Correlated Energy Generation and Imperfect State-of-Charge Knowledge in Energy Harvesting Devices

Nicolò Michelusi, Leonardo Badia, Ruggero Carli, Kostas Stamatiou, and Michele Zorzi
Department of Information Engineering, University of Padova, via Gradenigo 6B, 35131 Padova, Italy
{michelusi,badia,carli,kstamat,zorzi}@dei.unipd.it

***Abstract*—Nowadays, many devices in wireless sensor networks are provided with energy harvesting capability to allow for their continuous operation over long periods of time. In principle, the energy level within each sensor should be managed optimally to ensure the best performance. Network engineers, however, often consider optimality under the idealized assumption of perfect knowledge about the State-of-Charge (SOC) of the device. This information is not always realistic or accurate. In our previous work [1], we showed that optimal policies for sensing, transmission, and battery usage should rather consider uncertainty on the SOC of the device. In this paper, we extend that investigation, therein performed in the idealized scenario of i.i.d. energy arrivals, by considering a correlated energy generation process. We show that the knowledge of the SOC and that of the energy generation process are useful in a complementary manner, that is they can be traded for each other. Moreover, the knowledge on the state of the energy generation process can obviate the need for acquiring accurate SOC information. This investigation paves the road for a new line of research in wireless sensor networks, allowing a tighter interaction between the designers of energy harvesting and battery storage mechanisms on the one hand, and the engineers of network operation and control policies on the other.**

I. Introduction

Energy Harvesting Devices (EHDs) collect energy from the surrounding environment thanks to a number of harvesting techniques, including solar, motion, heat, aeolian [2], [3], and can thus be used in Wireless Sensor Networks (WSNs) to achieve autonomous network operation for long periods of time [4], [5]. However, a proper management of both tasks of continuous operation and prompt sensing is required [6]–[10].

From the modeling standpoint, EHDs can be thought of as *energy buffers*, where energy, generated according to a given statistical process, is stored and from where it is drawn by a controller to feed sensor microprocessors and transceiver equipment, whenever needed. Hence, an energy management algorithm decides on the most suitable device operation according to the energy level of the battery, commonly referred to as State-of-Charge (SOC) [7], [8]. In practice, EHDs store energy in electrochemical rechargeable batteries and/or super-capacitors. Thus, it may be questionable to assume that the SOC of these devices can be characterized with infinite precision and immediate availability at any time. For example, [3] states that a super-capacitor capacitance can fluctuate by approximately 30% with respect to the data sheet value. Moreover, it is possible to estimate the SOC online, but this comes at the price of an additional energy loss. Other algorithms [11] have been proposed to estimate the open circuit voltage, which is closely related to the SOC, but have non-negligible complexity, whose impact should be carefully evaluated, especially for resource-limited small devices. In general, we can conclude that the SOC estimation for electrochemical batteries is a non-trivial task and a precise knowledge of the SOC may be difficult.

For this reason, we claim that an energy management policy should better focus on scenarios where SOC knowledge is, if not totally unavailable, at least imperfect. In our previous work [1], the imperfect SOC knowledge is modeled by quantizing the SOC to a finite number of regions. Therein, optimal energy management policies are investigated, under the simplifying assumption that the energy arrival process is independent and identically distributed (i.i.d.) over time. In this paper, we aim at extending this evaluation by introducing correlation in the energy harvesting mechanism, modeling the energy arrival process as a hidden Markov chain. Especially, we envision that gaining knowledge on the statistics of the energy generation process and its correlation structure may be extremely beneficial for finding an efficient energy management policy. In other words, estimating that the EHD is undergoing a phase of correlated energy generation, which can be easily performed by tracking the history of the energy generation process via some supervised training [6], [9], can obviate the need for acquiring accurate SOC information. In this regard, we show that the knowledge of the SOC and that of the energy generation process are useful in a complementary manner, that is they can be traded for each other. In particular, the effect of having no information about the current SOC level but knowing the energy generation state is almost exactly equivalent to knowing the SOC perfectly, but neglecting the state of the energy harvesting mechanism.

The rest of this paper is organized as follows. In Section II, we describe the system model. In Section III, we present the optimization framework for the evaluation of the optimal resource management policy, which depends on the available information on the SOC and on the energy generation process. In Section IV, we present some numerical results. Finally, in Section V, we conclude the paper by highlighting possible further developments.

