Impact of Battery Degradation on Optimal Management Policies of Harvesting-Based Wireless Sensor Devices

Nicolò Michelusi, Leonardo Badia, Ruggero Carli, Luca Corradini and Michele Zorzi
Department of Information Engineering, University of Padova, Italy,
{michelusi,badia,carlirug,corradini,zorzi}@dei.unipd.it

***Abstract*—Harvesting-Based Wireless Sensor Devices are increasingly being deployed in today's sensor networks, due to their demonstrated advantages in terms 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 harvesting-based battery-powered sensor devices, focusing on the impact of the battery discharge policy on the irreversible degradation of the storage capacity. A general framework based on Markov chains which captures the battery degradation process is proposed. Based on such model, it is shown that a degradation-aware policy significantly improves the lifetime of the sensor compared to "greedy" operation policies, while guaranteeing the minimum required QoS.**

***Index Terms*—Battery management; Energy harvesting; Lifetime estimation; Markov processes; Wireless sensor networks.**

I. Introduction

Recent technological advances of consumer electronics have led to the widespread diffusion of networks of miniaturized wireless devices with sensing and communication capabilities, commonly referred to as Wireless Sensor Networks (WSNs) [1]. Prolonged and unsupervised WSN operation over time, especially in case of large-scale networks composed of tens to hundreds of nodes, poses the problem of *energy autonomy* of the sensor node. While the use of non-rechargeable batteries is currently widespread for powering WSN nodes, recent advances in the field of small-scale energy harvesting will enable the sensor to absorb ambient energy from solar, mechanical, thermal or RF sources, and locally store it on an on-board rechargeable battery [2]. The energy harvesting approach, combined with intelligent battery management, is envisioned to greatly prolong the WSN operating life [3].

An energy-aware *operation policy* is an algorithm that manages the energy stored in the battery, 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 [4], delay [5] or network sum rate [6]. In this context, many works in the literature assume *ideal* batteries, by neglecting any battery degradation issues related to battery usage, *e.g.*, see [7], [8]. Some notable exceptions are [9]–[11], which attempt to model different realistic battery imperfections and non-idealities, and [12], which presents a stochastic model to capture the recovery effect of electrochemical cells, based on which efficient battery management policies are designed.

This work has been supported in part by the Department of Information Engineering at the University of Padova 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/).

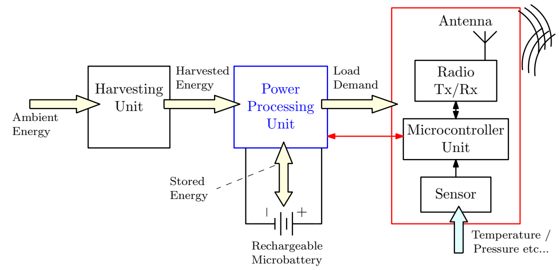

Fig. 1. Block diagram of a harvesting-based sensor node

The focus of this paper is on degradation effects causing the storage capability of the battery to diminish over time, depending on the battery usage policy. In the same spirit as [12], a Markov model which explicitly characterizes the degradation status of the battery and is suitable for policy optimization is proposed. Based on such model, battery operation policies are then designed by formulating an optimization problem which explicitly accounts for battery lifetime, while guaranteeing a minimum QoS.

The paper is organized as follows. Section II provides a system-level overview of a harvesting-based wireless sensor and outlines the fundamental issues concerning battery degradation. In Section III, we present the general framework and define the optimization problem, which is further developed in Section IV. In Section V, we provide numerical results.

II. Background

A block diagram of a harvesting-based sensor node is sketched in Fig. 1. The system *load* is the hardware that needs to be powered, *e.g.*, microcontroller, Rx/Tx transceiver and sensing unit. The *power processing unit* manages the ambient energy source and the on-board rechargeable battery to provide regulated energy to the load.

