Delay-Optimal Resource Scheduling of Energy Harvesting based Devices

Ibrahim Fawaz, *Student Member, IEEE,* Mireille Sarkiss, *Member, IEEE,*
and Philippe Ciblat, *Senior Member, IEEE*

Abstract—This paper investigates resource scheduling in a wireless communication system operating with Energy Harvesting (EH) based devices and perfect Channel State Information (CSI). The aim is to minimize the packet loss that occurs when the buffer is overflowed or when the queued packet is older than a certain pre-defined threshold. We so consider a strict delay constraint rather than an average delay constraint. The associated optimization problem is modeled as Markov Decision Process (MDP) where the actions are the number of packets sent on the known channel at each slot. The optimal deterministic offline policy is exhibited through dynamic programming techniques, i.e. Value Iteration (VI) algorithm. We show that the gain in the number of transmitted packets and the consumed energy is substantial compared to: i) a naive policy which forces the system to send the maximum number of packets using the available energy in the battery, ii) two variants of the previous policy that take into account the buffer state, and iii) a policy optimized with an average delay constraint. Finally, we evaluate our optimal policy under imperfect CSI scenario where only an estimate of the channel state is available.

I. INTRODUCTION

ENERGY harvesting (EH) technology has emerged recently as a promising solution to improve the energy efficiency and self-sustainability of 5G mobile and IoT networks. While relying on renewable energy sources in their surrounding environments, the mobile devices can harvest energy to perform their communications and operational tasks. In this way, they can extend their battery lifetimes by reducing their dependency on conventional battery and grid power, decreasing thus their carbon emissions. However, in contrast to conventional power supply where the available energy is fixed, harvested energy arrives randomly and sporadically due to environmental influence (weather, geolocation), rendering unpredictable the available energy behavior. To avoid the waste of energy excess and save it for future use, capacity-limited batteries is used to store the collected energy. The stochastic energy harvesting process and the energy storage constraints in addition to the time-varying nature of the wireless channels bring new design challenges in EH communications making the optimization of the transmission policies a more difficult task. Therefore, efficient resource

I. Fawaz is with CEA, LIST, Communicating Systems Laboratory, PC 173, 91191 Gif-sur-Yvette, France. M. Sarkiss is with Télécom SudParis, 9 Rue Charles Fourier, 91000 Evry, France. P. Ciblat is with Télécom ParisTech, 46 Rue Barrault, 75013 Paris, France. Contact: ibrahim.fawaz@cea.fr, mireille.sarkiss@telecom-sudparis.eu, philippe.ciblat@telecom-paristech.fr. Part of this work has been published in IEEE ICC 2018 conference [1]. This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 675891.

scheduling of mobile devices need to adapt the transmission rate and power to the dynamic levels of the available energy and the channels in order to ensure the users quality of service (QoS) and the system sustainability.

During the past decade, extensive research efforts have been devoted to investigate resource scheduling with EH capabilities at the transmitters [2]–[11]. Surveys can be read in [12], [13]. In these works, several performance criteria have been optimized such as throughput, completion time, average delay, outage probability, for various models of energy arrival rate, battery capacity, or fading channel. For instance, in [2], data amount transmitted during a pre-defined time was optimized and the transmission completion time was minimized by choosing carefully the transmit power when the channel is time-varying. The authors proposed optimal offline policies based on directional water-filling in a non-causal energy setting which means that the energy amount available at any time is known in advance. They also proposed online policies using continuous time stochastic dynamic programming in a causal energy setting. The throughput maximization problem was similarly investigated in [3] but for limited energy battery and limited data buffer, allowing thus buffer overflow. The optimal solutions were proposed by decoupling energy and data problems using a new variant of directional water-filling with added energy pumps, or applying recursively the shortest path algorithm. When only causal Energy State Information (ESI) and Channel State Information (CSI) are available, the same throughput maximization problem was modeled as Markov Decision Process (MDP) in [4] and related optimization techniques were used. In [5], an online algorithm maximizing the throughput is designed by assuming capacity-limited EH system. It relies on a new estimation method of future energy arrivals without any prior information. Both offline and online algorithms were also provided in [6] to maximize the throughput in finite-horizon scheduling with EH transmitter. The offline solution is expressed in terms of water levels and the online solution minimizes successively the expected throughput losses with respect to the offline optimal decision. Finite-horizon optimization problem was also considered in [7] to minimize the outage probability in a EH system. A low complexity fixed threshold transmission is proposed based on the offline mixed integer linear programming solution. In [8], an average delay optimal scheduling problem under energy consumption constraint was studied where the transmitter relies on hybrid energy

supplies. Actually, the data transmission is mainly powered by harvested energy and resorts to power grid as a backup. The problem was modeled as a two-dimensional Markov chain and an optimal policy depending on a critical threshold of the queue length is proposed using Linear programming formulations. In [9], optimal deterministic scheduling in EH-powered network satisfying an average delay constraint and an average consumed energy constraint was obtained by minimizing the packet blocking probability, due to non-transmission at the transmitter. The problem was formulated as an MDP and solved using dynamic programming Value Iteration (VI) algorithm. In [10], a weighted packet loss rate under an average delay constraint is minimized leading to a constrained MDP and solved by using a linear value iteration approximation that locally determines the energy allocation at every EH wireless node by multilevel water-filling. Near-optimal policy was also derived by applying online learning based on post-decision state framework. In [11], MDP modeling and online post-decision learning approach were derived to maximize the data arrival rate at the transmitter queue under delay and energy constraints. Two delay constraints were considered: average delay constraint or statistical delay constraint. This latter is a bounded delay with maximum acceptable delay-outage probability constraint.

In this paper, we address a resource scheduling for a point-to-point communication powered by energy harvesting at the transmitter side. It may correspond to an Uplink (UL) case where the transmitter is a node with energy harvesting ability and the receiver is a base station plugged on the grid. Unlike [8]–[10], the main novelty of this work is by considering a strict delay constraint on each queued packet in the transmitter buffer rather than an average delay constraint. We have initially introduced this hard constraint on the delay in [14] to find the optimal scheduling policy minimizing the average power consumption. Now, we incorporate energy harvesting aspects within the scheduling problem. Working with hard delay constraint is timely even if it involves a more complicated system description. It has especially led to a new way to think information theory by using short-length block codes as in [15] and by applying it on some resource allocation issues as in [16] via the notion of Ultra Reliable Low Latency Communications (URLLC). In this paper, we do not consider short-length block codes but we consider that the packets cannot stay in the buffer beyond a certain pre-defined duration.

Nevertheless, the aforementioned transmission policies rely on the perfect knowledge of the channel state information or an accurate estimation of the relevant statistics. In typical wireless systems, the receiver performs channel estimation and feeds back CSI on limited-capacity feedback channel to the transmitter in order to design channel-adapted transmission techniques. In practice, the acquired CSI have errors due to different factors such as time-varying channel, inaccurate channel estimation, quantization and feedback errors, which can inevitably cause performance degradation. Recently, some works have focused on new design strategies to deal with CSI imperfections in energy harvesting networks. For

instance, an optimal transmission power policy based on only 1-bit feedback was proposed in [17] for EH communications over Rayleigh fading channels. The receiver sends bit 1 if the channel realization is above a certain threshold. Then, the transmitter does not transmit if the bit is 0 or transmits with a certain pre-defined power. The related data rate is chosen according to the threshold and not to the true value of the channel realization. Consequently, the selected data rate always ensures a safe transmission but with a pessimistic rate. The paper found out the optimal feedback channel threshold and the optimal policy that maximizes the throughput based on finite-horizon constrained MDP formulation. In [18], the problem of data amount maximization within a fixed duration was studied assuming imperfect CSI at the transmitter (CSIT) in point-to-point communications with an EH transmitter. The authors proposed first a Markov process to model the energy arrivals and the channel impulse response with strong correlations and then derived the optimal online power scheduling policy using finite-horizon dynamic programming techniques. In addition, they studied the performance limits of EH systems under imperfect CSIT through an asymptotic analysis of the average throughput at low and high average energy recharge rates. In [19], they determined the optimal offline policy for a similar problem. In previous paper, they do not consider the cost to obtain the CSI even imperfectly, such as the energy consumption to send training sequence and the time spent to estimate and so not available anymore for doing data transmission.

In this paper, we investigate both perfect and imperfect channel state information at the transmitting EH device in our scheduling problem. In a first part, for perfect CSI scenario, we ideally assume that the channel is perfectly known at the transmitter without any cost. Taking into account sporadic energy arrivals, random data arrivals and time-varying channel states, we minimize the packet loss rate, i.e., the average number of discarded packets induced by strict delay constraint in addition to buffer overflow constraint. We formulate the problem as an MDP and solve it using Relative Value Iteration algorithm. We find an optimal offline stationary policy and compare it with a naive policy that performs immediate scheduling irrespective of energy and buffer states, and two variants of it taking into account the buffer state in the decision process. Then, we compare our proposed system with a similar one using average delay constraint. In this part, we mainly consider i.i.d EH process for sake of simplicity and clarity, but we compare also the results when time-correlated EH process is considered.

