Energy Management Policies for Harvesting-Based Wireless Sensor Devices with Battery Degradation

Nicolò Michelusi, *Member, IEEE,* Leonardo Badia, *Senior Member, IEEE,* Ruggero Carli, *Member, IEEE,* Luca Corradini, *Member, IEEE,* and Michele Zorzi, *Fellow, IEEE*

***Abstract*—Energy Harvesting Wireless Sensor Devices are increasingly being considered for deployment in sensor networks, due to their demonstrated advantages of prolonged lifetime and autonomous operation. However, irreversible degradation mechanisms jeopardize battery lifetime, calling for intelligent management policies, which minimize the impact of these phenomena while guaranteeing a minimum Quality of Service (QoS). This paper explores a mathematical characterization of these devices, focusing on the interplay between the battery discharge policy and the irreversible degradation of the storage capacity. We propose a stochastic Markov chain framework, suitable for policy optimization, which captures the degradation status of the battery. We present a general result of Markov chains, which exploits the timescale separation between the communication time-slot of the device and the battery degradation process, and enables an efficient optimization. We show that this model fits well the behavior of real batteries for what concerns their storage capacity degradation over time. We demonstrate that a degradation-aware policy significantly improves the lifetime of the sensor compared to "greedy" policies, while guaranteeing the minimum required QoS. Finally, a simple heuristic policy, which never discharges the battery below a given threshold, is shown to achieve near-optimal performance in terms of battery lifetime.**

***Index Terms*—Battery management systems; energy harvesting; lifetime estimation; Markov processes; sensor phenomena and characterization; wireless sensor networks.**

I. Introduction and Motivation

RECENT technological advances and enhancements of consumer electronics have led to the widespread diffusion of networks of miniaturized devices with sensing and communication capabilities, commonly referred to as Wireless Sensor Networks (WSNs) [2].

One key requirement of such networks is a prolonged and unsupervised sensor operation over time, which poses the problem of their energy autonomy. While the use of primary (non-rechargeable) batteries is currently widespread for powering WSN nodes, recent advances in the field of small-scale energy harvesting will enable the sensors to use ambient energy absorbed, for instance, from solar, mechanical, thermal or RF sources [3]. Present technologies require a local energy storage element to filter the intermittent harvested energy, as data sensing and processing, transmission/reception tasks, and higher layer operations (*e.g.*, routing) rely on a continuous and stable energy reserve. The energy harvesting approach, combined with an intelligent use of the local energy storage, is envisioned to greatly prolong the WSN operating life [4], and could in principle lead to perpetual operation.

From an abstract perspective, a rechargeable battery is typically modeled by an energy buffer, which stores the energy supplied by the energy harvesting process, and provides it to the device to perform sensing, processing and data communication tasks, whenever needed. An energy-aware *operation policy* is an algorithm for the management of the energy buffer, for example aimed at avoiding energy overflow or battery depletion, so as to provide a stable operation of the device over time, while guaranteeing a satisfactory Quality of Service (QoS), *e.g.*, throughput [5], delay [6] or network sum rate [7]. In this context, most works in the literature assume *ideal* and *perpetual* battery operation, and neglect any battery degradation issues related to battery usage, *e.g.*, see [8], [9].

In reality, batteries involve more complex mechanisms than just storing and drawing energy on-demand and without side effects. In this context, some works attempted to model realistic battery imperfections and non-idealities, and their impact on the performance of harvesting based devices and networks, *e.g.*, battery leakage [10], imperfect state of charge knowledge [11], non-linearity between energy storage level and power delivered by a battery [12], energy storage losses [13]. References [14], [15] provide an overview of several tractable mathematical models that capture battery discharge characteristics for battery powered devices (no EH), which can be employed to design optimal strategies that extract maximum charge. Reference [16] presents a stochastic model to capture the recovery effect of electrochemical cells, based on which efficient battery management policies can be designed.

The focus and novelty of this paper is on degradation effects, which cause the storage capability of a battery to diminish over time, depending on how the battery is used [17]. This is in contrast to the framework studied in [10], which accommodates a deterministic battery capacity degradation

Manuscript received January 9, 2013; revised July 30, 2013. The editor coordinating the review of this paper and approving it for publication was V. Wong.

This work has been supported in part by the Department of Information Engineering at the University of Padova, by PRAT grant no. CPDA112224, and by the European Commission through the FP7 EU project "Symbiotic Wireless Autonomous Powered system (SWAP)" (G.A. no. 251557, http://www.fp7-swap.eu/). Part of this work has been presented at *IEEE INFOCOM 2013* [1].

N. Michelusi is with the Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, USA (e-mail: michelus@usc.edu).

L. Badia, R. Carli, and M. Zorzi are with the Department of Information Engineering, University of Padova, Italy (e-mail: {badia, carlirug, corradini, zorzi}@dei.unipd.it). L. Badia and M. Zorzi are also with Consorzio Ferrara Ricerche, via Saragat, 1, 44100 Ferrara, Italy.

Digital Object Identifier 10.1109/TCOMM.2013.111113.130022

over time that is not influenced by how the battery is used. Degradation phenomena due to deep discharge are particularly strong for Lithium-Ion (Li-Ion) batteries, which represent the reference case of rechargeable batteries in consumer electronics. Importantly, the deeper the discharge of the battery, the *faster* the degradation. Thus, for example, an appropriate approach to enhancing the battery lifetime would be to have very frequent and shallow discharge periods, compatibly with the operating constraints of the network and the intermittent nature of the ambient energy supply. In contrast, fully exploiting the battery charge with deep discharge cycles should be avoided as it is detrimental to battery lifetime.

In this paper, we develop a Markov model which explicitly characterizes the degradation status of the battery and is suitable for policy optimization. Secondly, we show that the model accurately fits the behavior of real batteries for what concerns their storage capacity degradation over time. A discussion on the non-trivial consequences of this model on the battery operation policies is also provided by formulating an optimization problem which accounts for battery lifetime explicitly. We believe that this evaluation represents an important step towards the realistic characterization of rechargeable batteries and, by extension, of WSNs and their management strategies.

A strong suit of our approach is to join two different perspectives, namely, those of microelectronics and network engineering. Microelectronic characterizations of batteries often give a very detailed parametric description but fail to provide a behavioral analysis over time and in a broader context. Conversely, network models may be entirely flawed if they do not properly account for a correct physical characterization. This paper aims at bridging the gap between these two approaches.

The rest of this paper is organized as follows. In Sec. II, we survey the existing models of battery degradation available in the literature. In Sec. III, we present the general framework and define the optimization problem, which is further developed in Sec. IV. In Sec. V, we extrapolate the battery degradation probabilities from experimental data and models available in the literature. In Sec. VI, we provide numerical results. In Sec. VII, we discuss some possible extensions of the model presented in this paper. Sec. VIII concludes the paper.

II. Background

The block diagram of an energy harvesting sensor node is sketched in Fig. 1. From an energy flow standpoint, the system *load* is the hardware that needs to be powered: microcontroller, Rx/Tx transceiver and sensor fall in this category. The *power processing unit* manages the ambient energy source and the on-board battery to provide regulated energy to the load.

In such a system, the ambient energy source often provides most of the energy within certain periods of time, during which the on-board battery is recharged. In the remaining periods, little or no energy is available from the source, and the on-board battery is partially or totally discharged, depending on the load demand. The charge/discharge process of the battery is called *cycling*, and the percentage amount D of charge

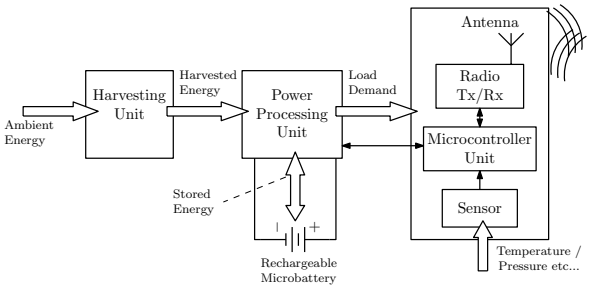


Fig. 1. Block diagram of a harvesting-based sensor node

withdrawn from the battery during discharge, with respect to its nominal capacity, is termed *Depth of Discharge* (DoD). In a photovoltaic scavenger, for instance, battery cycling is determined on a daily basis by the availability of solar energy. Other energy sources, such as RF, thermal or mechanical, may present different trends. In general, the target application and deployment scenario of the WSN play an important role in determining the cycling period and its degree of randomness.

Denoting with C_0 the nominal battery capacity in milliampere-hours (mAh) and with $Q(N_{\mathrm{cyc}})$ the total charge delivered by the battery after N_{cyc} cycles at DoD D, one might expect that

$$Q(N_{\mathrm{cyc}}) = N_{\mathrm{cyc}} \cdot C_0 \cdot D. \tag{1}$$

Two fundamental facts, however, complicate the deceptively simple scenario implied by (1). First, a rechargeable battery has a finite *cycle life*, *i.e.*, it cannot cycle indefinitely due to irreversible degradation mechanisms, which ultimately reduce C_0 to unacceptable levels [17]. Manufacturers typically define the battery cycle life N_{cyc} as the number of cycles a battery delivers at DoD $D = 1$ before C_0 drops below a given threshold, *e.g.*, 80% or 50% of the initial value [18]. Secondly, the foregoing degradation process is strongly dependent on *how* the battery is cycled. More precisely, shallow DoDs result in a slower degradation of C_0 and ultimately in increased cycle life [17], [19]–[21]. For instance, a microbattery rated with $N_{\mathrm{cyc}} = 100$ cycles at 100% DoD may last up to $N_{\mathrm{cyc}} = 1000$ cycles at 20% DoD, indicating that roughly twice the amount of energy is extracted from the battery in the latter case [18].

A simple empirical model for the N_{cyc} vs D dependence is

$$N_{\mathrm{cyc}}(D) = N_{\mathrm{cyc},0} \cdot e^{\alpha(1-D)}, \tag{2}$$

where $N_{\mathrm{cyc},0}$ represents the cycle life at 100% DoD, and α is a characteristic constant of the battery. Exponential-based models like (2) have been proven to be a good fit for data from a rather wide range of battery chemistries and sizes [19]–[22]. Equation (2) may therefore be taken as representative also for microbatteries targeted for low-power equipment. Note, however, that different $N_{\mathrm{cyc}}(D)$ relationships could be employed depending on the available experimental data and the desired accuracy.

Acknowledging the degradation of the battery capacity and the dependence of N_{cyc} on D opens up intriguing options for more advanced energy-aware policies, which represent the main focus of this work. In the next section, the foregoing qualitative discussion is formulated within the framework of a stochastic model which captures the essential features of the

problem, such as source pseudo-periodicity, battery cycling and cycle life vs DoD dependence found in commercial microbatteries.

III. System Model

We consider a slotted-time system, where slot k is the time interval $[k\Delta t, k\Delta t+\Delta t)$, $k \in \mathbb{Z}^+$, and Δt is the slot duration. The battery is modeled by a buffer with nominal capacity C_0, and is uniformly quantized to a number of charge levels, using a quantization step (charge quantum) $\Delta c \ll C_0$. The maximum number of quanta that can be stored at the nominal capacity is $q_{\max} = \left\lfloor \frac{C_0}{\Delta c} \right\rfloor$ and the set of possible charge levels is denoted by $\mathcal{Q} = \{0, 1, \ldots, q_{\max}\}$.

Due to the aforementioned battery degradation mechanisms, the nominal battery capacity $q_{\max}$ is not always entirely available, but rather decreases over time. Let $Q_{\max}(k)$ be the battery capacity at time k, with $Q_{\max}(k) \leq Q_{\max}(k-1)$ and $Q_{\max}(0) = q_{\max}$. Denote the (quantized) charge level of the battery at time k as Q_k. The evolution of Q_k is given by

$$Q_{k+1} = \min\left\{[Q_k - A_k]^+ + B_k, Q_{\max}(k+1)\right\}, \qquad (3)$$

where $[x]^+ = \max\{x, 0\}$ and:

- $\{B_k\}$ is the harvesting process, taking values in $\mathcal{B} \triangleq \{0, 1, \ldots, B_{\max}\}$, which models the randomness in the energy harvesting source, *e.g.*, due to an intermittent energy supply. We define an underlying *scenario process* $\{S_k\}$ [23], and we model it as an irreducible stationary Markov chain with transition probability $p_S(s_{k+1}|s_k) \triangleq \Pr(S_{k+1} = s_{k+1}|S_k = s_k)$ and steady state distribution $\pi_S(s)$, taking values in a finite state space $\mathcal{S}$. Given $S_k \in \mathcal{S}$, the energy harvest B_k is drawn from $\mathcal{B}$ according to the distribution $p_B(b_k|s_k) \triangleq \Pr(B_k = b_k|S_k = s_k)$. Then, we denote the average harvesting rate as

$$\bar{b} \triangleq \sum_{s\in\mathcal{S}} \pi_S(s) \sum_{b\in\mathcal{B}} b p_B(b|s). \qquad (4)$$

We assume that a new energy quantum harvested in slot k can only be used in a later slot.

- $\{A_k\}$ is the action process, which is governed by the Energy Harvesting Device (EHD) controller, as detailed in Sec. III-A, and takes values in $\mathcal{A} \triangleq \{0\} \cup \{A_{\min}, A_{\min}+1, \ldots, A_{\max}\}$. $A_{\min} \in \mathbb{N}$ and $A_{\max} \in \mathbb{N}$ represent the minimum and maximum load requirements, respectively. Action $A_k = 0$ accounts for the possibility to remain idle in a given time-slot, due to either a controller's decision or energy outage. A_k may include the energy cost of performing data acquisition and transmission.

Remark 1. Note that, if $A_k > Q_k$, then more energy is requested by the sensor node to run the task than the amount currently available in the battery. In this case, the device starts drawing energy from the battery to run the task, but energy depletion occurs before its completion. The new battery charge state becomes $Q_{k+1} = \min\{B_k, Q_{\max}(k+1)\}$, *i.e.*, it equals the amount of charge harvested in the current slot. In the following, perfect knowledge of Q_k is assumed at the EHD controller, so that the latter can (and will always) choose $A_k \leq Q_k$, and no energy is wasted. On the other hand, if $[Q_k - A_k]^+ + B_k > Q_{\max}(k+1)$, part of the harvested energy is lost due to energy overflow.

