On the optimal resource allocation for a wireless energy harvesting node considering the circuitry power consumption

Maria Gregori, *Student Member, IEEE,* and Miquel Payaró, *Senior Member, IEEE*

***Abstract*—In this paper, an energy harvesting transmitter operating in a point-to-point link through a discrete-time fading channel is considered, where symbols can be transmitted through several parallel independent streams. Taking into account the different sources of energy consumption at the transmitter, the resource allocation (in terms of power allocation and selection of the active streams and channel accesses) that asymptotically maximizes the mutual information is derived by assuming that the transmitter has non-causal knowledge of the harvested energy and channel state (offline approach). The *Boxed Water-Flowing* graphical interpretation is presented, which intuitively depicts the asymptotically optimal offline resource allocation. Moreover, an online algorithm is proposed for the case in which the transmitter only has causal (past and present) knowledge of the harvested energy and channel state. Finally, the performance of the proposed offline and online solutions is numerically evaluated and the associated computational complexities are assessed and compared.**

***Index Terms*—Energy harvesting, mutual information, power allocation, circuitry power consumption.**

I. Introduction

Energy availability is becoming the major operational bottleneck for devices with high mobility requirements (e.g., hand-held devices) or devices that have difficulties to access the power grid (e.g., sensor nodes). Energy harvesting, i.e., the process by which energy of different kinds (e.g., light, temperature, wind, etc.) is collected from the environment and converted into usable electric power, is a potential technology to increase the operational lifetime of battery powered devices. Traditional power allocation strategies (e.g., Classical Water-filling (CWF) [5]) are no-longer optimal when the transmitter has the ability to harvest energy from the environment since the transmitter must satisfy the Energy Causality Constraints (ECCs), which impose that the energy used by the node must be smaller than or equal to the energy harvested. Therefore, energy harvesting opens a new research paradigm in the design of online and offline resource allocation strategies. The *online* approach accounts for the information causality at the transmitter regarding both the dynamics of the energy harvesting process and channel state (i.e., the transmitter only knows the past and current occurrences of these processes). In some situations, the transmitter may also have statistical knowledge of the dynamics of these processes. The *offline* approach assumes that the node has full knowledge of the energy harvesting process and channel state, which is only a realistic assumption under very specific scenarios where the channel is static and the energy source is controllable (e.g., in wireless power transmission scenarios). In any case, the interest of studying the offline approach is that it provides analytical and intuitive solutions and an upper bound on the performance of any online algorithm. Therefore, the derivation of the optimal offline solution is a good first step to gain insight for the later design and evaluation of the online transmission strategy, which is the one that can potentially be deployed in practical set ups. In this context, [6]–[11] considered a point-to-point communication with an energy harvesting transmitter. The Directional Water-Filling (DWF), i.e., the power allocation strategy that maximizes the mutual information by a deadline was derived in [9] as[1]

$$P_n = \left(W_j - \tfrac{1}{h_n}\right)^+, \quad \forall n \in \tau_j, \tag{1}$$

where n is the channel use index, h_n is the channel gain, τ_j is a set that contains the consecutive channel accesses between two energy arrivals and W_j is the water level associated to τ_j. The difference between CWF and DWF is that in the latter the water level changes over time.

The result in (1) assumes that the radiated power is the unique source of energy consumption. This is a reasonable assumption when the link distance is large since the radiated power dominates over other sources of energy consumption. However, when energy efficient network topologies are considered, the trend is to reduce the transmission range by implementing multiple hops and, then, the other sources of energy consumption at the transmitter become relevant and may even dominate over the radiated power [12], [13]. A more realistic power consumption model is given in [1], [13], [14], where the total consumed power at the n-th channel access is modeled as

$$P_n^{(total)} = \begin{cases} \frac{\xi}{\eta} P_n + P_c & \text{if } P_n > 0, \\ P_o & \text{if } P_n = 0, \end{cases} \tag{2}$$

The authors are with the Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), 08860 - Castelldefels, Barcelona, Spain (e-mails: maria.gregori@cttc.cat and miquel.payaro@cttc.cat).

This work was partially supported by: the Catalan Government under grants 2013FI_B2 00075 and 2014 SGR 1551; the Spanish Ministry of Economy and Competitiveness under project TEC2011-29006-C03-01 (GRE3N-PHY); and the European Commission under project Network of Excellence in Wireless Communications (Newcom#, Grant Agreement 318306).

[1]**Notation:** Matrices and vectors are denoted by upper and lower case bold letters, respectively. $[\mathbf{v}]_n$ denotes the n-th component of the vector $\mathbf{v}$. $[\mathbf{A}]_{pq}$ is the component in the p-th row and q-th column of the matrix $\mathbf{A}$. $\preceq$ denotes the component wise "smaller than or equal to" inequality. Finally, $(x)^+ = \max\{0, x\}$.

Table I
STATE OF THE ART: CIRCUITRY ENERGY CONSUMPTION IN WIRELESS ENERGY HARVESTING NODES (WEHNS)

	Continuous channel model		Discrete channel model
	Static	Temporal fading	Temporal fading
SISO	[1]	[2]	[3], [4]
Frequency/space parallel channels	[1]: Circuitry cost per temporal access only	**This paper**: Circuitry cost per temporal and frequency/spatial accesses	

where P_n is the transmission radiated power; P_c is the power consumption of the different components of the Radio Frequency (RF) chain; ξ and η are the power amplifier output back-off and drain efficiency, respectively; and P_o models the circuitry consumption when the transmitter is silent, which is much smaller than P_c and it is commonly assumed to be zero.

In [15] and [16], the authors considered non-harvesting wireless nodes, analyzed how CWF is modified when the circuitry energy consumption is considered and showed that bursty transmission achieves capacity.