In a second part, for realistic imperfect CSI scenario, we consider that acquiring channel estimates incurs some time and energy costs on the system performance. We assess the previously obtained optimal policy under imperfect CSI conditions due to channel estimation errors. We also consider imperfect CSI assumption with Automatic Repeat ReQuest (ARQ) protocol, allowing thus packet re-transmission. Therefore, in these cases, the packet loss rate is affected twofold: on one side, with respect to the imposed strict delay because of a smaller transmission period of data packets,

or because of a longer duration of packets in the buffer for re-transmission (with ARQ protocols); and on the other side, with respect to the erroneous channel estimation which can lead to an increase in the number of discarded packets. We analyze the system taking into account these errors and show through numerical results that an appropriate trade-off is needed between the channel estimation accuracy and the transmission period in order to reduce the dropped packets depending on the available energy, energy arrivals and data arrivals.

The remainder of the paper is organized as follows. In Section II, we describe the system model. In Section III, we formulate the optimization problem as an MDP and solve it using value iteration algorithm. In Section IV, we present the framework of the imperfect CSI scenario. We provide and analyze numerical results in Section V. Finally, we give some concluding remarks and perspectives in Section VI.

II. SYSTEM MODEL

We consider a point-to-point communication over a fading channel with an energy harvesting transmitter. The transmitter is equipped with two queues: one corresponds to a capacity-limited battery to store harvested energy from an external source and the other is a finite buffer to store data packets arriving from the upper layer. The communication is slotted into consecutive epochs of equal duration T_s. At the beginning of each time slot, scheduling decisions are made to define the number of packets to be transmitted during the slot depending on energy arrivals and data arrivals during previous slot as well as channel states at the current time.

A. Energy model

Due to the random nature of energy harvesting sources, we model the EH process as an independent identically distributed (i.i.d.) Poisson distribution with an average arrival rate λ_e. We assume that the energy arrives in multiple packets of energy units (e.u) of $\mathcal{E}_U$ Joules (J) [1]. The received energy is stored in a battery of finite capacity B_e, and is lost when it exceeds B_e. At the beginning of time slot n, let e_n denote the harvested incoming energy (counting as a number of the energy units). Its probability distribution is given by

$$p(e_n = e) = e^{-\lambda_e} \cdot \frac{(\lambda_e)^e}{e!}.$$

We assume that the processing energy is negligible compared to the transmission energy, thus the energy stored in the battery is only used for communication. We also consider the energy causality constraint where the system can only transmit if a sufficient amount of energy is available in the battery. Let b_n denote the energy level of the battery at the beginning of time slot n, $b_n \in \{0, ..., B_e\}$, and E_n the energy consumed to send packets during time slot n, then $E_n \leq b_n$. In addition,

we suppose perfect energy state information at the transmitter (ESIT).

B. Data queue model and strict delay constraint

The transmitter receives also data packets and store them for future transmission in a data buffer of size B_d packets. We model the data arrival process as an i.i.d. process following a Poisson distribution with an average arrival rate λ_d. We assume that all packets are of the same size L bits. At the beginning of time slot n, let q_n denote the queue length in the buffer, $q_n \in \{0, ..., B_d\}$, and a_n the received packets with probability distribution

$$p(a_n = a) = e^{-\lambda_d} \cdot \frac{(\lambda_d)^a}{a!}.$$

A packet is discarded from the buffer

- if there is a **buffer overflow**, i.e., if the sum of packets in the queue and arrival packets exceeds the buffer size. In that case, we discard the arrival packets in overflow;
- if there is a **delay violation**, i.e., it stays in the queue more than K_0 slots. This can occur if the system decides not to transmit for a long period due to energy shortage or bad channel conditions.

In order to describe the delay violation, we need to introduce a new variable $k_i(n)$ counting the time spent in the buffer of the i-th packet at time n. By definition, we have $k_i(n) \in \{-1, ..., K_0\}, \forall i, k$ and $k_i(n) = -1$ for an empty space in the buffer (i.e., when the i-th packet does not exist). In Fig. 1, we provide a buffer state at time n. Notice that $k_j(n) \leq k_i(n)$, $\forall i \leq j$.
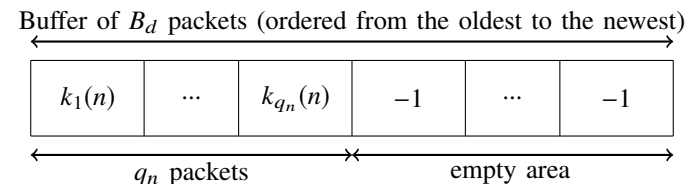
Buffer of B_d packets (ordered from the oldest to the newest)

Fig. 1. Buffer configuration at slot n.

C. Channel model

We consider a single user flat-fading channel with signal bandwidth W (Hz) and additive white Gaussian noise with zero mean and variance N_0. During time slot n, the channel remains constant with complex-valued amplitude h, and varies in a i.i.d. manner across time slots. We define the channel gain as $g_n = |h_n|^2$, where g_n is a continuous random variable distributed exponentially with probability density function $p(g) = \frac{1}{\xi} e^{-\frac{g}{\xi}}$ with mean ξ. For the sake of simplicity, we assume only quantized channel state $x_n = Q_g(g_n)$, where $Q_g(.)$ represents the quantization process [2]. Fixing a sequence of fading power quantization thresholds, the channel gain x_n is then a discrete variable taking values from a finite channel state space $\mathcal{X}$.

---

[1]There is a huge amount of literature assuming i.i.d EH processes. We adopted this approach for sake of clarity. Nevertheless, this work can be easily extended to time-correlated EH processes. This is done in Section V to plot Fig. 11.

[2]This assumption is generally justified in practice to account for the capacity-limited feedback and has also been adopted in the literature [20]–[28].

In order to define the discrete channel states, let M be the number of quantization levels, $\{t_m\}_{m=0}^{M-1}$ the set of thresholds and $\{\mathcal{L}_m\}_{m=0}^{M-1}$ the set of quantization levels for Q_g. The quantization regions of the channels are then given by the intervals $I_m = [t_m, t_{m+1}[$ with t_0 is fixed such that the transmission of 1 packet using 1 e.u. is guaranteed and $t_M = +\infty$. In our model, we consider a *Uniform* quantizer. So, let $t_{\max} = t_{M-1}$ be the maximal threshold such that the transmitter can send U_0 packets using B_e e.u., where U_0 is the maximal value of scheduled packets. By applying $E(x_n, u_n) = B_e$ in (1) and $u_n = U_0$ in (2) (equations (1) and (2) are defined in next section II-A), we obtain the corresponding value for x_n which is forced to $t_{\max}$. The uniform quantization thresholds are given by $t_{m+1} = t_m + \delta$ with $m = 0 \ldots M - 2$ and $\delta = \frac{t_{\max}}{M-1}$. We select the quantization levels as the lower bound of the regions, which is the worst case scenario. Thus, $\mathcal{L}_m = t_m$ for $m = 0 \ldots M - 1$ and a channel is said to be in state $x_n = \mathcal{L}_m$ if $g_n \in I_m = [t_m, t_{m+1}[$.

Note that the defined quantization process and parameters are used by default for the perfect CSI scenario, thus the values of x correspond to the perfect discrete channel states. However, for the imperfect CSI, the channel is first estimated before being quantized. Let $\hat{h}_n$ and $\hat{g}_n = |\hat{h}_n|^2$ denote the estimated channel and the estimated channel gain. Then, the estimated discrete (quantized) channel states are defined by $\hat{x}_n$ accordingly. In this case, a channel is said to be in state $\hat{x}_n \neq x_n$ if $\hat{g}_n \in I_{m'}$ while $g_n \in I_m$ with $m' \neq m$.

D. Consumed energy

We denote u_n ($u_n \leq q_n$) the number of packets to be transmitted during time slot n of period T_s, through the channel of gain x_n for perfect CSI and the channel of gain $\hat{x}_n$ for imperfect CSI. In the former case, the consumed energy to transmit these packets is expressed as an integer multiple of the energy unit. It is given by

$$E(x_n, u_n) = \left\lceil \frac{P(x_n, u_n).T_s}{\mathcal{E}_U} \right\rceil, \tag{1}$$

where

$$P(x_n, u_n) = \frac{WN_0}{x_n}\left(2^{\frac{u_n L}{WT_s}} - 1\right) \tag{2}$$

is the required power for this transmission. In the latter case, similar expressions are obtained by replacing x_n by the estimated channel gain $\hat{x}_n$ and T_s by $T_s - \tau$ where τ is the time required to perform channel estimation.

