This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2654454, IEEE Access

Performance Limits of Online Energy Harvesting Communications with Noisy Channel State Information at the Transmitter

Mohamed Ridha Zenaidi, *Student Member, IEEE,* Zouheir Rezki, *Senior Member, IEEE,* and Mohamed-Slim Alouini, *Fellow, IEEE*

***Abstract*—In energy harvesting communications, the transceivers have to adjust the data transmission to the energy scavenged during the course of communication. The performance of the transmission depends on the channel conditions which vary randomly due to mobility and environmental changes. In this paper, we consider the problem of power allocation taking into account the energy arrivals over time and the quality of channel state information (CSI) measured at the transmitter, in order to maximize the throughput. Differently from previous work, we focus on energy harvesting communications where the CSI at the transmitter is not perfect and may include estimation errors. In the present paper, we introduce a Markov process that models the energy arrival process. Indeed, we solve the throughput maximization problem with respect to energy harvesting constraints. We show that the optimal online power policy can be found using dynamic programming. Furthermore, we study the asymptotic behavior of the communication system at low and high average recharge rate (ARR) regime. Selected numerical results are provided to support our analysis.**

***Index Terms*—Asymptotic analysis, channel estimation, channel state information, energy harvesting, Markov process, optimal power allocation, throughput maximization.**

I. Introduction

DURING this era, wireless communication systems are used tremendously to run plenty of applications. On the other hand, these devices are required to guarantee the network connectivity. This challenge and those resulted from the remarkable advance in telecommunications technology and applications, make the wireless communication systems suffer from the problem of the energy inefficiency. Therefore, massive search for green technologies contributed to the emergence of self-powered systems such as energy harvesting (EH) communication systems [1]. The performance of such systems is determined by adapting the transmission power to energy availability. The analysis of this performance leads to several power allocation policies that were presented in recent work.

A. Related work

The recent state-of-the-art advances in point-to-point EH communications can be classified into two main approaches. The first approach considers an offline model for the energy arrival [2]-[5]. However, the assumption of having a perfect knowledge of the amounts of energy does not hold in practice due to the random nature of the energy harvested from the surrounding environment. The second approach considers a more realistic scenario that suggests an online model for the energy arrival process [6]-[11]. In the first model (offline), the transmitter (TX) has a perfect and causal knowledge of the energy arrivals, whereas in the second scenario (online) a random EH model is considered.

In [2], the throughput maximization problem is considered for an EH system assuming perfect CSI-T and CSI-R. The work was extended in [3] considering battery with energy leakage. In [4] and [5], the transmission completion time minimization problem was solved. It is shown in [4] that the solution for the optimal scheduling is the same as the solution for the throughput maximization problem. In [5],

Part of this paper has been accepted for presentation in the IEEE Global Communications Conference (GLOBECOM'16).

Mohamed Ridha Zenaidi and Mohamed-Slim Alouini are with the Electrical Engineering Program, Computer, Electrical, and Mathematical Science and Engineering (CEMSE) Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia. E-mails: {mohamed.znaidi,slim.alouini}@kaust.edu.sa.

Zouheir Rezki is with the Department of Electrical and Computer Engineering, University of Idaho, Idaho, USA. E-mail: zrezki@uidaho.edu.

the authors solve the problem taking into account both the arrivals of the harvested energy and the data over time. The proposed model, in the previous references, for the energy harvested process can be used in applications where the amounts can be estimated quite accurately prior data transmission e.g., solar energy during the day or energy harvested from heartbeat. Since this offline model does not take into consideration the randomness and the uncertainty of the amounts of the energy harvested, it is not relevant for some applications. This motivates other researchers to consider an online model that captures the limitations of the first energy model.

One way to solve the problem of power allocation scheduling taking into account unknown amounts of energy, is to design a stochastic model that considers the imperfect knowledge of the energy arrival process. For this class of the energy harvested model, several work adopt the Markov process tool, although in a different context, e.g., in [12], to solve the throughput maximization problem. In [6], the authors used a first-order Markov process and dynamic programming to build a framework to solve the problem in the context of single-input-single-output (SISO) EH system. Reference [7] proposed a solution for increasing the lifetime of the energy storage devices. In fact, the authors proposed a dynamic power management policy to stabilize data queues. In order to prolong lifetime of wireless network, optimal transmission policies taking into account both transmission and processing energy costs are studied in [8]. Reference [9] suggested sub-optimal power policies for perfect and imperfect knowledge of the energy. In [10], a Markov process was proposed to solve the throughput maximization problem with 1-bit channel feedback. The authors in [11] proposed a data-driven stochastic EH model to solve the throughput maximization problem.

B. Motivation and Contribution

The majority of the recent research in EH communications has considered full CSI-T, which can not be achieved in most operating scenarios where the feedback link is noisy. In the present paper, we focus on more practical systems in order to design the optimal power policy during transmission assuming an imperfect CSI-T [13]-[15]. More specifically, the present paper considers the problem of optimal power allocation with the objective of maximizing the throughput under energy harvesting (EH) constraints. Unlike [16], this paper considers an online model for the energy arrival process. The problem is solved for point-to-point communication system where the transmitter (TX) is equipped with energy harvester node assuming an imperfect channel state information at the transmitter (CSI-T). Meanwhile, the channel state information at the receiver (CSI-R) is assumed to be perfectly known during the communication. In fact, we assume that TX has an estimated version of the actual channel gain obtained, for instance, through a minimum mean square error (MMSE) filter of a feedback link. This model of CSI-T is widely accepted by the community and has been extensively adopted, e.g, [14] and [15].

We investigate the optimal power scheduling for maximum throughput by a deadline T assuming a statistic knowledge about the energy arrival process. Our contribution in this paper is as follows:

- We develop a stochastic model to capture the dynamics of the EH communication system under the assumption of imperfect CSI-T.
- We provide the optimal power policy for communication system powered by energy harvester nodes under the assumption of imperfect CSI-T, depending on the knowledge of the amounts of energy harvested.
- We study the performance of EH systems by analyzing the asymptotic behavior of the average throughput where the energy harvested amounts are very small and when they are very high for, different cases of CSI-T.

C. Outline of the Paper

This paper is organized as follows. Section II provides the system model. In Section III, we solve the optimal power allocation problem. In Section IV, we investigate the performance of the communication system. Selected numerical results that support our analysis are presented in Section V. Section VI concludes the paper.

D. Notations

We use $\underline{v}$ to denote a vector of length N, i.e., $\underline{v} = (v_1, v_2, ..., v_N)$. The symbol "$\approx$" designates asymptotic equality, i.e., $f(x) \overset{a}{\approx} g(x)$ if and only if $\lim\limits_{x \to a} \frac{f(x)}{g(x)} = 1$. When it is clear from the context, we omit a in $\overset{a}{\approx}$, for convenience. For a Random Variable X, $f_x(\cdot)$ and $\mathbb{E}_X(.)$ denote the probability density function and the expectation operation, respectively. The set $M_n(\mathbb{R})$ represents the set of real matrices $n \times n$. We use $\mathbb{P}$ to denote the probability.