The energy source often provides most of the energy within certain periods of time, during which the on-board battery is charged. 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 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 may present different trends. Denoting with C_0 the nominal battery capacity in milliampere-hours (mAh) and with $Q(N_{\mathrm{cyc}}, D)$ the total charge delivered by the battery after N_{cyc} cycles at DoD D, one might expect

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

Two 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 unrecoverable levels [13]. Secondly, the degradation process is strongly dependent on *how* the battery is cycled, with shallow DoDs resulting in a slower degradation of C_0 and increased cycle life [13]. A microbattery rated with $N_{\rm cyc} = 100$ at 100% DoD may last up to $N_{\rm cyc} = 1000$ at 20% DoD, with roughly twice the energy extracted in the latter case [14]. A heuristic model for the $N_{\rm cyc}$ vs. D dependence is

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

where $N_{\rm cyc,0}$ represents the cycle life at 100% DoD, and α is a characteristic constant of the battery. Exponential-based models like (2) have been found to be a good fit for data from a rather wide range of battery chemistries and sizes (*e.g.*, see [15]), including also microbatteries.

III. System Model

We consider a slotted-time system, where slot k is the time interval $[kT, kT+T)$, $k \in \mathbb{Z}^+$, and T 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,[1] 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 fully usable. Let $Q_{\max}(k)$ be the usable battery capacity at time k, with $Q_{\max}(k) \le Q_{\max}(k-1)$ and $Q_{\max}(0) = q_{\max}$. Denote the (quantized) charge level of the battery at time k as Q_k. Letting $[x]^+ = \max\{x, 0\}$, the evolution of Q_k is given by

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

- $\{B_k\}$ is the energy harvesting process, taking values in $\mathcal{B} \triangleq \{0, 1, \ldots, B\}$, which models the randomness in the energy harvesting mechanism, *e.g.*, due to an intermittent energy supply. We define an underlying *energy harvesting state* process $\{S_k\}$, and we model it as an irreducible stationary Markov chain with transition probabilities $p_S(s_2|s_1) \triangleq \Pr(S_{k+1} = s_2|S_k = s_1)$ and steady state distribution $\pi_S(s)$, taking values in the 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 $p_B(b|s) \triangleq \Pr(B_k = b|S_k = s)$. Then, we denote the average harvesting rate as $\bar{b} \triangleq \sum_{s\in\mathcal{S}} \pi_S(s) \sum_{b\in\mathcal{B}} b p_B(b|s)$. 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 Section III-A, and takes values in $\mathcal{A} \triangleq \{0\} \cup \{A_{\min}, \ldots, A_{\max}\}$. $A_{\min}$ and $A_{\max}$ 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.

We model the battery degradation process as follows. We define the *battery health state*, H_k, taking values in $\mathcal{H} \equiv \{0, 1, \ldots, H_{\max}\}$. Given H_k, the battery capacity at time k is given by $Q_{\max}(k) = \left\lfloor \frac{H_k}{H_{\max}} q_{\max} \right\rfloor$, and the set of available charge levels is denoted by $\mathcal{Q}(H_k) = \{0, 1, \ldots, Q_{\max}(k)\}$. 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 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

$$p_H(h; q) \triangleq \Pr(H_{k+1} = h-1 | H_k = h, Q_k = q). \quad (3)$$

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) $p_H(h_1; q_1) \ge p_H(h_2; q_2), \ \forall h_2 \ge h_1, \ q_2 \ge q_1$.

Ass. 1.a) implies that the battery health state will eventually reach $H_k = 0$, 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 which are much longer than the cycling period; Ass. 1.c) means that the more discharged and degraded the battery, the faster the battery degradation process [13].

At time k, $\boldsymbol{Z}_k = (Q_k, H_k, S_{k-1})$ is the EHD state, taking values in the state space $\mathcal{Q} \times \mathcal{H} \times \mathcal{S}$. In practice, $\boldsymbol{Z}_k$ should be inferred and estimated from measurements of the battery state of charge, capacity, and input energy flows. For simplicity, we assume that $\boldsymbol{Z}_k$ is perfectly known to the EHD controller. Note that the harvesting state S_k is unknown at time k, as reflected by the state $\boldsymbol{Z}_k$, since B_k has not been observed yet. On the other hand, the posterior distribution of S_{k-1} can be inferred from the observed harvesting sequence $\{B_0, \ldots, B_{k-1}\}$. For example, for a solar harvesting source, we may have $\mathcal{S} = \{\text{day}, \text{night}\}$. The state $S_k \in \mathcal{S}$ may then be accurately estimated from $\frac{1}{N}\sum_{i=k-N+1}^{k} B_i$ by appropriately choosing a threshold and the window N.

