IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 26, 2020, accepted January 5, 2021, date of publication January 18, 2021, date of current version January 25, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3052218

Transmission Power Rate Control for EHD With Temporal and Complete Deaths

YUN LI[1,2], SHENGDA TANG[3], AND ZHICHENG TAN[2,4]
[1]School of Information and Statistics, Guangxi University of Finance and Economics, Nanning 530003, China
[2]College of Electronic Information and Automation, Guilin University of Aerospace Technology, Guilin 541004, China
[3]School of Mathematics and Statistics, Guangxi Normal University, Guilin 541004, China
[4]School of Informatics, Xiamen University, Xiamen 361005, China

Corresponding author: Shengda Tang (tangsd911@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61761008 and Grant 61861014, and in part by the Natural Science Foundation of Guangxi under Grant 2018GXNSFAA281238.

ABSTRACT This paper proposes a continuous-time communication model of an energy harvesting device (EHD) in the scenario that such EHD suffers from the temporal death caused by the energy depletion and completion death caused by the destruction of its hardware or software, respectively. The harvested energy is modeled as continuous fluid process, it arrives continuously and varies in time. The data is assumed to be infinite backlog, and the EHD transmits it via a wireless channel fluctuates randomly due to fading. We reformulate the system model into an equivalent piece-deterministic Markov process (PDMP) based on the imbedded discrete-time decision epoch sequence of the system model, and then an infinite horizon discrete-time Markov decision process (MDP) is built. We show the existence of the stationary deterministic optimal transmission power rate (TPR) policy, and an algorithm for computing the TPR policy and the maximum total expected throughput is developed. Finally, numerical examples are provided to confirm the analytical findings. The effects of some system parameters to the optimal TPR policy and the maximum expected throughput are investigated numerically.

INDEX TERMS EH-WSN, transmission power rate (TPR), throughput maximization, temporal death, complete death, piece-deterministic markov process(PDMP).

I. INTRODUCTION

Conventional energy-constrained wireless sensor networks (WSNs) in general have limited lifetime. To alleviate these energy scarcity issue, numerous energy harvesting technologies emerge, such as mini solar panel, piezoelectric transducers, and cognitive radio. The energy harvesting (EH) technique enables the WSN nodes to harvest energy from the natural sources, and store the collected energy in a storage for future use. By embedding the energy harvesting modules, energy harvesting WSNs (EH-WSNs) can exploit renewable energy sources in the environment continuously to provide energy consumption for the EH-WSN nodes. Harvesting energy from ambient sources becomes a promising approach to prolong the lifetime of wireless networks.

Since the EH-WSN nodes harvests energy from the ambient energy sources asynchronously or periodically, the harvested energy varies dramatically over time due to the varying environmental conditions. In many applications, the collected energy cannot meet the power requirement of the EH-WSN node uninterruptedly. Thus, the effective utilizations of the harvested energy have received much attention from the industry, and it becomes an important research area.

In contrast to the energy management strategy for the traditional sensor nodes, additional considerations to the impacts from the energy replenishment process are required. The energy management policy of the nodes need to adapt to the harvested energy status. Obviously, an overly conservative energy expenditure may limit the transmitted data by failing to take the full advantage of the harvested energy, while an overly aggressive use of energy may result in an energy outage, which prevents nodes from functioning properly in the future.

Recent years there has been a significant amount of literature is focused on the efficient optimal transmission

The associate editor coordinating the review of this manuscript and approving it for publication was Guangjie Han.

and energy allocation policies with different objectives in EH-WSNs. For example, [1] studies the joint energy allocation and routing problem for network utility maximization, and an online solution is developed to achieve asymptotic optimality. In [2], the authors propose a throughput-optimal energy allocation algorithm for a times-slotted system under time-varying fading channel and energy source by using Markov Decision Process (MDP). To maximize the aggregate average importance of the transmitted data, the authors of [3] numerically show that the Balanced Policy (BP) and the Heuristic Constrained Energy Independent Policy (HCEIP) achieve near optimal performance in most cases of interest. In [4], the authors develop optimal energy scheduling algorithms for N-user fading multiple-access channels with EH to maximize the channel sum-rate. In [5], the authors consider wireless powered communication networks, and they propose a transmission scheme for optimal allocation of the BS broadcasting power and time sharing among the wireless nodes to maximize the overall network throughput. The authors of [6] propose a multi-layer Markov fluid queue (MLMFQ) model to model the EH-WSN node with temporal death, an algorithm for the optimal transmission policy to maximize the steady-state average reward rate of the reported data packets is proposed. The paper [7] explores joint power allocation and route selection in a multi-hop cognitive radio network. A new frame structure of radio frequency EH-cooperation-transmission is considered, and closed form expressions of the optimal time duration for EH and power allocation factor on each relay are developed. An overview of recent exciting achievements related to the EH wireless communication system are provided in [8]–[10] and the references therein.

In this paper, we develop a novel wireless communication of the EH-WSN node, and investigate the optimal control policy of the transmission power rate (TPR) to maximal the expected throughput. Compared to the literature mentioned above, our work have the following main contributions.

- Firstly, we develop a continuous-time hybrid-state stochastic system model to describe the wireless communication of the EH-WSN node. Especially, two discrete-state stochastic processes are used to capture the EH process and channel fading process respectively. Unlike previous work, we use a continuous fluid process to catch the dynamic of the energy level in the EH-WSN node. In the later we will see that continuous fluid model of the energy can characterise the energy evolution at a more finer granularity.
- Secondly, two peculiar phenomena, *i.e.,* temporal death and complete death, of the EH-WSN node are considered simultaneously in our model. More specifically, when the node runs out of the energy, it will be dead temporally, and it will restart again after the accumulated energy reaches a predefined threshold. When the node suffers from hardware or software damage and can not self-healing, the complete death will be occurred, and the node will be dead permanently. We believe that the current framework is more practical and more natural to guide the design and optimization of the communication of the EH WSNs.
- Thirdly, with the system state information, *i.e.,* the current energy harvesting state, fading channel status, and the current residual energy level, we propose an optimal control policy to achieve the maximum expected amount of data transmitted during the lifetime of the EH-WSN node. And the TPR control policy can be computed offline and stored in a lookup table for implementation.

The remainder of this paper is organized as follows. The system model is given in Section II, and the control problem is reformulated in Section III. In Section IV, we reformulate our control problem, and the optimal policy and the corresponding algorithm are derived. In Section V, we present numerical examples to illustrate the analytical results, and close with conclusions in Section VI.

II. MODEL DESCRIPTION

Now, we develop a continuous time communication system model to describe the data transmission from a terminal EH device (EHD) to a receiver. In such a scenario, we assume the EHD has a back-logged data queue for transmission, and it transmits the data via a point-to-point wireless communication channel fluctuates randomly due to fading. The EHD harvests energy continuously from the ambient environment to meet the energy requirement. The energy arrives continuously and varies in time. Furthermore, the EHD suffers from the temporal death and complete death, which will be described in more detail later. Our main goal is to develop an optimal transmission scheduling policy for the transmit power to achieve the maximum total expected throughput in its lifetime.

The following is the detail description of the working mechanism and control schedule of the EHD. The main nomenclatures are listed in Table 1.

TABLE 1. The main nomenclatures of the model.

J^H	Energy harvesting process.
J^C	Channel fading process.
B	Background process.
X	Energy level process.
q_{ij}^h	Transmit rate of energy harvesting state from i to j.
q_{ij}^c	Transmit rate of channel fading state from i to j.
q_{ij}^b	Transmit rate of background state from state from i to j.
h_t	Energy harvesting rate at time t.
h^i	Energy harvesting rate at background state i.
c_t	Channel gain rate at time t.
c^i	Channel gain rate at background state i.
π_t	Transmission power rate at time t.
θ	Capacity of the energy buffer.
T	Lifetime of the EHD.
L	Duration of the temporal death of the EHD.
η	Restart energy threshold.
α	Probability rate of the complete death of the EHD.
ψ	Maximal TPR of the EHD.