II. System model

A. System description

In this paper, we consider point-to-point EH communication system where TX is powered by an EH sensor as shown in Fig.1. The data transmission is performed during a deadline T. We assume that there is always data available, in the buffer, for transmission. The channel between TX and the receiver (RX) is a Rayleigh fading channel. The base band received signal y over a bandwidth W, is given by:

$$y = hx + n, \tag{1}$$

where $h \sim \mathcal{CN}(0, \sigma^2)$ is a zero mean circularly symmetric complex variable with variance σ^2, x is the channel input, and n is a zero-mean additive white Gaussian noise with spectral density N_0, and is independent of h. If we consider a Gaussian signaling for transmission over a complex Gaussian channel, with a cost of Lp units of energy, TX sends $L\ \log(1+\gamma p)$ bits of data, where p is the power used for transmission, L the duration of the time slot (TS) and γ is the fading level given by $\gamma = |h|^2$. Thus, the energy available in the battery determines the feasible bits that can be transmitted during each (TS). During each TS, TX encodes the bits to be sent as data symbols, where the block length of each symbol is assumed to be large enough so that we can guarantee the reliability of the decoding process. A feedback link is considered between the RX and TX, the CSI feedback is sent, at the beginning of TSs, from RX to TX.

B. Energy Model

Differently from the offline model of the energy arrival process, in this paper we take into account the randomness and the uncertainty of the amounts of the energy harvested. In order to capture applications where TX does not have a deterministic knowledge of the amounts of EH, we design a more sophisticated online model where we develop a Markov process that takes into consideration the unpredictable amounts of the energy harvested. We assume that TX is equipped with a single EH sensor node. This node senses periodically in a discrete time, during the deadline T, the amount of energy available in the environment at beginning of TS $i \in \{1, 2, ..., N\}$. At t_i, E_i units of energy is harvested, and the battery is assumed to be empty initially. The transmission begins at $t_0 = 0$ when the energy E_0 is scavenged by TX, see Fig.2. Since the sensing is considered periodic, the length of TSs is assumed to be constant, i.e., $L_i = t_{i+1} - t_i = L = \frac{T}{N}$. Different from most previous work that consider a deterministic EH process, we propose a first-order stationary Markov chain for the energy arrival process. In our Markov model, we suggest a finite-state Markov chain, with N_s states, and the probabilities of transition from one state to another. Each state of the model depends mainly on the environment and the amount of the energy harvested in the past. Therefore, TX disposes of a statistical knowledge of energy arrival amounts. This statistical knowledge can be learned from system history or measurements, for instance.

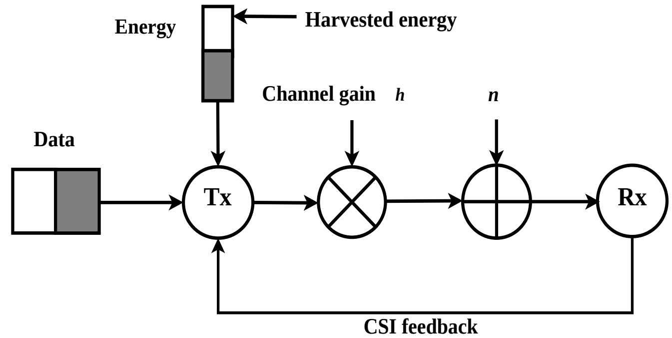

Fig. 1. An energy harvesting communication system model with CSI feedback.

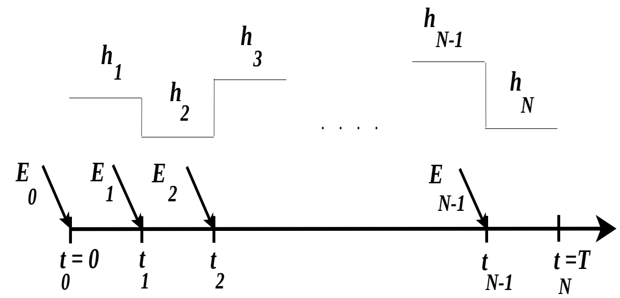

Fig. 2. Energy arrival and fading channel during a period T for the online model.

where E_j^h designates the amount of energy avail-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2654454, IEEE Access

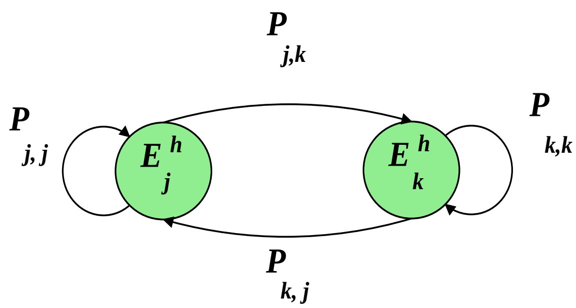

Fig. 3. Two state from the N_s-state first order Markov chain.

able at state j and $P_{j,k}$ designates the probability of transition from state j to state k, $j, k \in N_s$.

The energy harvested amount E_i takes values in a finite set of states $\mathcal{E} = \{E_1^h = 0, E_2^h, ..., E_{N_s}^h\}$. In Fig.3, we show a model for two states chosen randomly from the N_s states. In fact, the states contain the zero state, i.e., $E_1^h = 0$ which corresponds to the case where no energy is available in the environment. The transition probabilities between states are denoted as follows:

$$\mathbb{P}(E_j^h \rightarrow E_k^h) = P_{j,k}, \forall j, k \in [\![1, N_s]\!]. \qquad (2)$$

The matrix of the transition probabilities is denoted $\mathbf{M} \in M_{N_s}(\mathbb{R})$, defined by $M_{jk} = P_{j,k}$. The steady state probability in this case is given by $\pi^* = [\mathbb{P}_{E_1^h} \cdots \mathbb{P}_{E_{N_s}^h}]$. Since the EH process is modeled as first-order stationary Markov chain, the transition probability of the random variable E_i is defined as follows: $\forall i \in [\![1, N-1]\!]$

$$\mathbb{P}(E_i | E_0, E_1, \cdots, E_{i-1}) = \mathbb{P}(E_i | E_{i-1}). \qquad (3)$$

Also one can write : $\forall i \in [\![1, N-1]\!]$

$$\mathbb{P}(E_i, E_{i-1}, \cdots, E_1 | E_0) = \prod_{t=1}^{i} \mathbb{P}(E_t | E_{t-1}). \qquad (4)$$

Based on the statistical knowledge provided as a look-up table to TX and the amount of energy harvested in the previous TS, TX has to adapt its transmission scheme to allocate optimally the energy across time in order to maximize the throughput.

C. Channel Model

In the present paper, we consider the throughput maximization problem over a deadline T. Optimal powers are obtained according to the channel conditions. Depending on the availability of the information about these conditions, we propose an analysis of the performance of the proposed communication system.

We assume that the wireless channel varies slowly over time, so the fading level is constant at each TS during the transmission of data. That is, the fading changes in discrete time instants $t_1, t_2, ..., t_{N-1}$ which represent the sensing periods of the energy harvester node as shown in Fig.2. The fading level during TS i, is assumed to be equal to h_i. We assume that throughout the communication, perfect CSI-R is available. Here, we assume a noisy CSI-T and its estimate follows the same linear model suggested in the offline analysis. This model is given by:

$$h = \sqrt{1-\alpha}\ \widehat{h} + \sqrt{\alpha}\ \tilde{h}, \qquad (5)$$

where α is the error variance, $\alpha \in [0, 1]$. TX estimates the actual channel estimation as $\hat{h}$ which is $\mathcal{CN}(0, \sigma^2)$ independent of the channel estimation error denoted as $\tilde{h}$ which is also $\mathcal{CN}(0, \sigma^2)$.