A. Policy definition and Problem statement

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.*, given that $\boldsymbol{Z}_k = (Q_k, H_k, S_{k-1})$, $\mu_{H_k}(a; Q_k, S_{k-1})$ is the probability of requesting a charge quanta from the battery, when operating at health state H_k.[2] Under any policy μ, the state process $\{\boldsymbol{Z}_k\}$ is a Markov chain, so that the whole decision problem can be modeled as a Markov Decision Process [16].

[1] Note that 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 assumed constant throughout the device operating life. Modeling battery voltage dynamics is out of the scope of this paper and can be considered as a future refinement.

[2] For the sake of maximizing a long-term average reward function of the state and action processes, it is sufficient to consider only state-dependent stationary policies [16].

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

$$g(A_k, Q_k) = \begin{cases} 0, & A_k > Q_k, \\ g^*(A_k), & A_k \leq Q_k, \end{cases} \tag{4}$$

where $g^*(A_k)$ is a concave increasing function of A_k with $g^*(0) = 0$. 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

$$K_h = \min\{k \geq 0 : H_k = h\}, \ h \in \mathcal{H}. \tag{5}$$

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

$$G_\mu^{\text{tot}}(h, \boldsymbol{Z}_0) = \mathbb{E}\left[\sum_{k=K_h}^{K_{h-1}-1} g(A_k, Q_k) \,\middle|\, \boldsymbol{Z}_0\right], \tag{6}$$

$$T_\mu(h, \boldsymbol{Z}_0) = \mathbb{E}\left[K_{h-1} - K_h \mid \boldsymbol{Z}_0\right], \tag{7}$$

$$G_\mu(h, \boldsymbol{Z}_0) = \frac{G_\mu^{\text{tot}}(h, \boldsymbol{Z}_0)}{T_\mu(h, \boldsymbol{Z}_0)}. \tag{8}$$

where the expectation is taken with respect to $\{(B_k, S_k, H_k, A_k)\}$ and A_k is drawn according to μ.

Let $\mathcal{G}^* > 0$ be a minimum QoS requirement, which is met in health state h if $G_\mu(h, \boldsymbol{Z}_0) \geq \mathcal{G}^*$. We define the following.

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

$$T_\mu(\mathcal{G}^*, \boldsymbol{Z}_0) = \sum_{h \geq h_\mu^*} T_\mu(h, \boldsymbol{Z}_0), \ \text{where} \tag{9}$$

$$h_\mu^* = \max\{h : G_\mu(h, \boldsymbol{Z}_0) < \mathcal{G}^*\} + 1 \tag{10}$$

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

The condition $G_\mu(H_{\max}, \boldsymbol{Z}_0) \geq \mathcal{G}^*$ guarantees that the problem is feasible; 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 $h \geq h_\mu^*$. 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 $\sum_{h \geq h_\mu^*} G_\mu^{\text{tot}}(h, \boldsymbol{Z}_0) / \sum_{h \geq h_\mu^*} T_\mu(h, \boldsymbol{Z}_0)$. 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 QoS requirement $\mathcal{G}^* = 1.5$ and $H_{\max} = 2$ and a policy μ such that

$$G_\mu(h, \boldsymbol{Z}_0) = h, \ T_\mu(h, \boldsymbol{Z}_0) = 10^6, \ \forall h \in \{0, 1, 2\}. \tag{11}$$

Then, according to Def. 1, we have $T_\mu(\mathcal{G}^*, \boldsymbol{Z}_0) = 10^6$, since the QoS $\mathcal{G}^*$ can be supported only at health state 2. However, an *average* QoS of

$$\frac{G_\mu^{\text{tot}}(2, \boldsymbol{Z}_0) + G_\mu^{\text{tot}}(1, \boldsymbol{Z}_0)}{T_\mu(2, \boldsymbol{Z}_0) + T_\mu(1, \boldsymbol{Z}_0)} = 1.5 = \mathcal{G}^* \tag{12}$$

can be supported over a time-interval of duration 2×10^6, 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.*,

$$\mu^* = \arg\max_\mu T_\mu(\mathcal{G}^*, \boldsymbol{Z}_0) = \arg\max_\mu \sum_{h \geq h_\mu^*} T_\mu(h, \boldsymbol{Z}_0). \tag{13}$$

IV. Optimization

We develop problem (13), showing that it can be recast as an independent Linear Program (LP) on each health state, under Ass. 1.b) on $p_H(h; q)$. We give the following definition.