A. TRANSMISSION MODEL

We propose a finite-state continuous-time Markov chain (CTMC), which is denoted by $J^C = \{J_t^C, t \geq 0\}$, to model the channel fading process similar to [13], [14]. Specifically, we assume that channel state space is divided into n^c non-overlapping state, which is denoted by $\Omega^C = \{1, 2, \ldots, n^c\}$. Let transition rate matrix of J^C be $Q^C = [q_{ij}^c]$, where q_{ij}^c, $i, j \in \Omega^C$, is the transmit rate of the channel from i to $j(\neq i)$. We denote the channel gain rate by c_t at time t, and assume that the channel gain rate in each state is fixed and known.

Consider a continuous-time communication of the EHD that similar to that presented in [12]. At time t, let the transmission power rate (TPR) of the EHD be π_t, then the instantaneous rate of the EHD at time t is

$$r[c_t, \pi_t] = Wlog_2\left(1 + \frac{c_t \pi_t}{N_0 W \Gamma}\right), \tag{1}$$

where W is the bandwidth of the channel, N_0 is the power spectral density, and Γ is the SNR gap.

B. ENERGY EVOLUTION AND DEATH

We adopt the view of [15] and establish a finite-state CTMC to describe the EH process. More specifically, let $J^H = \{J_t^H, t \geq 0\}$ describe the evolution of the ambient energy source. The energy source has n^h states, which is denoted by $\Omega^H = \{1, 2, \ldots, n^h\}$. We denote the transition rate matrix of J^H be $Q^H = [q_{ij}^h]$, where q_{ij}^h, $i, j \in \Omega^H$, means the transmit rate of the resources from i to $j(\neq i)$. The energy harvesting rate (EHR), which is denoted by h_t at time t, is assumed to be constant in each fixed state. The energy harvested by the EHD is stored in an energy buffer and can be used immediately. The energy buffer is with finite capacity which is denoted by θ, when the collected energy reaches its capacity, the excess energy will be discarded.

To characterise the energy evolution at a more finer granularity, we use a continuous fluid level process to model the energy dynamic in the energy buffer. Denote the energy evolution process by $X = \{X_t, t \geq 0\}$, and the state space of X is denoted by $\Omega^X = [0, \theta]$. Let $\Delta_{\pi_t}^{h_t} = h_t - \pi_t$, and we refer to $\Delta_{\pi_t}^{h_t}$ as the energy net increment rate of the EHD at time t. We can see that when $\Delta_{\pi_t}^{h_t} < 0$, the energy in the buffer decreases, as soon as the EHD runs out of its energy, *i.e.*, $X_t = 0$, the EHD would be dead temporally due to energy depletion, and we refer to this type of death as temporal death in this paper.

Different from the general energy outage, when the temporal death is occurred, the device becomes inactive and its communication functionalities are consequently lost. However, the EH module of the device is still harvesting energy from the environment independently, which results in the accumulating of the energy in the energy buffer. The EHD has to be kept in the dead state until the amount of harvested energy meets a predefined threshold energy level, which is denoted by $\eta \in [0, \theta]$, then the device is triggered to wake up and switch to a normal working fashion. We denote the temporal death duration by L, it can be seen that L is a random variable, it is independent of the TPR and completely determined by the energy harvesting process. A more detailed analysis of the temporal death can be found in our previous work [6]. A sample path of the energy level with temporal death is described in Fig. 1.

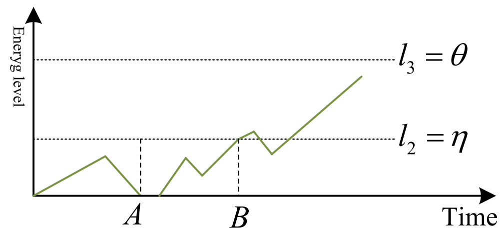

FIGURE 1. A sample path of the energy level in the energy buffer of the EHD, where the interval *AB* is one of the temporal death duration.

The temporal death can not be avoid completely in a stochastic communication system without backup energy, because the energy supply is determined completely by the ambient environment, and the basal energy expenditure for the EHD is necessary.

Moreover, we assume that the EHD is typically deployed in hostile environments, and it is highly susceptible to the random physical destructions. When it suffers unrecoverable hardware or software destruction, it will be dead permanently. We refer to this type of death as complete death. We let $\alpha > 0$ be the complete death rate, and α is assumed to be constant [16], [17]. We denote the lifetime of the device by T, and we can see T is a random variable and it is exponential distributed with mean α^{-1}.

In summary, the evolution of the energy level process X of the EHD can be described as

$$\frac{\partial X_t}{\partial t} = \begin{cases} \min\{0, \Delta_{\pi_t}^{h_t}\} & X_t = \theta, t \notin L, \\ \Delta_{\pi_t}^{h_t} & 0 < X_t < \theta, t \notin L, \\ h_t & 0 \leq X_t < \eta, t \in L. \end{cases} \tag{2}$$

Now, we obtain our stochastic system model, which is denoted by

$$[J^H, J^C, X] = \{[J_t^H, J_t^C, X_t], t \geq 0\}.$$

It is important to note that the evolution of J^H only depends on the environment factors, such as climate, time, location, and other ambient conditions. Thus we assume that the process J^H operates independently of the process J^C.

III. CONTROL FORMULATION

A. CONTROL MODE

When the system state changes, a new TPR is taken to adapt to the new system state. More specifically, the TPR will be not rescheduled when the state of the system keep constant, and the decision of choosing a new TPR is only driven by one of the following events: (i) Energy harvesting state changes; (ii) Channel fading status changes; (iii) Energy overflow occurs; (vi) EHD wakes up from a temporal death.

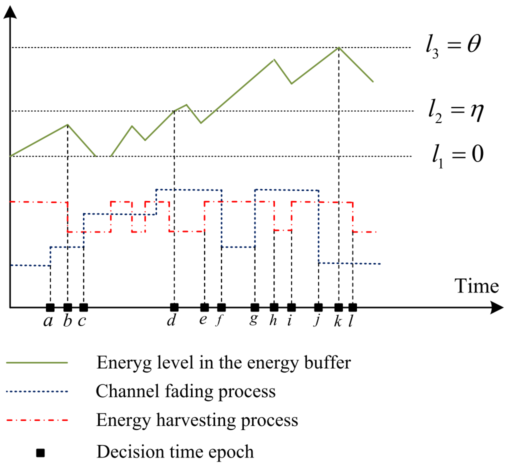

FIGURE 2. The sample path of the decision epoch sequence, where the points a, c, f, g, j are the time epochs that the channel fading changes, b, e, h, l are the energy harvesting state changing points, points c is temporal death occurrence point, and d is the restart time epoch, k is the energy overflow occurrence point. It should be noted that during the time interval $[c, d]$, there is no decision epochs due to the temporal death.

We consider the optimal TPR control policy with system state information. That is, a TPR is chosen only based on the current known system state, *i.e.,* energy harvesting state, the channel status, the energy level in the buffer, and no other more information are known.

We index the decision time epochs by $t_0, t_1, t_2, \ldots$ with the convention that $t_0 = 0$, and denote decision epoch set by $\{t_n\} = \{t_n, n = 0, 1, 2, \ldots\}$. At each decision epoch t_n, one of the TPR should be chosen to transmit the current data. In practice, the EHD just has finite different TPR levels, we assume that there are n^a available discrete TPR levels, which is denoted by $\mathcal{A} = \{a_1, a_2, \ldots, a_{n^a}\}$, where $0 = a_1 < a_2 < \ldots < a_{n^a} = \psi < \infty$, and ψ is the maximal TPR that determined by the configuration of the EHD. We refer to a_n, $n = 1, 2, \ldots, n^a$, as an action, and $\mathcal{A}$ as the action space. At time epoch t_n, let a TPR $\widehat{\sigma}_n \in \mathcal{A}$ is chosen, we then define

$$\pi_t = \widehat{\sigma}_n \in \mathcal{A}, \quad t \in [t_n, t_{n+1}), \; n = 0, 1, \ldots, \tag{3}$$

and π is referred to as a TPR policy. Since the TPR must be kept constant in each epoch, thus TPR control policy π_t is a piecewise linear function. It is intuitive that, when the EHD is in the temporal death, we need not to control the TPR. That is, we can assume that $\pi_t = 0$ when $t \in L$.