The works that study the impact of the circuitry energy consumption in point-to-point links with WEHNs are listed in Table I. In [1] and [2], a continuous-time channel was considered: the authors of [2] studied a Single-Input Single-Output (SISO) channel and showed that the mutual information maximization problem is convex when the channel is continuous in time; whereas in [1], a system composed by multiple parallel Additive White Gaussian Noise (AWGN) channels was studied, but the channel was considered static along time, which substantially simplifies the analysis since, when the channel is static, there is no tradeoff between channel gain and energy availability (see Section II). Additionally, Xu et al. [1] considered a power consumption model that has a fixed cost for activating the transmitter in a given time instant independently of the number of active parallel channels. Due to this, their model is only applicable to a limited set of transmitter architectures as it is argued in Section II. In opposition to [1] and [2], we consider a WEHN operating in a discrete-time channel, composed of multiple parallel streams at each channel use, and that is affected by temporal and spatial/frequency fading. The fact of considering a discrete-time channel model is key because it is the actual channel model that is being used in current digital communication systems, e.g., in Orthogonal Frequency Division Multiplexing (OFDM). As it is later shown, the discreteness of the channel and the temporal variations of the channel coefficients substantially complicate the problem since it is no longer convex. In prior preliminary works [3], [4], we studied a simplified scenario where each channel use is modeled as a SISO AWGN channel. In contrast, the major contributions of this paper are:

- Generalizing the power consumption model in (2) to consider multiple parallel AWGN channels and showing its applicability in practical transmitter architectures.
- Studying the resource allocation that maximizes the mutual information over N channel accesses when there are multiple parallel data streams by jointly considering energy harvesting and the different sources of energy consumption at the transmitter.
- Deriving an upper bound of the achievable mutual information and two asymptotically optimal solutions of the offline maximization problem, i.e., solutions that tend to the optimal when the number of streams or channel accesses grows without bound.
- Proposing an intuitive graphical representation of the asymptotically optimal offline solution, named *Boxed Water-Flowing*.
- Implementing an online algorithm that achieves a mutual information that is close to the one achieved by the optimal offline solution.
- Evaluating and comparing the computational complexities of the proposed strategies.

This paper is structured as follows. In Section II, the system model and problem formulation are presented. The offline resource allocation problem is studied in Section III from two perspectives: integer relaxation (Section III-A) and through the dual problem (Section III-B). A graphical interpretation of the offline solution is presented in Section III-C. In Section IV, the online solution is presented. The mutual information and computational requirements of the different algorithms are evaluated in Section V. Finally, the paper is concluded in Section VI.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a WEHN transmitting in a point-to-point link in which, at each channel access, the communication channel can be decomposed into a set of K parallel non-interfering streams by performing some joint signal processing at the transmitter and receiver, e.g., by using OFDM or by diagonalizing a Multiple-Input Multiple-Output (MIMO) channel. Let $y_{k,n}$ be the channel output of the k-th stream at the n-th channel access, i.e., $y_{k,n} = \sqrt{P_{k,n}} g_{k,n} x_{k,n} + w_{k,n}$, where $x_{k,n}$ is the input symbol with $E\{|x_{k,n}|^2\} = 1$, $P_{k,n}$ is the radiated power, $g_{k,n}$ is the complex channel response with $h_{k,n} = |g_{k,n}|^2$ being the channel power gain, and $w_{k,n} \sim CN(0,1)$ is the noise. First, in Section III, we assume that the transmitter has non-causal knowledge of all the channel gains. This assumption is removed in Section IV for the design of the online algorithm.

The energy harvesting process at the transmitter is commonly characterized by a packetized model, e.g., [7], [9], where the node is able to collect a packet of energy containing E_j Joules at the beginning of the e_j-th channel access[2] for some $e_j \in [1, N]$. J denotes the total number of harvested energy packets. The initial battery level of the node E_1 is modeled as the first harvested packet, thus, $e_1 = 1$ and the battery

[2]We assume that the transmission strategy can only be changed in a channel use basis. Thus, we can consider that, independently of the energy packet arrival instant, it becomes available for the transmitter at the beginning of the following channel use.

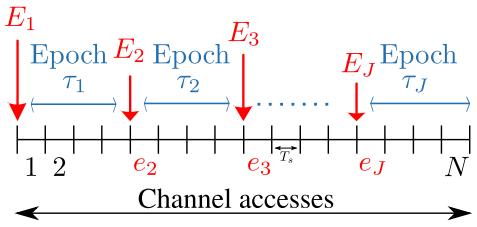

Figure 1. Temporal representation of energy arrivals.

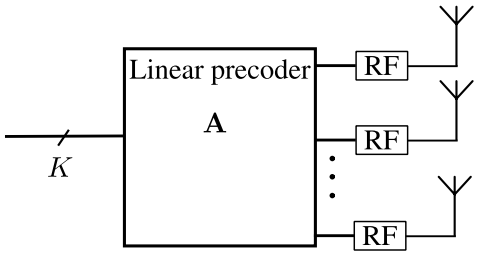

Figure 2. Example of MIMO and V-BLAST ($\mathbf{A} = \mathbf{I}_K$) transmitter architectures.

capacity is assumed to be infinite. We use the term *epoch* τ_j, $j = 1, \ldots, J$, to denote the set of channel accesses between two energy arrivals, i.e., $\tau_j = \{e_j, e_j + 1, \ldots, e_{j+1} - 1\}$ with $e_{J+1} = N + 1$ so that the last epoch is well defined. A temporal representation is given in Figure 1. Throughout the paper, $j \in [1, J]$ is the epoch index, $n \in [1, N]$ denotes the channel access index, and $k \in [1, K]$ is the stream index.

The power consumption at the transmitter depends on its hardware and software architecture. As we focus on architectures with multiple data streams, it naturally follows that the transmitter may experience either a power consumption associated with the channel access activation, P_c, or a power consumption associated with the activation of each of the streams, P_d, or both simultaneously. In this context, we propose a power consumption model that, as shown later, can be applied to several transmitter architectures, which is a generalization of (2) to scenarios in which there exist several parallel streams. The total consumed power at the n-th channel access is modeled as:

$$P_n^{(total)} = (1-\rho_n)P_o + \rho_n \left(P_c + \sum_{k=1}^{K} \left(\frac{\xi}{\eta} P_{k,n} + \psi_{k,n} P_d \right) \right),$$

where $\psi_{k,n}$ is the stream indicator variable that denotes whether a certain stream is active, i.e., $\psi_{k,n} = 1$ if $P_{k,n} > 0$ or $\psi_{k,n} = 0$ if $P_{k,n} = 0$. Similarly, the channel access indicator variable, ρ_n, denotes when a certain channel access is active, i.e., $\rho_n = 1$ if $\sum_k P_{k,n} > 0$ or $\rho_n = 0$ if $\sum_k P_{k,n} = 0$. The constants P_o, P_c and P_d model the idle state power consumption and the power consumptions associated with the channel access and stream activation, respectively. These constants have to accurately capture the different sources of energy consumption and are dependent on the transmitter hardware/software architecture. For example, consider the following architectures:

MIMO linear precoding: The information of the different streams is linearly processed by a precoding matrix, $\mathbf{A}$, and transmitted over the different antennas (see Figure 2). If the channel access is active, the precoding operation usually[3] activates all the RF chains at the transmitter independently of the number of active streams. Thus, P_c would account for the circuitry power consumption of all RF chains, whereas, the number of active streams affects on the number of products and summations required for the linear precoding. In this context, P_d models the power consumption at the base band processing boards of one of these products and additions.

