Operation Policies for Energy Harvesting Devices with Imperfect State-of-Charge Knowledge

Nicolò Michelusi*, Kostas Stamatiou†, Leonardo Badia* and Michele Zorzi*†
{michelusi,kstamat,badia,zorzi}@dei.unipd.it
*Department of Information Engineering, University of Padova - via Gradenigo, 6b, 35131 Padova, Italy
†Consorzio Ferrara Ricerche (CFR) - via Saragat, 1, 44122 Ferrara, Italy

***Abstract*—As Energy Harvesting Devices (EHD) become more widely deployed in sensor network platforms, the need arises for "smart" operation policies which can ensure long-term, autonomous and reliable operation. Existing research has relied on the implicit assumption of perfect knowledge of the energy available in the EHD. However, estimating the energy level of the batteries or super-capacitors employed in real-world EHDs, commonly known as State-Of-Charge (SOC), is a non-trivial task. In this paper, we design operation policies that maximize the long-term reward under *imperfect* knowledge of the SOC. Through an array of simulation results, we quantify the performance degradation due to imperfect SOC knowledge, and show that it increases with decreasing storage capacity and increasing variance in the energy arrival process. In the particular case of a two-state controller, *i.e.*, a controller which knows only if the SOC is HIGH or LOW, we prove that, for a linear reward function, there is no performance loss, while, for a logarithmic reward function, simulations show that the loss is typically less than 5%.**

I. Introduction

Due to their ability to operate autonomously for long periods of time by collecting, or "harvesting", energy from the environment, Energy Harvesting Devices (EHD) have been steadily gaining popularity in Wireless Sensor Network (WSN) deployments, and have been the topic of intense investigation in diverse research communities in the last few years [1], [2]. In particular, a significant body of research has focused on how to best manage the available energy of the EHD, with the general objective of optimizing the long-term performance in regard to sensing and data-communication tasks [3]–[8]. Setting aside the particular details of the energy harvesting mechanism (*e.g.*, solar, motion, heat, aeolian), an EHD is typically modeled as consisting of an energy buffer which is supplied from an energy arrival process. At each time, a controller decides how much energy should be drawn from the energy buffer, according to the desired operation policy.

This paper tackles an issue which, to the best of our knowledge, has not been addressed in existing work, namely, the design of EHD operation policies when the controller has only imperfect knowledge of the available amount of energy. In practical EHDs, energy is stored in electrochemical batteries and/or super-capacitors, whose energy level at any given time, commonly referred to as State-Of-Charge (SOC), needs to be estimated. In [2], it is stated that variations in the super-capacitor capacitance relative to the data-sheet value, due to age or temperature fluctuations, may be of the order of 30%. An online SOC estimation algorithm based on controlled discharge is thus proposed and shown to perform well, albeit at a small energy loss. In [9], [10], different algorithms are designed for the estimation of the open circuit voltage of an electrochemical battery, which is linearly related to the SOC. The complexity of the algorithms signifies that SOC estimation for electrochemical batteries is not a trivial task, hence precise knowledge of the SOC may be unreliable or too expensive.

Motivated by the aforementioned real-world concerns, we consider an EHD where the controller knows the SOC only to within a certain degree of precision. The problem formulated in Section III is that of determining the optimal amount of energy to be drawn from the energy buffer, given the knowledge of the SOC interval, with the objective of maximizing a long-term expected reward of the consumed energy. In the particular case of a linear reward function and two intervals, *i.e.*, LOW and HIGH, it is shown under loose assumptions that there is no performance loss with respect to a policy that possesses perfect SOC knowledge. The intuition behind this result is that, due to the linearity of the reward function, a policy which avoids energy *outage* by staying idle when the SOC is LOW, and energy *overflow* by being aggressive when the SOC is HIGH, achieves optimal performance. In Section IV, we consider the case of a logarithmic reward function, and demonstrate with an array of simulation results that the performance penalty due to imperfect SOC knowledge increases with decreasing energy storage capacity and increasing second order moment of the energy arrival process. Conversely, when the energy storage capacity goes to infinity, a balanced policy, which does not require SOC knowledge, is optimal. Overall, for typical parameter values, we find that the performance penalty due to imperfect SOC knowledge is at most 5%. This points to the conclusion that avoiding situations of energy outage/overflow provides the bulk of the performance gain in EHDs with finite energy storage capacity, and precise knowledge of the SOC yields only a marginal additional benefit.