Definition 2. (Steady State 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[3]

$$\pi_{\mu_h}^h(q, s) = \lim_{K \to \infty} \frac{1}{K} \sum_{k=0}^{K-1} P^{(k)}(q, s | \boldsymbol{Z}_0), \tag{14}$$

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)

$$\hat{G}_{\mu_h}(h) = \sum_{(q,s) \in \mathcal{Q}(h) \times \mathcal{S}} \pi_{\mu_h}^h(q, s) \mathbb{E}_{\mu_h(\cdot; q, s)}\left[g(A, q)\right], \tag{15}$$

$$\hat{T}_{\mu_h}(h) = \left(\sum_{(q,s) \in \mathcal{Q}(h) \times \mathcal{S}} \pi_{\mu_h}^h(q, s) p_H(h; q)\right)^{-1}, \tag{16}$$

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), induced by policy μ_h.

Remark: Note that $\pi_{\mu_h}^h$ in (14) 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 (16) is the actual degradation probability. $\hat{G}_{\mu_h}(h)$ can be interpreted as the long-term average reward per time-slot in health state h, whereas $\hat{T}_{\mu_h}(h)^{-1}$ can be interpreted as the long-term average probability of degradation to the lower health state $h-1$. Such observations are formalized in the following lemma, whose proof is provided in [18].

[3] We assume that μ_h induces a Markov chain with a single closed communicating class, so that $\pi_{\mu_h}^h(q, s)$ is independent of $\boldsymbol{Z}_0$ [17].

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

$$G_\mu(h, \boldsymbol{Z}_0) = \hat{G}_{\mu_h}(h) + \mathcal{O}(p_H^*(h)), \tag{17}$$

$$T_\mu(h, \boldsymbol{Z}_0) = \hat{T}_{\mu_h}(h) + \mathcal{O}(1), \tag{18}$$

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

Lemma 1 states that, 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 $\max_q p_H(h;q) \to 0$, (18) is a good approximation. On the other hand, the average reward 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 (14) induced by policy μ_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, and substitute (18) in (9), yielding

$$T_\mu(\mathcal{G}^*, \boldsymbol{Z}_0) \simeq \sum_{h \geq h_\mu^*} \hat{T}_{\mu_h}(h), \tag{19}$$

where $h_\mu^* \simeq \max\{h : \hat{G}_{\mu_h}(h) < \mathcal{G}^*\} + 1$ from (10) and (17). Finally, substituting (19) in (13), we obtain the approximation

$$\mu^* = \arg\max_\mu \sum_{h \geq h_\mu^*} \hat{T}_{\mu_h}(h). \tag{20}$$

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, (20) 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`

$$\mu_h^* = \arg\min_{\mu_h} \sum_{(q,s)\in\mathcal{Q}(h)\times\mathcal{S}} \pi_{\mu_h}^h(q,s) p_H(h;q) \tag{21}$$

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

If the problem is infeasible, set `REP=false`, $h_{\mu^*}^* = h + 1$. If it is feasible and $h = 1$, set $h_{\mu^*}^* = 1$. Otherwise, update $h := h - 1$. `END WHILE`
3) `RETURN` the optimal policy $\mu^* = (\mu_h^*)_{h \geq h_{\mu^*}^*}$.

Remark: Step 2) is equivalent to

$$\mu_h^* = \arg\max_{\mu_h} \hat{T}_{\mu_h}(h), \quad \text{s.t. } \hat{G}_{\mu_h}(h) \geq \mathcal{G}^*, \tag{22}$$

and is obtained by substituting the expressions of $\hat{T}_{\mu_h}(h)$ and $\hat{G}_{\mu_h}(h)$ (see Def. 3) in (22), which can be solved numerically with standard tools [16]. Thus, the optimal policy μ_h^* maximizes the lifetime (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. Step 2) also determines $h_{\mu^*}^*$ in (10), for the optimal policy μ^*. In step 3) the optimal policy is found by concatenating the sub-policies $\{\mu_h^*\}$, and the corresponding lifetime (1) is computed using (18). The main advantage of this approach over a standard approach which solves the original optimization problem (13) jointly is that (13) is decomposed into a sequence of independent sub-problems (21) for each health state h, thus reducing the overall computational complexity.

