}_{n}$ of the right hand side is independent of $d$. By the induction hypothesis, we know that $V_{n+1}(c, d)$ is a concave function in $d$, in addition, $[d - \mu(a[c], a)]^{+}$ is concave in $d$, then $\tilde{V}_{n+1}(c', [d - \mu(a[c], a)]^{+})$ is a concave in $d$. Therefore the second term of the right hand side of (22) is concave in $d$ since it is a weighted sum of concave functions. Thus we have proved that $W^{[a]}_{n}(c, d)$ is concave in $d$.

In the following, we show that $\tilde{V}_{n}(c, d)$ is concave in $d$.

Let $\lambda \tilde{V}_{n}(c, d') + (1 - \lambda) \tilde{V}_{n}(c, d)$

$$\geq g_{n}^{[a]} + \sum_{c' \in \Omega_{c}} P^{c'}_{cc} \tilde{V}_{n+1}(c', [d - \mu(a[c], a)]^{+}) \times \left[ \tilde{V}_{n+1}(c', \lambda[d' - \mu(a[c], a)]^{+} + (1 - \lambda)[d - \mu(a[c], a)]^{+} \right]$$

$$\tilde{V}_{n}(c, d) \geq g_{n}^{[a]} + \sum_{c' \in \Omega_{c}} P^{c'}_{cc} \tilde{V}_{n+1}(c', [d - \mu(a[c], a)]^{+})$$

Then we prove that $\tilde{V}_{n}(c, d)$ is concave in $d$.

C. PROOF OF THEOREM 1

Proof: To prove that the optimal energy supplementary policy $\pi^{*}_{n}(c, d)$ is non-decreasing in $d$, we just need to prove that $W^{[a]}_{n}(c, d)$ is a submodular function in $(a, d)$ [23]. That is, let $a' < a, d' > d$, where $a', a \in \mathcal{A}$ and $d', d \in [0, \varepsilon]$, we need to prove

$$W^{[a]}_{n}(c, d') - W^{[a]}_{n}(c, d') - W^{[a]}_{n}(c, d) + W^{[a]}_{n}(c, d) \leq 0.$$  

(23)

To obtain the above inequality, we substitute (22) into the left hand side of (23) and rearrange it, we then have

$$W^{[a]}_{n}(c, d') - W^{[a]}_{n}(c, d') - W^{[a]}_{n}(c, d) + W^{[a]}_{n}(c, d)$$

$$= \sum_{c' \in \Omega_{c}} P^{c'}_{cc} \times \left[ \tilde{V}_{n+1}(c', [d' - \mu(a[c], a)]^{+}) - \tilde{V}_{n+1}(c', [d' - \mu(a[c], a)]^{+}) \right]$$

$$- \tilde{V}_{n+1}(c', [d' - \mu(a[c], a)]^{+}) + \tilde{V}_{n+1}(c', [d - \mu(a[c], a)]^{+})$$

Noting $\tilde{V}_{n}$ is a concave function in $d$, then

$$\tilde{V}_{n}(c', [d' - \mu(a[c], a)]^{+}) - \tilde{V}_{n}(c', [d' - \mu(a[c], a)]^{+})$$

$$- \tilde{V}_{n+1}(c', [d - \mu(a[c], a)]^{+}) + \tilde{V}_{n+1}(c', [d - \mu(a[c], a)]^{+}) \leq 0,$$
which lead to (23). Then we prove that the optimal energy supplementary policy \( \pi^*_{\text{sup}}(c, d) \) is non-decreasing in \( d \).

By the similar method, we can prove \( \pi^*_{\text{com}}(c, d) \) is a submodular function in \((a, n)\), then the optimal energy supplementary policy \( \pi^*_{\text{com}}(c, d) \) is non-decreasing in \( n \). Then the results are proved.

REFERENCES

[1] L. Ruan and V. Lau, “Power control and performance analysis of cognitive radio systems under dynamic spectrum access and imperfect knowledge of system state,” IEEE Trans. Wireless Commun., vol. 8, no. 9, pp. 4616–4622, Sep. 2009.

[2] Z. Hasan, G. Bansal, E. Hossain, and V. K. Bhargava, “Energy-efficient power allocation in OFDM-based cognitive radio systems: A risk-return model,” IEEE Trans. Wireless Commun., vol. 8, no. 12, pp. 6078–6088, Dec. 2009.