II. System model

We consider a slotted-time system, where slot k is the time interval $[k, k+1)$, $k \in \mathbb{N}_0$, and $\mathbb{N}_0$ denotes the set of non-negative integers. Energy is stored in the EHD in an energy buffer, in the form of energy quanta, whose absolute value depends on the application-specific scenario, and is not considered herein (this model has been widely employed in existing work, *e.g.*, see [5], [8]). The energy level, *i.e.*, the SOC, at time k is denoted by E_k and takes values in the discrete set $\mathcal{E} = \{0, 1, \dots, e_{\max}\}$, where $e_{\max} \geq 1$ is the buffer capacity. Starting from the initial condition E_0, the evolution of the random variable E_k is governed by the following equation

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

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

- $\{B_k\}$ is the *energy arrival process*, which models the randomness in the energy harvesting mechanism, *e.g.*, due to an erratic energy supply. We model $\{B_k\}$ as i.i.d. stationary, taking values in $\mathcal{B} = \{0, 1, \dots, b_{\max}\}$ with probability mass function $p_B(b), b \in \mathcal{B}$, and mean $\bar{b} = \mathbb{E}[B_k]$. We refer to $\bar{b}$ as the *average harvesting rate*.
- Q_k is the number of energy quanta requested by the controller of the EHD in slot k to perform a certain task. We define the *action space* of the controller as $\mathcal{Q} = \{0, \dots, q_{\max}\}$, for some $0 < q_{\max} \leq e_{\max}$, so that $Q_k \in \mathcal{Q}, \forall k$. The parameter $q_{\max}$ reflects a physical constraint on the maximum amount of energy that can be drawn from the buffer at any given time.

We assume that only partial knowledge of E_k is available at the controller, *e.g.*, due to uncertainty in its estimation. Let $\{\mathcal{I}(n), n = 0, \dots, \tilde{n} - 1\}$ be a partition of the state space $\mathcal{E}$ defined as $\mathcal{I}(n) = \{\tilde{e}_n, \dots, \tilde{e}_{n+1} - 1\}$, $n \in \{0, \dots, \tilde{n} - 1\}$, where $0 = \tilde{e}_0 < \tilde{e}_1 < \dots < \tilde{e}_{\tilde{n}} = e_{\max} + 1$ define the interval boundaries. Suppose that, at time k, $E_k \in \mathcal{I}(N_k)$, for some $N_k \in \{0, \dots, \tilde{n} - 1\}$. We assume that the 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. To this end, we define the *interval index* process $\{N_k, k \geq 0\}$, taking values in $\{0, \dots, \tilde{n}-1\}$. Note that, if $\tilde{n} = e_{\max}+1$, we obtain the special case of perfect knowledge, *i.e.*, $E_k = N_k$.

Given the SOC E_k and the decision Q_k, the following phenomena may occur due to (1):

1) *Energy outage*: if $Q_k > E_k$, an energy outage occurs, since the node runs out of energy before the completion of the executed task. An energy outage is a consequence of the imperfect knowledge of E_k, due to which the controller may attempt to draw more energy than what is available.
2) *Energy overflow*: If $B_k > e_{\max} - [E_k - Q_k]^+$, the energy buffer is unable to store all of the harvested energy B_k. This is a consequence of the limited capacity of the energy buffer.

III. Optimization

A. Policy definition and optimization problem

In general, given the interval index N_k and the history $\mathcal{H}_k = \{N_0, \dots, N_{k-1}, Q_0, \dots, Q_{k-1}, \text{past outage events}\}$ at time k, a controller *policy* μ decides on the amount of energy Q_k to be drawn from the energy buffer. Formally, μ is a probability measure on the action space $\mathcal{Q}$, parameterized by the state $(N_k, \mathcal{H}_k)$, *i.e.*, given $(N_k, \mathcal{H}_k)$, $\mu(q; (N_k, \mathcal{H}_k))$ is the probability of choosing action $Q_k = q \in \mathcal{Q}$ in slot k.

We define the reward function $g : \mathcal{Q} \times \mathcal{E} \mapsto \mathbb{R}^+$ as

$$g(Q_k, E_k) = \begin{cases} 0 & Q_k > E_k \\ \tilde{g}(Q_k) & Q_k \leq E_k, \end{cases} \tag{2}$$

where $\tilde{g} : \mathcal{Q} \mapsto \mathbb{R}^+$ is a concave increasing function of Q_k, with $\tilde{g}(0) = 0$. If $Q_k > E_k$ the reward is 0, which models the inability of the sensor node to complete the requested task, when there is energy outage. As an example, if the reward function is the transmission rate, then, according to the Shannon formula, $\tilde{g}(Q_k) \propto \ln(1 + \alpha Q_k)$, where $\alpha > 0$ is an SNR scaling factor. The controller spreads the energy Q_k over the entire codeword. If $Q_k > E_k$, the EHD runs out of energy when only a fraction E_k/Q_k of the codeword has been transmitted, hence the codeword is discarded.