V-BLAST: The information of the different streams is directly sent over the channel without performing any linear precoding, i.e., in Figure 2, set $\mathbf{A} = \mathbf{I}_K$. In this architecture, the number of active RF chains is equal to the number of active streams and thus P_d accounts for the power consumption of each RF chain and $P_c = 0$ since there is no additional cost per channel access.

OFDM: The number of operations required for the Inverse Fast Fourier Transform (IFFT) depends on the number of active streams. Thus, the cost per operation of the IFFT can be modeled through P_d and the cost of performing the serial to parallel conversion is mapped in P_c.

As mentioned in the introduction, the previous results for multi-stream communications with WEHNs [1], only considered architectures in which $P_c \neq 0$ and $P_d = 0$. Moreover, the results in [1] only apply to time-static channels. Thus, we extend the results obtained in [1] to a broader class of transmitter architectures and to consider time-varying fading channels.

Let $\mathbf{P} = [\mathbf{p}_1, \ldots, \mathbf{p}_N]$ and $\mathbf{\Psi} = [\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_N]$ be the $K \times N$ matrices that contain the radiated powers and stream indicator variables, where $\mathbf{p}_n = [P_{1,n}, \ldots, P_{K,n}]^\mathsf{T}$ and $\boldsymbol{\psi}_n = [\psi_{1,n}, \ldots, \psi_{K,n}]^\mathsf{T}$ stack the transmission powers and stream indicator variables of the different streams at the n-th channel access. Similarly, $\boldsymbol{\rho} = [\rho_1, \ldots, \rho_N]$ is a $1 \times N$ vector that contains the channel access indicator variables. The optimal resource allocation must fulfill the ECCs, which impose that the energy spent by the end of the ℓ-th epoch, $\ell = 1, \ldots, J$,

$$\phi_\ell(\mathbf{P}, \boldsymbol{\rho}, \mathbf{\Psi}) = T_s \sum_{j=1}^{\ell} \sum_{n \in \tau_j} \left(P_c \rho_n + \sum_{k=1}^{K} (P_{k,n} + P_d \psi_{k,n}) \right), \tag{3}$$

cannot be greater than the energy harvested up to the beginning of the epoch, $\beta_\ell = \sum_{j=1}^{\ell} E_j$. To simplify the problem notation and without loss of generality we have assigned $\xi/\eta = 1$ and $P_o = 0$ (note that we can scale the constants P_c, P_d, and β_ℓ of the ECCs to have $\xi/\eta = 1$ and $P_o = 0$).[4]

Assuming Gaussian distributed input symbols, the mutual information at the k-th stream of the n-th channel use is $\mathsf{log}(1 + h_{k,n} P_{k,n})$ and the accumulated mutual information is $I(\mathbf{P}, \boldsymbol{\rho}, \mathbf{\Psi}) = \sum_{j=1}^{J} \sum_{n \in \tau_j} \rho_n \sum_{k=1}^{K} \psi_{k,n} \mathsf{log}(1 + h_{k,n} P_{k,n})$. The goal of this paper is to derive the resource allocation $\mathbf{P}$,

[3]Excluding some specific precoder designs such as $\mathbf{A} = \mathbf{I}_K$, the precoder is generally designed to transmit through the channel eigenmodes and as a result, if one stream is active, all the elements at the output of the precoding matrix are "active" (different than zero).

[4]We assume that the node has enough energy to at least be kept in the idle state during the whole transmission duration, i.e., $T_s \sum_{j=1}^{\ell} \sum_{n \in \tau_j} P_o \leq \sum_{j=1}^{\ell} E_j, \forall \ell = 1, \ldots, J$, as otherwise the problem would not have a feasible solution.

$\boldsymbol{\rho}$, and $\boldsymbol{\Psi}$ that maximizes the mutual information $I(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi})$ while satisfying the ECCs,

$$\mathcal{I}^\star = \max_{\mathbf{P},\boldsymbol{\rho},\boldsymbol{\Psi}\in\mathcal{F}} \quad I(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}) \tag{4a}$$

$$\text{subject to} \quad \boldsymbol{\phi}(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}) \preceq \boldsymbol{\beta}, \tag{4b}$$

where $\boldsymbol{\phi}(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}) = [\phi_1(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}), \ldots, \phi_J(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi})]^\mathsf{T}$ and $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_J]^\mathsf{T}$. The feasible set of the optimization variables is $\mathcal{F} = \{P_{k,n} \geq 0, \rho_n \in \{0,1\}, \psi_{k,n} \in \{0,1\}, \forall k, n\}$.[5] To simplify the notation and without loss of generality, the remainder of the paper assumes that, within a given time index n, the streams are sorted with non-increasing channel gains, i.e., $h_{1,n} \geq h_{2,n} \cdots \geq h_{K,n}$, $\forall n$.

Note that (4) is not a convex optimization problem as the feasible set $\mathcal{F}$ is not a convex set and the objective function is not jointly convex in the optimization variables. The complexity of (4) lies in the selection of the active channel accesses and streams ($\boldsymbol{\rho}$ and $\boldsymbol{\Psi}$), because once $\boldsymbol{\rho}$ and $\boldsymbol{\Psi}$ are fixed, the optimal power allocation in the active streams is given by the DWF in (1) [3].