Remark 2. Note that the harvesting model considered in this paper is a special instance of the *Generalized Markov model* presented in [23]. Therein, the scenario process is modeled as a first-order Markov chain with two states, whereas the harvest B_k statistically depends on $(B_{k-1}, B_{k-2}, \ldots, B_{k-L})$ and on the scenario S_k, for some order $L \geq 0$. In particular, it is shown that, by quantizing the harvest B_k with 20 states, the order $L = 0$ models well a piezoelectric energy source, whereas the order $L = 1$ models well a solar energy source. In this paper, we employ $L = 0$ for simplicity. However, the analysis can be extended to the case $L = 1$, as detailed in Sec. VII.

Remark 3. The interaction between the EHD and the battery is here modeled in terms of normalized *charge* quanta. In reality, both the harvesting and the action, or load, processes are *energy*-driven rather than *charge*-driven. Exchanged charge and energy are proportional only as long as the battery voltage is constant throughout the device operating life, which is assumed to hold in this paper. Modeling battery voltage dynamics is out of the scope of this paper and can be considered as a future refinement.

We model the battery degradation process, which causes the battery capacity $Q_{\max}(k)$ to diminish irreversibly over time, as follows. We define the *battery health state*, H_k, taking values in $\mathcal{H} \equiv \{0, 1, \ldots, H_{\max}\}$, where $H_{\max} > 0$. For a given H_k, the battery capacity at time k, *i.e.*, the total amount of charge which can be delivered by a fully charged battery over a discharge phase, is given by

$$Q_{\max}(k) = \left\lfloor \frac{H_k}{H_{\max}} q_{\max} \right\rfloor, \qquad (5)$$

and the set of available charge levels is denoted by $\mathcal{Q}(H_k) = \left\{0, 1, \ldots, \left\lfloor \frac{H_k}{H_{\max}} q_{\max} \right\rfloor\right\}$. We assume that $\{\text{History up to time } k-1\} \to (H_k, Q_k) \to H_{k+1}$ forms a Markov chain, *i.e.*, H_{k+1} is conditionally independent of the history up to time $k-1$, given (H_k, Q_k). We denote the transition probability from health state $H_k = h$ to health state $H_{k+1} = h-1$ as

$$p_H(h; q) \triangleq \Pr(H_{k+1} = h-1 | H_k = h, Q_k = q). \qquad (6)$$

Moreover, $\Pr(H_{k+1} = \tilde{h}|H_k = h, Q_k = q) = 0$ if $\tilde{h} \notin \{h-1, h\}, \forall q \in \mathcal{Q}(h)$, so that no transition is possible between two non-consecutive health states, or to a higher health state. As a consequence, the probability of remaining in health state h is $1 - p_H(h; q)$. We further make the following assumptions on $p_H(h; q)$:

Assumption 1. a) $p_H(h; q) > 0,\ \forall h \in \mathcal{H}, q \in \mathcal{Q}(h)$,
b) $p_H(h; q) \ll 1,\ \forall h \in \mathcal{H}, q \in \mathcal{Q}(h)$,
c) $p_H(h_1; q_1) \geq p_H(h_2; q_2),\ \forall h_2 \geq h_1,\ q_2 \geq q_1$.

Ass. 1.a) implies that the battery health state will eventually reach state $H_k = 0$, so that the lifetime, defined in Def. 1 in Sec. III-A, is finite; Ass. 1.b) expresses the fact that aging processes taking place in the battery operate over time scales that are much longer than the cycling period and the communication time-slot of the EHD; Ass. 1.c) means that the more discharged and degraded the battery, the faster the battery degradation process [17].

Remark 4 (Choice of the number of health states $H_{\max}$). Note that, in practice, the battery capacity degrades continuously over the set $[0, q_{\max}]$. The model (5) approximates such behavior via quantization of the battery capacity to discrete values, which brings the following trade-off: if the number of health states $H_{\max}$ is small, then the model is less likely to closely replicate a realistic battery degradation profile. On the other hand, if $H_{\max}$ is large, then Ass. 1.b) becomes weaker, so that $H_{\max}$ should not be too large either. In fact, using a fine-grained quantization of the health states, the optimal policy, defined via (23) for each health state, does not vary significantly over subsequent health states, leading to over-parameterization.

Remark 5. Unlike the empirical degradation model (2), where the battery lifetime is affected by the number of cycles at a given DoD, our proposed model does not characterize degradation via the cycle count. This is because one cycle cannot be defined under the stochastic setting considered in this paper. On the other hand, such choice of the health model is suitable for policy optimization and analysis via the well developed theory of Markov Decision Processes (MDPs) [24], and captures a salient feature of battery based systems, *i.e.*, the dependence of the battery degradation on the discharge depth of the latter via the transition probability $p_H(h; q)$.

The EHD state at time k is $\boldsymbol{Z}_k = (Q_k, H_k, S_{k-1})$, taking values in the state space $\mathcal{Q} \times \mathcal{H} \times \mathcal{S}$. Note that the scenario S_k is unknown at time k, as reflected by state $\boldsymbol{Z}_k$, since B_k has not yet been observed. In practice, $\boldsymbol{Z}_k$ should be inferred and estimated from measurements of the battery state of charge, capacity, and input energy flows. A cost-effective technique for accurate estimation of the state-of-charge Q_k is presented in [25]. The health state H_k can be determined by estimating the current battery capacity $Q_{\max}(k)$ [26]. The posterior distribution of the scenario S_{k-1} can be inferred from the observed harvesting sequence $\{B_0, \dots, B_{k-1}\}$, and S_{k-1} can be estimated via MAP, $\hat{S}_{k-1} = \arg\max_s \mathbb{P}(S_{k-1} = s|B_0, \dots, B_{k-1})$. In this work, for simplicity, we assume that $\boldsymbol{Z}_k$ is perfectly known to the EHD controller. However, the analysis can be extended to the case where $\boldsymbol{Z}_k$ is unknown, by using the framework of POMDPs [27]. This extension is left for future work.

A. Policy definition and Optimization problem

Given $\boldsymbol{Z}_k = (Q_k, H_k, S_{k-1})$, the EHD controller determines $A_k \in \mathcal{A}$ at time k according to a given *policy* μ_{H_k}. Formally, μ_{H_k} is a probability measure on the action space $\mathcal{A}$, parameterized by the state (Q_k, S_{k-1}), *i.e.*, $\mu_{H_k}(a; Q_k, S_{k-1})$ is the probability of requesting a charge quanta from the battery, when operating in state $\boldsymbol{Z}_k$.[1] Under any policy μ, the state process $\{\boldsymbol{Z}_k\}$ is a Markov chain, so that the whole decision problem is a MDP [24].

The instantaneous reward accrued in time-slot k, in state $\boldsymbol{Z}_k = (Q_k, H_k, S_{k-1})$ under action A_k, is defined as

$$g(A_k, Q_k) = \begin{cases} 0, & A_k > Q_k, \\ g^*(A_k), & A_k \le Q_k, \end{cases} \tag{7}$$

where $g^*(A_k)$ is a non-decreasing function of A_k with $g^*(0) = 0$, so that a larger reward is accrued by using a larger amount of energy. As discussed in Remark 1, when the amount of charge requested by the controller exceeds that available in the battery (case $A_k > Q_k$), the task cannot be successfully completed, and the battery is depleted while no reward is earned.

We define the hitting times of the health states as

$$K_h = \min\{k \ge 0 : H_k = h\},\ h \in \mathcal{H}. \tag{8}$$

K_h is a random variable, which depends on the realization of $\{(B_k, A_k, H_k)\}$. Given an initial state $\boldsymbol{Z}_0 = (Q_0, H_{\max}, S_{-1})$ and a policy μ, we define the *total average reward* $G_\mu^{\text{tot}}(h, \boldsymbol{Z}_0)$, the *battery lifetime* $T_\mu(h, \boldsymbol{Z}_0)$ and the *average reward per time-slot* $G_\mu(h, \boldsymbol{Z}_0)$ of health state h as

$$G_\mu^{\text{tot}}(h, \boldsymbol{Z}_0) = \mathbb{E}\left[\sum_{k=K_h}^{K_{h-1}-1} g(A_k, Q_k) \,\middle|\, \boldsymbol{Z}_0\right], \tag{9}$$

$$T_\mu(h, \boldsymbol{Z}_0) = \mathbb{E}\left[K_{h-1} - K_h \mid \boldsymbol{Z}_0\right], \tag{10}$$

$$G_\mu(h, \boldsymbol{Z}_0) = \frac{G_\mu^{\text{tot}}(h, \boldsymbol{Z}_0)}{T_\mu(h, \boldsymbol{Z}_0)}, \tag{11}$$

where the expectation is taken with respect to $\{(B_k, S_k, H_k, A_k)\}$ and A_k is drawn according to μ. In particular, $G_\mu^{\text{tot}}(h, \boldsymbol{Z}_0)$ is the expected cumulative reward earned over health state h; $T_\mu(h, \boldsymbol{Z}_0)$ is the expected number of time-slots spent in health state h; and $G_\mu(h, \boldsymbol{Z}_0)$ represents the expected reward per time-slot accrued in health state h.

Remark 6. The choice of $g^*(A_k)$ depends on the specific application considered. A widely used model in the literature is to assume that $g^*(A_k)$ is a concave increasing function of A_k, *e.g.*, see [28]. In particular, if $g^*(A_k) = \log_2(1 + \sigma A_k/\bar{b})$, then $g^*(A_k)$ is the Shannon capacity of the static Gaussian channel, where σ is an SNR scaling parameter [29], and $G_\mu(h, \boldsymbol{Z}_0)$ is the expected throughput in health state h. If $g^*(A_k) = \chi(A_k \ge \bar{A})$, for some $\bar{A}$, where $\chi(\cdot)$ is the indicator function, then $G_\mu(h, \boldsymbol{Z}_0)$ is the probability that the load $\bar{A}$ is successfully supported in health state h, and may be used to model real time traffic where the QoS needs to be satisfied (probabilistically) in each slot.

With these definitions at hand, let $\mathcal{G}^*$ be a minimum QoS requirement, which is met in health state h if $G_\mu(h, \boldsymbol{Z}_0) \ge \mathcal{G}^*$. We give the following definition.

Definition 1. (Battery Lifetime) If $G_\mu(H_{\max}, \boldsymbol{Z}_0) \ge \mathcal{G}^*$, the battery lifetime $T_\mu(\mathcal{G}^*, \boldsymbol{Z}_0)$ under policy μ is defined as

$$T_\mu(\mathcal{G}^*, \boldsymbol{Z}_0) = \sum_{h \ge h_\mu^*} T_\mu(h, \boldsymbol{Z}_0), \tag{12}$$

$$\text{where } h_\mu^* = \max\{h : G_\mu(h, \boldsymbol{Z}_0) < \mathcal{G}^*\} + 1 \tag{13}$$

is the index of the lowest health state in which the QoS is met. If $G_\mu(H_{\max}, \boldsymbol{Z}_0) < \mathcal{G}^*$, then $T_\mu(\mathcal{G}^*, \boldsymbol{Z}_0) = 0$.

Note that the lifetime is defined for a specific policy μ and

[1] For the sake of maximizing an average long-term reward function of the state and action processes, it is sufficient to consider only state-dependent stationary policies [24].

QoS requirement $\mathcal{G}^*$. The condition $G_\mu(H_{\max}, \boldsymbol{Z}_0) \geq \mathcal{G}^*$ implies that policy μ yields a satisfactory reward in the healthiest state $H_{\max}$, so that the lifetime under μ is non-zero; otherwise, the lifetime is zero as there is no satisfactory reward even in the healthiest state. The lifetime is defined such that the QoS requirement $\mathcal{G}^*$ is guaranteed at each health state $h \geq h^*_\mu$, *i.e.*, $G_\mu(h, \boldsymbol{Z}_0) \geq \mathcal{G}^*$. In particular, the QoS constraint inherently assumes that the battery degradation processes taking place in the battery operate over time scales which are much longer than the communication time-slot (Ass. 1.b)), so that the system approaches a steady state operation in each health state. For the lower health state $h^*_\mu - 1$, we have $G_\mu(h^*_\mu - 1, \boldsymbol{Z}_0) < \mathcal{G}^*$, *i.e.*, the EHD can no longer sustain the required QoS requirement, and battery failure is declared. Note that a QoS requirement on each health state $h \geq h^*_\mu$ is stricter than an *average* QoS requirement over the entire lifetime, defined as $\sum_{h \geq h^*_\mu} G^{\text{tot}}_\mu(h, \boldsymbol{Z}_0) / \sum_{h \geq h^*_\mu} T_\mu(h, \boldsymbol{Z}_0)$. The latter may induce policies that exhibit wide performance variability across the health states, as made clear in the following example.

Example 1. Consider a system with slot duration 1s, QoS constraint $\mathcal{G}^* = 1.5$ and $H_{\max} = 2$, and a policy μ such that

$$G_\mu(h, \boldsymbol{Z}_0) = h, \; T_\mu(h, \boldsymbol{Z}_0) = 15 \times 10^6 \text{ slots} \simeq 6 \text{ months},$$

$\forall h \in \{0, 1, 2\}$. Then, according to Def. 1, the lifetime is $T_\mu(\mathcal{G}^*, \boldsymbol{Z}_0) = 15 \times 10^6$, hence the QoS $\mathcal{G}^*$ can be supported only at health state 2 for 6 months. However, an *average* QoS of

$$\frac{G^{\text{tot}}_\mu(2, \boldsymbol{Z}_0) + G^{\text{tot}}_\mu(1, \boldsymbol{Z}_0)}{T_\mu(2, \boldsymbol{Z}_0) + T_\mu(1, \boldsymbol{Z}_0)} = 1.5 = \mathcal{G}^* \tag{14}$$

can be supported over a time-interval of duration $30 \times 10^6 \simeq 1$ year, which is twice as long as $T_\mu(\mathcal{G}^*, \boldsymbol{Z}_0)$, despite the fact that a poor performance is attained in health state 1. ■

The optimization problem at hand is to determine the optimal μ^* such that the battery lifetime is maximized, under a given constraint on the minimum QoS $\mathcal{G}^*$, *i.e.*,

$$\mu^* = \arg\max_\mu T_\mu(\mathcal{G}^*, \boldsymbol{Z}_0) = \arg\max_\mu \sum_{h \geq h^*_\mu} T_\mu(h, \boldsymbol{Z}_0), \tag{15}$$

where h^*_μ is given in (13). The solution to (15) is carried out in the next section.