B. CONTROL PROBLEM

Let the initial state of the system be $[J_0^H, J_0^C, X_0] = [j, k, x]$, which means that the EH state is $j \in \Omega^H$, channel status is $k \in \Omega^C$, and the energy level in the EHD is $x \in \Omega^X$. We define

$$V^\pi[j, k, x] = \mathbb{E}^\pi_{[j,k,x]}\Big[\int_0^T r[c_t, \pi_t]dt\Big], \tag{4}$$

where $\mathbb{E}^\pi_{[j,k,x]}[\cdot]$ represents the conditional expectation under control π given the initial state $[j, k, x]$, $I_{\{\cdot\}}$ is an indicator function, $r[\cdot,\cdot]$ is given in (1) and T is the lifetime of the EHD. Then $V^\pi[j, k, x]$ is total expected throughput of the EHD during its lifetime when TPR policy π is taken. We also call V as value function in this paper.

We rewrite (4) equivalently as

$$\begin{aligned} V^\pi[j, k, x] &= \mathbb{E}^\pi_{[j,k,x]}\Big[\int_0^\infty \int_0^s \alpha e^{-\alpha s} r[c_t, \pi_t]dtds\Big] \\ &= \mathbb{E}^\pi_{[j,k,x]}\Big[\int_0^\infty e^{-\alpha t} r[c_t, \pi_t]dt\Big], \end{aligned}$$

It means that $V^\pi[j, k, x]$ can be considered to be a total expected α-discounted amount of the throughput of the EHD, and α can be interpreted as a discount factor in the value functions after reformulating the system model as an infinite horizon MDP later. Now, we obtain our optimization problem as follows.

For given $[j, k, x]$ under (1), (2), (3), (4), find a TPR policy π^* defined in (3) such that,

$$V[j, k, x] = V^{\pi^*}[j, k, x] = \max_\pi \{V^\pi[j, k, x]\} \tag{5}$$

Remark 1: In our control mode, we choose four events as the trigger event. In fact, we can add some other events to improve the control intensity in the system, for example, the event when the energy level reaches some predefined warning threshold, to be a trigger event without changing our theoretical analytical framework.

IV. CONSTRUCTION FORMULATION

In our control mode, we first note that the decision time sequence $\{t_n\}$ is a random time sequence, the evolution of the system is not deterministic even if some (optimal or not) policy is given due to the random nature of the system model. Moreover, since we control the system at the occurrence time epochs of the energy overflow and temporal death, the system is not a Markov (fluid) process any more, and the markov theory cannot be used directly.

To deal with the challenges, it is important to note that the markov property of the system still exists on some special time epochs, *i.e.,* the decision epoch sequence $\{t_n\}$. Then based on this important observation, we consider the sequence of states at these decision epochs, and using the similar method provided in [18], [19], and [20], we can reformulate our system model into a piecewise-deterministic Markov process (PDMP) based on the system state information. Ultimately we can obtain a discrete-time infinite-horizon MDP, which results in a solution of the control problem.

A. BACKGROUND PROCESS

To simplify the presentation in this paper, we first superimpose the two processes J^H and J^C. Let

$$\begin{aligned} B &= \{B_t, t \geq 0\} = \{[J_t^H, J_t^C], t \geq 0\}, \\ \Omega^B &= \{[i, j] | i \in \Omega^H, j \in \Omega^C\}. \end{aligned}$$

We refer to B and Ω^B as the background process and its state space, respectively.

To simplify the notation, we put the state $[i, j]$ in Ω^B in lexicographic order, and label all states in Ω^B from 1 to $n^h \times n^c$. Let $n^b = n^h \times n^c$, then state space Ω^B can be simply written as $\Omega^B = \{1, 2, \ldots, n^b\}$. In the remainder of the paper, we use the second symbols for simplification. The EH rate and the channel gain rate of B are denoted respectively by h^i and c^i when the background state is $i \in \Omega^B$. That is, when $B_t = i$, we have $h_t = h^i$ and $c_t = c^i$, respectively.

In order to introduce the complete death into our system model, let us imagine that there is an additional artificial cemetery state in the state space of the background process, and assume that the process enters this state with probability rate α, once the process B enters in the cemetery state, the system will be killed and terminated. Then B becomes a terminating CTMC on the state space Ω^B, and its transition rate matrix is given as $\mathbf{Q}^B = [q^b_{ij}] = Q^H \oplus Q^C - \alpha I$, where $\oplus$ denotes the Kronecker sum, and I represents unit matrix with an appropriate dimension [21].

It should be note that the control policy cannot change the evolution of the background process B. Now, we reduce our system model to the following simple form

$$[B, X] = \{[B_t, X_t], t \geq 0\}$$

on state space $\Omega^B \times \Omega^X$.

Since the randomness of the lifetime of the EHD are injected in the background process, in the reminder of this paper, we think equivalently that the EHD has a infinite lifetime, such equivalent assumption will make our presentation more easier.

B. CONTROL REFORMULATION

Now, we reformulate our system model as an PDMP. For conveniently describing the trajectory of the system model, we define the following first passage times (FPTs). Let

$$\nu_i = inf\{t > 0 : B_t \neq i,\ B_0 = i\} \quad i \in \Omega^B,$$
$$\kappa^y_x = inf\{t > 0 : X_t = y, X_0 = x\} \quad x \in \Omega^X.$$

We can see that ν_i is the time epoch that the background process B first leaves state i, and κ^y_x is the time epoch that the energy level process X starts from level x, and it first hits the energy level y. In particular, when $\Delta^{h^i}_a > 0$, the FPT κ^θ_x means the first occurrence time of the energy overflow of the EHD with initial energy level x and action a is taken, and we have

$$\kappa^\theta_x = \frac{\theta - x}{\Delta^{h^i}_a} \quad 0 \leq x \leq \theta.$$

Similarly, when $\Delta^{h^i}_a < 0$, the FPT κ^0_x is the first occurrence time of the temporal death due to the energy depletion, and we have

$$\kappa^0_x = \frac{x}{-\Delta^{h^i}_a} \quad 0 \leq x \leq \theta.$$

In the sequel, we reconstruct the sample path of the system model $[B, X]$ iteratively based on the imbedded discrete time decision epoch sequence $\{t_n\}$.

Let the current time epoch be $t_0 = 0$, and the current system state be $[B_0, X_0] = [i, x]$. We assume the action $\widehat{\sigma}_0 = a \in \mathcal{A}$ is chosen. The evolution of the system from t_0 is determined by the following three stochastic cases.