The optimal stream and channel access selection depends on the tradeoff between the magnitude of the *channel gain* and the *energy availability* and is a hard combinatorial problem [3], [4]. For instance, assuming that we knew in advance that a single channel access and stream is active, then we could wonder which would be the pair of indices (k, n) among the KN possibilities that provides the highest mutual information: the one with the best gain or some other pair that has the highest energy availability but worse gain? The answer depends on the specific values of the channel gains and the energy arrival distribution and, hence, the derivation of the optimal solution to (4) is not straight forward.

Remark 1. If the transmitter does not have energy harvesting capabilities (which means that it is only powered by the initial energy in the battery), then the presented system model still applies by particularizing $J = 1$. To the best of our knowledge, even for the particular case $J = 1$, this paper is the first work to derive an asymptotically optimal power allocation for battery operated nodes in a fading channel by considering both the channel access and stream activation costs (P_c and P_d).

Remark 2. The system model and problem formulation could also include: (i) instantaneous mask constraints on the transmission power; and (ii) concave non-linearities of the RF amplifier.[6] Although the structure of the solution and its graphical interpretation depend on the considered scenario, the numerical algorithms proposed in the remaining of the paper can be trivially extended to include (i) and (ii).

In this context, in the following section we study two different offline feasible solutions that perform close to $\mathcal{I}^\star$, whereas, in Section IV we propose an online resource allocation.

[5]Note that, by using non-continuous indicator functions on the ECCs, one could formulate an equivalent problem to (4) in which $\mathbf{P}$ is the unique optimization variable. In this paper, we have introduced the binary variables $\boldsymbol{\rho}$ and $\boldsymbol{\Psi}$, to avoid the discontinuities associated to the indicator functions.

[6]The objective function should be modified to $\sum_{j=1}^{J}\sum_{n\in\tau_j}\rho_n\sum_{k=1}^{K}\psi_{k,n}\mathsf{log}(1 + h_{k,n}g(P_{k,n}))$, where $g(\cdot)$ is the non-linear concave function that returns the output power at the RF amplifier as a function of the input power. Thus, the design variable would be the input power at the RF amplifier.

III. Offline resource allocation

In this section, we analyze the offline resource allocation from two different perspectives, namely, integer relaxation (Section III-A) and duality (Section III-B). In Section III-C, we present a graphical interpretation of the asymptotically optimal offline resource allocation.

A. Integer relaxation

In this section, we relax the original problem in (4) and formulate a similar convex optimization problem whose solution upper bounds the solution to (4). Moreover, from the optimal resource allocation of the relaxed problem, we derive a feasible solution to the original problem in (4) whose mutual information is close to $\mathcal{I}^\star$.

In this context, we have modified the objective function in (4) so that the new objective function, i.e., $\tilde{I}(\mathbf{P}, \boldsymbol{\Psi}) = \sum_{j=1}^{J}\sum_{n\in\tau_j}\sum_{k=1}^{K}\psi_{k,n}\mathsf{log}(1 + h_{k,n}P_{k,n}/\psi_{k,n})$, is jointly concave in the optimization variables and we have relaxed the binary constraint in the indicator variables, i.e., letting ρ_n and $\psi_{k,n}$ to be in the interval $[0, 1]$.[7] The relaxed problem to (4) is mathematically expressed as

$$\tilde{\mathcal{I}}^\star = \max_{\mathbf{P},\boldsymbol{\rho},\boldsymbol{\Psi}} \quad \tilde{I}(\mathbf{P}, \boldsymbol{\Psi}) \tag{5a}$$

$$\text{subject to} \quad \boldsymbol{\phi}(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}) \preceq \boldsymbol{\beta}, \tag{5b}$$

$$\psi_{k,n} \leq \rho_n, \qquad \forall k, n, \tag{5c}$$

$$\psi_{k,n} \leq 1, \ -\psi_{k,n} \leq 0, \qquad \forall k, n, \tag{5d}$$

$$-P_{k,n} \leq 0, \qquad \forall k, n. \tag{5e}$$

Note that in order to have a jointly concave objective function, we have removed the dependency on ρ_n from the objective function; however, we have included the channel access activation constraint in (5c) to ensure that the channel access indicator variable is at least as large as the pointwise maximum of the stream indicator variables. Therefore, with (5c), we force that if any stream is active, $\psi_{k,n} > 0$, the associated channel access circuitry consumption $\rho_n P_c$ is correctly accounted for in the ECCs in (5b), where $\phi_\ell(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi})$ is defined in (3). In (5d) and (5e), we ensure that $\psi_{k,n}$ and $P_{k,n}$ lie in their respective feasible sets. We do not have any constraint in the feasible set of ρ_n; however, note that the value of the optimal ρ_n is always in $[0, 1]$ since, in order to reduce the term $\rho_n P_c$ in the ECCs in (5b) (see (3)), the optimal ρ_n takes the minimum allowed value in (5c) that is the pointwise maximum of the optimal $\psi_{k,n}$, i.e., $\tilde{\rho}_n^\star = \max\{\tilde{\psi}_{1,n}^\star, \ldots, \tilde{\psi}_{K,n}^\star\}$, therefore, $\tilde{\rho}_n^\star$ is also in the interval $[0, 1]$. Moreover, observe that if $\psi_{k,n} = \{0, 1\}$ and $\rho_n = \{0, 1\}$, $\forall k, n$, then the value of the new objective function $\tilde{I}(\mathbf{P}, \boldsymbol{\Psi})$ is equal to the value of the original function $I(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi})$. This implies that the optimal solution to (4) is a feasible solution to (5) and, hence, the solution

[7]In this section, ρ_n can be viewed as if it represented the n-th channel access usage fraction rather than just indicating if the channel access is "on" or "off" and, similarly, $\psi_{k,n}$ can be interpreted as the usage fraction of the k-th stream.