II. System model

We consider an Energy Harvesting Device (EHD), which scavenges energy from the environment (*e.g.*, solar, kinetic, wind, RF). We assume a slotted-time system, where time-slot $k \geq 0$ corresponds to the time interval $[kD_{\text{ts}}, (k+1)D_{\text{ts}})$, and D_{ts} is the time-slot duration. The harvested energy is stored in an energy buffer, in the form of energy quanta, whose capacity is denoted by $e_{\max}$ (quanta). The energy available in the buffer at time kD_{ts}, *i.e.*, the State-of-Charge (SOC), is denoted by $E_k \in \mathcal{E}$, taking values in the set $\mathcal{E} \equiv \{0, \ldots, e_{\max}\}$. At the beginning of the kth time-slot, the EHD controller requests a number of energy quanta $Q_k \in \mathcal{Q}$ to be drawn from the buffer to perform a certain task, chosen from the *action space* $\mathcal{Q} = \{0, \ldots, q_{\max}\}$, where $0 < q_{\max} \leq e_{\max}$. During the time-slot duration, the EHD harvests $B_k \in \mathcal{B}$ energy quanta from the environment, which are stored in the buffer, where $\mathcal{B} = \{0, 1, \ldots, b_{\max}\}$ is the set of energy arrival values. Starting from the initial SOC level $E_0 \in \mathcal{E}$ available at time 0, the temporal evolution of the SOC E_k is governed by

$$E_{k+1} = \min\left\{[E_k - Q_k]^+ + B_k, e_{\max}\right\}, \quad k \geq 0, \tag{1}$$

where $[\cdot]^+ \triangleq \max\{\cdot, 0\}$.

The energy harvesting/consumption mechanism described by (1) entails two important phenomena. The first, referred to as *energy outage*, corresponds to the EHD running out of energy before the completion of the requested task, which happens when $Q_k > E_k$. In this case, the task cannot be completed, and the battery is depleted. If perfect knowledge of E_k is available at the EHD controller, then outage can always be avoided by choosing $Q_k \leq E_k$. However, herein we assume that the EHD controller is provided with imperfect knowledge of E_k, hence energy outage is possible, since the controller may attempt to draw more energy than what is actually available in the buffer. Alternatively, *energy overflow* may occur if $B_k > e_{\max} - [E_k - Q_k]^+$, *i.e.*, the energy buffer is unable to store all of the harvested energy B_k, due to the limited battery capacity. In this case, some of the energy is lost, hence, it cannot be used in the future for reward accrual.

We model the energy arrival process $\{B_k\}$ as a stationary hidden Markov process, taking values in the set $\mathcal{B}$. We define S_k as the (hidden) state of the energy arrival process at time-slot k. The process of the energy arrival states, $\{S_k\}$, is modeled as a stationary irreducible Markov chain taking values in the finite set $\mathcal{S}$, with transition probabilities $p_S(s_{k+1}|s_k) \triangleq \Pr(S_{k+1} = s_{k+1}|S_k = s_k)$, $s_k, s_{k+1} \in \mathcal{S}$. Given the energy arrival state $S_k = s$, the energy harvest B_k is drawn with probability mass function $p_B(b|s) \triangleq \Pr(B_k = b|S_k = s), b \in \mathcal{B}, s \in \mathcal{S}$. We define $\pi_S(s),\ s \in \mathcal{S}$, as the steady state distribution of the energy arrival states. We refer to $\bar{b} = \mathbb{E}[B_k] = \sum_{s \in \mathcal{S}} \pi_S(s) \sum_{b \in \mathcal{B}} b p_B(b|s)$ as the *average harvesting rate*.

We assume that, at the beginning of the kth time-slot, *i.e.*, at time kD_{ts}, the EHD controller knows the energy arrival state S_{k-1}. Notice that S_{k-1} can be estimated by measuring the past energy arrivals $\{B_j, j \leq k-1\}$ available up to time kD_{ts}. In fact, the posterior distribution of state S_{k-1} can be inferred recursively as

$$\Pr(S_{k-1} = s|B_0, \ldots, B_{k-1}) = \frac{p_B(B_{k-1}|s) \sum_{\sigma \in \mathcal{S}} p_S(s|\sigma) \Pr(S_{k-2} = \sigma|B_0, \ldots, B_{k-2})}{\sum_{\tilde{s} \in \mathcal{S}} p_B(B_{k-1}|\tilde{s}) \sum_{\sigma \in \mathcal{S}} p_S(\tilde{s}|\sigma) \Pr(S_{k-2} = \sigma|B_0, \ldots, B_{k-2})}, \tag{2}$$

where $\Pr(S_{k-2} = \sigma|B_0, \ldots, B_{k-2})$ is the posterior distribution inferred in the previous time-slot. Then, the current state S_k can be estimated by using, *e.g.*, Maximum A Posteriori (MAP) [12], given by $\hat{S}_{k-1} = \arg\max_{s \in \mathcal{S}} \Pr(S_{k-1} = s|B_0, \ldots, B_{k-1})$. For simplicity, we assume that S_{k-1} is estimated without error, *i.e.*, $\hat{S}_{k-1} = S_{k-1}$, hence perfect knowledge of S_{k-1} is available at the EHD controller at time kD_{ts}. The validity of such choice will be further discussed in the numerical results, Section IV. Conversely, we assume that S_k is unknown at time kD_{ts}, since the energy arrival B_k, which provides information about S_k through its distribution $p_B(B_k|S_k)$, is unknown at time kD_{ts}.