IV. Optimization

In this section, we develop problem (15), showing that it can be recast as an independent Linear Program (LP) on each health state, under Ass. 1.b) on $p_H(h; q)$. The solution to the optimization problem relies on the timescale separation between the communication time-slot of the EHD and the battery degradation process, so that the EHD achieves a steady state operation in each health state. In this light, we give the following definition.

Definition 2. (Steady State distribution of the non-absorbed chain) Assume that the EHD operates indefinitely at health state $h \in \mathcal{H}$ without being absorbed by the lower health state, *i.e.*, $p_H(h; q) = 0, \forall q \in \mathcal{Q}(h)$. Denote the steady state distribution of $(q, s) \in \mathcal{Q}(h) \times \mathcal{S}$ in health state h under policy μ_h as[2]

$$\pi^h_{\mu_h}(q, s) = \lim_{K \to \infty} \frac{1}{K} \sum_{k=0}^{K-1} P^{(k)}(q, s | \boldsymbol{Z}_0), \tag{16}$$

where $\boldsymbol{Z}_0 = (Q_0, h, S_{-1})$ is the initial state and

$$P^{(k)}(q, s | \boldsymbol{Z}_0) = \Pr\left(Q_k = q, S_{k-1} = s | \boldsymbol{Z}_0, p_H(h; \cdot) = 0\right).$$

We define the following quantities.

Definition 3. (Approximate reward per stage and lifetime of health state h)

$$\hat{G}_{\mu_h}(h) = \sum_{(q,s) \in \mathcal{Q}(h) \times \mathcal{S}} \pi^h_{\mu_h}(q, s) \mathbb{E}_{\mu_h(\cdot; q, s)}\left[g(A, q)\right], \tag{17}$$

$$\hat{T}_{\mu_h}(h) = \left(\sum_{(q,s) \in \mathcal{Q}(h) \times \mathcal{S}} \pi^h_{\mu_h}(q, s) p_H(h; q) \right)^{-1}, \tag{18}$$

where $\mathbb{E}_{\mu_h(\cdot; q, s)}\left[g(A, q)\right] = \sum_{a \in \mathcal{A}} \mu_h(a; q, s) g(a, q)$ is the expected reward in state (q, s).

Remark 7. Note that $\pi^h_{\mu_h}$ in (16) is computed under the assumption that the EHD operates indefinitely in health state h, *i.e.*, $p_H(h; q) = 0, \; \forall q$, whereas the term $p_H(h; q)$ in (18) is the actual degradation probability. $\hat{G}_{\mu_h}(h)$ can be interpreted as the average long-term reward per time-slot in health state h, whereas $\hat{T}_{\mu_h}(h)^{-1}$ can be interpreted as the average long-term probability of making a transition to the lower health state $h - 1$. Such observations are formalized in the following lemma. Its proof is provided in the appendix as a general result of Markov chains.

Lemma 1. Let $p^*_H(h) = \max_q p_H(h; q)$. For $p^*_H(h) \to 0$,

$$G_\mu(h, \boldsymbol{Z}_0) = \hat{G}_{\mu_h}(h) + \mathcal{O}(p^*_H(h)), \tag{19}$$

$$T_\mu(h, \boldsymbol{Z}_0) = \hat{T}_{\mu_h}(h) + \mathcal{O}(1), \tag{20}$$

where $f(x) = \mathcal{O}(v(x))$ for $x \to 0$ denotes a quantity such that $\limsup_{x \to 0} \left| \frac{f(x)}{v(x)} \right| < +\infty$.

From Lemma 1, when $\max_q p_H(h; q) \ll 1$, the duration of health state h, $T_\mu(h, \boldsymbol{Z}_0)$, can be approximated by $\hat{T}_{\mu_h}(h)$, up to a bounded additive factor. Since $T_\mu(h, \boldsymbol{Z}_0) \to +\infty$ for $\max_q p_H(h; q) \to 0$ (in fact, the smaller $\max_q p_H(h; q)$, the less likely the health process to be absorbed by the lower health state $h - 1$, hence the longer the amount of time spent in health state h), (20) is a good approximation. On the other hand, the average reward per time-slot in health state h, $G_\mu(h, \boldsymbol{Z}_0)$, can be approximated by $\hat{G}_{\mu_h}(h)$ up to an additive factor, which decays to zero at least as quickly as $\max_q p_H(h; q)$. Both approximations are independent of the initial state $\boldsymbol{Z}_0$, and solely depend on the steady state distribution (16) induced by policy μ_h in health state h, which is approached in each health state.

Since $\max_q p_H(h; q) \ll 1$ by Ass. 1.b), we use Lemma 1

[2]We assume that μ_h induces a Markov chain with a single closed communicating class, so that $\pi^h_{\mu_h}(q, s)$ exists and is independent of $\boldsymbol{Z}_0$ [30].

and replace (19-20) in (12), yielding

$$T_\mu(\mathcal{G}^*, \boldsymbol{Z}_0) \simeq \sum_{h \geq h^*_\mu} \hat{T}_{\mu_h}(h),$$
$$\text{where } h^*_\mu = \max\left\{h : \hat{G}_{\mu_h}(h) < \mathcal{G}^*\right\} + 1. \quad (21)$$

Finally, substituting (21) in (15), we obtain the approximation

$$\mu^* = \arg\max_{\mu} \sum_{h \geq h^*_\mu} \hat{T}_{\mu_h}(h). \quad (22)$$

Note that $\hat{T}_{\mu_h}(h)$ and $\hat{G}_{\mu_h}(h)$ are independent of the policy $\mu_{\tilde{h}}$ for $\tilde{h} \neq h$. Therefore, (22) can be solved independently for each health state h, yielding the following algorithm.

Algorithm 1. 1) `INIT`: set $h = H_{\max}$, `REP=true`
2) `WHILE REP=true AND h>0 SOLVE`

$$\mu^*_h = \arg\min_{\mu_h} \sum_{(q,s)\in\mathcal{Q}(h)\times\mathcal{S}} \pi^h_{\mu_h}(q,s) p_H(h;q) \quad (23)$$
$$\text{s.t.} \sum_{(q,s)\in\mathcal{Q}(h)\times\mathcal{S}} \pi^h_{\mu_h}(q,s)\left(\mathbb{E}_{\mu_h(\cdot;q,s)}\left[g(A,q)\right] - \mathcal{G}^*\right) \geq 0.$$

If the problem is infeasible, set `REP=false`, $h^*_{\mu^*} = h+1$. If it is feasible and $h = 1$, set `REP=false`, $h^*_{\mu^*} = 1$. Otherwise, update $h := h - 1$. `END WHILE`
3) `RETURN` the optimal policy $\mu^* = (\mu^*_h)_{h \geq h^*_{\mu^*}}$, with lifetime $T_{\mu^*}(\mathcal{G}^*, \boldsymbol{Z}_0) \simeq \sum_{h \geq h^*_{\mu^*}} \hat{T}_{\mu^*_h}(h)$.

Remark 8. Step 2) is equivalent to

$$\mu^*_h = \arg\max_{\mu_h} \hat{T}_{\mu_h}(h), \quad \text{s.t. } \hat{G}_{\mu_h}(h) \geq \mathcal{G}^*, \quad (24)$$

and is obtained by substituting the expressions of $\hat{T}_{\mu_h}(h)$ and $\hat{G}_{\mu_h}(h)$ (see Def. 3) in (24). It can be solved numerically via standard stochastic optimization tools, such as LP [24]. Thus, the optimal policy μ^*_h maximizes the lifetime of health state h (equivalently, it minimizes the long-term probability of battery degradation to the lower health state $h-1$) with a constraint on the minimum average QoS in health state h. Step 2) also determines $h^*_{\mu^*}$ in (13), for the optimal policy μ^*. Finally, in step 3) the optimal policy is found by concatenating the sub-policies μ^*_h for $h \geq h^*_{\mu^*}$, and the corresponding lifetime defined in Def. 1 is computed using (20) and (12). The main advantage of this approach over a standard approach which solves the original optimization problem (15) jointly is that (15) is decomposed into a sequence of independent sub-problems (23) for each health state h, thus reducing the overall computational burden.

V. Extrapolation of the Degradation Probabilities from Experimental Data

The battery degradation probabilities can be evaluated from manufacturer-provided data [18] by employing the empirical model (2). These probabilities should be denoted as $p_H(h;q)$, depending on the health h and the charge q, as in (6). The dependence of $p_H(h;q)$ on h is quite difficult to capture; however, in our numerical evaluations we found that its effect is generally very mild. Even by neglecting it entirely, one can still obtain a very good match with manufacturer data. Therefore, we drop any dependence on h and we denote the degradation probability as $p_H(q)$, *i.e.*, just depending on q.

In Sec. V-A, (2) is used to simulate an experiment where the battery is cyclically discharged and recharged at a given DoD until its capacity degrades to a fraction of the nominal capacity. First, the number of cycles as a function of the DoD and of the battery degradation rate function is derived. Then, the battery degradation rate function is found by matching the theoretical curve for the number of cycles to manufacturer data and the exponential model (2). In Sec. V-B, the $p_H(q)$'s are found by matching the *deterministic* degradation times derived in Sec. V-A with the *average* degradation times in the proposed stochastic, discrete time model.

A. Deterministic Degradation Model

We employ model (2) for the relationship between number of cycles and DoD, where the constants $N_{\text{cyc},0}$ and $\alpha > 0$ depend on the specific battery model employed. In particular, $N_{\text{cyc}}(D)$ is counted until the battery capacity degrades to a fraction $x \in (0,1)$ of the initial capacity (*e.g.*, $x \in \{0.5, 0.8\}$), so that, in general, $N_{\text{cyc},0}$ and α also depend on x.

Herein, we assume that the degradation process is a function of the instantaneous state of charge of the battery only, as discussed in the introduction to this section, and is described by the *rate of capacity degradation function* $\rho(q\Delta c/C_0)$ (in mAh/s) at the charge level $q\Delta c \in [0, C_0]$, where C_0 is the nominal capacity, Δc is the charge quantum and q is the charge level normalized to the quantum Δc. Then, if the battery operates at charge level $q\Delta c$ for δ seconds, its capacity degrades by $\delta\rho(q\Delta c/C_0)$ mAh. Moreover, we assume that, for proper coefficients $\theta > 0$, $\zeta > 0$,

$$\rho(q\Delta c/C_0) = \zeta e^{\theta(1-q\Delta c/C_0)}. \quad (25)$$

In the following analysis, and by simulation in Sec. VI, we show that this choice fits well the exponential model (2) for typical values of D (*e.g.*, $D \in [0.2, 1]$). Let $C_n,\ n \geq 0$ be the battery capacity at the beginning of the nth discharge/recharge cycle. In the nth cycle, the battery discharges from C_n to $C_n - C_0 D$ (with DoD D), and it then recharges from $C_n - C_0 D$ to C_{n+1}. Note that $C_{n+1} \leq C_n$, *i.e.*, the capacity at the end of the nth cycle cannot be larger than at the beginning of the cycle, due to irreversible degradation mechanisms.

The battery degradation in the nth cycle as a function of ρ and D is denoted by $\Delta_\rho(D, C_n) = C_n - C_{n+1}$. Assuming that $\Delta_\rho(D, C_n) \ll 2DC_0$, *i.e.*, the battery degradation is much smaller than the amount of charge exchanged by the battery over each cycle (this is a good approximation for typical values of D), and the discharge/recharge current is I, the duration of the nth discharge/recharge cycle is denoted by $T_n = [2DC_0 - \Delta_\rho(D, C_n)]/I \simeq 2DC_0/I$, and the discharge and recharge phases have the same duration (within one cycle). The charge level over the nth cycle, $Q_n(t)\Delta c$, $t \in (0, T_n)$, evolves as

$$\text{Discharge phase: } Q_n(t)\Delta c = C_n - It,\ t \in (0, T_n/2), \quad (26)$$
$$\text{Recharge phase: } Q_n(t)\Delta c = C_n - DC_0 + I(t - T_n/2),$$
$$t \in (T_n/2, T_n). \quad (27)$$

Moreover, due to the ongoing degradation, the instantaneous battery capacity in the nth cycle, denoted by $C_n(t),\ t \in$

$(0, T_n)$, obeys

$$\frac{\mathrm{d}C_n(t)}{\mathrm{d}t} = C_n'(t) = -\rho\left(\frac{Q_n(t)\Delta c}{C_0}\right),\ t \in (0, T_n), \quad (28)$$

with the boundary conditions $C_n(0) = C_n$, $C_n(T_n) = C_{n+1}$. By integrating the charge flows in one cycle, we then have

$$C_{n+1} = C_n + \int_0^{T_n/2} C_n'(\tau)\mathrm{d}\tau + \int_{T_n/2}^{T_n} C_n'(\tau)\mathrm{d}\tau, \quad (29)$$

and, substituting (28) in (29) and using the expression of ρ given in (25) and those for $Q_n(t)$ given in (26) and (27) for the two integrals, we obtain

$$\Delta_\rho(D, C_n) = \frac{2C_0\zeta}{I\theta} e^{\theta(1-C_n/C_0)}(e^{\theta D} - 1). \quad (30)$$