Given $E_0 = e_0$, the *long-term average reward per time-slot* under policy μ is defined as

$$G(\mu, e_0) = \lim_{K \to \infty} \inf \frac{1}{K} \mathbb{E}\left[\sum_{k=0}^{K-1} g(Q_k, E_k) \,\middle|\, E_0 = e_0\right], \tag{3}$$

where the expectation is over $\{B_k, Q_k, k = 0, \dots, K-1\}$. The general problem is to obtain a policy μ^* (possibly dependent on the initial state e_0) such that

$$\mu^* = \arg\max_{\mu} G(\mu, e_0). \tag{4}$$

We denote the respective optimal reward as $G(\mu^*, e_0) = G^*(e_0)$. Due to partial knowledge of the SOC, the above problem can only be solved approximately, using numerical optimization tools for Partially Observable Markov Decision Process (POMDP) [11]. Given the limited computational capabilities of EHDs, in this paper we focus on suboptimal policies, *i.e.*, policies that do not take into account the history $\mathcal{H}_k$, and depend solely on the current interval index N_k. In this light, we define $\mu(q; N_k)$ as the probability that the sensor node decides on action $Q_k = q$, given that $E_k \in \mathcal{I}(N_k)$. We now discuss the solution of (4), considering separately the cases of perfect/imperfect SOC knowledge.

B. Optimization with perfect SOC knowledge

Under perfect SOC knowledge, the controller selects action $Q_k = q$ when the SOC is E_k with probability $\mu(q; E_k)$. The sequence $\{(E_k, Q_k), k \geq 0\}$ constitutes a Markov Decision Process [12], so (4) is maximized by a stationary, state-dependent policy obtained by solving a linear program [13]. Note that the long-term reward under perfect SOC knowledge represents an upper bound to the performance of any policy under SOC uncertainty.

C. Optimization with SOC uncertainty

Under SOC uncertainty, by definition of the policy μ, Q_k is the same for all $E_k \in \mathcal{I}(N_k)$. This constraint is not linear with respect to the joint steady-state distribution of the SOC/action pair, (e, q). Hence, unlike in the scenario with perfect SOC knowledge, (4) cannot be solved via a linear program [13], and we opt to find the optimal policy via an exhaustive search. Furthermore, in order to reduce the complexity, we consider only deterministic policies, *i.e.*,

$$\begin{cases} \mu_\rho(q;n) = 1, \ q = \rho(n), \\ \mu_\rho(q;n) = 0, \ q \in \mathcal{Q} \setminus \{\rho(n)\}, \end{cases} \tag{5}$$

where $\rho : \{0, \dots, \tilde{n}-1\} \mapsto \mathcal{Q}$ is a function which maps the interval index $n \in \{0, \dots, \tilde{n}-1\}$ to the action $q \in \mathcal{Q}$. As a result, (4) can equivalently be written as

$$\rho^* = \arg\max_\rho G(\mu_\rho, e_0), \tag{6}$$

where, from (3),

$$G(\mu_\rho, e_0) = \sum_{n=0}^{\tilde{n}-1} \sum_{e \in \mathcal{I}(n)} \pi_\rho(e; e_0) g(\rho(n), e) \tag{7}$$

and $\pi_\rho(e; e_0)$ is the asymptotic distribution of the SOC $e \in \mathcal{E}$, given that the initial state is $E_0 = e_0$, *i.e.*,

$$\pi_\rho(e; e_0) = \lim_{N \to +\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathrm{Pr}_\rho \left(E_n = e | E_0 = e_0 \right), \tag{8}$$

where $\mathrm{Pr}_\rho \left(E_n = e | E_0 = e_0 \right)$ is the n-step transition probability of the chain under the policy ρ. In most practical cases, the asymptotic distribution can be evaluated as the unique solution of the system of steady-state equations [14]

$$\begin{cases} \sum_{n=0}^{N-1} \sum_{e \in \mathcal{I}(n)} \pi_\rho(e) = 1, \quad \text{(normalization)}, \\ \\ \pi_\rho(e) \geq 0, \quad \forall e \in \mathcal{E}, \quad \text{(non-negativity)}, \\ \\ \sum_{n=0}^{N-1} \sum_{s \in \mathcal{I}(n)} \pi_\rho(s) \mathrm{Pr}_\rho \left(E_1 = e | E_0 = s \right) = \pi_\rho(e), \\ \quad \forall e \in \mathcal{E}, \ \text{(steady-state equations)} \end{cases} \tag{9}$$

(Note that in this case, π_ρ is independent of the initial state $E_0 = e_0$).