III. PROBLEM FORMULATION AND RESOLUTION IN PERFECT CSI SCENARIO

In this section, we assume first perfect CSI at the transmitter without any cost. Our main objective now is to ensure reliable communication by minimizing the number of discarded packets due to strict delay and buffer overflow constraints. This can be achieved by finding an optimal policy that specifies the number of packets u to be scheduled at each time slot based on the past system states and actions. The optimization problem

can be formulated as MDP problem [30]. We characterize in this section the appropriate states, actions and reward of this MDP.

A. State Space

The state space $\mathcal{S}$ is the set of $\mathbf{s} = (\mathbf{k}, b, x)$ where
- $\mathbf{k} = [k_1, \cdots, k_{B_d}]$ is the vector indicating the age of each packet in the data buffer,
- b is the battery level, and
- x is the channel gain.

Notice that in the previous works [9], [10], the queue length q describes the data buffer states. In our work, q is replaced with $\mathbf{k}$ due the strict delay constraint. In fact, q is unnecessary when $\mathbf{k}$ is given since

$$q_n = \max\{i \mid k_i(n) \geqslant 0\}. \tag{3}$$

The state space is finite, and the total number of possible states is $|\mathcal{S}|$ which is upper-bounded by $(K_0 + 2)^{B_d}.|B_e + 1|.|\mathcal{X}|$. The state space can be significantly reduced by assuming that packets are queued in an increasing order of time spent in the buffer, i.e. $k_1(n) \geqslant k_2(n) \geqslant \cdots \geqslant k_{q_n}(n)$. For instance, if we consider $B_d = 6$, $K_0 = 3$, $B_e = 4$ and $|\mathcal{X}| = 5$, the upper-bound is 390625 while our system only has 5250 states by removing all the impossible combinations in $\mathbf{k}$.

B. Action Space

The action space $\mathcal{U}$ denotes the number of packets u that the transmitter can send during a time slot. This space is finite and the number of actions is $|\mathcal{U}| = U_0 + 1$.

C. Markov Decision Process

On one hand, during time slot n, $w_n = \max(u_n, m_n)$ packets leave the buffer, either transmitted and/or discarded where u_n is the number of transmitted packets and m_n is the number of packets with delay K_0 slots in the buffer[3]. The age of the remaining packets in the buffer is incremented by 1. Moreover, a_{n+1} new packets arrive to the buffer with age 0. Therefore, the vector $\mathbf{k}$ can be updated from slot n to slot $n+1$ according to the following rule.

1: **for** $i = 1$ **to** $q_n - w_n$ **do**
    $k_i(n+1) = k_{w_n+i}(n) + 1$
    end for
2: **for** $i = q_n - w_n + 1$ **to** $q_n - w_n + a_{n+1}$ **do**
    $k_i(n+1) = 0$
    end for
3: **for** $i = q_n - w_n + a_{n+1} + 1$ **to** B_d **do**
    $k_i(n+1) = -1$
    end for

On the other hand, during time slot n, e_{n+1} e.u are harvested and stored in the battery and E_n e.u are removed from the

---

[3]Indeed, if $m_n \geq u_n$, u_n packets are sent on the channel and only $(m_n - u_n)$ packets will be discarded before time $n + 1$ since their delay will be $K_0 + 1$: therefore, m_n packets are removed from the buffer. If $m_n < u_n$, any packet with an age of K_0 will be sent during this slot which implies that the packets removing from the buffer are only the u_n transmitted packets.

battery to schedule u_n packets. Therefore, at time slot $n + 1$, the battery state is updated according to

$$b_{n+1} = \min \{b_n - E_n + e_{n+1}, B_e\}. \tag{4}$$

We thus remark that $\mathbf{k}_{n+1}$ (resp. b_{n+1}) only depends on previous state $\mathbf{k}_n$ (resp. b_n), action u_n (resp. E_n) and external perturbation a_{n+1} (resp. e_{n+1}). Therefore, we can define $p(\mathbf{s}'|\mathbf{s}, u)$ as the transition probability to fall in the future state $\mathbf{s}' = (\mathbf{k}', b', x')$ after taking action u in the current state $\mathbf{s} = (\mathbf{k}, b, x)$. Assuming that the buffer, battery and channel states are independent and channel states are not correlated, the transition probability satisfies the following equation.

$$p(\mathbf{s}'|\mathbf{s}, u) = p(\mathbf{k}'|\mathbf{k}, b, u).p(b'|b, x, u).p(x'), \tag{5}$$

where $p(x')$ is the distribution of the channel states, $p(\mathbf{k}'|\mathbf{k}, b, u)$ indicates the probability transitions between buffer states, and $p(b'|b, x, u)$ indicates the probability transitions between battery states. After tedious but simple derivations, we obtain the transitions between the buffer states and the battery states according to the following respective rules.

1: **if** $u > q$ **or** $k_i' > k_i + 1$ **or** $q' < q - w$ **then**
   $p(k_i'|k_i, b, u) = 0$
2: **else if** $k_i' \neq k_{i+u} + 1$ **and** $k_{i+u} \neq -1$ **then**
   $p(k_i'|k_i, b, u) = 0$
3: **else if** $k_i' > 0$ **and** $k_{i+u} = -1$ **then**
   $p(k_i'|k_i, b, u) = 0$
4: **else if** $q = B_d$ **and** $u \neq 0$ **and** $k_i' > 0, \forall i \in \{q - w + 1, ..., B_d\}$ **then**
   $p(k_i'|k_i, b, u) = 0$
5: **else if** $q' < B_d$ **then**
   $p(k_i'|k_i, b, u) = e^{-\lambda_d} . \frac{(\lambda_d)^a}{a!}$
6: **else**
   $p(k_i'|k_i, b, u) = 1 - Q(B_d - q + w, \lambda_d),$

and

1: **if** $E > b$ **then**
   $p(b'|b, x, u) = 0$
2: **else if** $b' < b - E$ **then**
   $p(b'|b, x, u) = 0$
3: **else if** $b' < B_e$ **then**
   $p(b'|b, x, u) = e^{-\lambda_e} . \frac{(\lambda_e)^e}{e!}$
4: **else**
   $p(b'|b, x, u) = 1 - Q(B_e - b + E, \lambda_e).$

where $Q(\bullet, \bullet)$ is the regularized Gamma function.

D. Markov Decision Problem and its Resolution

In the context of infinite-horizon MDP, we consider time-averaged cost, where at a given time slot $n \in \{0, \cdots, N\}$, the system state is denoted by $\mathbf{s}_n = (\mathbf{k}_n, b_n, x_n)$ and $\mu(\mathbf{s}_n) = u_n$ is the action deciding the number of packets to be transmitted. We aim at finding the optimal policy $\mu^\star$ that minimizes the average number of dropped packets. The cost function of this infinite-horizon MDP problem is given by

$$\overline{D}(\mu) = \lim_{N \to +\infty} \frac{1}{N} \mathbb{E}^\mu \left[\sum_{n=1}^{N} \left(\varepsilon_d(\mathbf{s}_n, u_n) + \varepsilon_o(\mathbf{s}_n, u_n) \right) \right], \tag{6}$$

where $\mathbb{E}$ is the expectation with respect to the policy μ and where $\varepsilon_d(\mathbf{s}_n, u_n)$ is the instantaneous number of discarded packets due to delay violation and $\varepsilon_o(\mathbf{s}_n, u_n)$ is the expected number of discarded packets due to buffer overflow. According to [29], we know that finite-state MDP without additional constraint exhibits an optimal deterministic policy. Thus, the function μ is a deterministic policy and $\mu^\star$ is the optimal deterministic policy to be found.

At a given slot n, when the system state is $\mathbf{s}_n$ and the performed action is u_n, the number of discarded packets due to delay violation is given by

$$\varepsilon_d(\mathbf{s}_n, u_n) = \begin{cases} 0 & \text{if } m_n = 0 \text{ or } m_n \leqslant u_n \\ m_n - u_n & \text{otherwise.} \end{cases} \tag{7}$$

The buffer overflow occurs when $q_n - w_n + a_{n+1} > B_d$, thus the number of discarded packets due to buffer overflow is obtained as follows

$$\begin{aligned} \varepsilon_o(\mathbf{s}_n, u_n) &= \sum_{a=B_d-q_n+w_n+1}^{+\infty} (q_n - w_n + a - B_d).e^{-\lambda_d}. \frac{(\lambda_d)^a}{a!} \\ &= \lambda_d.(1 - Q(B_d - q_n + w_n, \lambda_d)) \\ &+ (q_n - w_n - B_d) \\ &\times (1 - Q(B_d - q_n + w_n + 1, \lambda_d)). \end{aligned} \tag{8}$$

We need to consider an expected reward for the buffer overflow since at the beginning of the slot (when the decision is made), the number of incoming packets is only known statistically.

Finally, our MDP optimization problem can be stated as
   Problem 1:

$$\mu^\star = \arg \min_\mu \overline{D}(\mu) \tag{9}$$

We know that $\mu^\star$ exists [29] and can be found via an offline dynamic programming approach using, for instance, the so-called VI algorithm [30]. Exploring statistical *a priori* knowledge of energy arrival and data arrival dynamics and channel states at the EH transmitter, the offline approach can accurately model the state transition probabilities of the MDP and provide an optimal solution. The optimal offline deterministic policy $\mu^\star$ for Problem 1 can be computed through Algorithm 1.