The term $N_{\mathrm{cyc}}(D)$ is equivalently defined as $N_{\mathrm{cyc}}(D) = \min\{n : C_n < xC_0\}$, since the number of cycles is counted until the battery capacity degrades to a fraction x of the nominal capacity. Herein, based on the fact that the battery capacity slowly degrades from the nominal value C_0 to the target xC_0, and that the number of cycles to obtain a small capacity degradation $\mathrm{d}C \ll C_0$ from $C \in (0, C_0]$ to $C - \mathrm{d}C$ are $\mathrm{d}C/\Delta_\rho(D, C)$, we approximate $N_{\mathrm{cyc}}(D)$ with the integral expression

$$N_{\mathrm{cyc}}(D) \simeq \int_{xC_0}^{C_0} \frac{1}{\Delta_\rho(D, C)}\mathrm{d}C. \quad (31)$$

Substituting (30) in (31), we thus obtain

$$N_{\mathrm{cyc}}(D) \simeq \left(\frac{I}{2\zeta}\frac{1 - e^{-\theta(1-x)}}{1 - e^{-\theta D}}\right) e^{-\theta D}. \quad (32)$$

Note that the term within the parentheses is a decreasing function of D, hence we obtain

$$N_{\mathrm{cyc}}(D) \geq \frac{I}{2\zeta}\frac{1 - e^{-\theta(1-x)}}{1 - e^{-\theta}} e^{-\theta D} \triangleq \hat{N}_{\mathrm{cyc}}(D), \quad (33)$$

where equality holds for $D = 1$. Finally, by approximating $N_{\mathrm{cyc}}(D)$ with its lower bound $\hat{N}_{\mathrm{cyc}}(D)$ and by matching this expression to the exponential model (2), we obtain

$$\alpha = \theta \text{ and } \zeta = \frac{I}{2N_{\mathrm{cyc},0}}\frac{1 - e^{-\alpha(1-x)}}{e^{\alpha} - 1} \text{ in (25).}$$

Remark 9. Note that the approximation (32) does not follow the exponential model (2). In particular, for $D \to 0$, in (32) we have $N_{\mathrm{cyc}}(D) \to \infty$. This is due to the fact that, in the derivation of (32), we have assumed that $\Delta_\rho(D, C_n) \ll 2DC_0$, *i.e.*, the DoD D is *large* with respect to the battery degradation in each cycle. However, this is a good approximation for typical values of D which the exponential model (2) has been fitted to [19]–[22], *e.g.*, $D \in [0.2, 1]$.

B. Stochastic Degradation Model

Based on the deterministic battery degradation model analyzed in the previous section, we now derive the degradation probabilities $p_H(q)$ for the stochastic model. To this end, we compute the deterministic time it takes for the battery to degrade from health state h, with capacity $\frac{h}{H_{\max}}C_0$, to the next lower health state $h - 1$, with capacity $\frac{h-1}{H_{\max}}C_0$. Then, we relate the *deterministic* degradation times to the *average* degradation times in the discrete-time stochastic model, and derive the corresponding transition probability.

Assume that the battery operates indefinitely at charge level $q\Delta c$ in the deterministic model studied in Sec. V-A. The initial battery capacity is $C(0) = \frac{h}{H_{\max}}q_{\max}\Delta c$. From (28), the battery capacity as a function of time is given by $C(t) = C(0) - \rho(q\Delta c/C_0)t$ and degrades to the next health state with capacity $\frac{h-1}{H_{\max}}q_{\max}\Delta c$ over a time-interval of duration

$$T_{\mathrm{det}}(q) = \frac{q_{\max}\Delta c}{H_{\max}\rho(q\Delta c/C_0)}. \quad (34)$$

On the other hand, in the stochastic, discrete-time model, assuming that the battery operates indefinitely at charge level q, measured in charge quanta, the *average* amount of time (in s) it takes for the battery to degrade to the lower health state is

$$T_{\mathrm{stoc}}(q) = \frac{\Delta t}{p_H(q)}, \quad (35)$$

where Δt is the time-slot duration. By forcing $T_{\mathrm{stoc}}(q) = T_{\mathrm{det}}(q)$, we finally obtain the relation

$$p_H(q) = \gamma \exp\left\{\alpha\left(1 - \frac{q}{q_{\max}}\right)\right\}, \quad (36)$$

where $\gamma = \frac{H_{\max}}{q_{\max}}\frac{\zeta\Delta t}{\Delta c}$ is a dimensionless constant. We note that (36) obeys Ass. 1.a) (as long as $\gamma \neq 0$) and Ass. 1.c) (since $\alpha > 0$). Moreover, if $\gamma \ll 1$, also Ass. 1.b) holds.

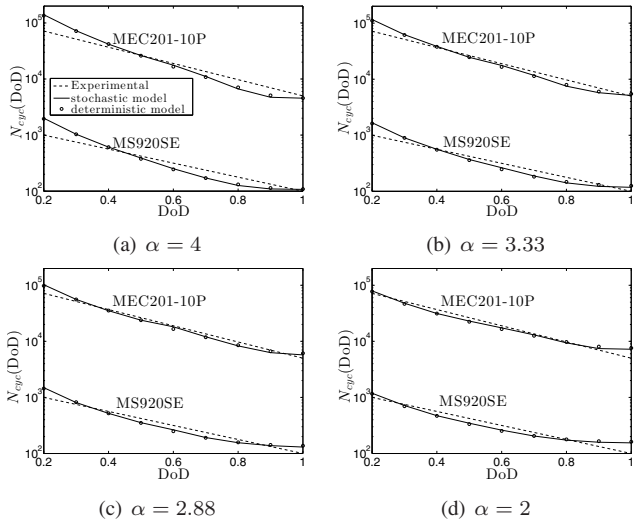
Remark 10. It is worth noting that the absolute value of γ does not affect the solution of the optimization problem (23), which, under the relationship (36), becomes

$$\mu_h^* = \arg\min_{\mu_h} \sum_{(q,s)\in\mathcal{Q}(h)\times\mathcal{S}} \pi_{\mu_h}^h(q, s) \exp\left\{\alpha\left(1 - \frac{q}{q_{\max}}\right)\right\}$$
$$\text{s.t.} \sum_{(q,s)\in\mathcal{Q}(h)\times\mathcal{S}} \pi_{\mu_h}^h(q, s)\left(\mathbb{E}_{\mu_h(\cdot;q,s)}\left[g(A, q)\right] - \mathcal{G}^*\right) \geq 0.$$

VI. Numerical Results

In this section, we present numerical results. In particular, we validate the proposed stochastic framework to model the battery degradation process, and we assess the performance of the proposed lifetime aware policies in terms of maximizing the battery lifetime, while guaranteeing a target QoS to the system. We consider a battery with capacity $q_{\max} = 500$ charge levels and $H_{\max} = 50$ health states. The parameter α, which determines the degradation probabilities $p_H(q)$ in (36), is obtained by interpolating the data-sheet values of two different Li-Ion rechargeable micro batteries: the battery device MS920SE [18], which is declared to provide 100 cycles at 100% DoD until the battery capacity degrades to 50% of the initial capacity C_0, and 1000 cycles at 20% DoD; and the battery device MEC201-10P [31], which is declared to provide 5000 cycles at 100% DoD until the battery capacity degrades to 80% of the initial capacity C_0, and 100000 cycles at 10% DoD. Using (2), we obtain $N_{\mathrm{cyc},0} = 100$, $\alpha \simeq 2.88$ for the device MS920SE, and $N_{\mathrm{cyc},0} = 5000$, $\alpha \simeq 3.33$ for the device MEC201-10P.

These values are then used to compute the degradation probabilities $p_H(q)$ as in (36). As discussed in Sec. V-B, the



(a) $\alpha = 4$ (b) $\alpha = 3.33$

(c) $\alpha = 2.88$ (d) $\alpha = 2$

Fig. 2. Number of cycles versus DoD. The curve for the stochastic model is obtained by averaging the number of cycles over 100 iterations. The storage capacity of MS920SE is degraded up to 50% of the nominal value. The storage capacity of MEC201-10P is degraded up to 80% of the nominal capacity.

constant γ in (36) does not affect the optimization problem (23), hence we choose a small value $\gamma = 2.5 \cdot 10^{-5}$ so as to satisfy Ass. 1.b) and Lemma 1.

In Fig. 2, we validate the proposed stochastic model against the experimental curve (2) for the $N_{\rm cyc}(D)$ versus DoD D dependence, for the battery models considered. In particular, these curves are obtained by cyclically discharging and recharging the battery with different values of the DoD D. The curves associated with the *stochastic* model are obtained by employing the stochastic model proposed in this paper to generate the health state process $\{H_k\}$, which determines the battery capacity via (5). The curves associated with the *deterministic* model, instead, are obtained by employing the deterministic degradation model developed in Sec. V-A to generate the battery degradation process. The number of cycles for a specific value of the DoD D are counted until the capacity degrades to 50% and 80% of the initial capacity C_0, for the two devices MS920SE and MEC201-10P, respectively. We notice that there is a good match between the deterministic and stochastic models, which gives evidence of the fact that the proposed Markov model captures the fundamental behavior of real batteries for what concerns their storage capacity degradation over time. Moreover, the stochastic model exhibits a good fit to the experimental curve, which validates our analysis in Sec. V. The values $\alpha = 2.88$ and $\alpha = 3.33$ give the best match to the experimental curves for the two devices MS920SE and MEC201-10P, respectively, and are employed in the following numerical evaluations (we have verified that these values minimize the mean square error with respect to the experimental curve, in the logarithmic domain).

In the following figures, the scenario process $\{S_k\}$ is modeled as a two state Markov chain with state space $\mathcal{S} = \{\mathrm{G}, \mathrm{B}\}$ [23] and transition probabilities $p_S(\mathrm{G}|\mathrm{G}) = p_S(\mathrm{B}|\mathrm{B}) = 0.96$, where G and B denote the "good" and "bad" scenarios, respectively. In the "bad" scenario ($S_k = \mathrm{B}$), no energy is harvested, *i.e.*, $B_k = 0$; in the "good" scenario ($S_k = \mathrm{G}$), the harvested energy is $B_k = 20$ deterministically. The average harvesting rate is thus given by $\bar{b} = 10$. In this case, we have a one-to-one mapping between S_k and B_k, so that, by measuring B_k, the state S_k is known exactly.

We employ the reward function $g^*(A_k) = \log_2(1+\sigma A_k/\bar{b})$, with $\sigma = 10$, which models the Shannon capacity of the static Gaussian channel, where σ is an SNR scaling parameter [29]. The action space is $\mathcal{A} \triangleq \{0\} \cup \{A_{\min}, \ldots, A_{\max}\}$ with $A_{\min} = 10$ and $A_{\max} = 20$. Note that a large value of $A_{\max} - A_{\min}$ implies a highly adaptive system, and this value depends on the specific application and system considered.

We consider the *Constant Load Lifetime Unaware Policy* (CLLUP), which supports a constant load of $A_k = \bar{A}$ charge quanta, for some $\bar{A} \in \mathcal{A}$, irrespective of the charge level available in the battery, and remains idle under energy outage. This policy requires minimal communication between the EHD controller and the power processing unit (Fig. 1), since the current charge level Q_k need not be known (except when $Q_k < \bar{A}$, in which case $A_k = 0$ as in (7)).

Moreover, we consider the *Lifetime Unaware Policy* (LUP), which greedily maximizes the average long-term reward (17) for the actual value of the battery capacity, without taking into account the impact of the policy on the battery lifetime. It is found via the Policy Iteration algorithm [24] as the solution of $\mu_h^* = \arg\max_{\mu_h} \hat{G}_{\mu_h}(h),\ \forall h \in \mathcal{H}$. This policy requires full knowledge of the current charge level, hence communication between the EHD controller and the power processing unit.

Finally, we consider the following policies, which explicitly take into account battery lifetime:

- *Lifetime Aware Optimal Policy* (LAOP): this is the optimal policy solution of problem (15), found via Algorithm 1.
- *Constant Load Lifetime Aware Policy* (CLLAP): This policy supports a constant load of $A_k = \bar{A}$ charge quanta, when the battery charge level is above a given DoD (with respect to the nominal capacity), and remains idle otherwise ($A_k = 0$). If the battery capacity degrades to a value such that the required DoD cannot be supported anymore, battery failure is declared. For both CLLUP and CLLAP, we choose $\bar{A} = \bar{b}$. However, other values may be used, depending on the load requirement, with no additional insights.

In the following plots, we evaluate the trade-off between lifetime and QoS, by comparing, respectively, LUP with LAOP and CLLUP with CLLAP. For a given policy and QoS $\mathcal{G}^*$, the battery lifetime is computed according to (12), using standard results on absorbing Markov Chains, see [30]. The corresponding *minimum reward* supported by policy μ over the battery lifetime is defined as $G_{\min}(\mu, \mathcal{G}^*) = \min_{h \geq h_\mu^*} G_\mu(h, \boldsymbol{Z}_0)$, where h_μ^* and $G_\mu(h, \boldsymbol{Z}_0)$ are defined in (13) and (11), respectively. The minimum reward represents the average reward per slot (averaged over a timescale much larger than the communication time-scale, but smaller than the battery degradation process) that is *guaranteed* over the entire battery lifetime.