Table II
KKT OPTIMALITY CONDITIONS OF PROBLEM (5)

$$\forall\, k,n: \quad \frac{\partial \tilde{\mathcal{L}}}{\partial P_{k,n}} = \frac{\psi_{k,n} h_{k,n}}{\psi_{k,n} + h_{k,n} P_{k,n}} + \xi_{k,n} - T_s \sum_{\ell=j}^{J} \lambda_\ell = 0. \tag{6a}$$

$$\forall\, k,n: \quad \frac{\partial \tilde{\mathcal{L}}}{\partial \psi_{k,n}} = \log\left(1 + \frac{h_{k,n} P_{k,n}}{\psi_{k,n}}\right) - \frac{h_{k,n} P_{k,n}}{\psi_{k,n} + h_{k,n} P_{k,n}} - T_s \sum_{\ell=j}^{J} \lambda_\ell P_d - \mu_{k,n} - \check{\eta}_{k,n} + \hat{\eta}_{k,n} = 0. \tag{6b}$$

$$\forall\, n: \quad \frac{\partial \tilde{\mathcal{L}}}{\partial \rho_n} = -T_s \sum_{\ell=j}^{J} \lambda_\ell P_c + \sum_{k=1}^{K} \mu_{k,n} = 0. \tag{6c}$$

$$\forall\, \ell = 1,\ldots,J: \quad \lambda_\ell(\phi_\ell(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}) - \beta_\ell) = 0. \tag{6d}$$

$$\forall\, k,n: \quad \xi_{k,n} P_{k,n} = 0, \tag{6e} \qquad \hat{\eta}_{k,n} \psi_{k,n} = 0, \tag{6f} \qquad \check{\eta}_{k,n}(\psi_{k,n} - 1) = 0, \tag{6g} \qquad \mu_{k,n}(\psi_{k,n} - \rho_n) = 0. \tag{6h}$$

$$\forall\, \ell = 1,\ldots,J, \quad \forall\, k,n: \quad \lambda_\ell \geq 0, \quad \xi_{k,n} \geq 0, \quad \hat{\eta}_{k,n} \geq 0, \quad \check{\eta}_{k,n} \geq 0, \quad \mu_{k,n} \geq 0. \tag{6i}$$

to (5) upper-bounds the solution to (4). The Lagrangian of (5) is $\tilde{\mathcal{L}} = \tilde{I}(\mathbf{P}, \boldsymbol{\Psi}) - \sum_{\ell=1}^{J} \lambda_\ell(\phi_\ell(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}) - \beta_\ell) + \sum_{j=1}^{J} \sum_{n \in \tau_j} \sum_{k=1}^{K} -\mu_{k,n}(\psi_{k,n} - \rho_n) - \check{\eta}_{k,n}(\psi_{k,n} - 1) + \hat{\eta}_{k,n}\psi_{k,n} + \xi_{k,n} P_{k,n}$, where λ_ℓ and $\mu_{k,n}$ are the Lagrange multipliers associated with the ECCs and the channel access activation constraints, respectively; $\hat{\eta}_{k,n}$, $\check{\eta}_{k,n}$ are the multipliers associated with the feasible set of $\psi_{k,n}$; and $\xi_{k,n}$ is the multiplier associated with the feasible set of $P_{k,n}$.

Since $\tilde{I}(\mathbf{P}, \boldsymbol{\Psi})$ is jointly concave (it is easy to check that its Hessian matrix is negative semidefinite) and the constraints are affine, (5) is a convex optimization problem and can be solved by, e.g., interior point methods [17]. In the following lines, we study the KKT sufficient optimality conditions, which are given in Table II, to gain some knowledge on the structure of the optimal solution to (5). This structural knowledge of the solution is later used in Section III-C to devise the graphical interpretation of the asymptotically optimal solution and in Section IV to design the online resource allocation algorithm. From (6a), we obtain that

$$\begin{aligned} P_{k,n} &= \psi_{k,n}\left(\frac{1}{-\xi_{k,n} + T_s \sum_{\ell=j}^{J} \lambda_\ell} - h_{k,n}^{-1}\right) \\ &= \psi_{k,n}\left(W_j - h_{k,n}^{-1}\right)^+, \end{aligned} \tag{7}$$

where $W_j = (T_s \sum_{\ell=j}^{J} \lambda_\ell)^{-1}$ is the j-th epoch water level that is equal for all the active streams contained in some channel access $n \in \tau_j$ and where we have used the slackness condition in (6e).

The j-th epoch water level, W_j, is related to the available energy at the transmitter through the dependence with the Lagrange multipliers λ_ℓ, $\ell = j,\ldots,J$. As it is later shown in Lemma 1, when the available energy is very low, then $W_j \to 0$ to satisfy the ECCs, the n-th channel access is "off" ($\rho_n = 0$), and the mutual information of the n-th channel access is zero. Then, if the available energy grows, W_j increases and there is a point that we refer to as the n-th *channel access cutoff water level*, $\hat{W}_n(M_n^\star)$, in which the obtained reward in terms of mutual information becomes higher than the activation cost. As it is shown next, $\hat{W}_n(M_n^\star)$ depends on P_c, P_d, the number of streams that contribute to the channel access activation, $M_n^\star \in [1, K]$, which is a priori unknown, and the channel gains of these streams. Thus, when the available energy and the other system parameters are such that $W_j = \hat{W}_n(M_n^\star)$, the n-th channel access becomes "partially active", i.e., $\rho_n \in (0,1)$. Finally, if the available energy is very high, the channel access is completely active, i.e., $\rho_n = 1$. In the following lemma, we derive the expression of the channel access cutoff water level as a function of $M_n^\star$ and later, in Proposition 1, we propose a low complexity method to obtain $M_n^\star$.