D. Battery Model

During the epoch i of length L, TX sends the data symbol i amplified by the power p_i. Then, TX consumes $L \times p_i$ units of energy during this epoch from the energy stored in the battery in this epoch B_i. Just after this time slot, the energy harvester node scavenges an amount of energy of E_i. Hence, in the epoch $i+1$ the stored energy in the battery is updated as follows:

$$B_{i+1} = min\{B_i - Lp_i + E_i, E_{max}\},\ \forall i \in [\![1, N-1]\!] \quad (6)$$
$$B_{N+1} = B_N - L_N p_N. \quad (7)$$

From (6), one can conclude that the energy stored in the battery follows a first-order Markov process that depends on the immediate past energy harvested and the transmitted power. The initial amount of the energy saved in the battery B_1 is equal to the amount of energy harvested E_0 by the node that engenders the beginning of the transmission of data, i.e., $B_1 = min\{E_0, E_{max}\}$.

E. Energy Harvesting Constraints:

With the use of EH technology, the region of feasible power policies is constrained by EH constraints [2]. In fact, the optimal scheduling is constrained by the energy harvested profile. To design the feasible region, we consider the causality constraint which means that TX can only use the energy

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2654454, IEEE Access

that has arrived. Following recent works (e.g., [2] and [6]), we assume that the transmitted power must be kept constant in each epoch. We denote the power consumed in each epoch i by p_i. So, the causality constraint is defined as follows:

$$\sum_{j=1}^{i} Lp_j \leq \sum_{j=0}^{i-1} E_j, \forall i \in [\![1, N]\!]. \quad (8)$$

Also, one can understand this constraint as if TX can transmit only energy which is already available in the battery, i.e.,

$$0 \leq L \times p_i \leq B_i, \quad \forall i \in [\![1, N]\!]. \quad (9)$$

In order to investigate the optimal power policy, we consider the energy storage constraint which means that the energy level in the battery can not exceeds E_{max}. This constraint can be expressed as:

$$\sum_{j=0}^{i} E_j - \sum_{j=1}^{i} Lp_j \leq E_{max}, \quad \forall i \in [\![1, N-1]\!]. \quad (10)$$

Also, this constraint can be reflected by (6) in the battery dynamics, which states that the energy stored in the battery during TS i must be less than E_{max}.

F. Overall Markov Model:

In this subsection, we investigate a model that can capture the energy and the battery model, jointly. Using an appropriate Markov chain to model the EH dynamics (3) and the first-order Markov process that captures the battery dynamics (6). Specifically, we design a new first-order Markov process whose states are defined as the joint state between the EH's state and the battery's state. More precisely, the state at the initial TS: $S_1 \triangleq (B_1)$ which is assumed to be known to TX, the states during TSs are defined by: $\forall j \in [\![2, N]\!]$, $S_j \triangleq (E_{j-1}, B_j)$ and a last state $S_{N+1} \triangleq (B_{N+1})$ which describes the state of battery by the end of transmission. Hence, one can deduce from (3) and (6) that:

$$\mathbb{P}(S_i|S_1, S_2, ..., S_{i-1}) = \mathbb{P}(S_i|S_{i-1}), \forall i \in [\![2, N+1]\!]. \quad (11)$$

Consequently, the state transition probability can be rewritten as follows: $\forall i \in [\![2, N+1]\!]$

$$\mathbb{P}(S_i, S_{i-1}, ..., S_2|S_1) = \prod_{k=2}^{i} \mathbb{P}(S_k|S_{k-1}). \quad (12)$$

Therefore, our objective is to formulate the throughput maximization problem given a deadline T subject to EH constraints under the assumption of imperfect CSI-T.

G. Nomenclature & Definitions

In regard of the large number of notations and parameters in this paper, it is convenient to define the notation used in this subsection.

Parameter	Description
B_i	amount of energy in the battery during TS i
E_i	energy harvested during epoch i
$\mathcal{E}$	set of possible energy states
E_j^h	amount of energy in state j of the Markov chain
E_{max}	maximum stored energy in the battery
h	channel fading
$\hat{h}$	channel fading estimation
$\tilde{h}$	channel fading estimation error
i	TS index
L	length of TSs
M	matrix of the transition probabilities
N_0	noise variance
N	number of TSs
N_s	cardinal of the Markov chain
p_i	transmitted power during TS i
$P_{j,k}$	transition probability from state j to state k
$\mathcal{S}^i$	set of possible states in TS $i+1$ given the state S_i
S_i	state at TS i
t_i	time index of energy sensing
T	deadline of communication
V_i	reward function during TS i
W	signal bandwidth
α	error variance
γ	fading level
$\hat{\gamma}$	fading level estimation
$\tilde{\gamma}$	fading level estimation error
π^*	steady state probability
σ^2	channel variance

Definitions: For the rest of the paper, we will use these definitions:

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2654454, IEEE Access

- The average throughput (AT): the throughput of the communication system per second. i.e.,

$$AT = \frac{1}{T}\sum_{i=1}^{N} R_i, \tag{13}$$

where R_i denotes the throughput at epoch i.
In this paper, we use these notations for AT:
• AT_{Stat}: AT when the channel is static.
• AT_{α}: AT when the CSI-T is imperfect with an error variance equal to α. One can remark that by setting $\alpha = 0$ and $\alpha = 1$, we retrieve the AT in the case where the CSI-T is perfectly known and the AT in the case where the CSI-T is unavailable, respectively.

- The average recharge rate (ARR): the average of energy harvested over the deadline T. i.e.,

$$ARR = \frac{1}{T}\mathbb{E}_{\underline{E}}\left(\sum_{i=0}^{N-1} E_i\right), \tag{14}$$

where E_i denotes the amount of energy harvested at epoch i and the averaging is over the distribution of energy arrivals.

III. Optimal Power Policy

In this section, our objective is to formulate the throughput maximization problem given a deadline T subject to EH constraints under imperfect CSI-T. Then, we determine the online optimal power policy that maximizes the throughput of the EH communication system.

A. Problem Formulation

Our throughput maximization problem can be classified as a discrete-time Markov Decision Process (MDP). In fact, at the initiation of the communication, TX has the information about the initial energy harvested. Then, TX decides to transmit with power p_i at the beginning of the TS i, taking into consideration the matrix of the transition probabilities. Our aim is to find an optimal power policy that maximizes the expected cumulative data rate during N TSs, i.e.,

$$(P): \begin{cases} \max_{\{p_i\}_{i=1}^{N}} \mathbb{E}_{S_2^N}\left\{\sum_{i=1}^{N} L\ \mathbb{E}_{\gamma|\hat{\gamma}_i}\left[\log(1+\gamma p_i^*(\underline{\hat{\gamma}})) \mid \hat{\gamma}_i\right] \mid \mathbf{M}, S_1\right\} & (15a) \\ \text{s.t.} \begin{cases} 0 \le Lp_i \le B_i, \quad \forall i \in [\![1,N]\!] \\ B_{i+1} = min\{B_i - Lp_i + E_i, E_{max}\}, \forall i \in [\![1,N-1]\!] \\ B_{N+1} = B_N - Lp_N. \end{cases} & (15b) \end{cases}$$

where $\mathbb{E}_{S_2^N}$ designates the statistical expectation over all possible states during TSs $i = 2,...,N$. $\mathbb{E}_{\gamma|\hat{\gamma}_i}$ designates the statistical expectation over the random variable $\gamma|\hat{\gamma}_i$. Generally, the optimization problem in (15) cannot be solved independently for each time slot due to the dependence of constraints along TSs. For instance, the energy consumed currently affects the energy stored in the battery in the next TS, and therefore affects the future power allocation. Consequently, such sequential optimization problem with random EH amounts can be solved optimally using finite-horizon dynamic programming (DP) [17]. Recall that DP is a general approach for solving multi-stage optimization problems over a finite time horizon. The underlying idea is to use backward recursion to solve such problems.