To further validate the stochastic model proposed in this paper, in Figs. 3 and 4 we plot the result of a simulation, where the battery degradation process follows either the stochastic model of Sec. III, or the deterministic model of Sec. V-A. However, notice that, in the latter case, the term *deterministic* is only referred to the fact that, in each time-slot, the battery capacity degrades by a deterministic quantity, which depends on the charge level, as in Sec. V-A. On the other hand,

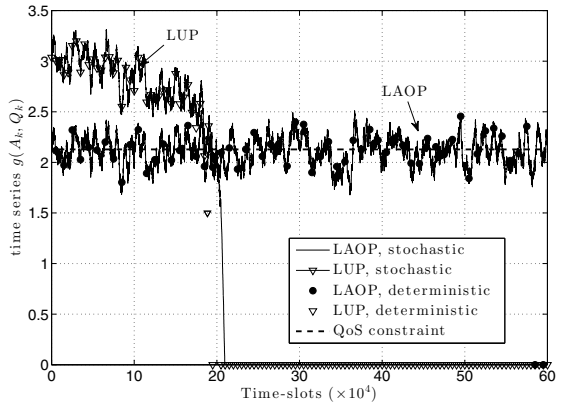


Fig. 3. Time series of the reward $g(A_k, Q_k)$, comparison of stochastic and deterministic degradation models for the battery device MS920SE. Each point in the curve is obtained by a moving-average window of 5000 time-slots. QoS requirement $\mathcal{G}^* = 2.13$ (corresponding to 70% of the maximum reward $\max_{\mu_{H_{\max}}} \hat{G}_{\mu_{H_{\max}}}(H_{\max}) \simeq 3.06$ in the maximum health state).

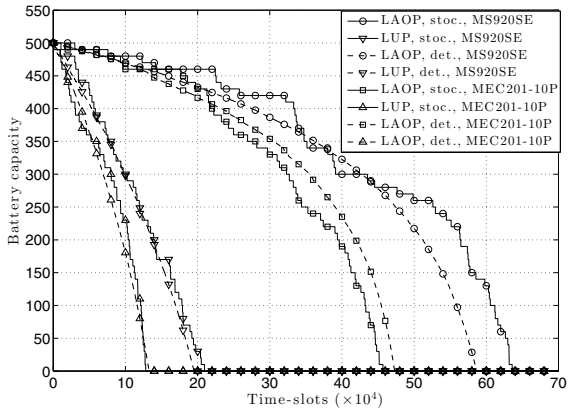


Fig. 4. Capacity degradation under the stochastic and deterministic degradation models. QoS requirement $\mathcal{G}^* = 2.13$.

the charge level is a stochastic process, induced by the stochastic energy arrival and decision processes. In particular, in Fig. 3, we plot the moving average curve associated with the reward sequence $\{g(A_k, Q_k)\}$ for the battery device MS920SE (similar considerations hold for the battery device MEC201-10P), and, in Fig. 4, we plot the time-series of the battery capacity for both devices. We notice a good match between the curves associated with the deterministic and stochastic models. It follows that a policy designed under the assumption of a stochastic battery degradation model (LAOP) achieves good performance even if the underlying degradation process is deterministic. For the device MS920SE, as shown in Fig. 3, LUP achieves a larger reward than LAOP in the time-horizon $[0, 18 \times 10^4]$, where the battery capacity is larger than ~ 50 (Fig. 4). This is because LUP exploits all the available charge levels to earn the maximum reward, by performing deep charge/discharge cycles. However, such behavior quickly deteriorates the battery capacity, which decays to zero much faster than LAOP. In contrast, LAOP performs close to the QoS requirement, and it intelligently manages the battery to prolong its lifetime. Finally, notice that the time-average reward sequence exhibits fluctuations around its mean, and

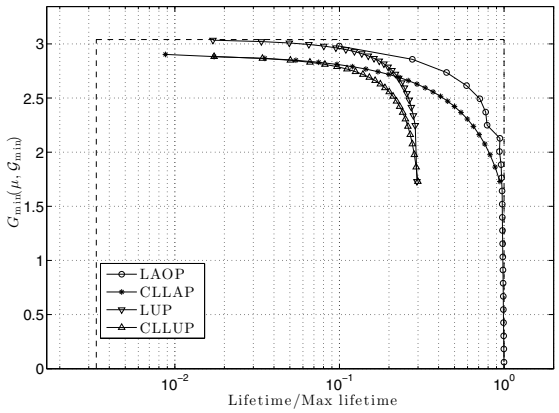


Fig. 5. Minimum reward over the battery lifetime versus normalized lifetime for the battery device MS920SE. The dashed lines represent the minimum and maximum lifetime and the maximum reward $\max_{\mu_{H_{\max}}} \hat{G}_{\mu_{H_{\max}}}(H_{\max})$.

often falls below the QoS constraint, due to the stochastic and time-correlated nature of the energy harvesting supply. A practical approach to reduce this effect would be to enforce a more stringent QoS constraint than the desirable one so as to provide some margin. Alternatively, as discussed in Remark 6, the reward function $g^*(A_k) = \chi(A_k \geq \bar{A})$ can be employed, and a constraint on the success probability $G_\mu(h, \boldsymbol{Z}_0) = \mathbb{P}(\bar{A} \leq A_k \leq Q_k) \geq \mathcal{G}^*$ may be enforced. Similar considerations hold for the device MEC201-10P.

In Fig. 5, we plot the minimum reward $G_{\min}(\mu, \mathcal{G}^*)$ versus the corresponding battery lifetime normalized to the maximum lifetime, which is defined as the lifetime when the battery is always fully charged, so that battery degradation mechanisms are slower, according to our extrapolated model (36) and Ass. 1.c). We note that, for a given minimum guaranteed QoS (a value in the y-axis of the figure), LAOP achieves a significant gain in terms of battery lifetime with respect to the "greedy" policy LUP, which does not take into consideration battery degradation mechanisms. In particular, the lifetime is increased by a factor ~ 3. The same observation holds when comparing CLLAP and CLLUP. Moreover, although CLLAP incurs a loss with respect to LAOP, it provides a good heuristic to enhance the battery lifetime, that is, battery lifetime can be significantly increased by allowing only shallow battery discharges, and by avoiding battery discharge below a predetermined DoD value. Finally, for all policies, the longer the lifetime, the smaller the minimum reward attained. This is due to the inherent trade-off between lifetime and reward. Namely, the battery lifetime is maximized by performing shallow charge/discharge cycles, which in turn considerably limits the usable charge levels, thus impairing the ability of the battery to filter out the fluctuations in the intermittent energy harvesting process, and to provide a satisfactory QoS over time. Conversely, the QoS is maximized by performing deep battery discharges, *e.g.*, during a long period of energy shortage, which inevitably shortens battery lifetime. This behavior is not captured by the models commonly used in the literature, which assume perpetual battery operation, *e.g.*, [5], [6], [8], [9], [32].

In Fig. 6, we plot the lifetime of each health state $h \in \mathcal{H}$, defined in (10) (lines). We also plot the lifetime approximation

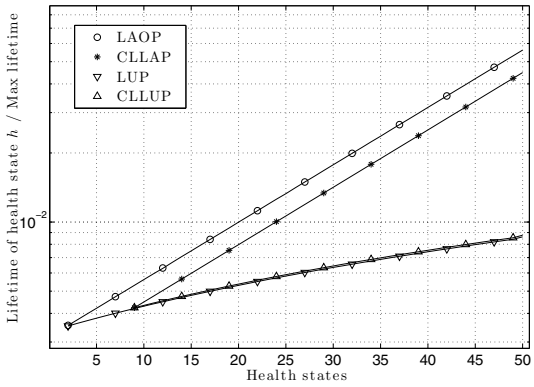


Fig. 6. Normalized lifetime of each health state. Exact lifetime (10) (lines). Approximation (18) (markers). QoS requirement $\mathcal{G}^* = 2.13$ for the battery device MS920SE.

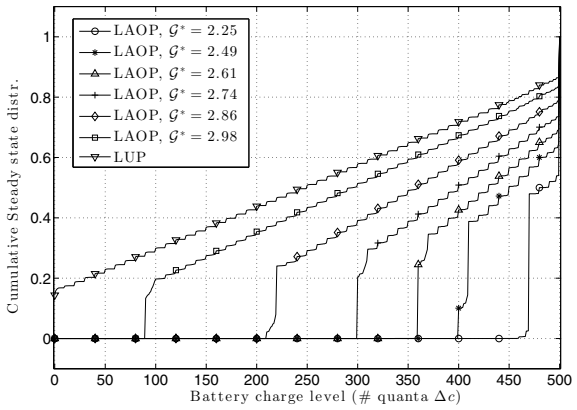


Fig. 7. Cumulative steady state distribution of charge levels at the maximum health state $H_{\max}$, for the battery device MS920SE.

(18) (markers). We notice that the exact lifetime expression (10) is closely approximated by (18), as proved in Lemma 1 when $\max_q p_H(h;q) \ll 1$. Moreover, LAOP maximizes the lifetime of all health states. In fact, LAOP is found using Algorithm 1, which, in step 2), determines the optimal policy that minimizes, for each health state $h \in \mathcal{H}$, the steady state probability of degradation (equivalently, it maximizes the lifetime of health state h), subject to a QoS constraint $\mathcal{G}^*$. Conversely, a much shorter lifetime is attained by LUP in each health state, since this policy greedily maximizes the reward, without taking into account its impact on the battery degradation. Similar considerations hold for CLLAP and CLLUP. Furthermore, CLLAP and CLLUP are unable to provide a satisfactory QoS when the health state falls below 8. In this case, adaptation of the action A_k becomes critical. In general, the more degraded the battery, the faster the degradation. This behavior is consistent with Ass. 1.c).

Finally, in Fig. 7, we plot the cumulative steady state distribution of the charge levels, for the maximum health state $H_{\max}$, for LUP and LAOP, for different QoS requirements (corresponding, in sequence, to 74%, 82%, 86%, 90%, 94% and 98% of the maximum reward $\max_{\mu_{H_{\max}}} \hat{G}_{\mu_{H_{\max}}}(H_{\max}) \simeq 3.06$ in the maximum health state). We note that the steady state distribution of LUP, which does not take into account the ongoing battery degradation mechanisms, is spread over all the battery charge levels. In particular, this policy operates for a significant fraction of the time at low charge levels, thus inducing a fast battery degradation. Conversely, LAOP spreads the steady state distribution over the upper charge levels only, thus slowing down battery degradation. Moreover, the more stringent the QoS requirement, the more spread the steady state distribution under LAOP over lower charge levels. This is because deeper discharge cycles need to be performed, in order to meet a stricter QoS requirement. Similar considerations hold for the device MEC201-10P.

VII. Extensions

In this section, we show how the model (3) can be extended to include other non-idealities, such as battery leakage, sensing, processing and activation costs. The impact of some of these phenomena has been analyzed from an information theoretic perspective in [10], [33] (battery leakage) and [6] (sensing and processing costs). Moreover, we show how this model can be further extended to accommodate a more general class of EH processes.

In particular, the model (3) can be extended to

$$Q_{k+1} = \min\left\{[Q_k - A_k - L_k]^+ + B_k, Q_{\max}(k+1)\right\}, \quad (37)$$

where L_k is the overall energy cost in slot k, not including the control A_k, which includes the costs of battery leakage, sensing, processing and activation of the circuitry after the node goes to sleep (if $A_{k-1} = 0$). We model L_k as a random variable with probability distribution $p_L(L_k|Q_k, A_k, I_k)$ taking values in the set $\mathcal{L} \triangleq \{0, 1, \ldots, L_{\max}\}$, possibly dependent on the charge level Q_k, action A_k, and idle state I_k. The idle state $I_k = \chi(A_{k-1} = 0)$ tracks the idle/active mode of the sensor node, so that, if $I_k = 0$, then the node was active in the previous slot $k-1$ ($A_{k-1} > 0$); otherwise, the node was idle ($A_{k-1} = 0$). The dependence of p_L on I_k may be used to model activation costs of the sensor circuitry, *i.e.*, $\mathbb{P}(L_k \geq l|Q_k, A_k, I_k = 1) \geq \mathbb{P}(L_k \geq l|Q_k, A_k, I_k = 0)$, $\forall l, \forall A_k > 0$, so that a higher energy cost is incurred when switching from idle to active mode ($I_k = 1$ and $A_k > 0$) than when staying active ($I_k = 0$ and $A_k > 0$).

The harvesting process $\{B_k\}$ can be generalized to include the *Generalized Markov model* with order $L = 1$, presented in [23]. In particular, as in (3), the scenario process $\{S_k\}$ is an irreducible stationary Markov chain with transition probability $p_S(s_{k+1}|s_k)$. Given the scenario $S_k \in \mathcal{S}$, the energy harvest B_k is drawn from $\mathcal{B}$ according to the distribution $p_B(b_k|s_k, b_{k-1}) \triangleq \Pr(B_k = b_k|S_k = s_k, B_{k-1} = b_{k-1})$, which is also a function of the previous energy harvest B_{k-1}. This model exhibits a good fit for the solar energy source [23].

For this more general model, the state in slot k is defined as $\boldsymbol{Z}_k = (Q_k, H_k, S_{k-1}, B_{k-1}, I_k)$, taking values in the state space $\mathcal{Q}\times\mathcal{H}\times\mathcal{S}\times\mathcal{B}\times\{0,1\}$, and the policy μ_{H_k} is a probability measure on the action space $\mathcal{A}$, parameterized by the state $(Q_k, S_{k-1}, B_{k-1}, I_k)$, *i.e.*, $\mu_{H_k}(a; Q_k, S_{k-1}, B_{k-1}, I_k)$ is the probability of requesting a charge quanta from the battery, when operating in state $\boldsymbol{Z}_k$. Note that also B_{k-1} is part of the state, since the statistics of B_k depends on both S_k and

B_{k-1}. The instantaneous reward accrued in time-slot k, in state $\boldsymbol{Z}_k$ under action A_k, is redefined as

$$g(A_k, Q_k, L_k) = \begin{cases} 0, & A_k > Q_k - L_k, \\ g^*(A_k), & A_k \leq Q_k - L_k, \end{cases} \tag{38}$$

Given the generality of the analysis in Sec. IV, Lemma 1 can be shown to hold for this more general model as well. The policy μ_{H_k} can then be optimized via Algorithm 1, where, letting $\mathcal{U}(h) \equiv \mathcal{Q}(h) \times \mathcal{S} \times \mathcal{B} \times \{0,1\}$ (23) is replaced with

$$\mu_h^* = \arg\min_{\mu_h} \sum_{(q,s,b,i) \in \mathcal{U}(h)} \pi_{\mu_h}^h(q,s,b,i) p_H(h;q) \tag{39}$$

$$\text{s.t.} \sum_{(q,s,b,i) \in \mathcal{U}(h)} \pi_{\mu_h}^h(q,s,b,i) \left(\mathbb{E}_{\mu_h(\cdot;q,s,b,i)} \left[g(A,q,L) \right] - \mathcal{G}^* \right) \geq 0.$$

VIII. Conclusions

In this paper, we have analyzed the impact of battery management policies on the irreversible degradation of the storage capacity of realistic batteries, affecting the lifetime of harvesting based Wireless Sensor Networks. We have proposed a general framework, based on Markov chains and suitable for policy optimization, which captures the degradation status of the battery. The proposed stochastic battery degradation model has been extrapolated from manufacturer-provided data and realistic deterministic models proposed in the literature, and has been shown to fit well the behavior of real batteries for what concerns their storage capacity degradation over time. Note, however, that different battery degradation models can be easily accommodated in the proposed framework, depending on the available experimental data and the desired accuracy. Based on the proposed model, we have formulated the policy optimization problem as the maximization of the battery lifetime, subject to a minimum guaranteed QoS in each battery degradation status. We have shown that this problem can be solved efficiently by a sequential linear programming optimization algorithm over the degradation states of the battery. The numerical evaluation gives evidence of the fact that a lifetime-aware management policy can significantly improve the lifetime of the sensor node with respect to a "greedy" operation policy, while guaranteeing the QoS.