Lemma 1. *The optimal channel access indicator variable satisfies that*

$$\tilde{\rho}_n^\star = \begin{cases} 1 & \text{if } W_j > \hat{W}_n(M_n^\star), \\ (0,1) & \text{if } W_j = \hat{W}_n(M_n^\star), \\ 0 & \text{if } W_j < \hat{W}_n(M_n^\star), \end{cases} \qquad n \in \tau_j,$$

where the n-th channel access cutoff water level reads as

$$\hat{W}_n(M_n) = \frac{\frac{1}{M_n}(P_c + M_n P_d - \sum_{k=1}^{M_n} h_{k,n}^{-1})}{\mathcal{W}_0\left(\frac{\prod_{k=1}^{M_n} h_{k,n}^{\frac{1}{M_n}}}{eM_n}(P_c + M_n P_d - \sum_{k=1}^{M_n} h_{k,n}^{-1})\right)}, \tag{8}$$

and depends on P_c, P_d, the number of streams that contribute to the channel access activation M_n and on the channel gains of these streams. $\mathcal{W}_0(\cdot)$ is the positive branch of the Lambert function [18]. Thus, the optimal resource allocation of the streams $k \in [1, M_n^\star]$ satisfies that

$$\tilde{P}_{k,n}^\star = \begin{cases} \left(W_j - h_{k,n}^{-1}\right) & \text{if } W_j > \hat{W}_n(M_n^\star), \\ \tilde{\psi}_{k,n}^\star\left(W_j - h_{k,n}^{-1}\right) & \text{if } W_j = \hat{W}_n(M_n^\star), \\ 0 & \text{if } W_j < \hat{W}_n(M_n^\star), \end{cases}$$

$$\tilde{\psi}_{k,n}^\star = \begin{cases} 1 & \text{if } W_j > \hat{W}_n(M_n^\star), \\ (0,1) & \text{if } W_j = \hat{W}_n(M_n^\star), \\ 0 & \text{if } W_j < \hat{W}_n(M_n^\star), \end{cases} \qquad n \in \tau_j.$$

Proof: See Appendix A-1. ∎

Remark 3. $\hat{W}_n(M_n)$ increases with both P_c and P_d, and decreases with $h_{k,n}$, $\forall k \in [1, M_n]$ (the proof follows from Lemma 4.b in Appendix B).

Note that, in Lemma 1, we have used that the $M_n^\star$ streams that contribute to the channel access activation are the ones with the best channel gains, i.e., $h_{1,n}, \ldots, h_{M_n^\star,n}$, because these streams are the ones that contribute the most to the objective function. Intuitively, the $M_n^\star$ streams that become active first share the cost of using the channel access P_c. Once the channel access is being used, the remaining streams, $k > M_n^\star$, may become active by just paying their own stream circuitry cost, P_d. As a result of this, the streams $k \in (M_n^\star, K]$ experience different activation water levels as shown in the following lemma:

Lemma 2. *The optimal resource allocation of the streams $k \in (M_n^\star, K]$ satisfies that*

$$\tilde{P}_{k,n}^\star = \begin{cases} \left(W_j - h_{k,n}^{-1}\right) & \text{if } W_j > \bar{W}_{k,n}, \\ \tilde{\psi}_{k,n}^\star \left(W_j - h_{k,n}^{-1}\right) & \text{if } W_j = \bar{W}_{k,n}, \\ 0 & \text{if } W_j < \bar{W}_{k,n}, \end{cases}$$

$$\tilde{\psi}_{k,n}^\star = \begin{cases} 1 & \text{if } W_j > \bar{W}_{k,n}, \\ (0,1) & \text{if } W_j = \bar{W}_{k,n}, \quad n \in \tau_j, \\ 0 & \text{if } W_j < \bar{W}_{k,n}, \end{cases}$$

where the k-th stream cutoff water level at the n-th channel use reads as

$$\bar{W}_{k,n} = \frac{P_d - h_{k,n}^{-1}}{\mathcal{W}_0\left(\frac{P_d h_{k,n} - 1}{e}\right)}, \quad k \in (M_n^\star, K], \tag{9}$$

and depends on the stream circuitry consumption, P_d, and the stream gain $h_{k,n}$.

Proof: See Appendix A-2. ∎

Remark 4. When $P_d \to 0$, the cutoff water level in DWF is recovered, i.e., $\bar{W}_{k,n} = h_{k,n}^{-1}$ (see (1)). Moreover, $\bar{W}_{k,n}$ increases with P_d and decreases with $h_{k,n}$ (this can be proved similarly as in the proof of Lemma 4.b in Appendix B).

Note that for coherence, $\bar{W}_{k,n} > \hat{W}_n(M_n^\star)$, $\forall k > M_n^\star$, which implies that the streams with higher gains are activated first. However, from the expressions (8) and (9), this is not obvious. Indeed, if $\bar{W}_{k',n} < \hat{W}_n(M_n^\star)$ for some $k' > M_n^\star$, then the stream k' would become active before the channel access was active, which is a logical contradiction. If such a situation happens, the stream k' should also contribute to activate the channel access, which means that actually $M_n^\star$ is not the optimal number of streams to activate the channel access. Since the mutual information of the n-th channel access is zero until the channel access becomes active, the optimal number of active streams at the channel access cutoff water level is the one that allows to activate the channel access with the lowest water level, i.e., $M_n^\star = \arg\min_{M_n} \hat{W}_n(M_n)$. To find $M_n^\star$, an exhaustive search over M_n could be performed. However, this may require a high computational complexity (especially when $K \gg 1$) that can be reduced by means of the following procedure:

Algorithm 1 Close to optimal solution to (4) from integer relaxation

Input: $\{\tilde{\mathbf{P}}^\star, \tilde{\boldsymbol{\rho}}^\star, \tilde{\boldsymbol{\Psi}}^\star\}$ and $\gamma \in (0,1)$