D. Special Cases

We now determine the optimal mapping ρ^* for some special scenarios. Our results are stated in Propositions 1, 2 and 3. For simplicity, we assume that $q_{\max} = e_{\max}$, *i.e.*, the EHD controller can draw all the available energy in the buffer, and $b_{\max} \leq e_{\max}$, *i.e.*, the range of values of the energy arrivals cannot exceed the buffer capacity. Similar results can be derived for general $q_{\max}$ and $b_{\max}$.

Proposition 1 (Deterministic Arrival Process) *If the energy arrival process is deterministic, i.e., $B_k = b_{\max}$ $\forall k$, then, for a general reward function, one optimal mapping ρ^* is*

$$\rho^*(n) = b_{\max}, \quad \forall n \in \{0, \dots, \tilde{n}-1\}, \tag{10}$$

and the optimal reward is $G^*(e_0) = \tilde{g}(b_{\max})$.

Proof: The energy harvesting mechanism (1) induces the long-term energy per time-slot constraint

$$\lim_{K \to +\infty} \sup \frac{1}{K} \mathbb{E} \left[\sum_{k=0}^{K-1} Q_k \right] \leq \bar{b} = b_{\max}. \tag{11}$$

Since $\tilde{g}(q)$ is concave and non-decreasing in q, we have that $G(\mu, e_0) \leq \tilde{g}(\bar{b}) = \tilde{g}(b_{\max})$, for all policies μ.

We now prove that this upper bound is achievable. Starting from the SOC $E_0 < b_{\max}$ and applying policy (10) to (1), we obtain the sequence

$$\begin{array}{lll} k=0: & E_0 < Q_0 & \to g(Q_0, E_0) = 0 \text{ (outage)} \\ k=1: & E_1 = b_{\max} = Q_1 & \to g(Q_1, E_1) = \tilde{g}(b_{\max}) \\ & \vdots & \\ k>1: & E_k = b_{\max} = Q_k & \to g(Q_k, E_k) = \tilde{g}(b_{\max}). \end{array}$$

Except for the first time-slot, in which outage occurs, in all other time-slots the reward $\tilde{g}(b_{\max})$ is accrued, so that, in the long-term, the upper bound is attained.

Similarly, starting from $E_0 \geq b_{\max}$ results in $Q_k = b_{\max}$, $E_k = E_0 \geq Q_k$, hence $g(Q_k, E_k) = \tilde{g}(b_{\max})$, $\forall k \geq 0$. Therefore, independently of the initial SOC E_0, in the long-term we have $G(\mu_{\rho^*}, E_0) = G^*(E_0) = \tilde{g}(b_{\max})$. ■

Note that a deterministic arrival process models approximately an energy source which exhibits slow fluctuations in time and predictable behavior, *e.g.*, solar energy in sunny days.

Proposition 2 (Linear Reward) *Under a reward $\tilde{g}(q) = q$, a general energy arrival process, and the following assumptions:*

(a) Two-interval SOC uncertainty, i.e., $\tilde{n} = 2$, $\mathcal{I}(0) = \{0, \dots, \tilde{e}_1 - 1\}$ and $\mathcal{I}(1) = \{\tilde{e}_1, \dots, e_{\max}\}$,

(b) $b_{\max} \leq \min\{\tilde{e}_1, e_{\max} + 1 - \tilde{e}_1\}$,

one optimal mapping ρ^* is

$$\begin{cases} \rho^*(0) = 0 \\ \rho^*(1) = \rho_1, \end{cases} \tag{12}$$

where ρ_1 is any value in the set $\{b_{\max}, \dots, \tilde{e}_1\}$, and the optimal reward is $G^*(e_0) = \bar{b}$. Moreover, this mapping is also optimal under perfect SOC knowledge.

Proof: For any policy μ, following the same steps as in the proof of Proposition 1, it can be shown that $G(\mu, e_0) \leq \tilde{g}(\bar{b}) = \bar{b}$. We now prove that policy (12) achieves this upper bound. Since the bound holds for any policy μ, (12) is also optimal under perfect SOC knowledge.