B. Online Power Policy with Fading Channels

In this section, we solve the problem (15) by using finite-horizon DP. In fact, the optimal power allocation is determined by backward induction method [17]. For instance, $\{p_1^*, p_2^*, ..., p_N^*\}$ is calculated in the time reversal order. The reward functions are calculated recursively as follows:

1) The last TS, $i = N$:

$$V_N(\hat{\gamma}_N, S_N, \alpha) = \max_{p_N} \mathbb{E}_{\gamma|\hat{\gamma}_N}\left[L\ \log(1+\gamma p_N) \mid \mathbf{M}, S_1\right] \tag{16}$$

$$\text{s.t. } 0 \le Lp_N \le B_N.$$

Proposition 1. The optimal last state S_{N+1}^* of the first-order Markov process is $S_{N+1}^* \triangleq (B_{N+1} = 0)$.

Proof. First, it can be easily seen that the reward function at the last TS is increasing in p_N. As the power consumed is bounded by the energy stored in battery, the optimal power is as follows $p_N = \frac{B_N}{L}$. Hence, $S_{N+1}^* \triangleq (B_{N+1} = 0)$. □

Proposition 1 states that TX transmits the last symbol with full energy available in the battery. This result is quite consistent with recent works where the causality constraint (10) is assumed to be satisfied with equality at the last time slot, e.g., [2]. Therefore, the reward function in the last time slot can be evaluated using $f_{\gamma|\hat{\gamma}}(\cdot)$ as follows:

$$V_N(\hat{\gamma}_N, S_N, \alpha) = L\ \mathbb{E}_{\gamma|\hat{\gamma}_N}\left[\log(1+\gamma\frac{B_N}{L})\right]. \tag{17}$$

As a consequence, by the end of the communication, we have:

$$\sum_{i=0}^{N-1} E_i = \sum_{i=1}^{N} L p_i^*. \tag{18}$$

2) The TSs, $i = N-1, N-2, \cdots, 1$: During the TS i, the optimal power is determined by maximizing the reward function corresponding to this TS which is expressed as follows:

$$V_i(\hat{\gamma}_i, S_i, \alpha) = \max_{p_i \leq \frac{B_i}{L}} \left\{ \mathbb{E}_{\gamma|\hat{\gamma}_i} \left(\frac{L}{2} \log_2(1 + \gamma p_i) \right) + \overline{V}_{i+1}(\hat{\gamma}_{i+1}, S_{i+1}, \alpha) \right\}, \tag{19}$$

where $\overline{V}_{i+1}(\widehat{\gamma}_{i+1}, S_{i+1}, \alpha)$ designates the expected values of the future reward functions and is defined as follows:

$$\overline{V}_{i+1}(\hat{\gamma}_{i+1}, S_{i+1}, \alpha) = \mathbb{E}_{\mathcal{S}^i} \left\{ \mathbb{E}_{\gamma|\hat{\gamma}_{i+1}} \left[V_{i+1}(\hat{\gamma}_{i+1}, S_{i+1}, \alpha) \right] | \mathbf{M}, S_i \right\}, \tag{20}$$

where $\mathbb{E}_{\mathcal{S}^i}$ designates the statistical expectation over the set of possible states in the TS $i+1$ given the state S_i and the transition probabilities. The reward function in (20) can be understood as the maximal sum of the data rate in the current TS. Taking into account the matrix $\mathbf{M}$, we maximize the expected cumulative data rate in the future TSs resulted from the current state and the current transmitted power. The reward functions $V_i(\widehat{\gamma}_i, S_i, \alpha)$ cannot be determined in a closed form.

Theorem 1. *Given a channel fading level estimation at different TSs denoted by the vector $\hat{\gamma} = (\hat{\gamma}_1, \hat{\gamma}_2, \cdots, \hat{\gamma}_N)$, the initial state $S_1 \triangleq (B_1) \triangleq (E_0)$ and the matrix of transition probabilities $\mathbf{M}$ related to Markov chain that models the EH process, the optimal power allocation for time slots $i = 1, 2, ..., N$ is given by the initial reward function $V_1(\hat{\gamma}_1, S_1, \alpha)$.*

Proof. The optimization problem is formulated as a recursive structure where the state during each TS depends only on the previous one. Based on this structure, the proof of Theorem 1 follows immediately by applying equations of Bellman [18] and the equations that resume the dynamic of the Markov process ((6) and (3)). □

1) Some properties of the optimal power policy:

Remark 1. Under the DP framework presented in this paper and the reward functions given by (19) and (20), $V_i(\widehat{\gamma}_i, S_i, \alpha)$ and $\overline{V}_i(\widehat{\gamma}_i, S_i, \alpha)$ are continuous, non-decreasing and concave in B_i. Indeed, the optimal cumulative data rate is concave and increasing in the initial state of the battery. This result is quite intuitive, the higher is the amount of energy harvested at the beginning of transmission, the higher is the maximum achievable expected data rate. We note that studying the continuity, the monotonicity and the convexity properties of the reward functions in the belief state is quite well known in the MDP literature [19] [20]. The proof of the statements in this remark follows similar lines as the proofs of Theorem 1 and Theorem 2 in [6].

Remark 2. One can easily remark that the optimal reward function $V_1(\widehat{\gamma}_1, S_1, \alpha)$ is non-increasing in α. This can be deduced immediately from the definition of the parameter α which measures the quality of the CSI-T.

We note that our framework can capture the perfect CSI-T and the no CSI-T cases, by setting $\alpha = 0$ and $\alpha = 1$, respectively.

2) Special cases:

a) Online optimal power policy with perfect CSI-T: Using the procedure described above, one can determine the optimal power allocation when TX has perfect knowledge of the channel during TSs. In order to capture this case, one can set $\alpha = 0$ and $f_{\gamma|\hat{\gamma}}(\gamma) = \delta(\gamma - \hat{\gamma})$.

b) Online optimal power policy without CSI-T: In some situations, having an estimate of the channel is challenging due to the huge amount of noise in the environment. Then, the optimal power allocation in such situations can also be determined by setting $\alpha = 1$ and $f_{\gamma|\hat{\gamma}}(.) = f_{\gamma}(.)$ in the above procedure.

c) Online optimal power policy with static channel: This work can be utilized also to retrieve the optimal reward function when the channel is equal to $\overline{\gamma}$ during TSs. For instance, one can perform the optimization problem under the setting $\alpha = 0$ and $\forall i \in [\![1, N]\!],\ \gamma_i = \overline{\gamma}$. This will provide a look up table where the optimal powers are stored.

Hence, our work is a generalization of maximizing the throughput considering the fading channel case.

IV. Asymptotic Analysis

In this section, the average throughput performance is studied when the CSI-T is partially available. Given an error variance α of the channel estimation, the average throughput can be written as follows: $\forall\ \alpha \in [0,1]$,

$$AT_{\alpha} = \mathbb{E}_{S_1^N}\left\{\mathbb{E}_{\underline{\hat{\gamma}}}\left(\sum_{i=1}^{N}\frac{L}{T}\mathbb{E}_{\gamma|\hat{\gamma}_i}\left[\log(1+\gamma p_i^*(\underline{\hat{\gamma}}))\mid\hat{\gamma}_i\right]\right)\mid \mathbf{M}\right\}. \tag{21}$$

Based on the developed optimal power policy, we present our asymptotic analysis in two extremes cases, the low ARR regime $(ARR \to 0)$ and the high ARR regime $(ARR \to +\infty)$, respectively.

A. Low ARR Regime

In this subsection, we evaluate the performance of the communication system when the energy harvested is scarce.