- The first case: $\nu_i \leq \min\{\kappa^\theta_x, \kappa^0_x\}$.
  In this case, the background changes its state first before the occurrences of energy overflow and the temporal death, then we have
  $$t_1 = \nu_i; \quad X_{t_1} = x + \Delta^{h^i}_a \nu_i,$$
  The transition probability density, which is denoted by $P^a[j, dy|i, x]$, in this case can be given as
  $$\begin{aligned}\widehat{P}^a&[j, dy|i, x]\\ &= P^a[B_{t_1} = j, X_{t_1} \in dy|i, x]\\ &= \int_0^{\min\{\kappa^\theta_x, \kappa^0_x\}} q^b_{ij} e^{-q^b_i t} I_{\{i \neq j,\ x + \Delta^{h^i}_a t \in dy\}} dt, \end{aligned} \quad (6)$$
  where we let $q^b_i = -q^b_{ii}$, $i \in \Omega^B$.
- The second case: $\Delta^{h^i}_a > 0$, and $\nu_i > \kappa^\theta_x$.
  In this case, the energy level increases, and the overflow is occurred first before the background changes its state, and we have
  $$t_1 = \kappa^\theta_x; \ X_{t_1} = \theta; \ B_{t_1} = i.$$
  The transition probability density is given as
  $$\begin{aligned}\widehat{P}^a[j, dy|i, x] &= \int_{\kappa^\theta_x}^\infty q^b_i e^{-q^b_i t} I_{\{i=j,\ \theta \in dy\}} dt\\ &= I_{\{i=j,\ \theta \in dy\}} e^{-q^b_i \kappa^\theta_x}. \end{aligned} \quad (7)$$
- The third case: $\Delta^{h^i}_a < 0$, and $\nu_i > \kappa^0_x$.
  In this case, the energy level decreases, and the temporal death is occurred first before the background changes its state, and the system is temporally dead until the energy level first reaches the restart energy threshold η. Thus, we have
  $$t_1 = \kappa^0_x + \kappa^\eta_0; \quad X_{t_1} = \eta,$$
  The transition probability density is given as
  $$\begin{aligned}\widehat{P}^a[j, dy|i, x] &= \int_{\kappa^0_x}^\infty q^b_i e^{-q^b_i t} g_{ij}[\eta] I_{\{\eta \in dy\}} dt\\ &= I_{\{\eta \in dy\}} g_{ij}[\eta] e^{-q^b_i \kappa^0_x}. \end{aligned} \quad (8)$$
  The probability $g_{ij}[\eta]$ is defined as
  $$g_{ij}[\eta] = P[B_{\kappa^\eta_0} = j, X_{\kappa^\eta_0} = \eta \mid B_0 = i, X_0 = 0],$$
  and we denote $G[\eta] = (g_{ij}[\eta])$. The derivation of $g_{ij}[\eta]$ is too cumbersome, and we give it in the appendix.

We are considering the throughput obtained during $[0, t_1)$ as a lump reward received at the beginning of the interval. Thus, during $[0, t_1)$ the throughput is given as follows

$$\begin{aligned}\widehat{R}^a&[i, x]\\ &= \mathbb{E}^\pi_{[i,x]}\Big[\int_0^{t_1} r[c_t, \pi_t] dt\Big]\end{aligned}$$

$$= I_{\{\Delta_a^{h^i}=0\}}\Big[\int_0^{\infty} q_i^b e^{-q_i^b t} r[c^i, a]tdt\Big] + I_{\{\Delta_a^{h^i}>0\}}\Big[\int_0^{\kappa_x^\theta} q_i^b e^{-q_i^b t} r[c^i, a]tdt + e^{-q_i^b\kappa_x^\theta} r[c^i, a]\kappa_x^\theta\Big] + I_{\{\Delta_a^{h^i}<0\}}\Big[\int_0^{\kappa_x^0} q_i^b e^{-q_i^b t} r[c^i, a]tdt + e^{-q_i^b\kappa_x^0} r[c^i, a]\kappa_x^0\Big].$$

A short calculation gives the expected throughput during $[0, t_1)$ as follows

$$\widehat{R}^a[i, x] = \frac{r[c^i, a]}{q_i^b}\Big[1 - e^{-q_i^b\kappa_x^\theta} I_{\{\Delta_a^{h^i}>0\}} - e^{-q_i^b\kappa_x^0} I_{\{\Delta_a^{h^i}<0\}}\Big]. \tag{9}$$

Now, we obtain the state at t_1 and the reward in $[0, t_1)$, which is viewed as the reward at time epoch 0. At time epoch t_1, a new decision $\widehat{\sigma}_1 \in \mathcal{A}$ is chosen, and the system model "restarts" from t_1 to reach the time epoch $t_2, t_3, \ldots$, in the same fashion described above until it enters the cemetery state, and then the process is killed and terminated.

We denote the new embedded discrete-time stochastic process in the PDMP as

$$[\widehat{B}_n, \widehat{X}_n] = \{[B_{t_n}, X_{t_n}], n = 0, 1, 2, \ldots\}.$$

It is noticeable that all the states, the transition probabilities, and the expected throughput of $[\widehat{B}_n, \widehat{X}_n]$ at the next time epoch are only dependent on the current state and the action, this important fact guarantees that we can reformulate the embedded process $[\widehat{B}_n, \widehat{X}_n]$ into an equivalent infinite horizon discrete-time MDP [24].

C. REFORMULATION OF THE CONTROL PROBLEM

Based on $[\widehat{B}_n, \widehat{X}_n]$, we reformulate the original control problem as an equivalent infinite horizon discrete-time MDP given as follows:

- State space: $\widehat{\Omega} = \Omega^B \times \Omega^X$.
- Action space: $\widehat{\mathcal{A}} = \mathcal{A}$.
- One-step reward: $\widehat{R}^a[i, x]$ which is given in (6), for all $[i, x] \in \widehat{\Omega}$ and $a \in \widehat{\mathcal{A}}$.
- Transition density: $\widehat{P}^a[j, dy|i, x]$, which is given in (6)-(8) for all $[i, x], [j, y] \in \widehat{\Omega}$ and $a \in \widehat{\mathcal{A}}$.

We denote this MDP by $\widehat{\mathcal{M}}$, and define the policy of $\widehat{\mathcal{M}}$ by

$$\widehat{\sigma} = \{\widehat{\sigma}_0, \widehat{\sigma}_1, \widehat{\sigma}_2, \ldots\}, \tag{10}$$

where $\widehat{\sigma}_n \in \widehat{\mathcal{A}}$, $n = 0, 1, 2, \ldots$.

Under the policy $\widehat{\sigma}$, the reward function of $\widehat{\mathcal{M}}$ is defined as

$$\widehat{V}^{\widehat{\sigma}}[i, x] = \mathbb{E}^{\widehat{\sigma}}_{[i,x]}[\sum_{n=0}^{\infty} \widehat{R}^{\widehat{\sigma}_n}[\widehat{B}_n, \widehat{X}_n]] \tag{11}$$

for all $[i, x] \in \widehat{\Omega}$.

The optimization problem of $\widehat{\mathcal{M}}$ is defined as

$$\widehat{V}[i, x] = \max_{\widehat{\sigma}} \widehat{V}^{\widehat{\sigma}}[i, x], \tag{12}$$

and $\widehat{V}[i, x]$ is the maximum expected throughput over a infinite horizon under policy $\widehat{\sigma}$.

The next lemma shows that the original control problem (5) can be treated as the discrete time control problem (9) equivalently.

Lemma 1: Let π be the control policy for $[B, X]$ defined in (3), and $\widehat{\sigma}$ be the control policy of $\widehat{\mathcal{M}}$ defined in (10). Then we have

$$\widehat{V}^{\widehat{\sigma}}[i, x] = V^{\pi}[i, x], \tag{13}$$

$$\widehat{V}[i, x] = V[i, x]. \tag{14}$$

Proof: (14) can be obtained directly from (13).