If $E_k \in \mathcal{I}(0)$, then the action $Q_k = \rho^*(0) = 0$ is chosen. From (1), we have $E_{k+1} = \min\{E_k + B_k, e_{\max}\}$. Since $B_k \leq b_{\max}$ and $E_k \leq \tilde{e}_1 - 1$ (from $E_k \in \mathcal{I}(0)$), from (b) we have that $E_k + B_k \leq \tilde{e}_1 - 1 + b_{\max} \leq e_{\max}$. This implies that neither overflow nor outage occurs when $E_k \in \mathcal{I}(0)$.

If $E_k \in \mathcal{I}(1)$, then the action $Q_k = \rho_1 \in \{b_{\max}, \dots, \tilde{e}_1\}$ is chosen. Since $E_k \geq \tilde{e}_1 \geq Q_k$, outage does not occur, hence $g(Q_k, E_k) = \tilde{g}(Q_k) = Q_k$. Moreover, since $B_k \leq b_{\max} \leq Q_k$, at any time-slot enough energy quanta are drawn from the buffer to make room for the new arrivals, hence overflow does not occur.

Since neither overflow nor outage occurs at any time, we have $E_{k+1} = E_k - Q_k + B_k$ and $g(Q_k, E_k) = \tilde{g}(Q_k) = Q_k$. All harvested energy contributes to reward accrual, hence

$$G(\mu_{\rho^*}, e_0) = \lim_{K\to\infty} \inf \frac{1}{K}\mathbb{E}\left[\sum_{k=0}^{K-1} Q_k \middle| E_0 = e_0\right] = \bar{b}, \quad (13)$$

which proves the achievability of the upper bound. ■

Note that, if the length of the intervals $\mathcal{I}(0), \mathcal{I}(1)$ differ by at most one unit, *i.e.*, $\tilde{e}_1 = \lceil e_{\max}/2 \rceil$, then assumption (b) simplifies to $b_{\max} \leq \lceil e_{\max}/2 \rceil$, *i.e.*, the buffer capacity is at least twice the maximum amount of energy that can be harvested in a time-slot.

It is worth noting that any policy avoiding energy outages and overflows is optimal in the linear reward case, so that, in general, (12) may not be unique to attain optimal performance. For example, in the deterministic case, one optimal mapping is $\rho^*(0) = \rho^*(1) = b_{\max}$ (Proposition 1), which clearly violates (12) since $\rho^*(0) > 0$. In the following proposition, we state without proof a condition under which (12) is necessary for optimality.

Proposition 3 (Necessity of (12)) *If, in addition to the assumptions of Proposition 2, $p_B(b) > 0, \forall b \in \{0, 1, b_{\max}\}$, then (12) is necessary, i.e., any policy violating it is strictly suboptimal.*

E. Discussion

Harvested energy is lost, thus incurring a performance degradation, when there is energy outage due to uncertain SOC knowledge, or energy overflow due to limited energy buffer capacity. When the energy arrival process is deterministic, the controller can forecast the future energy arrivals, and can avoid both outage and overflow, thus achieving optimal performance without knowledge of the SOC (Proposition 1). We thus expect that the performance degradation due to imperfect SOC knowledge is related to the randomness of the energy arrival process, which is verified by simulation in Section IV.

When the energy arrival process is random, the controller has limited knowledge about the future energy arrivals. In this case, overflow can be avoided by an *aggressive* policy, which draws $q_{\max} = b_{\max}$ energy quanta when the battery SOC approaches its capacity. This choice guarantees that enough energy quanta are drawn from the buffer, thus making room for the new energy arrival. Moreover, outage can be avoided by a *conservative* policy, which stays idle when the battery SOC approaches depletion.

If the reward function is linear and the assumptions of Proposition 2 hold, this policy is optimal, since no energy is lost, and all energy quanta contribute to the reward accrual. However, when the reward function is concave, a too aggressive or too conservative behavior is penalized, hence the optimal policy strives to be less conservative (aggressive) for LOW (HIGH) energy availability, thus inevitably incurring outage and overflow.

In the next section, we depart from the particular cases considered in this section, and determine numerically the optimal policy and respective long-term reward for a more general scenario.