Proposition 2. Given that all energies are very low, the performance of the studied system in the online model is similar to the one in the offline for the following cases.

- If the channel between TX and RX is a fading channel where the CSI-T is unavailable. In this case, the AT of the system grows linearly with ARR, i.e.,
$$AT_1(ARR) \approx \mathbb{E}_{\tilde{\gamma}}(\tilde{\gamma})ARR \tag{22}$$
- When the channel is static equal to $\overline{\gamma}$ during the communication. The AT of the system grows as follows
$$AT_{Stat}(ARR) \approx \overline{\gamma}ARR. \tag{23}$$

Proof. In the low ARR regime, we assume that all powers are very low. In the case where the CSI-T is unavailable, by setting $\alpha = 1$, we have:

$$AT_1(ARR) = \mathbb{E}_{S_1^N}\left\{\mathbb{E}_{\tilde{\gamma}}\left[\sum_{i=1}^{N}\frac{L}{T}\log(1+\tilde{\gamma}p_i^*)\right]\mid \mathbf{M}\right\} \tag{24}$$

$$\approx \mathbb{E}_{S_1^N}\left\{\mathbb{E}_{\tilde{\gamma}}\left[\sum_{i=1}^{N}\frac{L}{T}p_i^*\tilde{\gamma}\right]\mid \mathbf{M}\right\} \tag{25}$$

$$= \mathbb{E}_{S_1^N}\left\{\sum_{i=1}^{N}\frac{L}{T}p_i^*\mid \mathbf{M}\right\}\mathbb{E}_{\tilde{\gamma}}(\tilde{\gamma}). \tag{26}$$

Using (18), then we have:

$$AT_1(ARR) = \mathbb{E}_{\tilde{\gamma}}(\tilde{\gamma})\mathbb{E}_{S_1^N}\left\{\sum_{i=0}^{N-1}\frac{E_i}{T}\mid \mathbf{M}\right\} \tag{27}$$

$$= \frac{\mathbb{E}_{\tilde{\gamma}}(\tilde{\gamma})}{T}\sum_{i=0}^{N-1}\sum_{j=1}^{N_s}\mathbb{P}_{E_j^h}E_j^h \tag{28}$$

$$= \frac{\mathbb{E}_{\tilde{\gamma}}(\tilde{\gamma})}{T}\mathbb{E}_{\underline{E}}\left\{\sum_{i=0}^{N-1}E_i\right\} = \mathbb{E}_{\tilde{\gamma}}(\tilde{\gamma})\ ARR. \tag{29}$$

The expression of the AT_{Stat} can be proved along similar lines as the proof for AT_1 . □

Proposition 2 asserts that with no CSI-T, no gain is provided by fading in terms of AT for the same expected value of the fading level $\overline{\gamma}$. In fact, in the low ARR regime, the AT_{Stat} and AT_1 of the communication system studied in this work depend only on the ARR and the property of the channel. Hence, the model considered for the energy arrival process does not have an effect on the performance of the communication in terms of the data rate.

Conjecture 1. Given an error variance α, the AT in the low ARR regime is a linear combination of the AT when the fading level of the channel is known perfectly and the AT when the CSI is unavailable at TX, i.e., $\forall\ \alpha \in [0,1]$,

$$AT_{\alpha}(ARR) \approx (1-\alpha)\ AT_0(ARR) + \alpha\ AT_1(ARR). \tag{30}$$

This conjecture is based on numerical observations and is seemingly expected. Indeed, in the offline analysis [16], the authors have proved similar result after determining a closed form expression of the optimal powers. However, in this online analysis, since a closed form solution for the optimal powers could not be found, following a similar approach is not feasible. However, we noticed, from Proposition 2, that the AT_{Stat} and the AT_1 depend only on the ARR. Therefore, we have expected to have the same results since we remarked that the performance of the communication system is only a function of the ARR. Consequently, the AT at low ARR does not take into account the considered energy model and it depends only on the ARR.

B. High ARR Regime

In this subsection, we evaluate the performance of the communication system when the energy harvested is abundant. In such a scenario, the state $E_1^h = 0$ is omitted because the node can harvest a very high amount of energy at each sensing instant.

Proposition 3. Provided that all powers are very high in the high ARR regime, then the AT scales as follows: $\forall \alpha \in [0, 1]$

$$\lim_{ARR \to \infty} \left[AT_\alpha(ARR) - AT_{Stat}(ARR)\right] = \mathbb{E}\left[\log(\gamma)\right]. \tag{31}$$

Proof. See Appendix A. □

The interpretation of the result stated in Proposition 3 is that, at high ARR, the AT when the CSI-T is imperfect increases similarly as the AT in the case where the channel is static over all TSs, irrespective of the value of α. According to the previous result, the gap between $AT_{I.CSI-T}$ and AT_{Stat} is constant at high ARR, and the loss depends on the property of the fading distribution. Consequently, knowing the channel state when the ARR is high provides no gain. In fact, TX can achieve the same performance in terms of data transmitted regardless of its channel knowledge. Hence, TX can utilize during the communication the same power profile as the worst case ($\alpha = 1$) to achieve a similar performance as the best case ($\alpha = 0$).

V. Numerical Results

In this section, we present selected numerical results that illustrate our theoretical analysis. Indeed, we consider a point to point communication system in order to compare the average throughput of this system, for different scenarios of CSI-T availability.

A. Model for Simulation Results

In order to obtain the numerical results, the deadline of the transmission is fixed to $T = 10$ sec, the maximum energy that can be stored by the battery is $E_{max} = 100$ J. Also we consider a band-limited additive white Gaussian noise channel, with bandwidth W chosen $W = 1$ MHZ for simulations and the noise power spectral density is $N_0 = 10^{-19}$ W/Hz and $\sigma^2 = 1$. All simulations are performed for 5×10^4 channel realizations. As mentioned in our considered model, the data rate sent to RX at TS i is calculated as $C(h_i, p_i) = W \log_2(1 + \frac{|h_i|^2 p_i}{N_0 W})$ bits/sec. We assume that the energy takes values in a finite set, i.e., $N_s = 3$, for simulations we fix $\mathcal{E} = \{E_1^h = 0, E_2^h = \beta, E_3^h = \eta\}$, such that $0 < \beta \leq \eta$. For the EH process, we assume that the matrix of the transition probabilities $\mathbf{M}$ is defined by:

$$\mathbf{M} = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}. \tag{32}$$

It is clear that the steady state probability in this case is given by $\pi^* = [\mathbb{P}_{E_1^h}\mathbb{P}_{E_2^h}\mathbb{P}_{E_3^h}] = \frac{1}{3}[1\ 1\ 1]$. Thus, in this case, the average energy that can be harvested at instant t_i is defined by:

$$\overline{E} = \mathbb{E}(E_i) = \sum_{j=1}^{N_s=3} \mathbb{P}_{E_j^h} E_j^h = \frac{\beta + \eta}{3}. \tag{33}$$

B. Characteristics of the Optimal Power Policy

The optimal power policy is determined by induction. Fixing an error estimation variance α, the powers $p_1^*, p_2^*, \cdots, p_N^*$ are calculated recursively in the time reversal order. Starting by the evaluation of the reward function at the last TS N, using the closed form expression given in (17). Then, we calculate $\overline{V}_N(\widehat{\gamma}_N, S_N, \alpha)$ using (20), for different battery capacity B_N discretized in an adaptive step size depending on the regime where the system is running (low, medium and high ARR regime). We set the step size $\Delta B = 0.02$ J to plot figures 4, 6 and 7 and $\Delta B = 2 \times 10^{-7}$J to plot Fig.5. Also, $\overline{V}_N(\widehat{\gamma}_N, S_N, \alpha)$ is averaged over 5×10^4 channel realizations. Then, we calculate $V_{N-1}(\widehat{\gamma}_{N-1}, S_{N-1}, \alpha)$ solving (19) using a one line search method. The value of $V_{N-1}(\widehat{\gamma}_{N-1}, S_{N-1}, \alpha)$ is stored as a look-up table to be used for the next TSs. Then, the same procedure is applied for TSs $N-2, ..., 1$. At TS 1, TX determines the optimal reward function by evaluating $V_1(\underline{\widehat{\gamma}}, S_1, \alpha)$. To evaluate the optimal AT, the reward function is averaged over 5×10^4 independent realizations of $\hat{\gamma}$ and over all possible states for the initial state $\underline{S_1}$.