Appendix

Proof of Lemma 1: For the proof of the lemma, we present a general result of Markov chains. The relationship to the specific problem considered in this paper is provided at the end of the proof. Consider a finite Markov chain $\{Z_k\} \subseteq \mathcal{Z} \equiv \{1, 2, \ldots, N_\mathrm{t}+1\}$, where the state space $\mathcal{S}$ is partitioned into a set of transient states $\mathcal{Z}_\mathrm{t} \equiv \{1, \ldots, N_\mathrm{t}\}$ forming a communicating class, and the absorbing state $\mathcal{Z}_\mathrm{a} \equiv \{N_\mathrm{t}+1\}$, with transition matrix

$$\mathbf{P}_\epsilon = \begin{bmatrix} (\mathbf{I}_{N_t} - \epsilon \mathbf{P}_\mathrm{a}) \mathbf{P}_\mathrm{t} & \epsilon \mathbf{P}_\mathrm{a} \mathbf{1}_{N_t} \\ \mathbf{0}_{N_t}^T & 1 \end{bmatrix}, \tag{40}$$

where $\mathbf{0}_K$ is a $K \times 1$ vector with all entries equal to zero; $\mathbf{1}_K$ is an $K \times 1$ vector with all entries equal to one; $\mathbf{I}_K$ is the $K \times K$ identity matrix; $\mathbf{P}_\mathrm{t}$ is the $N_\mathrm{t} \times N_\mathrm{t}$ transition probability matrix associated with transitions in $\mathcal{Z}_\mathrm{t}$, given that the Markov chain is not absorbed by $\mathcal{Z}_\mathrm{a}$; $\mathbf{P}_\mathrm{a}$ is an $N_\mathrm{t} \times N_\mathrm{t}$ diagonal matrix with strictly positive diagonal elements, and $\epsilon \mathbf{P}_\mathrm{a}(i,i) \in (0,1)$ is the probability of moving from state i to the absorbing state $N_\mathrm{t}+1$, where the scaling parameter ϵ can take any value in $(0, 1/\max_i \mathbf{P}_\mathrm{a}(i,i))$ (we will be interested in $\epsilon \to 0$). In the following, $\mathbf{e}_{1,K}$ denotes the first column of $\mathbf{I}_K$. Moreover, for convenience we drop the dependence of $\mathbf{0}_K$, $\mathbf{1}_K$, $\mathbf{I}_K$ and $\mathbf{e}_{1,K}$ on K in the notation whenever the size K can be deduced from the context.

We assume that $\mathbf{P}_\mathrm{t}$ is a regular stochastic matrix (*i.e.*, the associated Markov chain is irreducible and aperiodic). Therefore, $\mathbf{X}_\epsilon = (\mathbf{I} - \epsilon \mathbf{P}_\mathrm{a}) \mathbf{P}_\mathrm{t}$ is a primitive matrix and, from the Perron-Frobenius Theorem [34], there is a real positive eigenvalue λ_ϵ of $\mathbf{X}_\epsilon$, with algebraic multiplicity 1, such that any other eigenvalue β of $\mathbf{X}_\epsilon$ has $|\beta| < \lambda_\epsilon$. Since $\mathbf{X}_\epsilon$ is continuous in ϵ, λ_ϵ is also continuous. We denote the corresponding right eigenvector as $\mathbf{v}_\epsilon$, *i.e.*,

$$(\mathbf{X}_\epsilon - \lambda_\epsilon \mathbf{I}) \mathbf{v}_\epsilon = \mathbf{0}. \tag{41}$$

We normalize the eigenvector $\mathbf{v}_\epsilon$ so that the sum of its elements equals N_t[3], *i.e.*, $\mathbf{1}^T \mathbf{v}_\epsilon = N_t$, so that $\mathbf{v}_\epsilon$ is uniquely defined for each $\epsilon > 0$, and is continuous in ϵ. Since $\mathbf{X}_0 = \mathbf{P}_\mathrm{t}$ is a regular stochastic matrix, we have $\lambda_0 = 1$ and $\lambda_\epsilon < 1$ for $\epsilon > 0$. Moreover, $\mathbf{v}_0 = \mathbf{1}$ and there exists a unique $\boldsymbol{\pi}_{\mathrm{t},\infty}$ such that $\boldsymbol{\pi}_{\mathrm{t},\infty} = \boldsymbol{\pi}_{\mathrm{t},\infty} \mathbf{P}_\mathrm{t}$. We can thus write $\mathbf{X}_0$ as

$$\mathbf{X}_0 = \mathbf{U}_0 \mathbf{D}_0 \mathbf{U}_0^{-1}, \tag{42}$$

where $\mathbf{D}_0$ is the Jordan normal form of $\mathbf{X}_0$, and $\mathbf{U}_0$ is the matrix whose columns are the corresponding generalized eigenvectors [35]. Without loss of generality, $\mathbf{D}_0$ is given by

$$\mathbf{D}_0 = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{J}_0 \end{bmatrix}, \tag{43}$$

where $\mathbf{J}_0$ is a block diagonal matrix, whose diagonal blocks are given by the Jordan blocks corresponding to the eigenvalues of $\mathbf{X}_0$ inside the unit circle. Therefore, $\mathbf{U}_0 \mathbf{e}_1 = \mathbf{1}$ and $\mathbf{e}_1^T \mathbf{U}_0^{-1} = \boldsymbol{\pi}_{\mathrm{t},\infty}$, since $\mathbf{1}$ and $\boldsymbol{\pi}_{\mathrm{t},\infty}$ are, respectively, the right and left eigenvectors of $\mathbf{X}_0$ associated to the eigenvalue 1.

Recall, from standard results on absorbing Markov Chains (see [30]), that the expected time until absorption is given by

$$\mathcal{T}_\epsilon(\boldsymbol{\pi}_{\mathrm{t},0}) = \boldsymbol{\pi}_{\mathrm{t},0} \left(\mathbf{I} - \mathbf{X}_\epsilon \right)^{-1} \mathbf{1}, \tag{44}$$

where $\boldsymbol{\pi}_{\mathrm{t},0}$ is an initial distribution over $\mathcal{Z}_\mathrm{t}$. Note that, when $\epsilon > 0$, the eigenvalues of $\mathbf{X}_\epsilon$ are all strictly inside the unit circle, so that $\mathbf{I} - \mathbf{X}_\epsilon$ is invertible and (44) is well defined. We prove that

$$\mathcal{T}_\epsilon(\boldsymbol{\pi}_{\mathrm{t},0}) = \frac{1}{\epsilon \boldsymbol{\pi}_{\mathrm{t},\infty} \mathbf{P}_\mathrm{a} \mathbf{1}} + \mathcal{O}(1), \text{ for } \epsilon \to 0, \tag{45}$$

or equivalently, by definition of $\mathcal{O}(x)$,

$$\lim_{\epsilon \to 0} \left| \mathcal{T}_\epsilon(\boldsymbol{\pi}_{\mathrm{t},0}) - \frac{1}{\epsilon \boldsymbol{\pi}_{\mathrm{t},\infty} \mathbf{P}_\mathrm{a} \mathbf{1}} \right| < \infty. \tag{46}$$

[3]This is always possible since the Perron-Frobenius Theorem guarantees that there always exists an eigenvector associated to the eigenvalue λ_ϵ with all positive elements [34].

We have

$$\begin{aligned}\mathcal{T}_\epsilon(\boldsymbol{\pi}_{\mathrm{t},0}) - \frac{1}{\epsilon\boldsymbol{\pi}_{\mathrm{t},\infty}\mathbf{P}_\mathrm{a}\mathbf{1}} &\overset{(a)}{=} \boldsymbol{\pi}_{\mathrm{t},0}\left(\mathbf{I}-\mathbf{X}_\epsilon\right)^{-1}\mathbf{1} - \frac{\boldsymbol{\pi}_{\mathrm{t},0}\mathbf{1}}{\epsilon\boldsymbol{\pi}_{\mathrm{t},\infty}\mathbf{P}_\mathrm{a}\mathbf{1}}\\ &= \frac{1}{\epsilon\boldsymbol{\pi}_{\mathrm{t},\infty}\mathbf{P}_\mathrm{a}\mathbf{1}}\boldsymbol{\pi}_{\mathrm{t},0}\left(\mathbf{I}-\mathbf{X}_\epsilon\right)^{-1}\left[\epsilon\mathbf{1}\boldsymbol{\pi}_{\mathrm{t},\infty}\mathbf{P}_\mathrm{a}\mathbf{1} - \left(\mathbf{I}-\mathbf{X}_\epsilon\right)\mathbf{1}\right]\\ &\overset{(b)}{=} \frac{1}{\boldsymbol{\pi}_{\mathrm{t},\infty}\mathbf{P}_\mathrm{a}\mathbf{1}}\boldsymbol{\pi}_{\mathrm{t},0}\left(\mathbf{I}-\mathbf{X}_\epsilon\right)^{-1}\left(\mathbf{1}\boldsymbol{\pi}_{\mathrm{t},\infty}-\mathbf{I}\right)\mathbf{P}_\mathrm{a}\mathbf{1}\\ &= \boldsymbol{\pi}_{\mathrm{t},0}\left(\mathbf{I}-\mathbf{X}_\epsilon\right)^{-1}\left(\mathbf{1}\boldsymbol{\pi}_{\mathrm{t},\infty}-\mathbf{I}\right)\mathbf{x}, \end{aligned} \tag{47}$$

where we have defined the vector $\mathbf{x} = (\boldsymbol{\pi}_{\mathrm{t},\infty}\mathbf{P}_\mathrm{a}\mathbf{1})^{-1}\mathbf{P}_\mathrm{a}\mathbf{1}$. In step (a), we have used the fact that $1 = \boldsymbol{\pi}_{\mathrm{t},0}\mathbf{1}$. In step (b), we have used the fact that $\mathbf{X}_\epsilon\mathbf{1} = \left(\mathbf{I}-\epsilon\mathbf{P}_\mathrm{a}\right)\mathbf{P}_\mathrm{t}\mathbf{1} = \left(\mathbf{I}-\epsilon\mathbf{P}_\mathrm{a}\right)\mathbf{1}$. Let

$$\mathbf{U}_\epsilon = \mathbf{U}_0 + (\mathbf{v}_\epsilon - \mathbf{1})\mathbf{e}_1^T. \tag{48}$$

Since $\mathbf{U}_0$ is invertible, there exists $\epsilon_\mathrm{th} > 0$ such that $\mathbf{U}_\epsilon$ is also invertible, for all $\epsilon \in (0, \epsilon_\mathrm{th})$, by continuity. For any such ϵ, we can thus write

$$\mathbf{X}_\epsilon = \mathbf{U}_\epsilon\mathbf{D}_\epsilon\mathbf{U}_\epsilon^{-1}, \text{ where } \mathbf{D}_\epsilon = \begin{bmatrix} \lambda_\epsilon & \mathbf{r}_\epsilon \\ \mathbf{0} & \mathbf{J}_\epsilon \end{bmatrix}, \tag{49}$$

and, using the fact that $\mathbf{e}_1^T[\mathbf{0},\mathbf{I}]^T = \mathbf{0}^T$, hence $\mathbf{U}_\epsilon[\mathbf{0},\mathbf{I}]^T = \mathbf{U}_0[\mathbf{0},\mathbf{I}]^T$,

$$\begin{bmatrix} \mathbf{r}_\epsilon \\ \mathbf{J}_\epsilon \end{bmatrix} = \mathbf{U}_\epsilon^{-1}\mathbf{X}_\epsilon\mathbf{U}_\epsilon\begin{bmatrix} \mathbf{0}^T \\ \mathbf{I} \end{bmatrix} = \mathbf{U}_\epsilon^{-1}\mathbf{X}_\epsilon\mathbf{U}_0\begin{bmatrix} \mathbf{0}^T \\ \mathbf{I} \end{bmatrix}. \tag{50}$$