IV. Numerical Results

In this section, we consider the maximization of (4) for $\bar{b} = 20$,[1] and (unless otherwise stated) a geometric energy arrival distribution truncated at $b_{\max} = 4\bar{b}$. The reward function is the normalized throughput

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

where α is an SNR scaling factor. Note that, in the limit $\alpha \to 0^+$, we obtain the linear reward scenario $\tilde{g}(q) = q/\bar{b}$.

Since (1) imposes the constraint

$$\lim_{K\to+\infty} \sup \frac{1}{K}\mathbb{E}\left[\sum_{k=0}^{K-1} Q_k\right] \leq \bar{b},$$

from the concavity of $\tilde{g}(q)$ we have

$$G(\mu, e_0) \leq \tilde{g}(\bar{b}) = 1. \quad (15)$$

This upper bound is asymptotically achievable for $e_{\max} \to +\infty$ by the *balanced policy*, defined as the policy which, in every time slot, attempts to draw $q = \bar{b}$ energy quanta from the buffer. Hence, as $e_{\max}$ increases (while all other parameters are kept constant), $G(\mu, e_0)$ is expected to converge to 1, independently of α and of the energy arrival distribution. Therefore, the choice of (14) as the reward function makes the performance comparison possible for different concavity levels, energy arrival distributions and buffer capacity values $e_{\max}$.

We consider the following policies: balanced policy (BP), 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-equal-interval uncertainty (P2, *i.e.*, $\mathcal{I}(0) = \{0, \ldots, \tilde{e}_1 - 1\}$ and $\mathcal{I}(1) = \{\tilde{e}_1, \ldots, e_{\max}\}$ with $\tilde{e}_1 = \lceil \frac{e_{\max}}{2} \rceil$). PP is obtained by solving (4) as a linear program, while P1 and P2 by solving (6) via exhaustive search. Note that, by definition, BP belongs to the class of policies with no SOC knowledge.

In Fig. 1, we plot $G(\mu)$ vs. the ratio of the buffer capacity over the average harvesting rate $e_{\max}/\bar{b}$. As expected, the best performance is achieved by PP, followed by P2, P1 and BP. At a buffer capacity $e_{\max} \approx 2\bar{b}$, the performance degradation of P2 with respect to PP is about 5%. As $e_{\max}$ increases, the degradation becomes smaller, since the impact of outage and overflow, which occur when the SOC approaches 0 and $e_{\max}$, respectively, becomes smaller. Also note that P1 performs better than BP, since it is the optimal policy with no SOC knowledge. We have verified that P1 is more conservative than BP for small $e_{\max}$ values. In the limit of large $e_{\max}$, P1 draws

[1] This value is large enough to allow for a sufficiently fine-grained quantization of the physical quantities of interest and, at the same time, it is small enough to guarantee a manageable computation time.

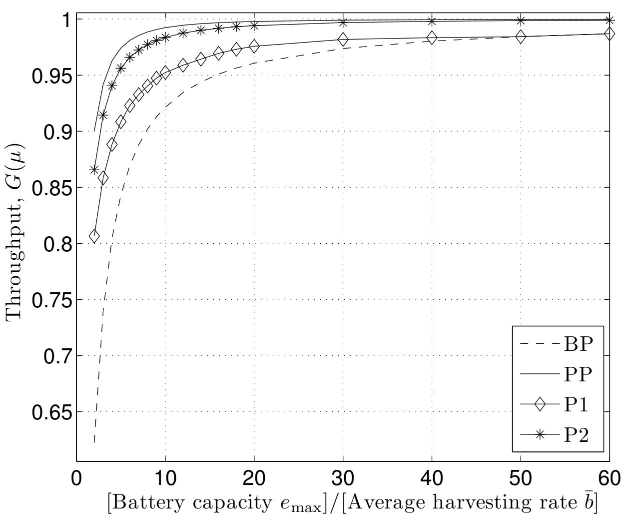

Figure 1. Throughput as a function of $e_{\max}/\bar{b}$, for different policies. ($\alpha = 1$, $\bar{b} = 20$)