We start by examining the AT with different values of α. In Fig.4, we can see that the AT is a non-increasing function in α as stated in our analytical claims. Also, we can observe that the EH system with a static channel equal to $\overline{\gamma} = 1$, performs better than the one experiencing fading.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2654454, IEEE Access

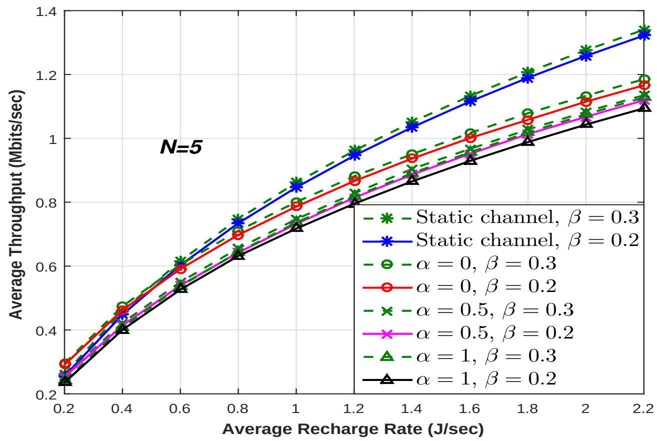

Fig. 4. Performance of the online policy with Rayleigh fading channels for various channel estimation errors, $E_{max} = 100$ J and $\sigma^2 = 1$.

However, in the low ARR regime, we remark that the fading channel has a better performance (red curve). This points out the benefit provided by the fading to the EH system at low ARR. This highlights the gain provided by the available CSI-T. Note that the performance depends on the states of the Markov chain associated to the EH process. In fact, having the same ARR does not mean having the same performance. In Fig.4, we compare two scenarios by setting the second state $\beta = 0.2$ J and $\beta = 0.3$ J and the third state is fixed by the ARR.

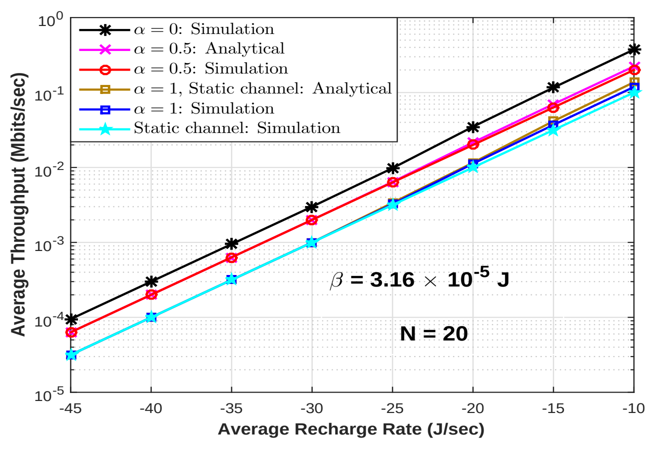

Fig. 5. Performance of the online policies at low ARR, $E_{max} = 100$ J and $\sigma^2 = 1$.

As discussed previously, simulations in Fig.4 stir us to boost our analysis on the low ARR regime presented in Fig.5. Such an analysis is reasonable and very useful for EH applications that are running at low ARR. We observe in Fig.5 that the results found are quite consistent with our theoretical claims. In fact, the EH system takes advantage of the partial knowledge of the channel to enhance its performance. We can see that when the channel experiences fading without CSI-T, the communication system has almost the same performance as the case where the channel is static with $\overline{\gamma} = 1$, (the brown and the blue curves). We use (22) to plot the cyan curve, simulation results in the case of the static channel and the case where $\alpha = 1$ confirm our theoretical claims. Also, we plot the AT using expression (30) (magenta curve) for $\alpha = 0.5$, and we remark that this curve is close to the simulated AT for this settings.

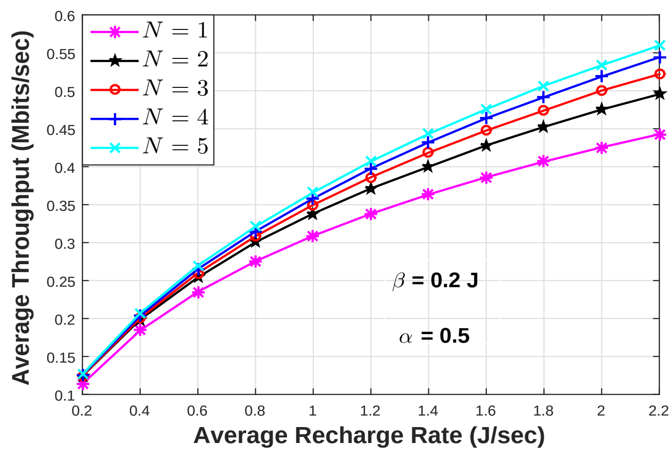

Fig. 6. Performance of the online policy with Rayleigh fading channels for various sensing periods, $E_{max} = 100$ J and $\sigma^2 = 1$.

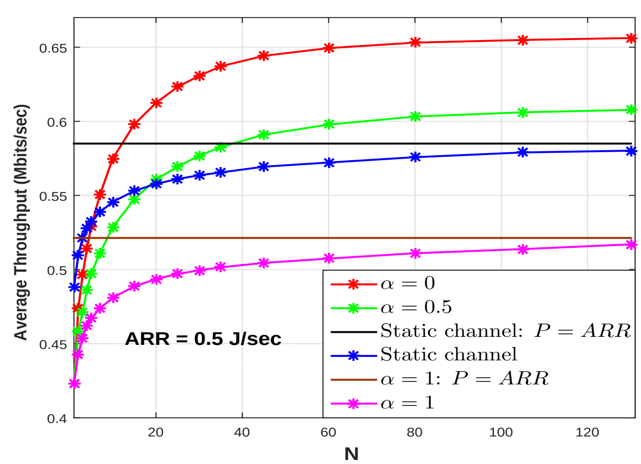

Fig. 7. Performance of the online policies for N very large, $E_{max} = 100$ J and $\sigma^2 = 1$.

In Fig.6, we investigate the effect of varying the number of TSs N on the AT. As we remark, the

EH communication system performs better as N increases. Simulations in this figure are performed for a fading channel with error variance $\alpha = 0.5$. Next, we investigate the performance of the EH for large sensing instants. We observe in Fig.7 that the EH system performs better when the channel experiences fading. This is reasonable because the ARR assumed for this simulation is relatively low. Furthermore, we note that beyond a certain value of N the AT saturates. For the two specific cases (static, no CSI-T), the upper bound is defined by transmitting symbols with constant power $P = ARR$ during all TSs. For instance, large number N corresponds to a higher sensing frequency, since the sensing is performed in the same space of states and the duration between two consecutive sensing $\frac{L}{T}$ sec is asymptotically equal to 0. Consequently, the EH sensing can be considered in this case as a continuous process in time. Hence, the communication system can always harvest energy during the deadline T, as if it has the ability to scavenge "infinitely" the energy.