IV. IMPERFECT CSI SCENARIO

In wireless communication systems, channel state information is not perfectly known at the transmitter and can include errors due to the channel estimation process. Indeed, in a Time Division Duplex (TDD) UL transmission between an EH device and a base station, the CSI can be obtained at the EH device by first estimating the channel at the base station via an UL training process and then feeding back a quantized version of the estimate to the transmitter. We assume that the feedback channel is error-free and instantaneous as soon as the receiver has estimated the channel. Therefore, accounting for the channel estimation phase, the time slot structure is

---
Algorithm 1 VI algorithm

1: Initialization
   Set $v^0(s) = 0 \; \forall s \in S$
   Fix a tolerance parameter $\epsilon > 0$
   Set $n = 1$
2: For each $s \in S$ compute

$$v^n(s) = \min_{u \in U} \left[c(s,u) + \sum_{s' \in S} P(s'|s,u).V^{n-1}(s') \right] \quad (10)$$

$$V^n(s) = v^n(s) - v^n(s_0) \quad (11)$$

   where $c(s,u)$ is the instantaneous cost and s_0 is a fixed
   state chosen arbitrarily.
3: If $\mathrm{sp}(V^n - V^{n-1}) < \epsilon$, where $\mathrm{sp}(V) = \max_{s \in S} V(s) - \min_{s \in S} V(s)$, let π^ϵ be the resulting policy that solves
   equation (10) and stop; else set $n = n + 1$ and go to step
   2.

---

divided into two parts: a duration of τ ms to acquire CSI at the
mobile device and the remaining $(T_s - \tau)$ ms to schedule data
packets. In particular, the EH device exploits the acquired CSI
to send data whenever scheduling decisions are made. In this
section, we aim at evaluating the optimal policy $\mu^\star$ obtained
with Algorithm 1 when the CSI are imperfect which means
that the current states used for computing the output of $\mu^\star$ are
not necessary correct.

A. Channel estimation

At τ ms after the beginning of time slot n, we consider that
the EH mobile device has an estimated discrete channel state
$\hat{x}_n$ as described in Section II-C. This estimated channel can be
obtained through a training sequence of η pilot symbols using
a total training power P_{tr} during the period τ of the time slot.
Then, the required energy to perform this channel estimation
is

$$E_{ce}(\hat{x}_n) = \left\lceil \frac{P_{\mathrm{tr}}.\tau}{\mathcal{E}_U} \right\rceil. \quad (12)$$

Due to the imperfect channel estimation, we have

$$\hat{h}_n = h_n + e_{h_n}, \quad (13)$$

where e_{h_n} is the estimation error independent of h_n and it
is a zero-mean i.i.d. complex-valued Gaussian process with
variance σ_e^2 per complex dimension. According to [31], this
error variance can be expressed in terms of energy per pilot
symbol E_s, the number of pilot symbols used for estimation
η and the Gaussian noise variance per complex dimension σ_w^2
as

$$\sigma_e^2 = \mathbb{E}[|\hat{h} - h|^2] = \frac{\sigma_w^2}{\eta E_s} = \frac{N_0}{\tau P_{\mathrm{tr}}}. \quad (14)$$

Given the channel gain g_n, the estimated channel gain $\hat{g}_n = |h_n + e_{h_n}|^2$ is a non central χ^2 random variable with 2 degrees
of freedom in which the Gaussian variables are independent
with common variance $\sigma_e^2/2$ and mean $g_n = |h_n|^2$. It has a
probability density function (PDF) of the form

$$P_{\hat{G}|G}(\hat{g}|g) = \frac{1}{\sigma_e^2} e^{-\frac{g+\hat{g}}{\sigma_e^2}} I_0\left(\frac{2}{\sigma_e^2} \sqrt{g\hat{g}} \right), \quad (15)$$

where I_0 is the zero-order modified Bessel function of the first
kind [32].

B. Error probability and packet loss rate

In this section, we analyze the impact of channel estimation on
the system performance, in particular on the packet loss rate.
In fact, channel estimation can affect the number of discarded
packets in three ways. Firs of all, the transmission period is
reduced which offers less time to transmit the same amount
of data. On one hand, if the channel estimate is smaller than
the actual channel, less packets can be scheduled at decision
instants. Thus, more packets can be queued in the data buffer
with higher delays, and may lead to more delay violation and
buffer overflow occurrences. On the other hand, if the channel
estimate is higher than the actual channel, the scheduled
packets are all dropped. This latter condition incurs additional
loss rate besides the delay violation and buffer overflow losses
given in equations (7) and (8). Therefore, we need to take into
account such errors in the total error probability. This extra
error probability (called, channel mismatch probability in the
rest of the paper) can be expressed as

$$P_{e,\mathrm{CSI}} = \mathrm{Prob}\ (\hat{x} > x) = \sum_{m'|m'>m} \mathrm{Prob}\ (\hat{g} \in I_{m'}, g \in I_m), \quad (16)$$

where $I_m = [t_m, t_{m+1}[$ and $I_{m'} = [t_{m'}, t_{m'+1}[$, $m = 0, \dots, M-1$,
$m' > m$ are the quantization regions of the perfect channel
state and the estimated channel state, respectively.

Then, using Bayes rule and some derivations, we can compute

$$
\begin{aligned}
\mathrm{Prob}(\hat{g} \in I_{m'}, g \in I_m) &= \int_{\hat{g} \in I_{m'}} \int_{g \in I_m} P(\hat{g}, g)\, d\hat{g}\, dg \\
&= \int_{\hat{g} \in I_{m'}} \int_{g \in I_m} P_{\hat{G}|G}(\hat{g}|g) P_G(g)\, d\hat{g}\, dg \\
&= \int_{g \in I_m} \left(\int_{\hat{g} \in I_{m'}} P_{\hat{G}|G}(\hat{g}|g)\, d\hat{g} \right) P_G(g)\, dg \\
&= \int_{t_m}^{t_{m+1}} \left(Q_1\left(\frac{\sqrt{2g}}{\sigma_e}, \frac{\sqrt{2t_{m'}}}{\sigma_e} \right) \right. \\
&\quad - \left. Q_1\left(\frac{\sqrt{2g}}{\sigma_e}, \frac{\sqrt{2t_{m'+1}}}{\sigma_e} \right) \right) P_G(g)\, dg \quad (17)
\end{aligned}
$$

where $P_{\hat{G}|G}(\hat{g}|g)$ is given in (15), Q_1 is the Marcum function,
and $P_G(g) = \frac{1}{\sigma_h^2} e^{-\frac{g}{\sigma_h^2}}$ for $g \geq 0$ and 0 otherwise is the PDF
of the channel gain.

At a given time slot n, when the action u_n is done by apply-
ing the optimal policy $\mu^\star$ (obtained for the perfect channel
knowledge case) on the estimated channel state $\hat{x}_n > x_n$, the
number of discarded packets due to CSI errors is computed as

$$\varepsilon_e(u_n, P_{e,\text{CSI}}) = u_n \times \mathbb{1}(P_{e,\text{CSI}} \neq 0), \tag{18}$$

and the cost function of our MDP problem under policy $\mu^\star$ and imperfect CSI is given by

$$\overline{D'}(\mu^\star) = \lim_{N \to +\infty} \frac{1}{N} \mathbb{E}^{\mu^\star} \left[\sum_{n=1}^{N} \Big(\varepsilon_d(\mathbf{s}_n, u_n) + \varepsilon_o(\mathbf{s}_n, u_n) + \varepsilon_e(u_n, P_{e,\text{CSI}}) \Big) \right]. \tag{19}$$

V. NUMERICAL RESULTS

We evaluate numerically the optimal policy obtained by resolving Problem 1. We consider a system as described in Section II with the following characteristics: the slot duration is $T_s = 1$ ms and the maximum delay is $K_0 = 3$ (i.e., in absolute time $K_0 T_s = 3$ ms). Energy arrivals follow a Poisson distribution with mean λ_e e.u per slot, where $\mathcal{E}_U = 100$ nJ. Energy units are stored in a battery of size $B_e = 4$ e.u. The maximum available power at the transmitter is $P_{\max} = 0.4$ mW. Data arrivals follow a Poisson distribution with mean λ_d packets, where packets are of equal size $L = 5000$ bits. Data packets are stored in a buffer of size $B_d = 6$ packets. Limited by the queue size, we fix $U_0 = 6$ packets per slot. The mean channel is $\xi = 1$ and the channel states x takes 5 possible values (expressed in dB) from the finite set $\mathcal{X} = \{-10, -3.98, -0.97, 0.792, 2.04\}$. These channel values are obtained according to Section II-C. The noise power spectral density is $N_0 = -87$ dBm/Hz and the allocated bandwidth is $W = 5$ MHz.