We assume that only partial knowledge of the SOC E_k is available, *e.g.*, due to uncertainty in its estimation. As in [1], we model the uncertainty on the SOC E_k by defining a partition of the SOC space, $\{\mathcal{I}(n), n = 0, \ldots, \tilde{n}-1\}$, where $\mathcal{I}(n) = \{\tilde{e}_n, \ldots, \tilde{e}_{n+1} - 1\}$ is the nth SOC interval, $n \in \{0, \ldots, \tilde{n}-1\}$, and $0 = \tilde{e}_0 < \tilde{e}_1 < \cdots < \tilde{e}_{\tilde{n}} = e_{\max} + 1$ define the interval boundaries. Suppose that, at time-slot k, $E_k \in \mathcal{I}(N_k)$, for some $N_k \in \{0, \ldots, \tilde{n}-1\}$. We assume that the EHD controller knows only the interval index N_k, *i.e.*, it knows that $E_k \in \mathcal{I}(N_k)$, rather than the exact SOC E_k. We define the *interval index* process $\{N_k, k \geq 0\}$, taking values in $\{0, \ldots, \tilde{n}-1\}$. The special case with perfect SOC knowledge is obtained by letting $\tilde{n} = e_{\max} + 1$, hence $E_k = N_k$.

Note that keeping track of the energy arrivals $\{B_j,\ j = 0, \ldots, k-1\}$ can, to some extent, provide knowledge of the SOC E_k. However, as the SOC is quantized, and since the battery is subject to non-idealities, the SOC measurement cannot be indefinitely precise; yet, keeping track of the energy arrivals can serve to identify whether the SOC state is generally HIGH or LOW. Instead, the current energy arrival state is much easier to acquire.

A. Policy definition and problem statement

At time kD_{ts}, the EHD controller knows the interval index N_k of the SOC E_k, the energy arrival state of the previous time-slot S_{k-1} (we recall that S_k is assumed to be unknown at time kD_{ts}) and the history $\mathcal{H}_k = \{(N_0, S_{-1}, Q_0), \ldots, (N_{k-1}, S_{k-2}, Q_{k-1}), \text{past outage events}\}$. Then, given $(N_k, S_{k-1}, \mathcal{H}_k)$, a control *policy* μ decides on the amount of energy Q_k to be requested from the buffer. In particular, $\mu(q; (N_k, S_{k-1}, \mathcal{H}_k))$ is the probability that action $Q_k = q \in \mathcal{Q}$ is chosen in time-slot k.

We define the reward function $g : \mathcal{Q} \times \mathcal{E} \mapsto \mathbb{R}^+$, where $g(Q_k, E_k)$ is the reward accrued in time-slot k, when the SOC level is $E_k \in \mathcal{E}$ and action Q_k is chosen, as

$$g(Q_k, E_k) = \begin{cases} 0 & Q_k > E_k \\ \tilde{g}(Q_k) & Q_k \leq E_k, \end{cases} \tag{3}$$

where $\tilde{g} : \mathcal{Q} \mapsto \mathbb{R}^+$ is a concave increasing function of Q_k, with $\tilde{g}(0) = 0$. Notice that, if $Q_k > E_k$, then $g(Q_k, E_k) = 0$, which models an energy outage event.

We define the *long-term average reward per time-slot* under policy μ, starting from the initial state $E_0{=}e_0, S_{-1}{=}s_{-1}$, as

$$G(\mu; e_0, s_{-1}) \triangleq \lim_{K\to\infty} \inf \frac{1}{K}\mathbb{E}\left[\sum_{k=0}^{K-1} g(Q_k, E_k) \middle| \begin{matrix} E_0 = e_0, \\ S_{-1} = s_{-1} \end{matrix}\right], \tag{4}$$

where the expectation is computed with respect to the random variables $\{B_k, S_k, Q_k, k = 0, \ldots, K-1\}$.

In this paper, we consider the following optimization problem, subject to imperfect knowledge of the SOC and, possibly, of the state of the energy generation process:

$$\mu^* = \arg\max_{\mu} G(\mu; e_0, s_{-1}). \tag{5}$$

Its numerical optimization is carried out in the next Section.

III. Optimization

Due to the partial knowledge of the SOC, (5) can be recast under the framework of the Partially Observable Markov Decision Processes (POMDPs) [13], and can be solved by using numerical optimization tools available in the literature [14]. However, due to the limited processing capability that typically characterizes practical EHDs, in this paper we focus on suboptimal policies, which neglect the history $\mathcal{H}_k$ available at time-slot k. Therefore, $\mu(q; N_k, S_{k-1})$ is the probability that the EHD controller decides on action q, given that $E_k \in \mathcal{I}(N_k)$ and the state of the energy arrival process was S_{k-1} in the previous time-slot.