1: Let $\mathcal{S}$ be the set that contain the partially used streams, i.e., $\mathcal{S} = \{\{k,n\} | [\tilde{\boldsymbol{\psi}}^\star]_{k,n} \in (0,1)\}$ and let $\mathcal{S}_1 = \{\{k,n\} | [\tilde{\boldsymbol{\Psi}}^\star]_{k,n} \in [\gamma, 1)\}$ and $\mathcal{S}_0 = \{\{k,n\} | [\tilde{\boldsymbol{\Psi}}^\star]_{k,n} \in (0,\gamma)\}$ be a partition of $\mathcal{S}$, where γ is a constant in $(0,1)$.
2: **if** $|\mathcal{S}| = 0$ **then**
3: $\{\hat{\mathbf{P}}, \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\Psi}}\} = \{\tilde{\mathbf{P}}^\star, \tilde{\boldsymbol{\rho}}^\star, \tilde{\boldsymbol{\Psi}}^\star\}$
▷ $\{\hat{\mathbf{P}}, \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\Psi}}\}$ is the optimal solution to (4).
4: **else**
5: $[\hat{\mathbf{P}}]_{k,n} = [\tilde{\mathbf{P}}^\star]_{k,n}$, $[\hat{\boldsymbol{\Psi}}]_{k,n} = [\tilde{\boldsymbol{\Psi}}^\star]_{k,n}$, $\forall \{k,n\} \notin \mathcal{S}$.
▷ Assign the resource allocation in $\{\tilde{\mathbf{P}}^\star, \tilde{\boldsymbol{\rho}}^\star, \tilde{\boldsymbol{\Psi}}^\star\}$ to all the streams not contained in $\mathcal{S}$.
6: $[\hat{\mathbf{P}}]_{k,n} = [\tilde{\mathbf{P}}^\star]_{k,n}$ and $[\hat{\boldsymbol{\Psi}}]_{k,n} = 1$, $\forall \{k,n\} \in \mathcal{S}_1$;
$[\hat{\mathbf{P}}]_{k,n} = 0$ and $[\hat{\boldsymbol{\Psi}}]_{k,n} = 0$, $\forall \{k,n\} \in \mathcal{S}_0$.
▷ Round up or down the stream indicator variables.
7: $[\hat{\boldsymbol{\rho}}]_n = \max_k [\hat{\boldsymbol{\Psi}}]_{k,n}$, $\forall n = 1, \ldots, N$.
▷ Compute the channel access indicator variables.
8: Ensure the feasibility of $\{\hat{\mathbf{P}}, \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\Psi}}\}$ by scaling down the transmission radiated power of the channel accesses that produce some ECC violation.
9: **end if**
10: **return** $\{\hat{\mathbf{P}}, \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\Psi}}\}$

Proposition 1. *The n-th channel access cutoff water level, $\hat{W}_n(M_n^\star)$, can be found by performing a forward search over M_n, i.e.,*

1) *Initially, set $M_n := 1$.*
2) *Compute $\hat{W}_n(M_n)$ and $\bar{W}_{M_n+1,n}$.*
3) *Check if $\hat{W}_n(M_n) < \bar{W}_{M_n+1,n}$: if the condition is true, then $M_n^\star = M_n$ and the algorithm ends; otherwise, increase M_n, i.e., $M_n := M_n + 1$ and go back to step 2.*

Proof: See Appendix B. ∎

Until now, we have derived Lemmas 1, 2 and Proposition 1 to gain some knowledge on the structure of the optimal solution to (5). As mentioned before, since (5) is a convex optimization problem, the resource allocation that maximizes (5), $\{\tilde{\mathbf{P}}^\star, \tilde{\boldsymbol{\rho}}^\star, \tilde{\boldsymbol{\Psi}}^\star\}$, can be found by, e.g., interior point methods [17]. In Algorithm 1, we propose a procedure to derive a feasible resource allocation of (4), $\{\hat{\mathbf{P}}, \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\Psi}}\}$, from the solution to (5), $\{\tilde{\mathbf{P}}^\star, \tilde{\boldsymbol{\rho}}^\star, \tilde{\boldsymbol{\Psi}}^\star\}$, whose mutual information, $\hat{\mathcal{I}} = I(\hat{\mathbf{P}}, \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\Psi}})$, performs close to $\mathcal{I}^\star$, as argued in the following lines.

Note that in general $\hat{\mathcal{I}} \leq \mathcal{I}^\star \leq \tilde{\mathcal{I}}^\star$. However, these inequalities are tight ($\hat{\mathcal{I}} = \mathcal{I}^\star = \tilde{\mathcal{I}}^\star$), when, in Algorithm 1, we have that $\mathcal{S} = \{\emptyset\}$, or, equivalently, if $W_j \neq \hat{W}_n(M_n^\star)$ and $W_j \neq \bar{W}_{k,n}$, $\forall n \in \tau_j, \forall j, \forall k > M_n^\star$. This means that $\{\tilde{\mathbf{P}}^\star, \tilde{\boldsymbol{\rho}}^\star, \tilde{\boldsymbol{\Psi}}^\star\}$ is the optimal resource allocation to (4). Alternatively, when $\mathcal{S} \neq \{\emptyset\}$, we know that the optimality gap, i.e., $\mathcal{I}^\star - \hat{\mathcal{I}}$, is at most $\tilde{\mathcal{I}}^\star - \hat{\mathcal{I}}$ and is closely related to the cardinality of $\mathcal{S}$. Since for most of the streams and channel accesses the water level is different to the cutoff water level, we know that $|\mathcal{S}| \ll KN$, which implies that the optimal resource allocation to (5) is used in the majority ($KN - |\mathcal{S}|$) of the streams. This discussion is later continued in Remark 7 once the graphical representation of the asymptotically optimal solution is presented.

Remark 5. Observe that $\{\tilde{\mathbf{P}}^\star, \tilde{\boldsymbol{\rho}}^\star, \tilde{\boldsymbol{\Psi}}^\star\}$ is the optimal solution

to the time continuous channel problem. Hence, if we particularize $K = 1$, then our solution reduces to the directional glue pouring algorithm introduced in [2].

In the following section, we solve the dual problem to (4). Interestingly, the concept of the cutoff water level also appears when solving the dual problem, which is indeed surprising due to the great difference between the relaxed and dual problem approaches.

B. Duality

In this section, we study the Lagrange dual problem to (4) and show that, even though (4) is not a convex optimization problem, the duality gap tends asymptotically to zero as the number of streams or channel accesses per epoch grows without bound.

The Lagrangian of (4) is $\mathcal{L}(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}, \boldsymbol{\lambda}) = I(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}) - \boldsymbol{\lambda}^{\mathsf{T}}(\boldsymbol{\phi}(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}) - \boldsymbol{\beta})$, where $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_J]^{\mathsf{T}}$ is the dual variable that contains the Lagrange multipliers associated with the ECCs. The dual function is defined for $\boldsymbol{\lambda} \succeq \mathbf{0}$ as $g(\boldsymbol{\lambda}) = \max_{\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi} \in \mathcal{F}} \mathcal{L}(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}, \boldsymbol{\lambda})$ (see [17]) and yields to upper bounds to the maximum achievable mutual information $\mathcal{I}^\star$ obtained by maximizing the primal problem (4), i.e., $\mathcal{I}^\star \leq g(\boldsymbol{\lambda})$. The Lagrange dual problem, $\mathcal{D}^\star = \min_{\boldsymbol{\lambda} \succeq \mathbf{0}} g(\boldsymbol{\lambda})$, is a convex program that determines the best upper bound on $\mathcal{I}^\star$ as $\mathcal{I}^\star \leq \mathcal{D}^\star \leq g(\boldsymbol{\lambda})$. The duality gap is defined as $\mathcal{D}^\star - \mathcal{I}^\star$ and it is zero if Slater qualification constraints are satisfied. However, in our problem the Slater qualification constraints are not satisfied as $\mathcal{F}$ is not a convex set and, therefore, the duality gap might not be zero. The *time-sharing condition* introduced in [19] provides a condition under which the duality gap is zero even though the primal optimization problem is not convex. In the following proposition, we demonstrate that the *time-sharing condition* is asymptotically satisfied as the number of streams or channel accesses per epoch grows without bound.