To prove (13), we denote the natural filtration of $[B, X]$ upon time t_n by $\mathfrak{F}_n$, and let $\mathbb{E}^{\widehat{\sigma}_n}_{\mathfrak{F}_n}[\cdot]$ denote the conditoinal expection $\mathbb{E}^{\widehat{\sigma}_n}_{\mathfrak{F}_n}[\cdot|\mathfrak{F}_n]$. Noting

$$\pi_t = \widehat{\sigma}_n, \quad t \in [t_n, t_{n+1}),\ n = 0, 1, \ldots. \tag{15}$$

Then in view of the strong Markov property, the value function $V^{\pi}[i, x]$ of (4) can be rewritten in terms of the process $(\widehat{B}_n, \widehat{X}_n)$ as

$$\begin{aligned} V^{\pi}[i, x] &= \mathbb{E}^{\pi}_{[i,x]}[\int_0^L r[c_t, \pi_t]dt] \\ &= \mathbb{E}^{\pi}_{[i,x]}[\sum_{n=0}^{\infty}\int_{t_n}^{t_{n+1}} r[c_t, \pi_t]dt] \\ &= \mathbb{E}^{\pi}_{[i,x]}\Big[\sum_{n=0}^{\infty}\mathbb{E}^{\widehat{\sigma}_n}_{\mathfrak{F}_n}[\int_{t_n}^{t_{n+1}} r[c_t, \pi_t]dt]\Big] \\ &= \mathbb{E}^{\pi}_{[i,x]}[\sum_{n=0}^{\infty}\widehat{R}^{\widehat{\sigma}_n}[\widehat{B}_n, \widehat{X}_n]] \\ &= \widehat{V}^{\widehat{\sigma}}[i, x] \end{aligned}$$

This completes the proof. □

From lemma 2, to obtain the control policy π for the original optimal problem (5), we just need to investigate the control policy $\widehat{\sigma}$ of $\widehat{\mathcal{M}}$ based on the connection (12), the existence of a policy $\widehat{\sigma}$ for $\widehat{\mathcal{M}}$ would imply the existence of a policy π for the optimal problem (5), and vice versa.

D. EXISTENCE OF OPTIMAL STATIONARY POLICIES

For the MDP $\widehat{\mathcal{M}}$ obtained above, the following theorem shows the optimal deterministic stationary policy exists in our control problem.

Theorem 1: In the MDP $\widehat{\mathcal{M}}$ described above, the reward function $\widehat{V}$ satisfies uniquely the following dynamic programming equation

$$\widehat{V}[i, x] = \max_{a\in\widehat{\mathcal{A}}}\Big\{\widehat{R}^a[i, x] + \sum_{j=1}^{n^b}\int_0^{\theta}\widehat{V}[j, y]\widehat{P}^a[j, dy|i, x]\Big\}. \tag{16}$$

If for each $[i, x] \in \widehat{\Omega}$, we choose $a^*[i, x]$ to be the smallest action in $\mathcal{A}$ to achieve the maximum in (16), then $\widehat{\sigma} = \{a^*, a^*, \ldots\}$ is an optimal policy.

Proof: By the principle of optimal programming, we have

$$\widehat{V}[i,x] = \max_{a\in\widehat{\mathcal{A}}} \Big\{ \widehat{R}^a[i,x] + \sum_{j=1}^{n^b} \frac{q_{ij}^b}{q_i^b} \times \int_0^\infty q_i^b e^{q_i^b t} \widehat{V}[j, \min\{(x+\Delta_a^{h^i})^+, \theta\}] dt \Big\}, \quad (17)$$

where the notation $[z]^+ = \max\{0, z\}$.

Substitute (6)-(8) into above equation (17), and we obtain the bellman equation (16) directly.

Now, we show that MDP $\widehat{\mathcal{M}}$ satisfies the following conditions:

(a) For each $[i,x] \in \Omega$, the feasible action set $\widehat{\mathcal{A}}$ is a non-empty and finite set.

(b) For all $[i,x] \in \Omega$, the value function $\widehat{V}[i,x] < \infty$.

The condition (a) is obvious, and we only need to check condition (b). Since complete death probability rate $\alpha > 0$ in the system model, then the sum in (11) has only finitely many non-zero terms, and then the discrete time optimal problem (12) is well defined. That is, for all $[i,x] \in \widehat{\Omega}$ and policy $\widehat{\sigma}$, we have $\widehat{V}[i,x] < \infty$. Under the conditions, by the similar proof of Th. 2.2 in [25], we can obtain the results. □

Remark 2: The assumption of non-empty and finiteness of the action set in our model simplify the mathematical theory needed for the existence of optimal solution. The proof of the existence of the optimal stationary solution in the continuous (compact) action set case can be found in [26] and [27].

E. OPTIMAL TPR ALGORITHM

In the following, we propose an optimal TPR control algorithm for the TPR control policy of the EHD based on theorem 1. In this paper, we use a state-space discretization to yield recursive approximations to MDP. This is a general method for stochastic control problems with hybrid-state space. The convergence and the complexity of this algorithm can be found in [27]. The TPR algorithm is summarized into Algorithm 1.

In Algorithm 1, we can adjust the parameters k and ε in step 1 and step (4-i) according to the accuracy required in practical application. In step (3-iv) the probability $\widehat{p}^a[j, x_v|i, x_u]$ can be calculated by the definite integral directly. For example, when $\Delta_a^{h^i} > 0$, $v > u$ and $i \neq j$, we have

$$\begin{aligned}\widehat{p}^a[j, x_v, k|i, x_u] &= \int_{\omega_v} \widehat{p}^a[j, dy|i, x_u]dy \\ &= \int_{\frac{u}{k}\theta}^{\frac{v}{k}\theta} \widehat{p}^a[j, dy|i, x_u]dy \\ &= \frac{q_{ij}^b}{q_i^b}(e^{-q_i^b \frac{v-u}{k}} - e^{-q_i^b \frac{v-u+1}{k}}).\end{aligned}$$

Algorithm 1 Computing the Optimal TPR Policy and Maximum Expected Throughput

- **Step 1.** Specify $k > 0$ large enough.
- **Step 2.** Partition of Ω^X.
  (2-i) Let $\Omega_k^X = \{\omega_1, \ldots, \omega_{k+1}\}$, where $\omega_i = ((i-1)\delta, i\delta]$, $i = 1, \ldots, k$, where $\delta = \frac{\theta}{k}$.
  (2-ii) Fix any point x_i in ω_i, $i = 1, \ldots, k+1$, and let $X_k = \{x_1, \ldots, x_{k+1}\}$.
- **Step 3.** Construction of MDP $\widehat{\mathcal{M}}$.
  (3-i) Build state space $\widehat{\Omega} = \Omega^B \times X_k$.
  (3-ii) Build action space $\widehat{\mathcal{A}} = \mathcal{A}$.
  (3-iii) Compute one-step reward $\widehat{R}^a[i, x_u]$ by (9), for all $[i, x_u] \in \widehat{\Omega}$, and $a \in \widehat{\mathcal{A}}$.
  (3-iv) Compute transition probabilities
  $$\widehat{p}^a[j, x_v|i, x_u] = \int_{\omega_v} \widehat{p}^a[j, dy|i, x_u]dy$$
  by (6)-(8) and (18)-(19) for all $[i, x_u]$, $[(j, x_v] \in \widehat{\Omega}$, and $a \in \widehat{\mathcal{A}}$,
- **Step 4.** Computation of MDP $\widehat{\mathcal{M}}$.
  (4-i) Specify $\varepsilon > 0$ small enough.
  For each $[i, x_u] \in \widehat{\Omega}$, generate a candidate of $\widehat{V}_0[j, x_v] = 0$, set $n = 0$.
  (4-ii) Compute iteratively $\widehat{V}_{n+1}[i, x_u]$ by
  $$\widehat{V}_{n+1}[i, x_u] = \max_{a\in\widehat{\mathcal{A}}} \Big\{\widehat{R}^a[i, x_u] + \sum_{j=1}^{n^b}\sum_{v=1}^{k} \widehat{V}_n[j, x_v]\widehat{p}^a[j, x_v|i, x_u]\Big\}.$$
  (4-iii) If $|\widehat{V}_{n+1} - \widehat{V}_n| \leq \varepsilon$, let the maximal throughput and the optimal TPR policy be
  $$\widehat{V}[i, x_u] = \widehat{V}_{n+1}[i, x_u], \ a^*[i, x_u] = a$$
  and stop. Otherwise, increase n by 1 and go to step (4-ii).

And other transition probabilities can be obtained similarly.