In the following subsections, we distinguish between the cases with perfect (*i.e.*, $\tilde{n} = e_{\max} + 1$) and imperfect SOC knowledge. Moreover, we refer to the case where perfect knowledge of the energy generation process is available at the EHD controller (*i.e.*, knowledge of the energy harvesting state S_{k-1}). The case where the Markov structure of the energy arrivals is neglected, hence the energy arrivals are treated as i.i.d., is obtained in the following optimization by replacing the state space of the energy arrival process $\mathcal{S}$ with $\tilde{\mathcal{S}} = \{1\}$, its transition matrix $\mathbf{P}_S$ with $\tilde{\mathbf{P}}_S = 1$, and the energy arrival distribution $p_B(b|s)$ with the marginal $\tilde{p}_B(b|s) = p_B(b) = \sum_{s\in\mathcal{S}} p_B(b|s)\pi_S(s)$.

A. Optimization with perfect SOC knowledge

When perfect SOC knowledge is available at the EHD controller, the policy μ maps the state of the system (E_k, S_{k-1}) to the probability of drawing q energy quanta from the buffer. The sequence $\{(E_k, S_{k-1}, Q_k), k \geq 0\}$ constitutes a Markov Decision Process (MDP), and the long term reward is maximized by a stationary, deterministic policy [15]. In this case, the optimal policy is found by using standard tools, such as policy iteration. Note that the long-term reward under perfect SOC knowledge represents an upper bound to the performance of any policy under SOC uncertainty.

B. Optimization with SOC uncertainty

Under SOC uncertainty, the sequence $\{(N_k, S_{k-1}, Q_k), k \geq 0\}$ does not constitute an MDP, hence the optimal policy μ^* cannot be found via the policy iteration algorithm. In order to reduce the complexity, we consider only the set of deterministic policies. To this end, we define the function $\rho : \{0, \ldots, \tilde{n}-1\} \times \mathcal{S} \mapsto \mathcal{Q}$, which maps the state pair (N_k, S_{k-1}) to the action $Q_k = \rho(N_k, S_{k-1})$. Then, letting μ_ρ be the deterministic policy associated with the action mapping ρ, defined as $\mu_\rho(\rho(n, s); n, s) = 1$, $\mu_\rho(q; n, s) = 0 \ \forall \ q \neq \rho(n, s)$, (5) reduces to

$$\rho^* = \arg\max_{\rho} G(\mu_\rho; e_0, s_{-1}). \tag{6}$$

where, from (4),

$$G(\mu_\rho; e_0, s_{-1}) = \sum_{n=0}^{\tilde{n}-1} \sum_{e\in\mathcal{I}(n)} \sum_{s\in\mathcal{S}} \pi_\rho(e, s; e_0, s_{-1}) g(\rho(n, s), e).$$

The term $\pi_\rho(e, s; e_0, s_{-1})$ is the asymptotic distribution of the state pair $(E_k, S_{k-1}) = (e, s) \in \mathcal{E} \times \mathcal{S}$, given that the initial state is $(E_0, S_{-1}) = (e_0, s_{-1})$, and is defined as

$$\pi_\rho(e, s; e_0, s_{-1}) = \tag{7}$$
$$\lim_{K\to+\infty} \frac{1}{K} \sum_{k=0}^{K-1} \mathrm{Pr}_\rho\left(E_k = e, S_{k-1} = s | E_0 = e_0, S_{-1} = s_{-1}\right),$$

where $\mathrm{Pr}_\rho\left(E_k = e, S_{k-1} = s | E_0 = e_0, S_{-1} = s_{-1}\right)$ is the k-step transition probability of the chain under the action mapping ρ. In most practical cases of interest, as well as in the cases we consider in this paper, $\pi_\rho(e, s; e_0, s_{-1})$ is unique [16] and is independent of the initial condition $(E_0, S_{-1}) = (e_0, s_{-1})$. It follows that the long term reward $G(\mu_\rho; e_0, s_{-1})$ is independent of (e_0, s_{-1}). In the following treatment, for notational convenience, we thus neglect the dependence of $G(\mu_\rho; e_0, s_{-1})$ on (e_0, s_{-1}).

Unlike [1], where an exhaustive search is used to solve (6), herein we resort to local search methods. In fact, the additional energy arrival state $S_k \in \mathcal{S}$ yields an exponential increase in the cardinality of the set of action mappings over which the reward is optimized, given by $|\{\rho : \{0, \ldots, \tilde{n}-1\} \times \mathcal{S} \mapsto \mathcal{Q}\}| = |\mathcal{Q}|^{\tilde{n}|\mathcal{S}|}$, rendering an exhaustive search method impractical.

The proposed local search algorithm is defined as follows:

A.1 (Local Search Algorithm)

1) *Let $\rho^{(0)} : \{0, \ldots, \tilde{n}-1\} \times \mathcal{S} \mapsto \mathcal{Q}$ be an initial mapping, and set the counter $i = 0$.*
2) *In stage i:*
   - *Set $\rho^{(i+1)} = \rho^{(i)}$.*
   - *Update the mapping $\rho^{(i+1)}$ sequentially, for $n = 0, \ldots, \tilde{n}-1$ and $s \in \mathcal{S}$, as*

$$\rho^{(i+1)}(n, s) := \arg\max_{\rho^{(i+1)}(n,s)\in\mathcal{Q}} G\left(\mu_{\rho^{(i+1)}}\right). \tag{8}$$