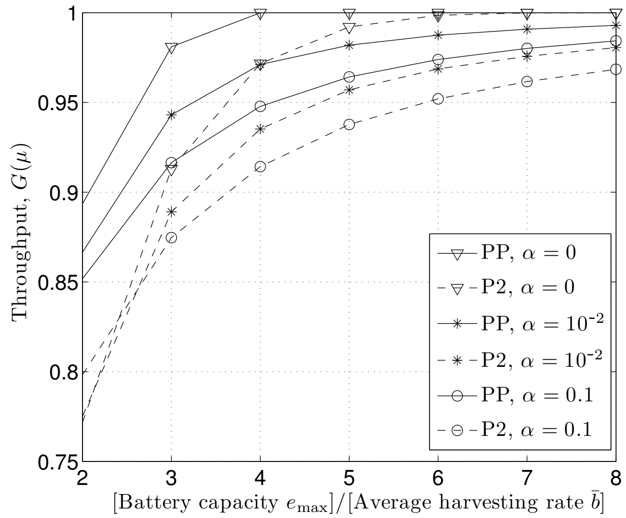

Figure 2. Throughput as a function of $e_{\max}/\bar{b}$, for different values of α. The reward function $\tilde{g}(q)$ is more concave for increasing values of α, and linear for $\alpha = 0$. ($\bar{b} = 20$)

energy with rate $\bar{b}$ and the performance of the two policies becomes identical.

In Fig. 2, we plot $G(\mu)$ vs. $e_{\max}/\bar{b}$ for PP and P2 and different values of the concavity level α. It is seen that, as $e_{\max}$ increases, the performance degradation due to imperfect SOC knowledge decreases and the throughput performance of PP and P2 approaches unity. The amounts of energy drawn by P2 in the LOW and HIGH SOC regimes for each α are plotted in Fig. 3. Note that, when $\alpha = 0$ (linear reward function), for $e_{\max}/\bar{b} \geq 8$, $\rho(0) = 0$ and $\rho(1) = 80$. A quick calculation setting $\tilde{e}_1 = e_{\max}/2$ (which is the optimal threshold in this case) and $b_{\max} = 4\bar{b} = 80$ shows that this observation is coherent with Proposition 2. Moreover, the optimal policy for increasing α is less conservative (aggressive) in the LOW (HIGH) regimes compared to the case of a linear reward, as discussed in Section III-E.

Fig. 4 examines the dependence of $G(\mu)$ on the energy arrival statistics. It is seen that, the larger the second order moment (equivalently, the variance, since the mean $\bar{b}$ is fixed), the larger the performance degradation, under both perfect and imperfect SOC knowledge. Interestingly, the geometric and Bernoulli energy arrival distributions, with the same mean $\bar{b} = 20$ and second-order moment $\mathbb{E}[B_k^2] = 722$, yield approximately the same performance. This suggests that the performance may be mainly affected by the second order statistics of the arrival process. As the capacity increases, the impact of an erratic energy source, in terms of outage and overflow, becomes smaller. Thus, the performance is determined by the average harvesting rate $\bar{b}$, and the curves approach the same value $G(\mu) = 1$ asymptotically.

Finally, Fig. 5 examines the dependence of $G(\mu)$ on the value of the threshold $\tilde{e}_1$ when P2 is employed, for different values of $e_{\max}$. It is seen that the best performance is achieved when $\tilde{e}_1/e_{\max} \approx 0.5$. For smaller (larger) thresholds, there is

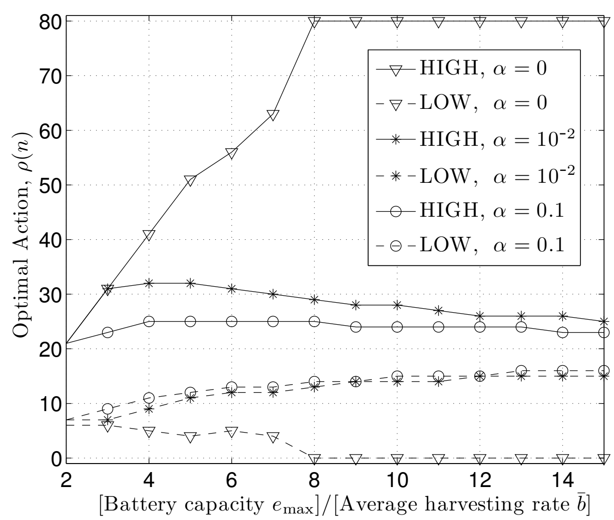

Figure 3. Actions $\rho(0)$ (LOW) and $\rho(1)$ (HIGH) of P2 corresponding to Fig. 2. ($\bar{b} = 20$)

less uncertainty in the LOW (HIGH) energy levels. However, the higher uncertainty in the HIGH (LOW) energy levels accounts for a performance degradation with respect to the case $\tilde{e}_1/e_{\max} \approx 0.5$. It is also observed that, when $\tilde{e}_1/e_{\max} \to 0$, the performance of P2 approaches that of P1. In contrast, when $\tilde{e}_1/e_{\max} = 1$, the knowledge of whether the buffer is full or not affords an advantage to P2 compared to P1, which becomes more substantial for decreasing $e_{\max}$.