V. NUMERICAL RESULTS AND OBSERVATIONS

In this section, we give numerical examples to illustrate the computation of the optimal TPR policy and the maximum expected throughput, we investigate the effect from the system parameters to the optimal TPR and the maximum expected throughput.

A. NUMERICAL SETTING AND DISCRETIZATION

In our numerical setting, we let the channel bandwidth be $W = 2$ MHz, and the noise power spectral density be $N_0 = 10^{-19}$ W/Hz. The SNR gap only depends on the error probability requirements, and we set the SNR gap be $\Gamma = 4$. The channel considered here has three states $\{0, 1, 2\}$, which represents the "worst", "fair", and "best" state of the channel. The channel gain rate of the state 0, 1, 2 are 0, 8×10^{-13} and 15×10^{-13}. Through statistical measurement, the transition rate matrix is given as

$$Q^C = \begin{bmatrix} -1 & 0 & 1 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{bmatrix}.$$

Similarly, we use {0, 1, 2} to represent the "bad", "normal" and "good" states of the energy source, and the EH rates at each state are 0, 0.05 J/s, and 0.1 J/s. The transition rate matrix is given by

$$Q^H = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 1 & -2 \end{bmatrix}.$$

Then we obtain a background process B, whose state space is denoted by $\{1, 2, \ldots, 9\}$ for simplification.

We let the action step-size and the energy step-size be $\delta_a = 0.05$ J/s and $\delta_e = 0.02$ J/s, respectively. And the action set and the discrete energy level can be denoted by $\mathcal{A} = \{i\delta_a,\ i = 0, 1, \ldots, n^a\}$, and $X_k = \{j\delta_e, j = 0, 1, \ldots, k\}$.

The system parameters: the maximal TPR $\psi = n^a\delta_a$, the capacity of the energy buffer θ and the completion death rate α will be given in various value to observation their impacts to the optimal policy and the maximal throughput later.

Then we obtain a infinite horizon discrete time MDP with finite state, and we implement the optimal TPR control algorithm in Algorithm 1, and our numerical results are plotted in Fig. 3 to Fig. 8.

B. OBSERVATIONS

In the following, we present some observations to highlight some characteristics of the EHD.

The optimal TPR control policies and the maximal throughput are plotted in Fig. 3 and Fig. 4 as functions with respect to various background states and the residual energy levels, respectively. From Fig. 3, we can see that at background state 1, 4 and 7, the optimal TPR is zero no matter what the initial energy level is. This observation is obvious since we note that the channel gain rate of state 1, 4 and 7 are all zeros, in this case, the optimal energy control policy is no transmission.

Given the background state, from Fig. 3 and 4, the TPR and the maximal throughput increase with the growth of the initial energy level. (In fact, use the mathematical induction, we can prove easily this result that the TPR and the maximal throughput are increase function in the initial energy level.) This fact illustrates that a more aggressive energy control policy can be taken when the residual energy is greater. Similarly, when energy level is given, the TPR and the maximum expected throughput with higher channel gain rate is larger than that of state with lower gain rate. The observation can be explained intuitively by the fact that, under the same EH rate and initial energy level, the better the channel status is, the greater the throughput is under the same energy consumption rate.

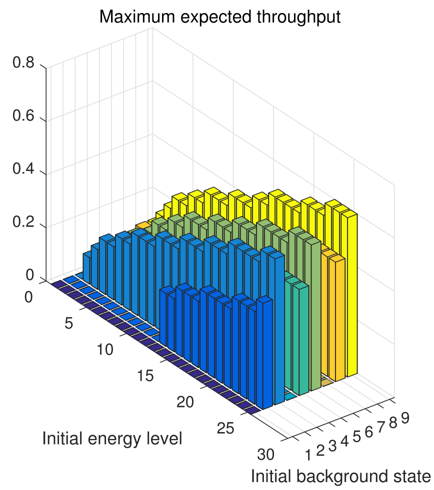

FIGURE 3. Optimal policy vs initial background states and energy levels when $\theta = 0.5$ J/s, $\eta = 0.2$ J/s, $\psi = 0.5$ J/s, $\alpha = 0.3$.

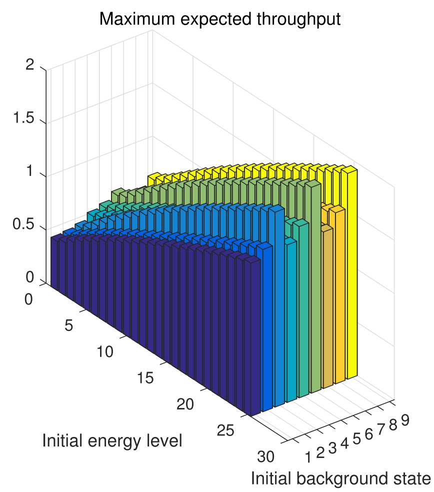

FIGURE 4. Maximum expected throughput vs initial background states and energy levels when $\theta = 0.5$ J/s, $\eta = 0.2$ J/s, $\psi = 0.5$ J/s, $\alpha = 0.3$.

Now, we investigate the impacts of the system parameters θ, ψ, α, and η on the maximum expected throughput, and the impacts on the optimal TPR can be analyzed similarly.

In Figure 5, we report the maximum expected throughput vs the energy buffer capacity θ when the initial background state is given as 5. From this figure, when the capacity θ increases, the maximum expected throughput increases too, however, the degree of the increment becomes weaker when the capacity increases to some degree. This fact is because that when the energy buffer is small, the energy overflow occurrence probability is large, which give rise to the much more energy waste and less expected throughput. When the energy buffer is bigger enough, the energy overflow occurrence is smaller or even becomes zero, in this case, the increase of the energy buffer makes less or even no impact to the maximal expected throughput.

Fig. 6 plots the maximum expected throughput vs various initial energy level when the background state is 5 when the

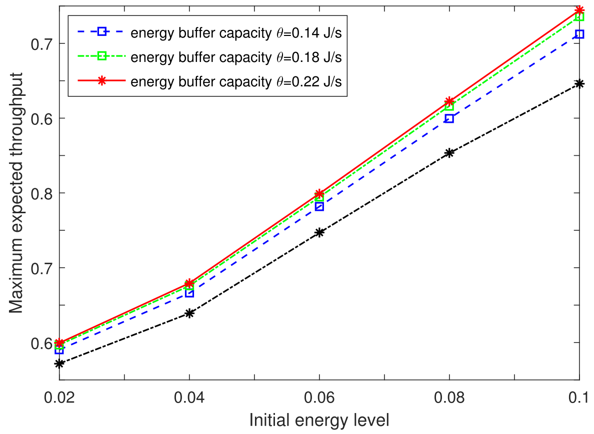

FIGURE 5. **Maximum expected throughput vs initial energy level when $\eta = 0.04$ J/s, $\psi = 0.4$ J/s, $\alpha = 0.3$.**

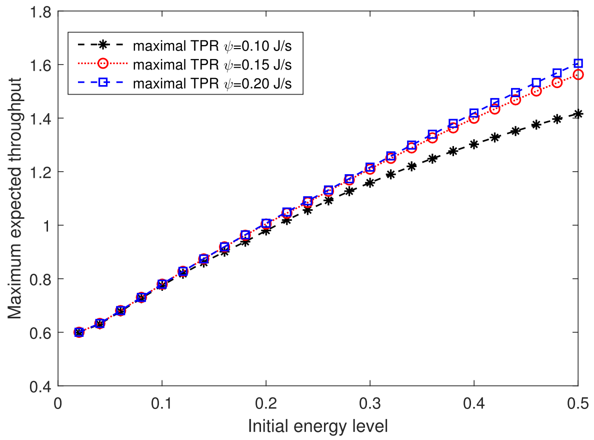

FIGURE 6. **Maximum expected throughput vs initial energy level when the background state is 5, $\theta = 0.5$ J/s, $\eta = 0.1$ J/s, $\alpha = 0.3$.**