VI. Conclusion

In this paper, we have focused on optimal power scheduling of point-to-point communication systems where TX is powered by the EH technique and has access only for an estimate version for the fading level of the channel. On the other hand, the CSI-R is assumed to be perfectly known. Indeed, we have proposed a stochastic model for the energy arrival process. Specifically, a first-order Markov has been defined as the joint process of the energy and the battery state. Furthermore, we have derived the optimal power policy that maximizes the throughput using the DP framework. Asymptotic analysis of the AT has been provided when ARR is either very low or very high. We remarked that the fading case outperforms the static case in terms of data transmitted in the low ARR regime. Unlike the low regime, the CSI-T does not provide an enhancement on the performance in high regime. This work provides an interesting study of the point-to-point communication system under EH constraints. Indeed, the analysis has been done from a throughput maximization point of view assuming imperfect CSI-T. Furthermore, the performance of the communication system has been studied in the asymptotic regimes of the ARR either low or high.

Appendix A
Proof of Proposition 3

Lemma 1. Given a random EH process and the matrix of transition probabilities $\mathbf{M}$, the optimal policy when the CSI is unavailable at TX is the same as the power policy in the case when the channel is static during communication.

Proof. Let us denote $(p_i^*)_i$ the optimal powers for the throughput maximization problem in the case where the channel is static, i.e., the powers that maximize the following objective function:

$$\max_{\{p_i\}_{i=1}^N} \mathbb{E}_{S_2^N}\left[\sum_{i=1}^{N} L\,\log(1+\overline{\gamma}p_i) \mid \mathbf{M}, S_1\right]. \quad (34)$$

The problem in the no CSI-T setting is:

$$\max_{\{p_i\}_{i=1}^N} \mathbb{E}_{S_2^N}\left\{\mathbb{E}_\gamma\left[\sum_{i=1}^{N} L\,\log(1+\gamma p_i)\right] \mid \mathbf{M}, S_1\right\}. \quad (35)$$

Since $(p_i^*)_i$ is the optimal power policy when the channel is static, and it is clear form the objective function that the optimal powers do not depend on the channel gain $\overline{\gamma}$. The optimal solution depends only on the energies harvested along time. Thus, $(p_i^*)_i$ does not depend on the static channel gain, then we have :

$$\sum_{i=1}^{N} L\,\log(1+\gamma p_i) \le \sum_{i=1}^{N} L\,\log(1+\gamma p_i^*),\ \forall\gamma \quad (36)$$

$$\mathbb{E}_{S_2^N}\left[\sum_{i=1}^{N} L\,\log(1+\gamma p_i) \mid \mathbf{M}, S_1\right] \le \mathbb{E}_{S_2^N}\left[\sum_{i=1}^{N} \frac{L}{2}\log(1+\gamma p_i^*) \mid \mathbf{M}, S_1\right] \quad (37)$$

$$\max_{\{p_i\}_{i=1}^N} \mathbb{E}_\gamma\left\{\mathbb{E}_{S_2^N}\left[\sum_{i=1}^{N} L\,\log(1+\gamma p_i) \mid \mathbf{M}, S_1\right]\right\} \le \mathbb{E}_\gamma\left\{\mathbb{E}_{S_2^N}\left[\sum_{i=1}^{N} L\,\log(1+\gamma p_i^*) \mid \mathbf{M}, S_1\right]\right\}. \quad (38)$$

So, $\mathbb{E}_\gamma\left\{\mathbb{E}_{S_2^N}\left[\sum_{i=1}^{N} L\,\log(1+\gamma p_i^*) \mid \mathbf{M}, S_1\right]\right\}$ is an upper bound for the problem (35) that can be achieved by the power policy $(p_i^*)_i$, also this policy is feasible because it satisfies the EH constraints. Consequently, the optimal power policy is $(p_i^*)_i$. □

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2654454, IEEE Access

Let $\alpha = 1$, the TX does not have information about the channel. We have just proved in Lemma 1 that the optimal power allocation is the same as the case where the channel is static. Thus, we have:

$$AT_1 = \mathbb{E}_{S_1^N}\left\{\mathbb{E}_{\tilde{\gamma}}\left(\sum_{i=1}^{N} L\ \log(1+\tilde{\gamma}p_i^*)\right) \mid \mathbf{M}\right\} \tag{39}$$

$$\approx \mathbb{E}_{S_1^N}\left\{\sum_{i=1}^{N}\frac{L}{T}\mathbb{E}_{\tilde{\gamma}}\left(\log(\tilde{\gamma}p_i^*)\right) \mid \mathbf{M}\right\} \tag{40}$$

$$= \mathbb{E}_{S_1^N}\left\{\sum_{i=1}^{N}\frac{L}{T}(\log(p_i^*)+\mathbb{E}_{\tilde{\gamma}}(\log(\tilde{\gamma}))) \mid \mathbf{M}\right\} \tag{41}$$

$$\approx AT_{Stat} + \mathbb{E}_{\tilde{\gamma}}\left[\log(\tilde{\gamma})\right]. \tag{42}$$

Let a fading channel with $\underline{\hat{\gamma}} = (\hat{\gamma}_1, \hat{\gamma}_2, ..., \hat{\gamma}_N)$ be an estimates of the fading level during transmission.

Lemma 2. Given the same initial state, the optimal power policy when the channel is perfectly known is related to the optimal power policy when the channel is static. In fact, the optimal powers satisfy the property:

$$\mathbb{E}_{\underline{\hat{\gamma}}}\left\{p_i^*(\underline{\hat{\gamma}})\right\} = p_{stat,i}^*, \tag{43}$$

where, $\underline{\hat{\gamma}}$ is a channel realization and $p_{stat,i}^*$ is the optimal power delivered during the TS i in the case where the channel is static.

Proof. Let us evaluate $\mathbb{E}_{\underline{\hat{\gamma}}}\left\{p_i^*(\underline{\hat{\gamma}})\right\}$ between two successive energy arrivals E_{i-1} and E_i. We have the average is over possible values of the channel, so $\mathbb{E}_{\underline{\hat{\gamma}}}\left\{p_i^*(\underline{\hat{\gamma}})\right\}$ depends only on the energy arrivals. Therefore, to maximize the throughput, the average energy (averaging with respect to the channel) consumed during the TS i between the two energy arrivals should be allocated optimally. Since, the average energy used does not depend on the channel, an optimal energy profile should be similar to the optimal energy profile when the CSI-T is unavailable. Using result stated in Lemma 2, the average energy consumed during this TS should be equal to the energy used when the channel is static between the two instants of energy arrivals. So, we have:

$$\mathbb{E}_{\underline{\hat{\gamma}}}\left\{Lp_i^*(\underline{\hat{\gamma}})\right\} = Lp_{stat,i}^*.$$