- *If $G\left(\mu_{\rho^{(i+1)}}\right)=G\left(\mu_{\rho^{(i)}}\right)$, return the policy $\rho^{(i+1)}$. Else, update the counter as $i:=i+1$ and repeat from 2).*

This algorithm determines a local optimum of (6) by sequentially optimizing the action performed on each pair (n,s), until convergence to a local maximum.[1] If $G\left(\mu_{\rho^{(i+1)}}\right)=G\left(\mu_{\rho^{(i)}}\right)$ for some $i\geq 0$, then a local maximum of the reward is found, since any unilateral change in the policy $\rho^{(i)}$ does not lead to an improved reward. The algorithm is thus terminated. The convergence of the algorithm is thus guaranteed within finite time, since the set of deterministic policies is finite, hence the local optimum is obtained within at most $|\mathcal{Q}|^{\tilde{n}|\mathcal{S}|}$ evaluations of the reward $G(\mu_\rho;e_0,s_{-1})$. However, convergence to the global optimum is not guaranteed, unless the reward function $G(\mu_\rho)$ is convex in ρ. A deeper discussion of this point is out of the scope of this paper and is left for future investigation.

IV. Numerical Results

In this section, we present numerical results for the optimal policy. We consider an energy arrival process $\{B_k\}$ with average harvesting rate $\bar{b}=10$, with 3 hidden states (state space $\mathcal{S}\equiv\{1,2,3\}$), which are defined as follows.

- In state 1, the arrival process follows a geometric distribution $p_B(b|1)=\kappa(\gamma)e^{-\gamma b}$, truncated at $b_{\max}=4\bar{b}$, with mean $\mathbb{E}[B_k|S_k=1]=\sum_{b=0}^{b_{\max}}bp_B(b|1)=\bar{b}$, where γ is the decay rate and $\kappa(\gamma)=\left(\sum_{b=0}^{b_{\max}}e^{-\gamma b}\right)^{-1}$ is the normalization factor. By imposing the constraint $\mathbb{E}[B_k|S_k=1]=\bar{b}$, γ is the unique solution of $\sum_{b=0}^{b_{\max}}(b-\bar{b})e^{-\gamma b}=0$.
- In state 2, B_k is deterministically equal to 0, *i.e.*, $p_B(0|2)=1$, $p_B(b|2)=0,\ \forall\ b>0$.
- Finally, in state 3, B_k is deterministically equal to $2\bar{b}$, *i.e.*, $p_B(2\bar{b}|3)=1$, $p_B(b|3)=0,\ \forall\ b\neq 2\bar{b}$.

The transition probabilities between the energy arrival states are defined via the $|\mathcal{S}|\times|\mathcal{S}|$ transition matrix $\mathbf{P}_S$, with entries $[\mathbf{P}_S]_{s_0,s_1}=\Pr(S_k=s_1|S_{k-1}=s_0)\triangleq p_S(s_1|s_0)$ given by

$$\mathbf{P}_S=\begin{bmatrix} x & \frac{1-x}{2} & \frac{1-x}{2} \\ 1-y & y & 0 \\ 1-y & 0 & y \end{bmatrix}, \tag{9}$$

where $x\triangleq p_S(1|1)$, $y\triangleq p_S(2|2)=p_S(3|3)$. For example, in the case of solar energy harvesting, this choice of the energy generation process can be interpreted as follows. Under direct sunlight, the arrival process can be approximated as being deterministic (state $s=3$). Conversely, in the dark, no energy is harvested, hence the arrival process is deterministic with value 0 (state $s=2$). State $s=1$ models a transient situation between the two states of direct light (3) and dark (2), where the arrival process exhibits a random behavior. We refer to [17] for a deeper discussion on energy generation models for solar and piezoelectric sources, based on empirical analysis.

[1]However, notice that the concept of local maximum is not well defined for a function with discrete inputs, and is used only as an approximation here. In fact, the reward function (4) maps discrete policies, hence it is not continuous

The steady state distribution of the energy arrival states is then given by

$$\pi_S(1)=\frac{1-y}{2-y-x}, \tag{10}$$
$$\pi_S(2)=\frac{1}{2}\frac{1-x}{2-y-x}, \tag{11}$$
$$\pi_S(3)=\frac{1}{2}\frac{1-x}{2-y-x}. \tag{12}$$

It is worth noting that the i.i.d. energy arrival scenario considered in [1] is obtained, $\forall\ y\in(0,1)$, by letting $x\to 1$. In this case, state 1 absorbs the Markov chain $\{S_k\}$, hence the energy arrival process $\{B_k\}$ exhibits an i.i.d. behavior with probability mass function $p_B(b|1),\ b\in\mathcal{B}$.

Note that, with this choice of the energy generation process, the state S_{k-1} can be estimated, based on B_{k-1} only, as follows:

- If $B_{k-1}=0$, then $\hat{S}_{k-1}=2$,
- If $B_{k-1}=2\bar{b}$, then $\hat{S}_{k-1}=3$,
- Otherwise ($B_{k-1}\neq 0$ and $B_{k-1}\neq 2\bar{b}$), $\hat{S}_{k-1}=1$.