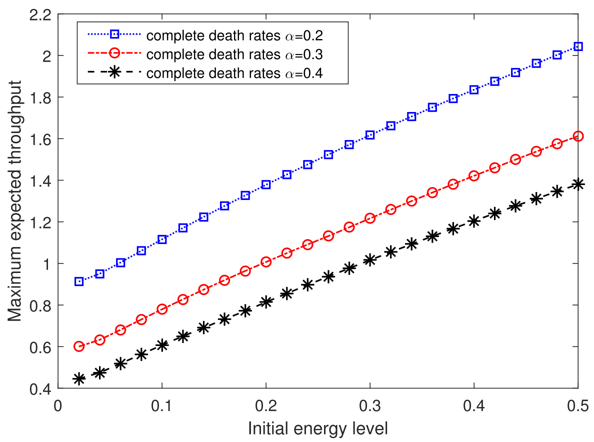

FIGURE 7. **Maximum expected throughput vs initial energy level when the background state is given as 5, $\theta = 0.5$ J/s, $\eta = 0.1$ J/s, $\psi = 1$ J/s.**

maximal TPR ψ is 0.1 J/s, 0.15 J/s and 0.2 J/s. We can see that when the initial energy level is low, the change of the maximal TPR ψ makes a negligible impact to the maximal TPR. However, when the initial energy level is high, the change of the maximal TPR ψ leads to more heavy impact to the maximum expected throughput with a higher energy level. This fact can be explained that when the initial energy is low, its corresponding optimal TPR is low, thus it is not sensitive to the changes of the TPR range. when the energy level becomes larger, the optimal TPR becomes larger, too. Thus, increasing the maximal TPR makes the range of the available TPR for selection becomes larger, which leads to the increase of the maximal throughput, when the global optimum of the TPR is included in this range, the increase of the maximal TPR ψ makes no impact to the throughput.

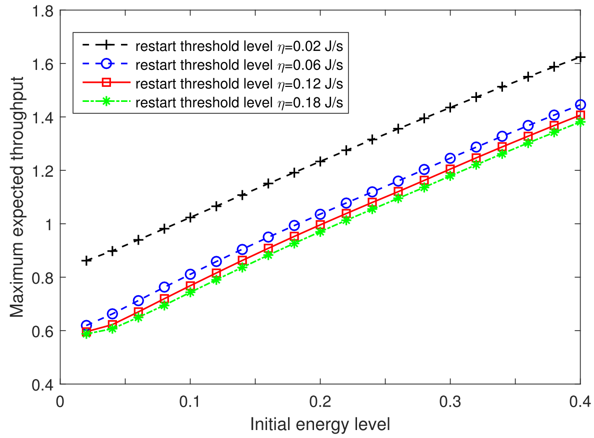

FIGURE 8. **Maximum expected throughput vs initial energy level when background state is 5, $\theta = 0.4$ J/s, $\psi = 0.4$ J/s, $\alpha = 0.3$.**

Fig. 7 plots the impacts of the different complete death rates α on the maximum expected throughput. As expected, the maximum expected throughput decrease with the increase of the complete death rate α, since larger death rate leads to a shorter lifetime of the EHD.

In Fig. 8, the maximum expected throughput is plotted as a function with respect to various restart threshold level η when the initial background state is given as 5. We can see that when increase the restart threshold level, the maximum expected throughput decreases. The reason for this observation is that increasing of the restart threshold level means the increase of the temporal death duration, which leads to the decrease of the working time of the EHD in its lifetime, and then the maximum throughput is decreasing.

VI. CONCLUSION

In this paper, we consider the wireless communication of an EHD with temporal and complete deaths in a wireless fading channel. We develop a continuous-time, hybrid-state stochastic communication system model of the EHD. We propose the TPR policy that maximizes the expected throughput in its lifetime, we optimize the objective by reformulating the control problem into a PDMDP by using the markov property of the system model. The space-space discretization algorithm is proposed, and the numerical examples are given to illustrate how to use our model and the algorithm to obtain the optimal policy and the value numerically. some observations are investigated based on the numerical examples. The model and the method can be extended to some specific communication application scenarios to analyze the transmission strategy, and thus help to provide guidance on designing more practical resource management schemes and determining optimal system parameters. One of the extension of this model and method that can be studied in the future is the optimal energy management to maximize the expected throughput under some additional Quality of Service(QoS) constraints, for example, the delay-constrained, the energy expenditure-constrained and so on.

APPENDIX A DERIVATION OF THE CONDITIONAL PROBABILITY MATRIX $G[\eta]$

We first note that the interval $[0, \kappa_0^\eta]$ is exactly the temporal death duration. During the temporal death, the EHD does not work but harvests energy from the environment until the accumulated energy reaches the level η. We further note that $\pi_t = 0$ for all t in the temporal death duration. That is, the background process B evolves independently in its own way. Therefore, during the temporal death, the process $[B, X]$ reduces to a standard markov fluid queue (MFQ) [28], [29]. For simplicity, we will drop the superscript "b" in the all notations.

First, we divide the state space of the background process Ω into two parts

$$\Omega = \Omega_0 \cup \Omega_+,$$

where $\Omega_0 = \{i \in \Omega | h^i = 0\}$ and $\Omega_+ = \{i \in \Omega | h^i > 0\}$. That is, when the background process enters the state in Ω_0, the EHD can not harvest energy, while the background process enters the state in Ω_+, the node can harvest energy from the environment.

We reorder and partition Q according to $\Omega = \Omega_0 \cup \Omega_+$ as

$$Q = \begin{bmatrix} Q_{00} & Q_{0+} \\ Q_{+0} & Q_{++} \end{bmatrix}.$$

We let $H = diag[h^j, j \in \Omega_+]$ be the EHR matrix.

During the temporal death duration, we define matrix $\mathbf{G}[\eta]$ of the MFQ $[B, X]$, where the $[i, j]$th entry of $\mathbf{G}[\eta]$ is given as

$$g_{ij}[\eta] = P[B_{\kappa_0^\eta} = j, X_{\kappa_0^\eta} = \eta \mid B_0 = i, X_0 = 0].$$

We assume that $\mathbf{G}[\eta]$ is also partitioned according to $\Omega_0 \cup \Omega_+$. Observing that the energy level first reaches the level η only in Ω_+, we have the block form of the matrix $\mathbf{G}[\eta]$ as given by

$$\mathbf{G}[\eta] = \begin{bmatrix} 0 & \mathbf{G}_{0+}[\eta] \\ 0 & \mathbf{G}_{++}[\eta] \end{bmatrix},$$

where 0 represents zero matrix with an appropriate dimension.

We determine the sub-block matrices in $\mathbf{G}[\eta]$. We have the following theorem. A similar proof can be found in [22], [23].

Theorem A.1: For the MFQ $[B, X]$ given above, we have

$$\mathbf{G}_{0+}[\eta] = -Q_{00}^{-1} Q_{0+} \mathbf{G}_{++}[\eta], \tag{18}$$

$$\mathbf{G}_{++}[\eta] = e^{K\eta}, \tag{19}$$

where

$$K = H^{-1}[Q_{++} - Q_{+0} Q_{00}^{-1} Q_{0+}].$$

Proof: For $i \in \Omega_0$ and $j \in \Omega_+$, let $q_i = -q_{ii}$, we have

$$g_{ij}[\eta] = \int_0^{+\infty} q_i e^{-q_i t} \sum_{l \neq i} \frac{q_{il}}{q_i} g_{lj}[\eta] dt$$
$$= q_i^{-1} \sum_{l \neq i} q_{il} g_{lj}[\eta].$$

Rewrite this in matrix form as follows

$$\mathbf{G}_{0+}[\eta] = \Lambda_0^{-1}[Q_{0+}\mathbf{G}_{++}[\eta] + [Q_{00} + \Lambda_0]\mathbf{G}_{0+}[\eta]],$$

where $\Lambda_0 = diag[q_j, \ j \in \Omega_0]$. Thus we have

$$\mathbf{G}_{0+}[\eta] = -Q_{00}^{-1} Q_{0+} \mathbf{G}_{++}[\eta].$$

Then we obtain (18).