A. Perfect CSI

In this section, we consider that the transmitter has a perfect knowledge of the channel state without any cost.

In Fig. 2, we plot the average number of discarded packets versus the number of iterations for evaluating the optimal policy obtained by the VI algorithm for various energy arrival rates λ_e where the data arrival rate λ_d is fixed to 1.5. We show that the VI algorithm converges rapidly within a few hundreds iterations for most cases. We can also notice that as λ_e increases, the average number of discarded packets considerably decreases. Indeed, when the available energy from the surrounding environment is larger, the system is able to send more packets, reducing thus the number of discarded packets.

In Fig. 3, we display the percentage of discarded packets versus the data arrival rate λ_d for different energy arrival rates for two policies. The first policy is the (deterministic offline) optimal one introduced in this paper and obtained after convergence of the VI algorithm. The second policy is a *naive* one in which we force the transmitter to send the maximum number of packets using the available energy in the battery. As we can observe, the proposed optimal policy provides significantly better performance than the naive one in terms of percentage of discarded packets. In
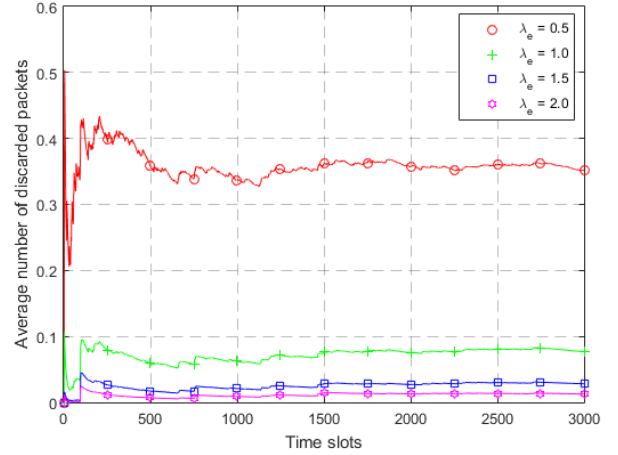

Fig. 2. Convergence analysis for the average number of discarded packets with different energy arrival rates.

fact, this policy enables us to adapt the transmission rate according to the buffer, battery and channel conditions. In addition, we remark that the number of discarded packets increases when the data arrival rate λ_d increases because the buffer overflow could happen more often. On the one hand, when the energy available to scavenge is low (small λ_e), an efficient energy management becomes crucial to ensure the sustainability of the system, and the gap between both policies increases. On the other hand, when a large amount of energy is available (large λ_e), the system can survive even without controlling relevantly the energy consumption which leads to similar performance between the optimal and naive policies.
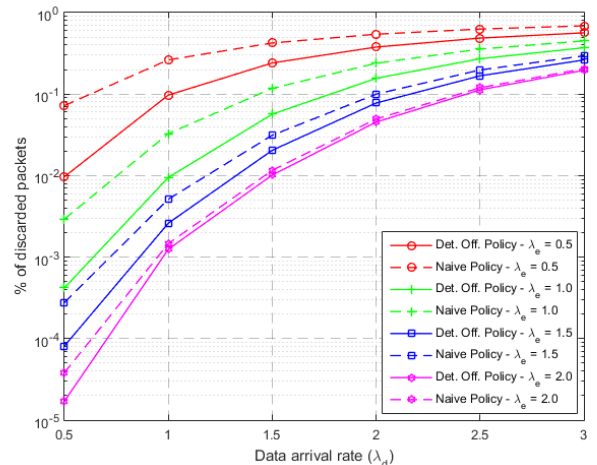

Fig. 3. Percentage of the discarded packets versus data arrival rate with different energy arrival rates.

Similar to Fig. 3, Fig. 4 compares the percentage of discarded packets of the optimal policy with two other variants of the naive policy. Unlike the naive policy that sends the maximum number of packets using the available energy in the battery, the introduced *p*-Naive policy restricts the number of packets

sent by the naive one by taking the buffer state into account through an additional parameter p in that way:

- **fixed** p: The policy sends only the packet i from the buffer if $k_i \geq p$
- **variable** p: The policy performs a first step similar to the previous case (fixed p). If no packet satisfies the condition, p is decreased by 1, and the first step is repeated, until $p = 0$.

The naive policy corresponds to a 0-Naive policy. Here, we choose $p = 2$ for the p-naive policy. As we can see, taking only the age of the packets into the buffer without adapting carefully the number of packets by the energy battery level and buffer state leads to decrease the number of sent packets, and therefore the naive policy remains much better.

Fig. 5. Percentage of the discarded packets due to delay violation versus data arrival rate and energy arrival rate with the optimal policy.

Fig. 4. Percentage of the discarded packets versus data arrival rate with different energy arrival rates and different naive policies.

Fig. 6. Average consumed energy versus data arrival rate with different energy arrival rates.

In Fig. 5, we show the percentage of discarded packets due to delay violation among all the discarded packets for the optimal policy with different data arrival rates λ_d and energy arrival rates λ_e. As explained before, a packet can be discarded due to either delay violation or buffer overflow. When the data arrival rate increases, the probability to discard a packet due to buffer overflow increases which decreases the contribution of the delay violation in the discarded packets. When the energy arrival rate decreases, the percentage of discarded packets due to the delay violation slightly increases because, in average, a packet remains more often in the buffer since there is no energy enough to transmit it. Hence, it is flushed from the buffer for latency's purpose.

In Fig. 6, we plot the average consumed energy versus the data arrival rate λ_d with different energy arrival rates λ_e. We observe that the optimal policy consumes less energy than the naive one while sending more packets because it adapts the number of transmitted packets per slot to the channel conditions and the battery state and thus, the transmission is done according to the energy it consumed.
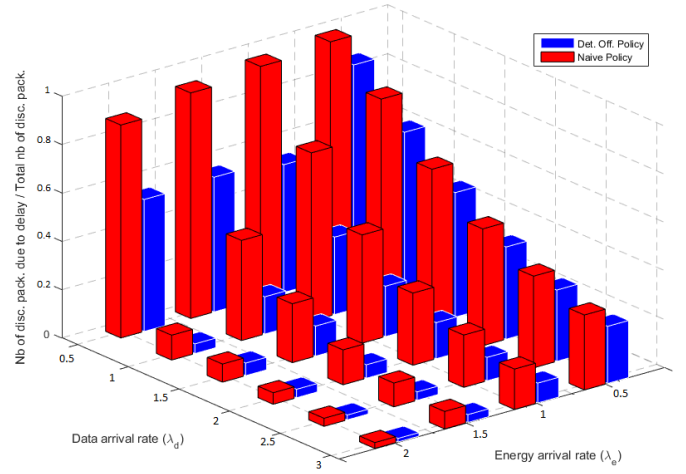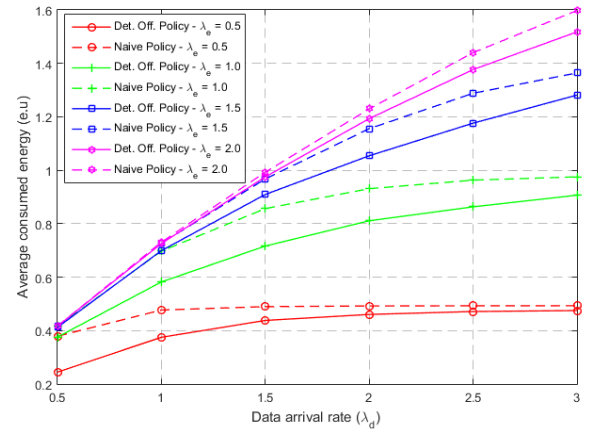
In Fig. 7, we show the average battery state versus the packet arrival rate λ_d with different energy arrival rates λ_e. As the optimal policy offers a lower energy consumption (see Fig. 6), the battery is less used and its energy level is thus higher. This ensures a better sustainable communication with less number of discarded packets.

In Fig. 8, 9 and 10, we compare the performance of our optimal policy to the optimal policy obtained by forcing the average (instead of the strict) delay to be less than a pre-defined threshold. Both policies are applied assuming buffer overflow and delay violation as the way to drop the packets, but the second policy is optimized just in order to minimize the buffer overflow and keep an average delay small enough. Therefore, this policy is obtained as follows: according to the Little's law, we propose to convert the average delay constraint D_{ct} into an average queue length constraint Q_{ct} since $Q_{ct} = \lambda_d.D_{ct}$ where λ_d is the data arrival rate. The policy ensuring a bounded average delay can be found by solving the following Constrained Markov Decision Problem
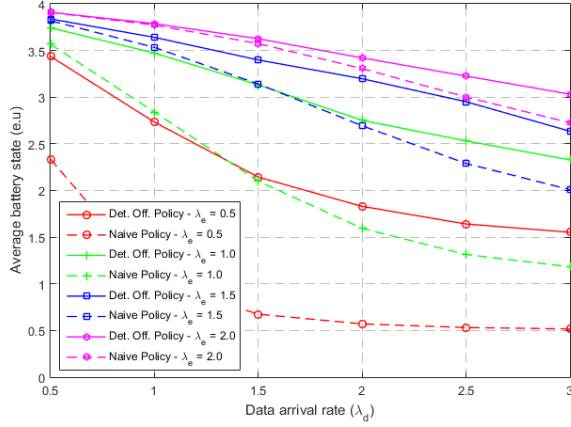
Fig. 7. Average battery state versus data arrival rate with different energy arrival rates.