Then, since in states $S_{k-1}=2$ and $S_{k-1}=3$ the arrival B_{k-1} is deterministic, a mis-detection error occurs if and only if $S_{k-1}=1$ and $B_{k-1}\in\{0,2\bar{b}\}$, in which case the state is mis-detected as $\hat{S}_{k-1}=2$ and $\hat{S}_{k-1}=3$ for $B_{k-1}=0$ and $B_{k-1}=2\bar{b}$, respectively. The probability of mis-detection error is then given by

$$\begin{aligned}\Pr\left(\hat{S}_{k-1}\neq S_{k-1}\right)&=\Pr\left(S_{k-1}=1,B_{k-1}\in\{0,2\bar{b}\}\right)\\&=\pi_S(1)\kappa(\gamma)\left[1+e^{-2\gamma\bar{b}}\right].\end{aligned} \tag{13}$$

As an example, if $x=y=0.95$, we obtain $\Pr\left(\hat{S}_{k-1}\neq S_{k-1}\right)\simeq 0.05$, so that the assumption of perfect knowledge of S_{k-1} is plausible.

We consider the reward function

$$\tilde{g}(q)=\frac{\ln(1+\alpha q)}{\ln(1+\alpha\bar{b})}, \tag{14}$$

which represents a normalized throughput, where the parameter α is an SNR scaling factor.

We define the *balanced policy* as the policy which, in every time-slot, attempts to draw $q=\bar{b}$ energy quanta from the buffer. Notice that this policy does not require knowledge of the SOC E_k nor of the energy arrival state S_{k-1}. Hence, we expect that any optimized policy with SOC uncertainty or which neglects the Markov structure of the energy arrival process outperforms the balanced policy, for any finite value of the battery capacity $e_{\max}$.

Similarly to [1], it can be shown that the long-term reward $G(\mu)$ is upper bounded by

$$G(\mu)\leq\tilde{g}(\bar{b})=1. \tag{15}$$

This upper bound is asymptotically achievable for $e_{\max}\to+\infty$ by the balanced policy. Then, any policy considered in this paper, which outperforms the balanced policy, is expected to approach asymptotically the upper bound $\tilde{g}(\bar{b})$.

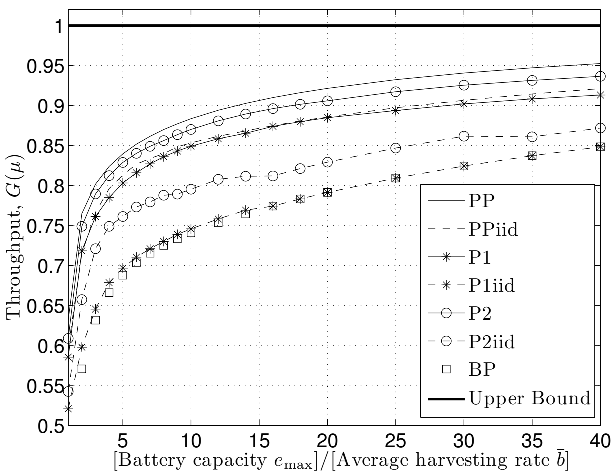

Figure 1. Throughput as a function of the battery capacity over harvesting rate ratio, $e_{\max}/\bar{b}$. ($\alpha = 1$, $\bar{b} = 10$, $x = y = 0.95$)

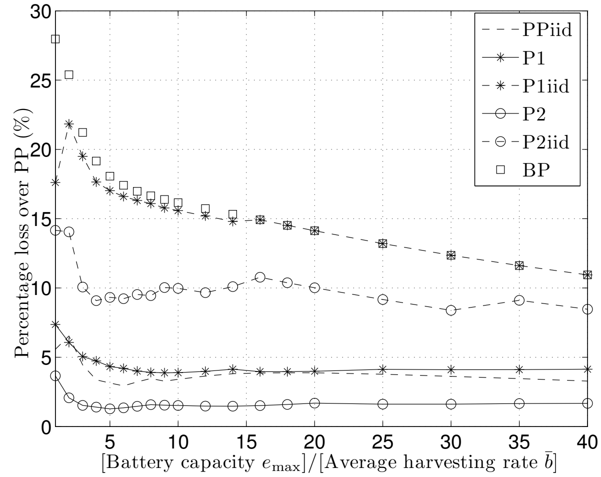

Figure 2. Percentage loss with respect to PP as a function of $e_{\max}/\bar{b}$. ($\alpha = 1$, $\bar{b} = 10$, $x = y = 0.95$)

In the simulations, we compare two classes of policies, assuming a different amount of information about the energy generation process:

1) policies with perfect knowledge of the energy arrival state, in particular: policy with perfect SOC knowledge (PP), policy with no SOC knowledge (P1, *i.e.*, one-interval uncertainty $\mathcal{I}(0) = \mathcal{E}$), and policy with two intervals uncertainty of equal size (P2, *i.e.*, $\mathcal{I}(0) = \{0, \ldots, \tilde{e}_1 - 1\}$ and $\mathcal{I}(1) = \{\tilde{e}_1, \ldots, e_{\max}\}$ with $\tilde{e}_1 = \lceil \frac{e_{\max}}{2} \rceil$).
2) policies which neglect the Markov structure of the energy arrivals, and treat them as i.i.d. with marginal distribution $p_B(b) = \sum_{s \in \mathcal{S}} \pi_S(s) p_B(b|s)$. In particular, we consider: balanced policy (BP), policy with perfect SOC knowledge (PPiid), policy with no SOC knowledge (P1iid, *i.e.*, one-interval uncertainty $\mathcal{I}(0) = \mathcal{E}$), and policy with two intervals uncertainty of equal size (P2iid, *i.e.*, $\mathcal{I}(0) = \{0, \ldots, \tilde{e}_1 - 1\}$ and $\mathcal{I}(1) = \{\tilde{e}_1, \ldots, e_{\max}\}$ with $\tilde{e}_1 = \lceil \frac{e_{\max}}{2} \rceil$).

Policies PP and PPiid are obtained via policy iteration as discussed in Section III-A. Policies P1, P2, P1iid and P2iid, on the other hand, are obtained using the local search algorithm A.1, defined in Section III-B.

In Fig. 1, we plot $G(\mu)$ as a function of the ratio $e_{\max}/\bar{b}$ for $x = y = 0.95$ and $\alpha = 1$. The best performance is achieved by PP, followed by P2, PPiid and P1. To better emphasize the performance degradation due to limited information about the current SOC and the Markov structure of the energy arrival process, the percentage loss incurred by the different policies with respect to the optimal policy PP is plotted in Fig. 2. As a general conclusion, we observe that the performance in the case of incomplete information is affected only by a limited loss. Especially, using only 2 quantization partitioning of the SOC but having perfect knowledge of the current energy arrival state (policy P2) incurs a small degradation in performance. In particular, with battery capacity $e_{\max} \approx \bar{b}$, the performance degradation of P2 with respect to PP is about 5%, and becomes smaller as $e_{\max}$ increases (around 2%), since the impact of outage and overflow, which occur when the SOC approaches 0 and $e_{\max}$, respectively, becomes smaller. Moreover, the effect of not knowing at all the current SOC level but being accurate in the energy arrival state (policy P1) is almost exactly equivalent to knowing the SOC perfectly, but neglecting the state of the energy harvesting mechanism (policy PPiid). Clearly, when accurate knowledge of neither SOC nor energy arrival process is available, the performance degrades substantially (policies P1iid and BP). Yet, if the battery capacity $e_{\max}$ is large enough, the performance degradation with respect to the optimum is below 15%.

Finally, in order to highlight the impact of the energy generation process on the system performance, in Figs. 3 and 4 we plot the throughput and the percentage loss with respect to PP, respectively, as a function of the transition probability $p_S(1|1)$, where $y = p_S(2|2) = p_S(3|3) = 0.95$. As discussed above, when $p_S(1|1) \to 1$, the energy arrival process becomes i.i.d., and therefore the policies P1iid and P2iid, which neglect the Markov structure of the process, become optimal as $p_S(1|1) \to 1$. However, these policies suffer for $p_S(1|1) < 0.8$, where the energy harvesting mechanism exhibits high correlation.

V. Conclusions and Future Work

In this paper, we investigated how to properly model EHDs and their optimal battery usage policies, depending on the information available about the SOC and the energy generation process. While it turned out that both these factors are key in identifying the correct energy management, neither of them is so critical that it must be known with full precision. Moreover, they complement each other, so that even a partial knowledge on the battery status (*i.e.*, which partitioning level the actual SOC value belongs to) can induce an energy management policy which is improved by the presence of accurate knowledge on the energy generation process, and vice versa.

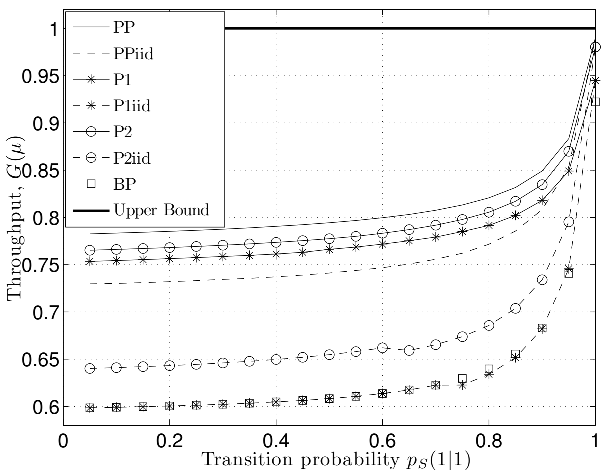

Figure 3. Throughput as a function of the transition probability $p_S(1|1)$. ($\alpha = 1$, $\bar{b} = 10$, $e_{\max} = 10\bar{b}$)