[3] W. L. Huang and K. B. Letaief, “Cross-layer scheduling and power control combined with adaptive modulation for wireless ad hoc networks,” IEEE Trans. Commun., vol. 55, no. 4, pp. 728–739, Apr. 2007.

[4] S. Gao, L. Qian, and D. R. Vaman, “Energy efficient adaptive modulation in wireless cognitive radio adhoc networks,” in Proc. IEEE SECON, Jun. 2008, pp. 1–8.

[5] 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.

[6] C. K. Ho and R. Zhang, “Optimal energy allocation for wireless communications with energy harvesting constraints,” IEEE Trans. Signal Process., vol. 60, no. 9, pp. 4808–4818, Sep. 2012.

[7] 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.

[8] K. Tutuncuoglu and A. Yener, “Optimum transmission policies for battery limited energy harvesting nodes,” IEEE Trans. Wireless Commun., vol. 11, no. 3, pp. 1180–1189, Mar. 2012.

[9] G. Huang, R. Zhang, and S. Cui, “Throughput maximization for the Gaussian relay channel with energy harvesting constraints,” IEEE J. Sel. Areas Commun., vol. 31, no. 8, pp. 1469–1479, Aug. 2013.

[10] J. Xie, W. Cai, L. Deng, and J. Yang, “QoS-constrained transmission policy in hybrid energy supply wireless communication system,” in Proc. IEEE 85th Veh. Technol. Conf. (VTC Spring), Sydney, NSW, Australia, Jun. 2013, pp. 1–6.

[11] 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.

[12] L. Liu, A. Xiong, P. Yu, L. Feng, W. Li, X. Qiu, and M. Wang, “Energy-saving management mechanism based on hybrid energy supplies in multi-operator shared LTE networks,” in Proc. IEEE/IFIP Netw. Oper. Manage. Symp. (NOMS), Taipei, Taiwan, Apr. 2018, pp. 1–5.

[13] D. Chang, X. Wen, Y. Chen, Z. Lu, and H. Shao, “Linear programming based energy cooperation in cellular networks powered by hybrid energy,” in Proc. IEEE Int. Conf. Energy Internet (ICEI), Beijing, China, Apr. 2017, pp. 238–243.

[14] N. A. T. Han, “Optimizing green energy utilization for mobile networks with hybrid energy supplies,” in Green Mobile Networks: A Networking Perspective. Piscataway, NJ, USA: Wiley-IEEE Press, 2016, pp. 91–128.

[15] Q. Zhang and S. A. Kassam, “Finite-state Markov model for Rayleigh fading channels,” IEEE Trans. Commun., vol. 47, no. 11, pp. 1688–1692, Nov. 1999.

[16] H. S. Wang and N. Moayeri, “Finite-state Markov channel—A useful model for radio communication channels,” IEEE Trans. Veh. Technol., vol. 44, no. 1, pp. 163–171, Feb. 1995.

[17] S. Mao, M. H. Cheung, and W. S. Wong, “Joint energy allocation for sensing and transmission in rechargeable wireless sensor networks,” IEEE Trans. Veh. Technol., vol. 63, no. 6, pp. 2862–2875, Jul. 2014.

[18] M. Ngo and V. Krishnamurthy, “Optimality of threshold policies for transmission scheduling in correlated fading channels,” IEEE Trans. Commun., vol. 57, no. 8, pp. 2474–2483, Aug. 2009.

[19] R. C. Chen and G. L. Blankenship, “Dynamic programming equations for discounted constrained stochastic control,” IEEE Trans. Autom. Control, vol. 49, no. 5, pp. 699–709, May 2004.

[20] A. B. Piunovskiy and X. Mao, “Constrained Markovian decision processes: The dynamic programming approach,” Oper. Res. Lett., vol. 27, pp. 119–126, Oct. 2000.

[21] A. B. Piunovskiy, “Dynamic programming in constrained Markov decision processes,” Control Cybern., vol. 35, no. 3, p. 645, 2006.

[22] R. K. Ritt and L. I. Sennott, “Optimal stationary policies in general state space Markov decision chains with finite action sets,” Math. Oper. Res., vol. 17, no. 4, pp. 901–909, Nov. 1992.

[23] R. Amir, “Supermodularity and complementarity in economics: An elementary survey,” Southern Econ. J., vol. 71, no. 3, pp. 636–660, 2005.

[24] O. Hernández-Lerma, Adaptive Markov Control Processes. New York, NY, USA: Springer, 2012.