V. Conclusions

Motivated by real-world EHD implementations, we have investigated the performance of different transmission policies for an EHD which operates under imperfect SOC knowledge.

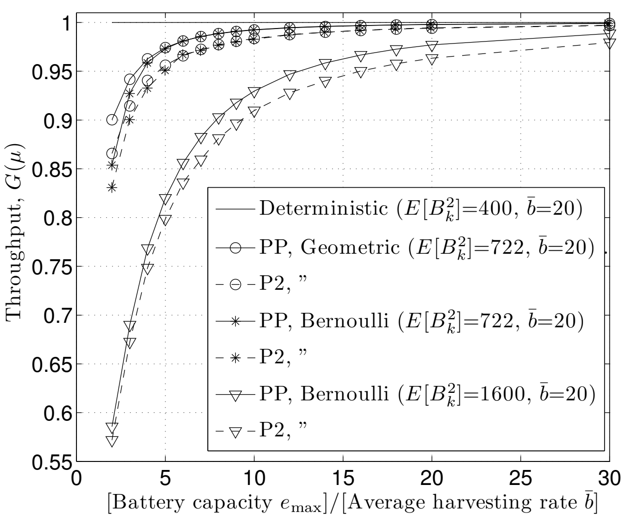

Figure 4. Throughput as a function of $e_{\max}/\bar{b}$, for different energy arrival statistics. ($\alpha = 1$, $\bar{b} = 20$)

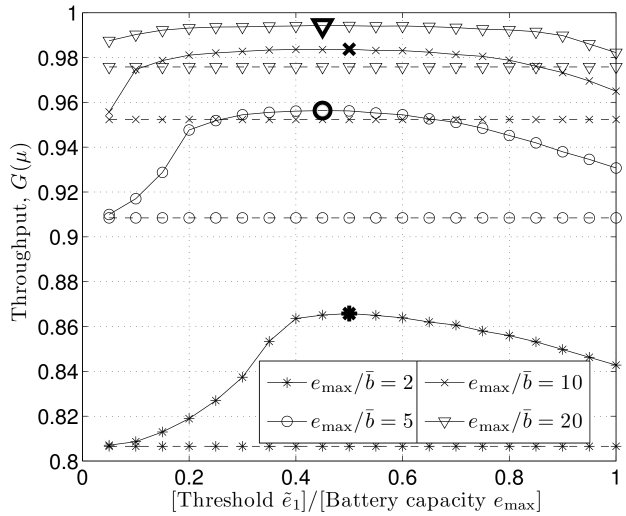

Figure 5. Throughput vs. $\tilde{e}_1/e_{\max}$ for P2 (solid), P1 (dashed) and different $e_{\max}$. The bold markers indicate the optimal choices of $\tilde{e}_1$. ($\alpha = 1$, $\bar{b} = 20$)

The SOC uncertainty incurs a performance loss which increases with decreasing buffer capacity, increasing variance of the energy arrival process and increasing concavity of the reward function. Our simulation results for a logarithmic reward function and a simple two-state controller, which knows only if the SOC is HIGH or LOW, indicate that the loss is less than 5% for typical parameter values. The design implication is that close-to-optimal performance may be achieved, as long as the EHD controller avoids energy outages and overflows. Future work will further investigate the interplay between the energy buffer capacity, the energy arrival distribution and the reward function, as regards the performance of EHD operation policies with imperfect SOC knowledge.

Acknowledgments

This work has been supported in part 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] J. A. Paradiso and T. Starner, "Energy scavenging for mobile and wireless electronics," *IEEE Pervasive Computing*, vol. 4, pp. 18–27, Jan. 2005.

[2] 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.

[3] 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.

[4] D. Niyato, E. Hossain, and A. Fallahi, "Sleep and wakeup strategies in solar-powered wireless sensor/mesh networks: performance analysis and optimization," *IEEE Transactions on Mobile Computing*, vol. 6, pp. 221–236, Feb. 2007.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] J. Chiasson and B. Vairamohan, "Estimating the state of charge of a battery," *IEEE Transactions on Control Systems Tech.*, vol. 13, pp. 465–470, May 2005.

[10] 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.

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

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

[13] K. W. Ross, "Randomized and past-dependent policies for Markov decision processes with multiple constraints," *Operations Research*, vol. 37, pp. 474–477, June 1989.

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