Then, using (49) and the fact that $\mathbf{I}-\mathbf{X}_\epsilon = \mathbf{U}_\epsilon(\mathbf{I}-\mathbf{D}_\epsilon)\mathbf{U}_\epsilon^{-1}$, we obtain

$$\begin{aligned}\left(\mathbf{I}-\mathbf{X}_\epsilon\right)^{-1} &= \mathbf{U}_\epsilon\begin{bmatrix} 1-\lambda_\epsilon & -\mathbf{r}_\epsilon \\ \mathbf{0} & \mathbf{I}-\mathbf{J}_\epsilon \end{bmatrix}^{-1}\mathbf{U}_\epsilon^{-1}\\ &= \mathbf{U}_\epsilon\begin{bmatrix} \frac{1}{1-\lambda_\epsilon} & \frac{1}{1-\lambda_\epsilon}\mathbf{r}_\epsilon(\mathbf{I}-\mathbf{J}_\epsilon)^{-1} \\ \mathbf{0} & (\mathbf{I}-\mathbf{J}_\epsilon)^{-1} \end{bmatrix}\mathbf{U}_\epsilon^{-1}\\ &= \frac{1}{1-\lambda_\epsilon}\mathbf{v}_\epsilon\mathbf{e}_1^T\mathbf{U}_\epsilon^{-1} + \frac{1}{1-\lambda_\epsilon}\mathbf{v}_\epsilon\mathbf{r}_\epsilon(\mathbf{I}-\mathbf{J}_\epsilon)^{-1}\left[\mathbf{0},\mathbf{I}\right]\mathbf{U}_\epsilon^{-1}\\ &\quad + \mathbf{U}_\epsilon[\mathbf{0},\mathbf{I}]^T(\mathbf{I}-\mathbf{J}_\epsilon)^{-1}[\mathbf{0},\mathbf{I}]\mathbf{U}_\epsilon^{-1}. \end{aligned} \tag{51}$$

In the last step, we have used the fact that $\mathbf{I} = \mathbf{e}_1\mathbf{e}_1^T + [\mathbf{0},\mathbf{I}]^T[\mathbf{0},\mathbf{I}]$ and $\mathbf{U}_\epsilon\mathbf{e}_1 = \mathbf{v}_\epsilon$, hence $\mathbf{U}_\epsilon = \mathbf{v}_\epsilon\mathbf{e}_1^T + \mathbf{U}_\epsilon[\mathbf{0},\mathbf{I}]^T[\mathbf{0},\mathbf{I}]$, $\mathbf{U}_\epsilon^{-1} = \mathbf{e}_1\mathbf{e}_1^T\mathbf{U}_\epsilon^{-1} + [\mathbf{0},\mathbf{I}]^T[\mathbf{0},\mathbf{I}]\mathbf{U}_\epsilon^{-1}$ and $(\mathbf{I}-\mathbf{D}_\epsilon)^{-1} = \frac{1}{1-\lambda_\epsilon}\mathbf{e}_1\mathbf{e}_1^T + \frac{1}{1-\lambda_\epsilon}\mathbf{e}_1\mathbf{r}_\epsilon(\mathbf{I}-\mathbf{J}_\epsilon)^{-1}[\mathbf{0},\mathbf{I}] + [\mathbf{0},\mathbf{I}]^T(\mathbf{I}-\mathbf{J}_\epsilon)^{-1}[\mathbf{0},\mathbf{I}]$; the result is then obtained by substituting these expressions, by expanding the products and by noting that $[\mathbf{0},\mathbf{I}]\mathbf{e}_1 = \mathbf{0}$. Since $\mathbf{J}_0$ is the Jordan matrix corresponding to eigenvalues of $\mathbf{X}_0$ within the unit circle, $\mathbf{I}-\mathbf{J}_0$ is invertible, hence, by continuity, $\mathbf{I}-\mathbf{J}_\epsilon$ is invertible for sufficiently small ϵ. By replacing (51) into (47), we thus get

$$\mathcal{T}_\epsilon(\boldsymbol{\pi}_{\mathrm{t},0}) - \frac{1}{\epsilon\boldsymbol{\pi}_{\mathrm{t},\infty}\mathbf{P}_\mathrm{a}\mathbf{1}} = A(\epsilon) + B(\epsilon) + C(\epsilon), \tag{52}$$

where we have defined

$$\begin{aligned} A(\epsilon) &= \frac{1}{1-\lambda_\epsilon}\boldsymbol{\pi}_{\mathrm{t},0}\mathbf{v}_\epsilon\mathbf{e}_1^T\mathbf{U}_\epsilon^{-1}\left(\mathbf{1}\boldsymbol{\pi}_{\mathrm{t},\infty}-\mathbf{I}\right)\mathbf{x}, \\ B(\epsilon) &= \frac{1}{1-\lambda_\epsilon}\boldsymbol{\pi}_{\mathrm{t},0}\mathbf{v}_\epsilon\mathbf{r}_\epsilon(\mathbf{I}-\mathbf{J}_\epsilon)^{-1}[\mathbf{0},\mathbf{I}]\mathbf{U}_\epsilon^{-1}\left(\mathbf{1}\boldsymbol{\pi}_{\mathrm{t},\infty}-\mathbf{I}\right)\mathbf{x}, \\ C(\epsilon) &= \boldsymbol{\pi}_{\mathrm{t},0}\mathbf{U}_\epsilon\left[\mathbf{0},\mathbf{I}\right]^T\left(\mathbf{I}-\mathbf{J}_\epsilon\right)^{-1}\left[\mathbf{0},\mathbf{I}\right]\mathbf{U}_\epsilon^{-1}\left(\mathbf{1}\boldsymbol{\pi}_{\mathrm{t},\infty}-\mathbf{I}\right)\mathbf{x}. \end{aligned}$$

We finally show that the limit of each term above exists and is finite for $\epsilon \to 0$, thus proving (46). Regarding the first term $A(\epsilon)$, since $\mathbf{e}_1^T\mathbf{U}_0^{-1}\left(\mathbf{1}\boldsymbol{\pi}_{\mathrm{t},\infty}-\mathbf{I}\right) = \boldsymbol{\pi}_{\mathrm{t},\infty}\left(\mathbf{1}\boldsymbol{\pi}_{\mathrm{t},\infty}-\mathbf{I}\right) = \mathbf{0}^T$, we obtain

$$A(\epsilon) = \frac{1}{1-\lambda_\epsilon}\boldsymbol{\pi}_{\mathrm{t},0}\mathbf{v}_\epsilon\mathbf{e}_1^T\left(\mathbf{U}_\epsilon^{-1}-\mathbf{U}_0^{-1}\right)\left(\mathbf{1}\boldsymbol{\pi}_{\mathrm{t},\infty}-\mathbf{I}\right)\mathbf{x}. \tag{54}$$

Moreover, from (48), we have

$$\begin{aligned}\mathbf{U}_\epsilon^{-1} - \mathbf{U}_0^{-1} &= \mathbf{U}_\epsilon^{-1}(\mathbf{U}_0-\mathbf{U}_\epsilon)\mathbf{U}_0^{-1}\\ &= -\mathbf{U}_\epsilon^{-1}(\mathbf{v}_\epsilon-\mathbf{1})\mathbf{e}_1^T\mathbf{U}_0^{-1} = -\mathbf{U}_\epsilon^{-1}(\mathbf{v}_\epsilon-\mathbf{1})\boldsymbol{\pi}_{\mathrm{t},\infty}. \end{aligned} \tag{55}$$

Substituting (55) in (54), we obtain $A(\epsilon) = 0$, since $\boldsymbol{\pi}_{\mathrm{t},\infty}\left(\mathbf{1}\boldsymbol{\pi}_{\mathrm{t},\infty}-\mathbf{I}\right) = \mathbf{0}^T$.

For the second term $B(\epsilon)$, substituting the expression of $\mathbf{r}_\epsilon = \mathbf{e}_1^T\mathbf{U}_\epsilon^{-1}\mathbf{X}_\epsilon\mathbf{U}_0[\mathbf{0},\mathbf{I}]^T$ given by (50) into (53), and using the fact that $\mathbf{e}_1^T\mathbf{U}_0^{-1}\mathbf{X}_0\mathbf{U}_0\left[\mathbf{0},\mathbf{I}\right]^T = \boldsymbol{\pi}_{\mathrm{t},\infty}\mathbf{U}_0\left[\mathbf{0},\mathbf{I}\right]^T = \mathbf{0}^T$, we obtain

$$\begin{aligned}B(\epsilon) =& \frac{1}{1-\lambda_\epsilon}\boldsymbol{\pi}_{\mathrm{t},0}\mathbf{v}_\epsilon\mathbf{e}_1^T(\mathbf{U}_\epsilon^{-1}\mathbf{X}_\epsilon - \mathbf{U}_0^{-1}\mathbf{X}_0)\mathbf{U}_0\\ &\times\left[\mathbf{0},\mathbf{I}\right]^T\left(\mathbf{I}-\mathbf{J}_\epsilon\right)^{-1}[\mathbf{0},\mathbf{I}]\mathbf{U}_\epsilon^{-1}\left(\mathbf{1}\boldsymbol{\pi}_{\mathrm{t},\infty}-\mathbf{I}\right)\mathbf{x}. \end{aligned} \tag{56}$$

Moreover, using (48) and (42), $\mathbf{U}_\epsilon\mathbf{U}_0^{-1}\mathbf{X}_0\mathbf{U}_0 = \mathbf{X}_0\mathbf{U}_0 + (\mathbf{v}_\epsilon-\mathbf{1})\mathbf{e}_1^T\mathbf{D}_0 = \mathbf{X}_0\mathbf{U}_0 + (\mathbf{v}_\epsilon-\mathbf{1})\mathbf{e}_1^T$, and therefore, since $\mathbf{X}_\epsilon = \left(\mathbf{I}-\epsilon\mathbf{P}_\mathrm{a}\right)\mathbf{P}_\mathrm{t}$,

$$\begin{aligned}&(\mathbf{U}_\epsilon^{-1}\mathbf{X}_\epsilon - \mathbf{U}_0^{-1}\mathbf{X}_0)\mathbf{U}_0 = \mathbf{U}_\epsilon^{-1}(\mathbf{X}_\epsilon\mathbf{U}_0 - \mathbf{U}_\epsilon\mathbf{U}_0^{-1}\mathbf{X}_0\mathbf{U}_0)\\ &= -\mathbf{U}_\epsilon^{-1}(\epsilon\mathbf{P}_\mathrm{a}\mathbf{X}_0\mathbf{U}_0 + (\mathbf{v}_\epsilon-\mathbf{1})\mathbf{e}_1^T). \end{aligned} \tag{57}$$

Therefore, by substituting (57) into (56), and noting that $\mathbf{e}_1^T[\mathbf{0},\mathbf{I}]^T = \mathbf{0}^T$, we obtain

$$\begin{aligned}B(\epsilon) =& -\frac{\epsilon}{1-\lambda_\epsilon}\boldsymbol{\pi}_{\mathrm{t},0}\mathbf{v}_\epsilon\mathbf{e}_1^T\mathbf{U}_\epsilon^{-1}\mathbf{P}_\mathrm{a}\mathbf{X}_0\mathbf{U}_0\\ &\times\left[\mathbf{0},\mathbf{I}\right]^T\left(\mathbf{I}-\mathbf{J}_\epsilon\right)^{-1}[\mathbf{0},\mathbf{I}]\mathbf{U}_\epsilon^{-1}\left(\mathbf{1}\boldsymbol{\pi}_{\mathrm{t},\infty}-\mathbf{I}\right)\mathbf{x}. \end{aligned} \tag{58}$$

Moreover, by left-multiplying each side of (41) by $\boldsymbol{\pi}_{\mathrm{t},\infty}$, for $\epsilon > 0$ we obtain

$$\frac{1-\lambda_\epsilon}{\epsilon} = \frac{\boldsymbol{\pi}_{\mathrm{t},\infty}\mathbf{P}_\mathrm{a}\mathbf{P}_\mathrm{t}\mathbf{v}_\epsilon}{\boldsymbol{\pi}_{\mathrm{t},\infty}\mathbf{v}_\epsilon} \to \boldsymbol{\pi}_{\mathrm{t},\infty}\mathbf{P}_\mathrm{a}\mathbf{1} > 0, \tag{59}$$

where the limit holds for $\epsilon \to 0$, since $\mathbf{v}_\epsilon \to \mathbf{1}$, $\mathbf{P}_\mathrm{t}\mathbf{1} = \mathbf{1}$ and $\boldsymbol{\pi}_{\mathrm{t},\infty}\mathbf{1} = 1$. Therefore, $B(\epsilon)$ for $\epsilon \to 0$ is bounded, since $\mathbf{I}-\mathbf{J}_0$ is invertible. Similarly, the limit of $C(\epsilon)$ for $\epsilon \to 0$ is bounded, and (45) is thus proved.

From [30], using a similar approach, the total cost/reward accrued before the process is absorbed by $\mathcal{Z}_\mathrm{a}$ is

$$\mathcal{C}_\epsilon^{\mathrm{tot}}(\boldsymbol{\pi}_{\mathrm{t},0}) = \boldsymbol{\pi}_{\mathrm{t},0}\left(\mathbf{I}-\mathbf{X}_\epsilon\right)^{-1}\mathbf{c}, \tag{60}$$

where $\mathbf{c} = [c(s)]_{s\in\mathcal{S}}$ is the cost/reward vector. We prove that

$$\frac{\mathcal{C}_\epsilon^{\mathrm{tot}}(\boldsymbol{\pi}_{\mathrm{t},0})}{\mathcal{T}_\epsilon(\boldsymbol{\pi}_{\mathrm{t},0})} = \boldsymbol{\pi}_{\mathrm{t},\infty}\mathbf{c} + \mathcal{O}(\epsilon). \tag{61}$$

Equivalently,

$$\lim_{\epsilon\to 0}\left|\frac{\mathcal{C}_\epsilon^{\mathrm{tot}}(\boldsymbol{\pi}_{\mathrm{t},0})}{\epsilon\mathcal{T}_\epsilon(\boldsymbol{\pi}_{\mathrm{t},0})} - \frac{1}{\epsilon}\boldsymbol{\pi}_{\mathrm{t},\infty}\mathbf{c}\right| < \infty. \tag{62}$$

Using (60) and (44), we obtain

$$\frac{\mathcal{C}_\epsilon^{\rm tot}(\boldsymbol{\pi}_{\rm t,0})}{\epsilon\mathcal{T}_\epsilon(\boldsymbol{\pi}_{\rm t,0})} - \frac{1}{\epsilon}\boldsymbol{\pi}_{\rm t,\infty}\mathbf{c} = \frac{\mathcal{C}_\epsilon^{\rm tot}(\boldsymbol{\pi}_{\rm t,0}) - \mathcal{T}_\epsilon(\boldsymbol{\pi}_{\rm t,0})\boldsymbol{\pi}_{\rm t,\infty}\mathbf{c}}{\epsilon\mathcal{T}_\epsilon(\boldsymbol{\pi}_{\rm t,0})} \quad (63)$$
$$= \frac{\boldsymbol{\pi}_{\rm t,0}\left(\mathbf{I}-\mathbf{X}_\epsilon\right)^{-1}\left(\mathbf{I}-\mathbf{1}\boldsymbol{\pi}_{\rm t,\infty}\right)\mathbf{c}}{\epsilon\mathcal{T}_\epsilon(\boldsymbol{\pi}_{\rm t,0})}.$$

We now compute the limit of the numerator and denominator of (63) separately. For the denominator $\epsilon\mathcal{T}_\epsilon(\boldsymbol{\pi}_{\rm t,0})$, from (45), $\epsilon\mathcal{T}_\epsilon(\boldsymbol{\pi}_{\rm t,0}) = \left(\boldsymbol{\pi}_{\rm t,\infty}\mathbf{P}_{\rm a}\mathbf{1}\right)^{-1} + \mathcal{O}(\epsilon)$, hence $\lim_{\epsilon\to 0}\epsilon\mathcal{T}_\epsilon(\boldsymbol{\pi}_{\rm t,0}) = \left(\boldsymbol{\pi}_{\rm t,\infty}\mathbf{P}_{\rm a}\mathbf{1}\right)^{-1}$, which is positive and bounded. Therefore, (62) holds as long as the numerator of (63) is bounded. This is directly shown since the numerator of (63) equals the last expression of (47) when $\mathbf{c} = -\mathbf{x}$, which, as previously shown, is bounded for $\epsilon \to 0$, for any bounded $\mathbf{x}$.