V. Numerical Results

We consider a battery with capacity $q_{\max} = 500$ charge levels and $H_{\max} = 50$ health states. The battery degradation probabilities $p_H(h;q)$ can be extrapolated from manufacturer-provided data by employing the deterministic, continuous time model (2). With this approach, described in detail in [18], we have found that a good match is given by

$$p_H(h;q) = \gamma \exp\left\{\alpha\left(1 - \frac{q}{q_{\max}}\right)\right\}, \tag{23}$$

where γ is a dimensionless constant. Note that γ does not affect the solution of (21). Therefore, we choose a small value $\gamma = 2.5 \cdot 10^{-5}$ so as to satisfy Lemma 1. The parameter α in (23) is obtained by interpolating the data-sheet values in [14] of battery type MS920SE, a Li-Ion rechargeable micro battery, which may be envisioned for applications in WSNs. A good fit is obtained with $\alpha \simeq 2.88$.

The underlying energy harvesting process $\{S_k\}$ is modeled as a two state Markov chain with state space $\mathcal{S} = \{\mathrm{G}, \mathrm{B}\}$ 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" harvesting states, respectively. In the "bad" state ($S_k = \mathrm{B}$), no energy is harvested, *i.e.*, $B_k = 0$; in the "good" state ($S_k = \mathrm{G}$), 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. The action space is $\mathcal{A} = \{0, \ldots, 20\}$. We consider the following policies:

- *Lifetime Unaware Policy* (LUP), which greedily maximizes the average long-term reward (15) 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 [16] as the solution of
$$\mu_h^* = \arg\max_{\mu_h} \hat{G}_{\mu_h}(h), \ \forall h \in \mathcal{H}; \tag{24}$$
- *Lifetime Aware Optimal Policy* (LAOP), which solves problem (13) via Algorithm 1.

In the following plots, for a given policy and QoS $\mathcal{G}^*$, the battery lifetime is computed according to (9), using standard results on absorbing Markov Chains, see [17]. The corresponding *minimum reward*[4] 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^* is defined in (10).

In Fig. 2, we plot the minimum reward versus the 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 and Ass. 1.c). We note that, for a

[4]The minimum reward represents the average reward (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.

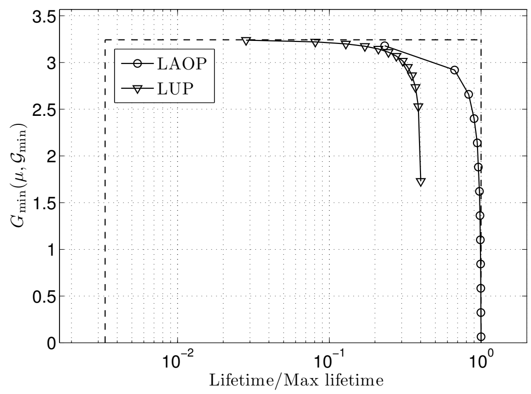

Fig. 2. Minimum reward over the battery lifetime versus normalized lifetime. The dashed lines represent the minimum and maximum lifetime and the maximum reward $\max_{\mu_{H_{\max}}} \hat{G}_{\mu_{H_{\max}}}(H_{\max})$.

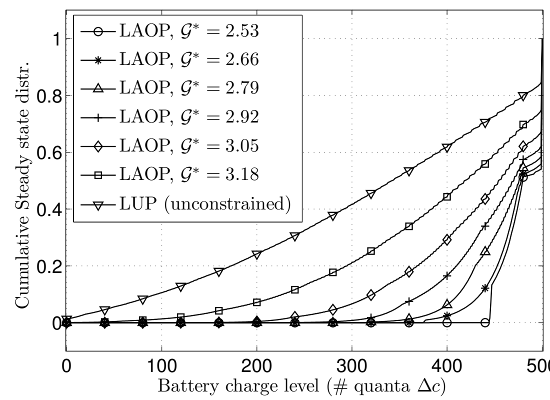

Fig. 3. Cumulative steady state distribution of charge levels at the maximum health state $H_{\max}$.

given minimum guaranteed QoS, LAOP achieves a significant gain in terms of battery lifetime with respect to the "greedy" policy LUP, which neglects battery degradation mechanisms. In particular, the lifetime is increased by a factor ~ 2.5. 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 energy shortage, which inevitably shortens the battery lifetime.