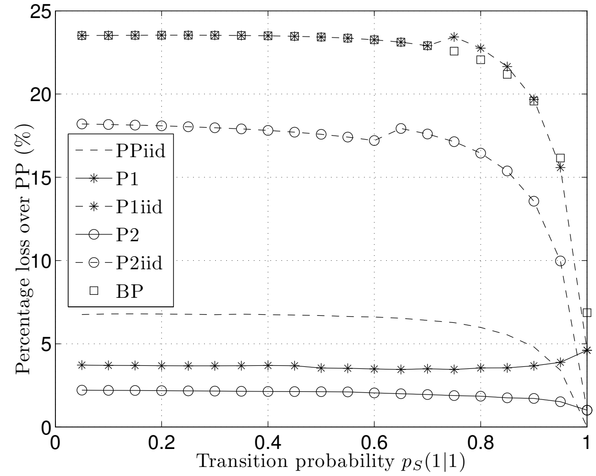

Figure 4. Percentage loss with respect to PP as a function of the transition probability $p_S(1|1)$. ($\alpha = 1$, $\bar{b} = 10$, $e_{\max} = 10\bar{b}$)

However, in spite of having extended the arrival process to more realistic scenarios with respect to the i.i.d. case, plenty of extensions are still possible to improve and generalize the analysis of the present paper. To start with, for real batteries, the SOC value is not only difficult to estimate with high accuracy, but also heavily influenced by the charge and discharge cycles. Moreover, the reward function of our analysis should be better matched to the evaluation of the long-term battery lifetime, instead of just assuming unlimited charge and discharge cycles. This would enable a better battery control in the long run [18], [19]. Moreover, the control policy was designed by neglecting the history of the energy harvesting process. Indeed, a better control can be designed both with a proper solution of the POMDP, and also by taking into account memory effects both in the energy arrival process and in the data traffic that the sensor must send out.

We believe that there are several challenges that can be faced in the future, by further developing the contributions made in the present paper. These investigations can be further framed in an interdisciplinary context, and have important consequences on the joint design of batteries, network elements, and control and actuation policies.

Acknowledgments

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/).

References

[1] N. Michelusi, K. Stamatiou, L. Badia, and M. Zorzi, "Operation policies for energy harvesting devices with imperfect state-of-charge knowledge," in *Proc. First IEEE International Workshop on Energy Harvesting for Communications (EHC)*, June 2012.

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

[3] C. Renner and V. Turau, "CapLibrate: self-calibration of an energy harvesting power supply with supercapacitors," in *International Conference on Architecture of Computing Systems (ARCS)*, Hannover, Germany, Feb. 2010, pp. 1–10.

[4] K. Lin, J. Yu, J. Hsu, S. Zahedi, D. Lee, J. Friedman, A. Kansal, V. Raghunathan, and M. Srivastava, "Heliomote: enabling long-lived sensor networks through solar energy harvesting," in *Proc. ACM SenSys*, 2005.

[5] K. Kar, A. Krishnamurthy, and N. Jaggi, "Dynamic node activation in networks of rechargeable sensors," *IEEE/ACM Transactions on Networking*, vol. 14, pp. 15–26, Feb. 2006.

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

[7] M. Gatzianas, L. Georgiadis, and L. Tassiulas, "Control of wireless networks with rechargeable batteries," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 581–593, Feb. 2010.

[8] V. Sharma, U. Mukherji, V. Joseph, and S. Gupta, "Optimal energy management policies for energy harvesting sensor nodes," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 1326–1336, Apr. 2010.

[9] A. Seyedi and B. Sikdar, "Energy efficient transmission strategies for body sensor networks with energy harvesting," *IEEE Trans. Commun.*, vol. 58, pp. 2116–2126, Jul. 2010.

[10] N. Michelusi, K. Stamatiou, and M. Zorzi, "On optimal transmission policies for energy harvesting devices," in *Information Theory and Applications Workshop (ITA)*, 2012.

[11] B. Xiao, Y. Shi, and L. He, "A universal state-of-charge algorithm for batteries," in *ACM/IEEE Design Automation Conference (DAC)*, Anaheim, CA, Jun. 2010, pp. 687–692.

[12] E. L. Lehmann and G. Casella, *Theory of Point Estimation*, 2nd ed. Springer, Aug. 1998.

[13] D. J. White, *Markov Decision Processes*. Wiley, 1993.

[14] J. M. Porta, N. Vlassis, M. T. Spaan, and P. Poupart, "Point-based value iteration for continuous POMDPs," *Journal of Machine Learning Research*, vol. 7, 2006.

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

[16] D. A. Levin, Y. Peres, and E. L. Wilmer, *Markov chains and mixing times*. American Mathematical Society, 2006.

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

[18] S. Hung, D. Hopkins, and C. Mosling, "Extension of battery life via charge equalization control," *IEEE Transactions on Industrial Electronics*, vol. 40, no. 1, pp. 96–104, Feb. 1993.

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