Hence, Lemma 2 follows immediately. □

Now, let us evaluate the AT when $\alpha = 0$:

$$AT_0 = \mathbb{E}_{S_1^N}\left\{\mathbb{E}_{\underline{\hat{\gamma}}}\left[\sum_{i=1}^{N}\frac{L}{T}\log(1+\hat{\gamma}_i p_i^*(\underline{\hat{\gamma}}))\right] \mid \mathbf{M}\right\} \tag{44}$$

$$\approx \mathbb{E}_{S_1^N}\left\{\mathbb{E}_{\underline{\hat{\gamma}}}\left[\sum_{i=1}^{N}\frac{L}{T}\log(\hat{\gamma}_i p_i^*(\underline{\hat{\gamma}}))\right] \mid \mathbf{M}\right\} \tag{45}$$

$$= \mathbb{E}_{S_1^N}\left\{\sum_{i=1}^{N}\frac{L}{T}\mathbb{E}_{\underline{\hat{\gamma}}}\left(\log\hat{\gamma}_i\right)\right\} + \mathbb{E}_{S_1^N}\left\{\mathbb{E}_{\underline{\hat{\gamma}}}\left[\sum_{i=1}^{N}\frac{L}{T}\log(p_i^*(\underline{\hat{\gamma}}))\right] \mid \mathbf{M}\right\}. \tag{46}$$

We have $(\hat{\gamma}_i)_{i\in[\![1,N]\!]}$ are i.i.d., so:

$$\mathbb{E}_{S_1^N}\left\{\sum_{i=1}^{N}\frac{L}{T}\mathbb{E}_{\underline{\hat{\gamma}}}\left(\log\hat{\gamma}_i\right)\right\} = \sum_{i=1}^{N}\frac{L}{T}\mathbb{E}_{\hat{\gamma}_i}\left(\log\hat{\gamma}_i\right) \tag{47}$$

$$= \sum_{i=1}^{N}\frac{L}{T}\mathbb{E}_{\hat{\gamma}}\left(\log\hat{\gamma}\right) \tag{48}$$

$$= \mathbb{E}_{\hat{\gamma}}\left(\log\hat{\gamma}\right). \tag{49}$$

Because the $\log$ function is concave, then:

$$\mathbb{E}_{S_1^N}\left\{\mathbb{E}_{\underline{\hat{\gamma}}}\left[\sum_{i=1}^{N}\frac{L}{T}\log(p_i^*(\underline{\hat{\gamma}}))\right] \mid \mathbf{M}\right\} \leq \mathbb{E}_{S_1^N}\left\{\sum_{i=1}^{N}\frac{L}{T}\log(\mathbb{E}_{\underline{\hat{\gamma}}}\left[p_i^*(\underline{\hat{\gamma}})\right]) \mid \mathbf{M}\right\}. \tag{50}$$

Then, we have:

$$AT_0 \leq \mathbb{E}_{S_1^N}\left\{\sum_{i=1}^{N}\frac{L}{T}\log(\mathbb{E}_{\underline{\hat{\gamma}}}\left[p_i^*(\underline{\hat{\gamma}})\right]) \mid \mathbf{M}\right\} + \mathbb{E}_{\hat{\gamma}}\left(\log\hat{\gamma}\right). \tag{51}$$

Using Lemma 2

$$AT_0 \leq \mathbb{E}_{S_1^N}\left\{\sum_{i=1}^{N}\frac{L}{T}\log(p_{stat,i}^*) \mid \mathbf{M}\right\} + \mathbb{E}_{\hat{\gamma}}\left(\log\hat{\gamma}\right) \tag{52}$$

$$= AT_{Stat} + \mathbb{E}_{\hat{\gamma}}\left(\log\hat{\gamma}\right). \tag{53}$$

Hence (31) follows immediately.

References

[1] D. Gündüz, K. Stamatiou, N. Michelusi, and M. Zorzi, "Designing intelligent energy harvesting communication systems," *IEEE Commun. Mag.*, vol. 52, no. 1, pp. 210-216, Jan. 2014.

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

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

[4] K. Tutuncuoglu, and A. Yener, "Optimum transmission policies for battery limited energy harvesting nodes," *IEEE Trans. Wireless Commun.*, vol 11, no. 3, pp. 1180-1189, Mar. 2012.

[5] J. Yang, and S. Ulukus, "Optimal packet scheduling in an energy harvesting communication system," *IEEE Trans. Commun.*, vol 60, no. 1, pp. 220-230, Jan. 2012.

[6] C. Ho, and R. Zhang, "Optimal energy allocation for wireless communications with energy harvesting constraints," *IEEE Trans. Signal. Process.*, vol 60, no. 9, pp. 4808-4818, Sep. 2012.

[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, no. 4, pp. 1326-1336, Apr. 2010.

[8] J. Xu and R. Zhang, "Throughput optimal policies for energy harvesting wireless transmitters with non-ideal circuit power," *IEEE J. Sel. Area. Commun.*, vol. 32, no. 2, pp. 322-332, Feb. 2014.

[9] Z. Wang, V. Aggarwal, and X. Wang, "Power allocation for energy harvesting transmitter with causal information," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 4080-4093, Nov. 2014.

[10] R. Ma, and W. Zhang, "Optimal power allocation for energy harvesting communications with limited channel feedback," *IEEE Global Conf. Signal Inf. Process.*, Atlanta, GA, USA, pp. 193-197, Dec. 3-5, 2014.

[11] M. L. Ku, Y. Chen, and K. J. R. Liu, "Data-driven stochastic models and policies for energy harvesting sensor communications," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 8, pp. 1505-1520, Aug. 2015.

[12] C. K. Ho, D. K. Pham, and C. M. Pang, "Markovian models for harvested energy in wireless communications," in *Proc. IEEE Int. Conf. on Commun. Syst. (ICCS), Singapore, Nov. 17-20 2010*, Singapore, Nov. 2010.

[13] J. Zhang, C. Yuen, C.K. Wen, S. Jin, K.K. Wong and H. Zhu "Large system secrecy Rate analysis for SWIPT MIMO wiretap channels," *IEEE Trans. Inf. Forensics Secur.*, vol. 11, no. 1, pp. 74-85, Jan. 2016.

[14] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inf. Theory.*, vol. 46, pp. 933-946, May 2000.

[15] Z. Rezki, and M.-S. Alouini, "Ergodic capacity of cognitive radio under imperfect channel-state information," *IEEE Trans. Veh. Technol.*,vol 61, no. 5, pp. 2108-2119, Jun. 2012.

[16] M. R. Zenaidi, Z. Rezki, H. Tembine, and M.-S. Alouini, "Performance Limits of Energy Harvesting Communications under Imperfect Channel State Information", in Proc. IEEE International Conference on Communications (ICC), May 2016, Kuala Lampur, Malaysia. An extension of this work has been submitted to *IEEE Transactions on Wireless Communications*.

[17] M. L. Puterman, *Markov Decision Process: Discrete Stochastic Dynamic Programming.* NY: John Wiley & Sons, 1994.

[18] D. P. Bertsekas, *Dynamic Programming and Optimal Control.* 2nd ed., vols. 1 and 2. Athena Scientific, 2001.

[19] S. C. Albright, *Structural results for partially observable Markov decision processes* Operations Research, vol. 27, no. 5, pp. 1041-1053. Oct. 1979.

[20] W. S. Lovejoy, *Some monotonicity results for partially observable Markov decision processes.* Operations Research, vol. 35, no. 5, pp. 736-743, Oct. 1987.

[21] M. R. Zenaidi, Z. Rezki, and M.-S. Alouini, *"On Communications under Stochastic Energy Harvesting with Noisy Channel State Information,"* in Proc. *IEEE Glob. Commun. Conf. (GLOBECOM'16)*, Dec 2016, Washington DC, USA.

[22] M. K. Simon, and M.-S. Alouini, *Digital communication over fading channels.* Wiley-Interscience, 2005, vol. 95.

[23] A. Goldsmith, and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inf. Theory,*, vol 43, pp. 1986-1992. Nov. 1997.