(CMDP):

$$\tilde{\mu}_{Q_{ct}} = \arg \min_{\mu} \quad \lim_{N \to +\infty} \frac{1}{N} \mathbb{E}^{\mu} \left[\sum_{n=1}^{N} \varepsilon_o(\mathbf{s}_n, u_n) \right] \quad (20)$$

$$\text{s.t.} \quad \lim_{N \to +\infty} \frac{1}{N} \mathbb{E}^{\mu} \left[\sum_{n=1}^{N} q_n \right] \leq Q_{ct} \quad (21)$$

where q_n is the queue length. Notice that we do not consider the delay violation for this optimization since the strict delay is not taken into account in this policy as we just force the average delay to be less than a threshold. So the policy $\tilde{\mu}_{Q_{ct}}$ is done to handle properly the average delay and not the strict delay.

Our optimal policy adapted to strict delay has been computed with $K_0 = 3$. In order to compare both policies in the strict delay constraint set up (it means that the packet is dropped if the delay is strictly larger than K_0 even if we apply the policy $\tilde{\mu}_{Q_{ct}}$), we need to choose properly D_{ct}. It makes sense to force $D_{ct} \leq 3$ in order to have a small amount of dropped packets due to delay violation. As $D_{ct} = 2$ or $D_{ct} = 3$ have led to similar performance, we have fixed $D_{ct} = 3$.

As we can see, our policy outperforms the policy considering only the average delay in terms of percentage of discarded packets, consumed energy (in most cases), and battery levels (in most cases). So, it was worth to do the effort to optimize the policy by taking into account the strict delay into the state model rather than just using the optimal policy adapted to the average delay with a well-tuned threshold.

We now consider that the EH process is time-correlated. In order to cast this assumption into an MDP framework, we need to add EH process e to the state of the system, i.e., $\mathbf{s} = (\mathbf{k}, b, e, x)$ instead of $(\mathbf{k}, b, x)$ as done previously. Then, a new optimal policy taking into account the EH correlation is re-computed by using the same tool, i.e., the VI algorithm. Here, we assume that the transition probability of the Markov Chain satisfies the following equation
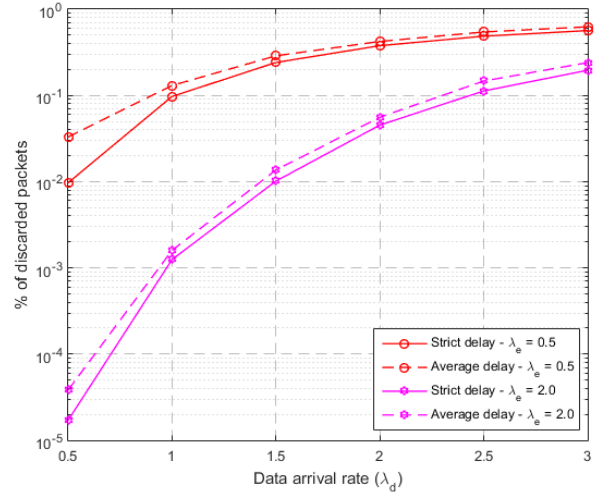

Fig. 8. Percentage of the discarded packets versus data arrival rate with different energy arrival rates between strict and average delay policies($\mu^\star$ and $\tilde{\mu}_3$, respectively).

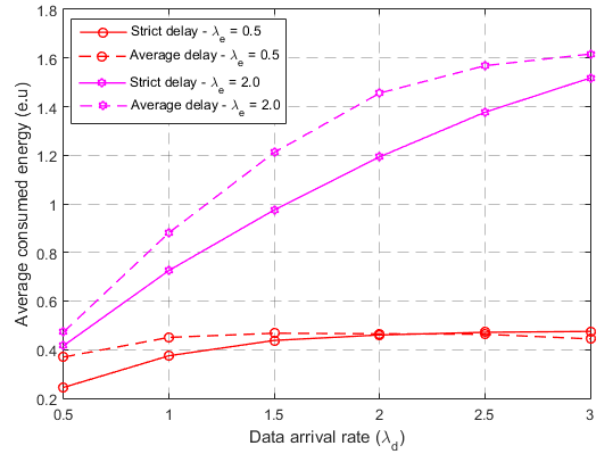Fig. 9. Average consumed energy versus data arrival rate with different energy arrival rates between strict and average delay policies ($\mu^\star$ and $\tilde{\mu}_3$ respectively).

$$p(\mathbf{s}'|\mathbf{s}, u) = p(\mathbf{k}'|\mathbf{k}, b, u) . p(b', e'|b, e, x, u) . p(x'), \quad (22)$$

where $p(b', e'|b, e, x, u)$ is obtained according to the following rules:

1: **if** $E > b$ **then**
   $p(b', e'|b, e, x, u) = 0$
2: **else if** $b' < b - E$ **then**
   $p(b', e'|b, e, x, u) = 0$
3: **else if** $b' = \min(B_e, b - E + e)$ **then**
   $p(b', e'|b, e, x, u) = p(e'|e)$
4: **else**
   $p(b', e'|b, e, x, u) = 0.$

In addition, the transition probability from a energy arrival state j at time slot n to another energy arrival state i at time slot $n + 1$ is given by
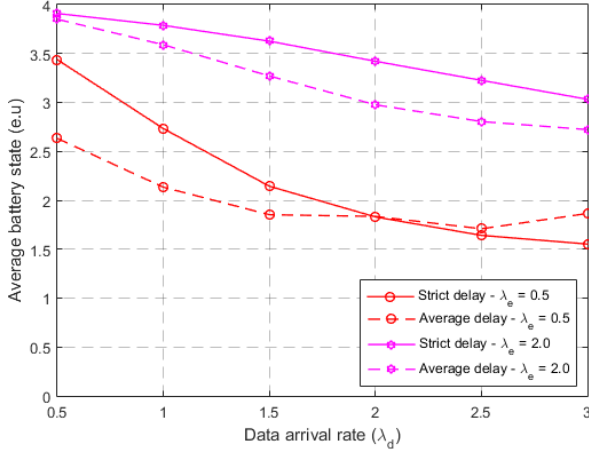
Fig. 10. Average battery state versus data arrival rate with different energy arrival rates between strict and average delay policies ($\mu^\star$ and $\tilde{\mu}_3$ respectively).

$$p(e_{n+1} = i|e_n = j) = \frac{(1 - \rho_e)^{|i-j|}}{\sum\limits_{k=0}^{|\mathcal{H}_e|-1} (1 - \rho_e)^{|k-j|}} \qquad (23)$$

where ρ_e is the so-called correlation factor and $\mathcal{H}_e$ is the set of potential energy units harvested during one slot.

In Fig. 11, we compare the performance of the optimal policy (adapted to time-correlated EH process) with the naive policy. We set $\mathcal{H}_e = \{0, 1, 2\}$ e.u. per slot.
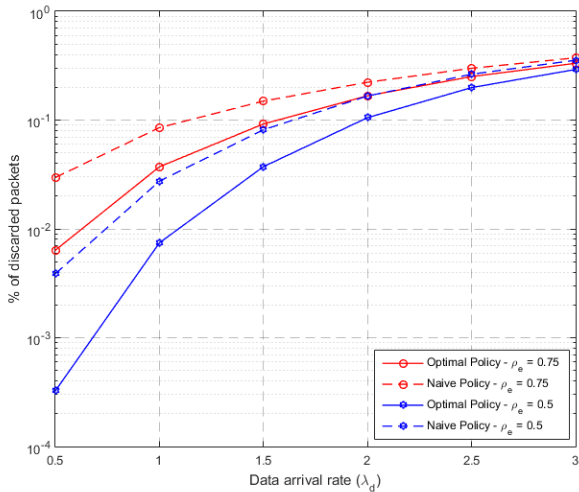

Fig. 11. Percentage of the discarded packets versus data arrival rate with different correlated energy arrival rates between strict and naive policies.

The proposed optimal policy is still better than the naive policy. The performance of the system decreases when ρ_e increases because the system will be trapped in the state $e = 0$ for a longer period of time, leading to more discarded packets.

B. Imperfect CSI

In this section, our goal is to evaluate the proposed optimal policy when the transmitter relies on an estimated version of the channel state. The estimation phase duration is equal to $\tau = 10$ μs (1% of T_s), and a power of $P_{tr} = 4$ mW is used. The corresponding energy consumption for the estimation phase is thus $E = 40$ nJ which can be neglected to the energy unit, and therefore we assume $E_{ce} = 0$ e.u..