Finally, in Fig. 3, we plot the cumulative steady state distribution of the charge levels, for the maximum health state $H_{\max}$. We note that the steady state distributions of LUP, which does not take into account the ongoing battery degradation mechanisms, are spread over all the battery charge levels. In particular, this policy operates for a significant amount of time at low charge levels, thus inducing a fast battery degradation. Conversely, LAOP spreads the steady state distribution over the upper charge levels only, and never discharges below a QoS-dependent charge level.

VI. Conclusions

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. 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, which 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 has the potential to significantly improve the lifetime of the sensor node with respect to a "greedy" operation policy, while guaranteeing the minimum required QoS.

References

[1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “A survey on sensor networks,” *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, Aug. 2002.

[2] J. A. Paradiso, T. Starner, “Energy scavenging for mobile and wireless electronics,” *IEEE Perv. Computing*, vol. 4, pp. 18–27, Jan. 2005.

[3] 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 Communications*, vol. 14, no. 4, pp. 90–96, Aug. 2007.

[4] M. Gatzianas, L. Georgiadis, and L. Tassiulas, “Control of wireless networks with rechargeable batteries,” *IEEE Transactions on Wireless Communications*, vol. 9, pp. 581–593, Feb. 2010.

[5] V. Sharma, U. Mukherji, V. Joseph, and S. Gupta, “Optimal energy management policies for energy harvesting sensor nodes,” *IEEE Transactions on Wireless Communications*, vol. 9, pp. 1326–1336, Apr. 2010.

[6] R.-S. Liu, P. Sinha, and C. Koksal, “Joint Energy Management and Resource Allocation in Rechargeable Sensor Networks,” in *Proceedings of IEEE INFOCOM*, Mar. 2010, pp. 1–9.

[7] A. Seyedi and B. Sikdar, “Energy efficient transmission strategies for body sensor networks with energy harvesting,” *IEEE Transactions on Communications*, vol. 58, pp. 2116–2126, July 2010.

[8] N. Michelusi, K. Stamatiou, and M. Zorzi, “Performance Analysis of Energy Harvesting Sensors with Time-Correlated Energy Supply,” in *50th Allerton Conf. on Comm., Control, and Computing*, Sep. 2012.

[9] B. Devillers and D. Gunduz, “Energy harvesting communication system with battery constraint and leakage,” in *IEEE GLOBECOM Workshops*, Dec. 2011, pp. 383–388.

[10] N. Michelusi, L. Badia, R. Carli, K. Stamatiou, and M. Zorzi, “Correlated energy generation and imperfect State-of-Charge knowledge in energy harvesting devices,” in *8th International Wireless Communications and Mobile Computing Conference (IWCMC)*, Aug. 2012, pp. 401–406.

[11] M. Gorlatova, A. Wallwater, and G. Zussman, “Networking Low-Power Energy Harvesting Devices: Measurements and Algorithms,” *IEEE Transactions on Mobile Computing*, no. 99, p. 1, 2012.

[12] C.-F. Chiasserini and R. Rao, “Energy efficient battery management,” *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 7, pp. 1235–1245, July 2001.

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

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

[15] S. Drouilhet and B. L. Johnson, “A Battery Life Prediction Method for Hybrid Power Applications,” 1997. [Online]. Available: http://www.nrel.gov/docs/legosti/fy97/21978.pdf

[16] D. Bertsekas, *Dynamic programming and optimal control*. Athena Scientific, Belmont, Massachusetts, 2005.

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

[18] N. Michelusi, L. Badia, R. Carli, L. Corradini, and M. Zorzi, “Energy Management Policies for Harvesting-based Wireless Sensor Devices with Battery Degradation,” 2013, submitted to IEEE Transactions on Communications.