The connection to the problem at hand is obtained as follows. In health state h, the set of transient states ($\mathcal{Z}_{\rm t}$ in the proof of the lemma) is $\mathcal{Q}(h)\times\{h\}\times\mathcal{S}$. The absorbing state $\mathcal{Z}_{\rm a}$ corresponds to the set $\mathcal{Q}(h-1)\times\{h-1\}\times\mathcal{S}$, so that $\mathcal{C}_\epsilon^{\rm tot}(\boldsymbol{\pi}_{\rm t,0})$ and $\mathcal{T}_\epsilon(\boldsymbol{\pi}_{\rm t,0})$ count, respectively, the expected total cumulative reward earned and total time spent by the process $\{\boldsymbol{Z}_k\}$ while in health state h, until it is absorbed by the lower health state $h-1$. The initial distribution $\boldsymbol{\pi}_{\rm t,0}$ corresponds to the state distribution in the set $\mathcal{Q}(h)\times\{h\}\times\mathcal{S}$, when the process $\{\boldsymbol{Z}_k\}$ first hits health state h (this event occurs at time K_h, as defined in (8)), as induced by policy μ, by (3) and by the energy harvesting process. The transition probability matrix $\mathbf{P}_{\rm t}$ is associated to transitions within the set of transient states $\mathcal{Q}(h)\times\{h\}\times\mathcal{S}$. $\mathbf{P}_{\rm t}$ is a function of the policy μ_h employed in health state h. The probability matrix $\mathbf{P}_{\rm a}$ has diagonal components given by the degradation probabilities $p_H(h;q)$. Therefore, $\mathcal{T}_\epsilon(\boldsymbol{\pi}_{\rm t,0})$ and $\frac{1}{\epsilon\boldsymbol{\pi}_{\rm t,\infty}\mathbf{P}_{\rm a}\mathbf{1}}$ correspond to (10) and (18), and $\frac{\mathcal{C}_\epsilon^{\rm tot}(\boldsymbol{\pi}_{\rm t,0})}{\mathcal{T}_\epsilon(\boldsymbol{\pi}_{\rm t,0})}$ and $\boldsymbol{\pi}_{\rm t,\infty}\mathbf{c}$ correspond to (11) and (17), respectively.

■

REFERENCES

[1] N. Michelusi, L. Badia, R. Carli, L. Corradini, and M. Zorzi, "Impact of battery degradation on optimal management policies of harvesting-based wireless sensor devices," in *Proc. 2013 IEEE INFOCOM*, pp. 590–594.

[2] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102–114, Aug. 2002.

[3] J. Paradiso and T. Starner, "Energy scavenging for mobile and wireless electronics," *IEEE Pervasive Computing*, vol. 4, no. 1, pp. 18–27, 2005.

[4] D. Niyato, E. Hossain, M. Rashid, and V. Bhargava, "Wireless sensor networks with energy harvesting technologies: a game-theoretic approach to optimal energy management," *IEEE Wireless Commun.*, vol. 14, no. 4, pp. 90–96, Aug. 2007.

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

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

[7] R.-S. Liu, P. Sinha, and C. Koksal, "Joint energy management and resource allocation in rechargeable sensor networks," in *Proc. 2010 IEEE INFOCOM*, pp. 1–9.

[8] A. Seyedi and B. Sikdar, "Energy efficient transmission strategies for body sensor networks with energy harvesting," *IEEE Trans. Commun.*, vol. 58, no. 7, pp. 2116–2126, 2010.

[9] N. Michelusi, K. Stamatiou, and M. Zorzi, "Transmission policies for energy harvesting sensors with time-correlated energy supply," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2988–3001, 2013.

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

[11] N. Michelusi, K. Stamatiou, L. Badia, and M. Zorzi, "Operation policies for energy harvesting devices with imperfect state-of-charge knowledge," in *Proc. 2012 IEEE International Conference on Communications*, pp. 5782–5787.

[12] M. Gorlatova, A. Wallwater, and G. Zussman, "Networking low-power energy harvesting devices: measurements and algorithms," *IEEE Trans. Mobile Computing*, vol. 12, no. 9, pp. 1853–1865, 2013.

[13] K. Tutuncuoglu and A. Yener, "Communicating using an energy harvesting transmitter: optimum policies under energy storage losses," arXiv, vol. abs/1208.6273, 2012.

[14] R. Rao, S. Vrudhula, and D. Rakhmatov, "Battery modeling for energy aware system design," *IEEE Computer Mag.*, vol. 36, no. 12, pp. 77–87, 2003.

[15] K. Lahiri, A. Raghunathan, S. Dey, and D. Panigrahi, "Battery-driven system design: a new frontier in low power design," in *Proc. 2002 ASP Design Automation Conference*, pp. 261–267.

[16] C.-F. Chiasserini and R. Rao, "Energy efficient battery management," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 7, pp. 1235–1245, July 2001.

[17] D. Linden and T. B. Reddy, *Handbook of Batteries*, 3rd ed. McGraw Hill, 2002.

[18] Seiko Instruments Inc., Micro Battery Product Catalogue, 2011-2012. Available: www.sii.co.jp/compo/catalog/battery_en.pdf

[19] H. N. Seiger, "Effects of depth of discharge on cycle life of near term batteries," in *Proc. 1981 Intersociety Energy Conversion Engineering Conference*, p. 102.

[20] L. H. Thaller, "Expected cycle life vs. depth of discharge relationships of well behaved single cells and cell strings," NASA, Tech. Rep., 1982. Available: http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19830002288_1983002288.pdf

[21] ——, "A prediction model of the depth-of-discharge effect on the cycle life of a storage cell," NASA, Tech. Rep., 1987. Available: http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19870012878_1987012878.pdf

[22] S. Drouilhet and B. L. Johnson, "A battery life prediction method for hybrid power applications," 1997. Available: http://www.nrel.gov/docs/legosti/fy97/21978.pdf

[23] C. K. Ho, P. D. Khoa, and P. C. Ming, "Markovian models for harvested energy in wireless communications," in *Proc. 2010 IEEE International Conference on Communication Systems*, pp. 311–315.

[24] D. Bertsekas, *Dynamic Programming and Optimal Control*. Athena Scientific, 2005.

[25] B. Buchli, D. Aschwanden, and J. Beutel, "Battery state-of-charge approximation for energy harvesting embedded systems," in *Wireless Sensor Networks*, ser. Lecture Notes in Computer Science, P. Demeester, I. Moerman, and A. Terzis, Eds. Springer Berlin Heidelberg, 2013, vol. 7772, pp. 179–196.

[26] J. Marcicki, F. Todeschini, S. Onori, and M. Canova, "Nonlinear parameter estimation for capacity fade in Lithium-ion cells based on a reduced-order electrochemical model," in *Proc. 2012 American Control Conference*, pp. 572–577.

[27] E. Sondik, "The optimal control of partially observable Markov processes," Stanford University, Tech. Rep. AD0730503, May 1971.

[28] 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, 2011.

[29] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, Inc., 2006.

[30] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*. Springer, 1960.

[31] Infinite Power Solutions, THINERGY MEC201. Available: http://www.infinitepowersolutions.com/images/stories/downloads/ips_thinergy_mec201_product_data_sheet_ds1012_v1-1_final_20110920.pdf

[32] N. Jaggi, K. Kar, and A. Krishnamurthy, "Rechargeable sensor activation under temporally correlated events," *Springer Wireless Networks*, vol. 15, pp. 619–635, July 2009.

[33] R. Rajesh, V. Sharma, and P. Viswanath, "Information capacity of energy harvesting sensor nodes," in *Proc. 2011 IEEE International Symposium on Information Theory*, pp. 2363–2367.

[34] C. D. Meyer, Ed., *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics, 2000.

[35] S. J. Axler, *Linear Algebra Done Right*. Springer, 1997.

Nicolò Michelusi (S'09, M'13) received the B.Sc. (with honors), M.Sc. (with honors) and Ph.D. degrees from the University of Padova, Italy, in 2006, 2009 and 2013, respectively, and the M.Sc. degree in Telecommunications Engineering from the Technical University of Denmark in 2009, as part of the T.I.M.E. double degree program. In 2011, he was at the University of Southern California, Los Angeles, USA, and, in Fall 2012, at Aalborg University, Denmark, as a visiting research scholar. He is currently a post-doctoral research fellow at the Ming Hsieh Department of Electrical Engineering, University of Southern California, USA. His research interests lie in the areas of wireless networks, cognitive radio networks, energy harvesting systems, stochastic optimization and distributed estimation.

Dr. Michelusi serves as a reviewer for the IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON INFORMATION THEORY, IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, and IEEE/ACM TRANSACTIONS ON NETWORKING.

Leonardo Badia (S'01, M'04, SM'13) received his MS (with honors) and PhD from the University of Ferrara, Italy. After having been with the University of Ferrara and the IMT Lucca Institute for Advanced Studies, Italy, in 2011 he joined the University of Padova, Italy, where is currently Assistant Professor. His research interests include protocol design for multihop networks, cross-layer optimization, transmission protocol modeling, and applications of game theory to radio resource management. He serves on the Editorial Boards of the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE WIRELESS COMMUNICATIONS LETTERS, and the *Wiley Journal of Wireless Communications and Mobile Computing*.

Ruggero Carli (M'11) received the Laurea Degree in Computer Engineering and the Ph.D. degree in Information Engineering from the University of Padova, Padova, Italy, in 2004 and 2007, respectively. From 2008 through 2010, he was a Postdoctoral Fellow with the Department of Mechanical Engineering, University of California at Santa Barbara. He is currently an Assistant Professor with the Department of Information Engineering, University of Padova. His research interests include control theory and, in particular, control under communication constraints, cooperative control, and distributed estimation.

Dr. Carli serves as reviewer for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, IEEE TRANSACTIONS ON SIGNAL PROCESSING, *Automatica*, *Systems and Control Letters*, *SIAM Journal on Control and Optimization* and *International Journal of Robust and Nonlinear Control*.

Luca Corradini (S'06, M'09) received the Laurea degree in electronic engineering and the Ph.D. degree in industrial electronics from the University of Padova, Italy, in 2004 and 2008 respectively. From 2008 to 2011 he was a postdoctoral Research Associate at the Colorado Power Electronics Center (CoPEC), University of Colorado at Boulder, USA. Since 2011 he has been Assistant Professor at the University of Padova, Italy. His research interests in power electronics include control and modeling of digital control techniques for switched-mode power converters, and power management solutions for small-scale energy harvesting devices. Dr. Corradini received the second 2008 Prize Paper Award from the Industrial Power Converter Committee of the IEEE Industry Applications Society.

Michele Zorzi (S'89, M'95, SM'98, F'07) was born in Venice, Italy, on December 6th, 1966. He received the Laurea and the PhD degrees in Electrical Engineering from the University of Padova, Italy, in 1990 and 1994, respectively. During the Academic Year 1992/93, he was on leave at the University of California, San Diego (UCSD) as a visiting PhD student, working on multiple access in mobile radio networks. In 1993, he joined the faculty of the Dipartimento di Elettronica e Informazione, Politecnico di Milano, Italy. After spending three years with the Center for Wireless Communications at UCSD, in 1998 he joined the School of Engineering of the University of Ferrara, Italy, where he became a Professor in 2000. Since November 2003, he has been on the faculty at the Information Engineering Department of the University of Padova. His present research interests include performance evaluation in mobile communications systems, random access in mobile radio networks, ad hoc and sensor networks, energy constrained communications protocols, broadband wireless access and underwater acoustic communications and networking.

Dr. Zorzi was the Editor-In-Chief of the *IEEE Wireless Communications Magazine* from 2003 to 2005 and the Editor-In-Chief of the IEEE TRANSACTIONS ON COMMUNICATIONS from 2008 to 2011, and currently serves on the Editorial Board of the *Wiley Journal of Wireless Communications and Mobile Computing*. He was also guest editor for special issues in the *IEEE Personal Communications Magazine* (Energy Management in Personal Communications Systems) *IEEE Wireless Communications Magazine* (Cognitive Wireless Networks) and the IEEE JUORNAL ON SELECTED AREAS IN COMMUNICATIONS (Multi-media Network Radios, and Underwater Wireless Communications Networks). He served as a Member-at-large of the Board of Governors of the IEEE Communications Society from 2009 to 2011.