In Fig. 12, we compare the percentage of discarded packets between perfect and imperfect CSI scenarios. For low data arrival rate λ_d, the gap between both scenarios is large. Indeed, in our set up, the smallest channel mismatch probability is between 10^{-3} and 10^{-2} which implies that the percentage of discarded packets is necessary worse since as soon as the channel is over-estimated, the packets are dropped. However, when the data arrival rate increases, the buffer overflow can happen more often and the channel mismatch probability has less impact, which lead both scenarios to behave similarly.
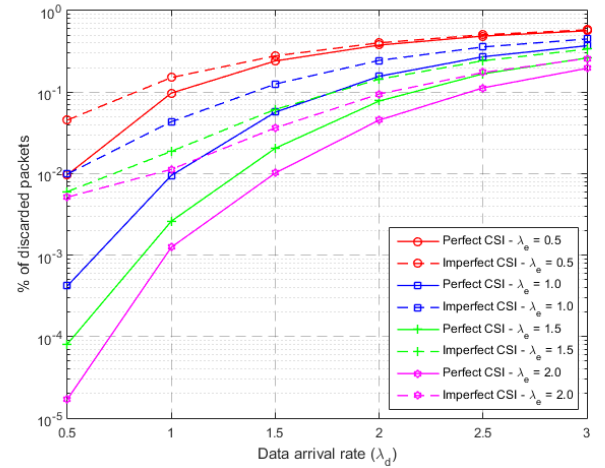

Fig. 12. Percentage of the discarded packets versus data arrival rate with different energy arrival rates between perfect and imperfect CSI scenarios.

In Fig. 13, we compare the optimal and naive policies under perfect and imperfect CSI scenarios. For small energy arrival rate λ_e, the optimal policy under imperfect CSI is better than the naive policy with perfect CSI, because the latter sends packets without any adaptation to the energy and data arrivals, so energy shortage can happen more often and the number of discarded packets increases. For high energy arrival rate, imperfect CSI has stronger impact since the energy has to be controlled in a smarter way and knowing the channel accurately is more required.

In Fig. 14, we compare the percentage of discarded packets for different estimation times τ (expressed in % of T_s). For low data arrival rate λ_d, increasing the estimation time leads to a better channel estimation, which slightly reduces the number of discarded packets since the impact of estimation error is high in this configuration (see Fig. 12). Nevertheless, after a certain threshold, for instance $\tau \approx 5\%$, the number
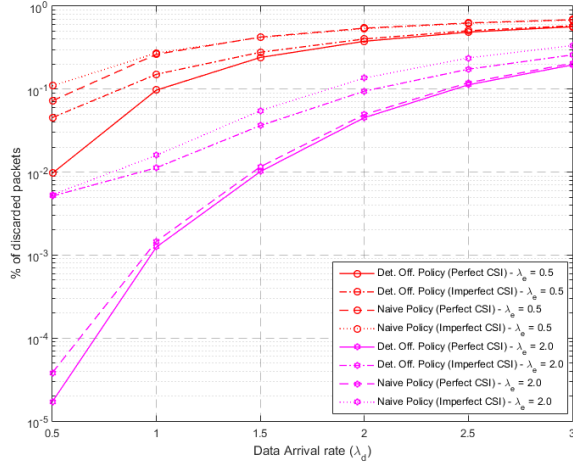
Fig. 13. Percentage of the discarded packets versus data arrival rate with different energy arrival rates for different policies and between perfect and imperfect CSI scenarios.

of discarded packets will increase because the remaining communication time of the slot is smaller. This leads to decrease the number of sent packets and so to increase the number of packets into the buffer, exhibiting thus more delay violation and buffer overflow. For high data arrival rate, we know that the estimation accuracy is not required (see Fig. 12). Therefore, increasing the estimation time directly decreases the performance since the system has less time for data packets transmission.
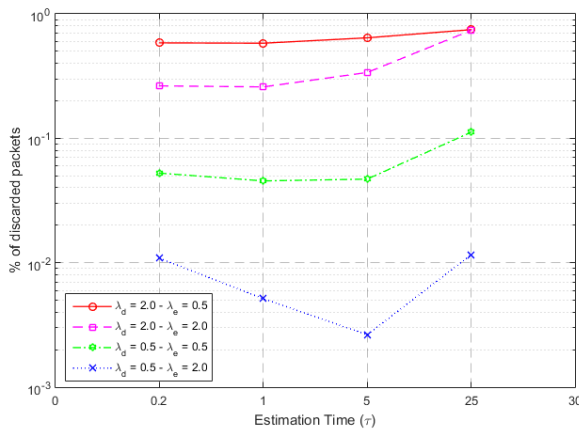

Fig. 14. Percentage of the discarded packets versus the estimation time τ (expressed in % of T_s) with different data and energy arrival rates.

In Fig. 15, we display the nature of discarded packets in percentage due to delay violation, buffer overflow and channel mismatch with different data and energy arrival rates. The number of discarded packets due to channel mismatch is significant for low data arrival rate because the delay violation or the buffer overflow can happen less often. However, for high data arrival rate, the number of packets discarded due to channel mismatch is negligible

and the policy behaves approximately in the same way for perfect and imperfect CSI. Nevertheless, the imperfect CSI degrades the whole system (on the delay violation and buffer overflow) since a part of the time slot is now devoted to perform the estimation rather than the transmission.

(a) $\lambda_d = 0.5$ - $\lambda_e = 0.5$ (b) $\lambda_d = 3.0$ - $\lambda_e = 0.5$

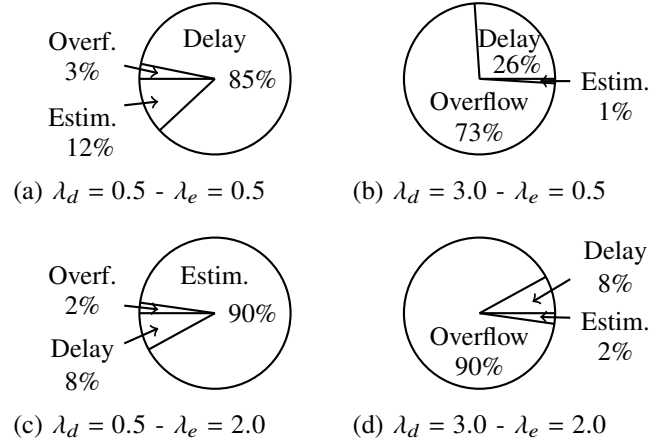(c) $\lambda_d = 0.5$ - $\lambda_e = 2.0$ (d) $\lambda_d = 3.0$ - $\lambda_e = 2.0$

Fig. 15. Percentage of the discarded packets due to delay violation, buffer overflow, and channel mismatch with different data and energy arrival rates.

Under imperfect CSI assumption, it is usual to allow packet re-transmission through an Hybrid Automatic Repeat ReQuest (HARQ) protocol instead of trashing the packet once sent [33]. But adapting our work to HARQ requires a huge modification of the MDP framework. Here, we just run our policy (the optimal one described in Section III) when ARQ and Chase Combining HARQ (CC-HARQ) protocols are carried out. The only modification is to keep the packet in the buffer at the end of the ARQ process instead of wasting it. So there is a trade-off between the higher probability for each packet to be correctly decoded at the receiver, the higher duration for the packet to stay in the buffer while waiting for the feedback, the higher energy consumed for re-transmitting the packet. In Fig. 16, ARQ and CC-HARQ are implemented with at most two transmissions (one re-transmission is allowed only). When λ_e is low, using ARQ and CC-HARQ is not efficient because re-transmitting the same packet twice consumes energy while it is not available in large quantities. However, when λ_e is large, these two protocols significantly reduce the number of discarded packets due to imperfect CSI.

VI. CONCLUSION

We have addressed resource scheduling problem under energy harvesting capabilities with strict delay constraint and perfect CSI. More precisely, we have solved the packet loss optimization problem using MDP framework and dynamic programming techniques. The optimal policy adapted the number of transmitted packets according to the channel conditions, the available energy in the battery, and the battery level such that the number of discarded packets is minimized. We have compared our proposed strict delay based policy with different variants of a naive policy and the state-of-the-art policy relying only on the average delay, showing significant savings in packet loss and energy consumption. Finally, we
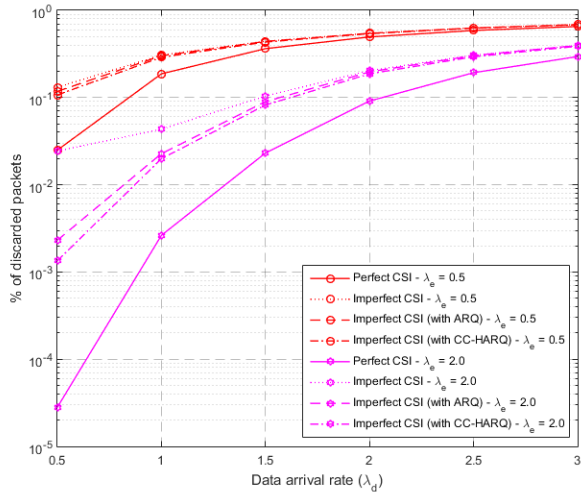
Fig. 16. Percentage of the discarded packets versus data arrival rate with different energy arrival rates between perfect and imperfect CSI scenarios.

have evaluated the impact of imperfect CSI without and with ARQ protocols on the optimal policy in terms of additional packet loss due to the channel estimation time and errors. As perspectives, we aim i) to use Deep Reinforcement Learning (DRL) techniques to deal with the curse of dimensionality, ii) to include offloading capabilities, where the system can choose to execute packets locally, offload it to nearby servers or base stations having more resources according to the buffer, available energy, and channel conditions under unknown CSI.