To prove (19), considering the first transition of the background process, for $i \in \Omega_+$ and $j \in \Omega_+$, we can write

$$g_{ij}[\eta] = I_{\{i=j\}} e^{-q_i \frac{\eta}{h^i}} + \int_0^{\frac{\eta}{h^i}} q_i e^{-q_i t} \sum_{l \neq i} \frac{q_{il}}{q_i} g_{lj}[\eta - h^i t] dt.$$

By substituting z for $\eta - h^i t$ into the above equation, we get

$$g_{ij}[\eta] = I_{\{i=j\}} e^{-q_i \frac{\eta}{h^i}} + \int_0^{\eta} e^{-q_i \frac{\eta - z}{h^i}} \sum_{l \neq i} q_{il} g_{lj}[\eta] \frac{1}{h^i} dz.$$

Multiplying with $e^{q_i \frac{\eta}{h^i}}$ on two sides, and differentiating with respect to η, the above equation can be written in a matrix form

$$\frac{\partial \mathbf{G}_{++}[\eta]}{\partial \eta} = H^{-1}[Q_{+0}\mathbf{G}_{0+}[\eta] + Q_{++}]\mathbf{G}_{++}[\eta].$$

We then have the differential equation

$$\frac{\partial \mathbf{G}_{++}[\eta]}{\partial \eta} = K\mathbf{G}_{++}[\eta],$$

where we denote

$$K = H^{-1}[[Q_{++}] - Q_{+0} Q_{00}^{-1} Q_{0+}].$$

Noting the condition $\mathbf{G}_{++}[0] = I$, we then have

$$\mathbf{G}_{++}[\eta] = e^{K\eta}.$$

This concludes the proof. □

REFERENCES

[1] S. Chen, P. Sinha, N. B. Shroff, and C. Joo, "A simple asymptotically optimal energy allocation and routing scheme in rechargeable sensor networks," in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 379–387.

[2] 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.

[3] A. Biason, D. D. Testa, and M. Zorzi, "Low-complexity policies for wireless sensor networks with two energy harvesting devices," in *Proc. 13th Annu. Medit. Ad Hoc Netw. Workshop (MED-HOC-NET)*, Jun. 2014, pp. 180–187.

[4] Z. Wang, V. Aggarwal, and X. Wang, "Iterative dynamic water-filling for fading multiple-access channels with energy harvesting," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 3, pp. 382–395, Mar. 2015.

[5] Z. Hadzi-Velkov, I. Nikoloska, G. K. Karagiannidis, and T. Q. Duong, "Wireless networks with energy harvesting and power transfer: Joint power and time allocation," *IEEE Signal Process. Lett.*, vol. 23, no. 1, pp. 50–54, Jan. 2016.

[6] S. Tang and L. Tan, "Reward rate maximization and optimal transmission policy of EH device with temporal death in EH-WSNs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 1157–1167, Feb. 2017.

[7] A. Banerjee, A. Paul, and S. P. Maity, "Joint power allocation and route selection for outage minimization in multihop cognitive radio networks with energy harvesting," *IEEE Trans. Cognit. Commun. Netw.*, vol. 4, no. 1, pp. 82–92, Mar. 2018.

[8] Y. He, X. Cheng, W. Peng, and G. L. Stuber, "A survey of energy harvesting communications: Models and offline optimal policies," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 79–85, Jun. 2015.

[9] K. Singh and S. Moh, "A comparative survey of energy harvesting techniques for wireless sensor networks," *Adv. Sci. Technol. Lett.*, vol. 142, no. 2016, pp. 28–33, 2016.

[10] A. Biason, "Stochastic optimization of energy harvesting wireless communication networks," M.S. thesis, Univ. Padua, Padua, Italy, 2018.

[11] B. M. Michelson, "Event-driven architecture overview," *Patricia Seybold Group*, vol. 2, no. 12, 2006, Art. no. 101571.

[12] 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.

[13] 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.

[14] 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.

[15] J. Ventura and K. Chowdhury, "Markov modeling of energy harvesting body sensor networks," in *Proc. IEEE 22nd Int. Symp. Pers., Indoor Mobile Radio Commun.*, Sep. 2011, pp. 2168–2172.

[16] P. Blasco, D. Gunduz, and M. Dohler, "A learning theoretic approach to energy harvesting communication system optimization," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1872–1882, Apr. 2013.

[17] R. D. Pietro and N. V. Verde, "Introducing epidemic models for data survivability in unattended wireless sensor networks," in *Proc. IEEE Int. Symp. World Wireless, Mobile Multimedia Netw.*, Jun. 2011, pp. 1–6.

[18] M. H. A. Davis, "Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models," *J. Roy. Stat. Soc., Ser. B (Methodol.)*, vol. 46, no. 3, pp. 353–376, Jul. 1984.

[19] M. H. A. Davis, *Markov Models and Optimization*. Boca Raton, FL, USA: CRC Press, 1993.

[20] M. H. A. Davis, "Control of piecewise-deterministic processes via discrete-time dynamic programming," in *Stochastic Differential Systems*. Berlin, Germany: Springer, 1986, pp. 140–150.

[21] S. Tang and X. Wang, "Battery-level-triggered transmit power control for energy harvesting communications," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 8011–8023, Dec. 2020.

[22] N. G. Bean, M. M. O'Reilly, and P. G. Taylor, "Hitting probabilities and hitting times for stochastic fluid flows," *Stochastic Processes Appl.*, vol. 115, no. 9, pp. 1530–1556, Sep. 2005.

[23] L. Tan and S. Tang, "Energy harvesting wireless sensor node with temporal death: Novel models and analyses," *IEEE/ACM Trans. Netw.*, vol. 25, no. 2, pp. 896–909, Apr. 2017.

[24] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Hoboken, NJ, USA: Wiley, 2014.

[25] 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.

[26] N. Bäuerle and U. Rieder, *Markov Decision Processes With Applications to Finance*. Berlin, Germany: Springer, 2011.

[27] O. Hernandez-Lerma, *Adaptwe Markov Control Processes*. New York, NY, USA: Springer-Verlag, 1989.

[28] N. G. Bean and M. M. O'Reilly, "Performance measures of a multi-layer Markovian fluid model," *Ann. Oper. Res.*, vol. 160, no. 1, pp. 99–120, Apr. 2008.

[29] S. Ahn and V. Ramaswami, "Efficient algorithms for transient analysis of stochastic fluid flow models," *J. Appl. Probab.*, vol. 42, no. 2, pp. 531–549, Jun. 2005.

YUN LI received the B.S. degree in electrical engineering and automation from Central South University, China, in 2002, the M.S. degree in communication and information systems from Sun Yat-sen University, China, in 2008, and the Ph.D. degree in information and communications engineering from Tianjin University, China, in 2018. She is currently a Professor with the Guangxi University of Finance and Economics. Her research interest includes underwater wireless sensor networks.

SHENGDA TANG received the Ph.D. degree in computer science from Central China Normal University, Wuhan, Hubei, China, in 2016. He is currently an Associate Professor with the School of Mathematics and Statistics, Guangxi Normal University. He was a Visiting Scholar with Columbia University, New York, USA, from January 2019 to January 2020. His research interests are broadly in the areas of stochastic processes and queueing models for communication networks, performance evaluation, control, optimization, and economic analysis of communication networks.

ZHICHENG TAN received the M.S. degree in electronic science and technology from the Guilin University of Electronic Technology, China, in 2017. He is currently pursuing the Ph.D. degree in communication and information systems with the School of Informatics, Xiamen University, Xiamen, China. His research interest includes underwater wireless sensor networks.

• • •