References

[1] I. Fawaz, M. Sarkiss and P. Ciblat, "Optimal Resource Scheduling for Energy Harvesting Communications under Strict Delay Constraint," IEEE International Conference on Communications (ICC), May 2018.

[2] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus and A. Yener, "Transmission with Energy Harvesting Nodes in Fading Wireless Channels: Optimal Policies," IEEE Journal on Selected Areas in Communications, vol. 29, no. 8, pp. 1732-1743, September 2011.

[3] B. Varan and A. Yener, "Energy Harvesting Communications with Energy and Data Storage Limitations," IEEE Global Communications Conference (GLOBECOM), December 2014.

[4] D. Gunduz, K. Stamatiou, N. Michelusi and M. Zorzi, "Designing Intelligent Energy Harvesting Communication Systems," IEEE Communications Magazine, vol. 52, no. 1, pp. 210-216, January 2014.

[5] W. Wu, J. Wang, X. Wang, F. Shan and J. Luo, "Online Throughput Maximization for Energy Harvesting Communication Systems with Battery Overflow," IEEE Transactions on Mobile Computing, vol. 16, no. 1, pp. 185-197, January 2017.

[6] B. T. Bacinoglu, E. U. Biyikoglu and C. E. Koksal, "Finite-Horizon Energy Efficient Scheduling with Energy Harvesting Transmitters over Fading Channels," IEEE Transactions on Wireless Communications, vol. 16, no. 9, pp. 6105-6118, September 2017.

[7] A. O. Isikman, M. Yuksel and D. Gunduz, " A Low Complexity Policy for Outage Probability Minimization with EH Transmitter," IEEE Communications Letters, vol. 21, no. 4, pp. 917-920, April 2017.

[8] J. Liu, H. Dai and W. Chen, "Delay Optimal Scheduling for Energy Harvesting based Communications," IEEE Journal on Selected Areas in Communications, vol. 33, no. 3, pp. 452-466, March 2015.

[9] R. A. Raghuvir and D. Rajan, "Delay Bounded Rate and Power Control in Energy Harvesting Wireless Networks," IEEE Wireless Communications and Networking Conference (WCNC), April 2011.

[10] L. Lei, Y. Kuang, X. Shen, K. Yang, J. Qiao and Z. Zhong, "Optimal Reliability in Energy Harvesting Industrial Wireless Sensor Networks," IEEE Transactions on Wireless Communications, vol. 15, no. 8, pp. 5399-5413, August 2016.

[11] I. Ahmed, K. T. Phan and T. L. Ngoc, "Optimal Stochastic Power Control for Energy Harvesting Systems with Statistical Delay Constraint," IEEE Journal on Selected Areas in Communications, vol. 34, no. 12, pp. 3512-3527, December 2016.

[12] S. Ulukus, A. Yener, E. Erkip, O. Simeone, M. Zorzi, P. Grover and K. Huang, "Energy Harvesting Wireless Communications: A Review of Recent Advances," IEEE Journal on Selected Areas in Communications, vol. 33, no. 3, pp. 360-381, March 2015.

[13] M. L. Ku, W. Li, Y. Chen and K. J. R. Liu, "Advances in Energy Harvesting Communications: Past, Present, and Future Challenges," IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp. 1384-1412, February 2016.

[14] I. Fawaz, P. Ciblat and M. Sarkiss, " Energy Minimization based Resource Scheduling for Strict Delay Constrained Wireless Communications," IEEE Global Conference on Signal and Information Processing (GlobalSIP), December 2017.

[15] Y. Polyanskiy, H. V. Poor and S. Verdú, "Channel coding rate in the finite blocklength regime," IEEE Transactions on Information Theory, vol. 56, no. 5, pp. 2307-2359, May 2010.

[16] Y. Hu and A. Schmeink, " Delay-Constrained Communication in Edge Computing Networks," IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), June 2018.

[17] R. Ma and W. Zhang, "Optimal Power Allocation for Energy Harvesting Communications with Limited Channel Feedback," IEEE Global Conference on Signal and Information Processing (GlobalSIP), December 2014.

[18] M. R. Zenaidi, Z. Rezki and M. S. Alouini, "Performance Limits of Online Energy Harvesting Communications with Noisy Channel State Information at the Transmitter," IEEE Access, vol. 5, pp. 1239-1249, March 2017.

[19] M. R. Zenaidi, Z. Rezki, H. Tembine and M. S. Alouini, "Performance Limits of Energy Harvesting Communications under Imperfect Channel State Information," IEEE International Conference on Communications (ICC), May 2016.

[20] H. Wang and N. B. Mandayam, "A Simple Packet-Transmission Scheme for Wireless Data Over Fading Channels," IEEE Transactions on Communications, vol. 52, no. 7, pp. 1055-1059, July 2004.

[21] Y. Chen and Q. Zhao, "An Integrated Approach to Energy-Aware Medium Access for Wireless Sensor Networks," IEEE Transactions on Signal Processing, vol. 55, no. 7, pp. 3429-3444, July 2007.

[22] D. V. Djonin and V. Krishnamurthy, "MIMO Transmission Control in Fading Channels — A Constrained Markov Decision Process Formulation With Monotone Randomized Policies," IEEE Transactions on Signal Processing, vol. 55, no. 10, pp. 5069-5083, October 2007.

[23] N. Salodkar, A. Bhorkar, A. Karandikar and V. S. Borkar, "An On-Line Learning Algorithm for Energy Efficient Delay Constrained Scheduling over a Fading Channel," IEEE Journal on Selected Areas in Communications, vol. 26, no. 4, pp. 732-742, May 2008.

[24] N. Salodkar, A. Karandikar and V. S. Borkar, "A Stable Online Algorithm for Energy-Efficient Multiuser Scheduling," IEEE Transactions on Mobile Computing, vol. 9, no. 10, pp. 1391-1406, October 2010.

[25] R. A. Berry, "Optimal Power-Delay Tradeoffs in Fading Channels — Small-Delay Asymptotics," IEEE Transactions on Information Theory, vol. 59, no. 6, pp. 3939-3952, June 2013.

[26] K. T. Phan, T. Le-Ngoc, M. Schaar and F. Fu, "Optimal Scheduling over Time-Varying Channels with Traffic Admission Control: Structural Results and Online Learning Algorithms," IEEE Transactions on Wireless Communications, vol. 12, no. 9, pp. 4434-4444, September 2013.

[27] R. A. Berry and R. G. Gallager, "Communication Over Fading Channels With Delay Constraints," IEEE Transactions on Information Theory, vol. 48, no. 5, pp. 1135-1149, May 2002.

[28] W. Labidi, M. Sarkiss and M. Kamoun, "Energy-Optimal Resource Scheduling and Computation Offloading in Small Cell Networks," IEEE International Conference on Telecommunications (ICT), April 2015.

[29] M. L. Puterman, "Markov Decision Processes," John Wiley & Sons, 1994.

[30] D. P. Bertsekas, "Dynamic Programming and Optimal Control," Third Edition, Athena Scientific, 2000.

[31] P. Ciblat and L. Vandendorpe, "On the Maximum-Likelihood based data-aided frequency offset and channel estimates," EURASIP European Signal Processing Conference (EUSIPCO), September 2002.

[32] J. G. Proakis and M. Salehi, "Digital Communications," Fifth Edition, McGraw-Hill, 2008.

[33] S. Lin and D. Costello, "Error Control Coding," Second Edition, Pearson, 2004.

Ibrahim Fawaz is an Early-Stage Researcher (ESR) within the European H2020-MSCA-ITN SCAV-ENGE project, pursuing the Ph.D. in Communications and Electronics within the French Alternative Energies and Atomic Energy Commission (CEA). He received the B.Sc. and M.Sc. degrees in Telecommunication Engineering in 2016 from the Lebanese University and Telecom ParisTech, respectively. His research interests include wireless communication, resource scheduling, computation offloading, energy harvesting, optimization, deep learning.

PLACE
PHOTO
HERE

Mireille Sarkiss Biography text here.

PLACE
PHOTO
HERE

Philippe Ciblat Biography text here.