Proposition 2. *The time-sharing condition is asymptotically satisfied when, within each epoch, every channel realization is observed a sufficiently large number of times.*[8]

Proof: See Appendix C. ∎

Thanks to the previous proposition, when the number of streams or channel accesses per epoch is high and the channel variations in one of the dimensions (time, space or frequency) are slow, the duality gap tends to zero and, consequently, the solution to $\mathcal{D}^\star$ asymptotically tends to $\mathcal{I}^\star$. At this stage, it is important to highlight that, in practice, it is not necessary that the number of streams or channel accesses per epoch grows without bound; a small duality gap is already observed for small values of these magnitudes as verified in the simulations results (see Section V) where $K = 8$ and the mean number of channel accesses per epoch is 5. This behavior was previously observed in scenarios without energy harvesting in, e.g., [19].

[8]The time sharing condition has been broadly used in different non-harvesting scenarios where the nodes have to satisfy a single sum-power constraint. In such non-harvesting scenarios, the requirement for the asymptotic fulfillment of the time sharing condition is that every channel realization must be observed a large number of times [19]. When energy harvesting is considered, the problem is constrained by a set of ECCs and the time-sharing condition is asymptotically satisfied if, within each epoch, every channel realization is observed a sufficiently large number of times. When $K = 1$, it is necessary that every channel realization is observed in a sufficiently large number of channel accesses. This situation happens, for instance, when the number of channel accesses per epoch is large, i.e., $e_{j+1} - e_j \gg 1$, $\forall j$, and $T_c \gg T_s$. Whereas when $K > 1$, this condition is more likely to be fulfilled due to the additional (space or frequency) dimension.

Algorithm 2 Projected subgradient

Initialization:
Set $q := 0$ and initialize $\boldsymbol{\lambda}^{(0)}$ to any value such that $\boldsymbol{\lambda}^{(0)} \succeq \mathbf{0}$.
For all $n = 1, \ldots, N$, compute $\hat{W}_n(M_n^\star)$ according to (8) with $M_n^\star$ obtained from the forward search in Proposition 1.
Step 1: If a termination condition is met, the algorithm stops.
Step 2: Compute the optimal primal variables at the q-th iteration that are $[\mathbf{P}^{(q)}, \boldsymbol{\rho}^{(q)}, \boldsymbol{\Psi}^{(q)}] = \arg\max_{\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi} \in \mathcal{F}} \mathcal{L}(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}, \boldsymbol{\lambda}^{(q)})$ by means of Algorithm 3 that requires $\boldsymbol{\lambda}^{(q)}$ and $\hat{W}_n(M_n^\star)$.
Step 3: Update the dual variable following the subgradient, i.e., $[\boldsymbol{\lambda}^{(q+1)}]_j = \lambda_j^{(q+1)}$, $j = 1, \ldots, J$, with

$$\lambda_j^{(q+1)} = \left(\lambda_j^{(q)} - s^{(q)} \left(\beta_j - \phi_j(\mathbf{P}^{(q)}, \boldsymbol{\rho}^{(q)}, \boldsymbol{\Psi}^{(q)})\right)\right)^+.$$

Step 4: Set $q := q + 1$ and go to Step 1.

To solve the dual problem we have implemented the projected subgradient method [20], presented in Algorithm 2, that guarantees convergence if the updating step size $s^{(q)}$ is correctly chosen. In this context, we have used $s^{(q)} = Q/(\sqrt{q}||\boldsymbol{\beta} - \boldsymbol{\phi}(\mathbf{P}^{(q)}, \boldsymbol{\rho}^{(q)}, \boldsymbol{\Psi}^{(q)})||)$ that satisfies the diminishing conditions $s^{(q)} \geq 0$, $\lim_{q\to\infty} s^{(q)} = 0$ and $\sum_{q=1}^{\infty} s^{(q)} = \infty$ [20], where $Q > 0$ is an arbitrary constant. When the algorithm converges to the optimal dual variable, $\boldsymbol{\lambda}^\star$, all the ECCs are satisfied, which is ensured by the termination condition in Step 1. In the next subsection, we explain Step 2 of Algorithm 2, i.e., how to obtain the primal variables $\mathbf{P}^{(q)}$, $\boldsymbol{\rho}^{(q)}$ and $\boldsymbol{\Psi}^{(q)}$ at the q-th iteration of the subgradient method.

1) Maximizing the Lagrangian for a given $\boldsymbol{\lambda}^{(q)}$: At every iteration of the subgradient algorithm, it is necessary to compute the optimal primal variables given the dual variables of the iteration, i.e., $\boldsymbol{\lambda}^{(q)}$. From the expression of $g(\boldsymbol{\lambda})$, the optimal primal variables at the q-th iteration are $[\mathbf{P}^{(q)}, \boldsymbol{\rho}^{(q)}, \boldsymbol{\Psi}^{(q)}] = \arg\max_{\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi} \in \mathcal{F}} \mathcal{L}(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}, \boldsymbol{\lambda}^{(q)})$. Note that the maximization of the Lagrangian is not a convex problem as $\mathcal{F}$ is not a convex set. To solve the maximization of the Lagrangian we apply decomposition as shown in (10).

We have reordered the sums over j and ℓ to decompose the Lagrangian maximization in N independent maximization problems, one for each channel use, where the objective function is $g_n(\mathbf{p}_n, \rho_n, \boldsymbol{\psi}_n) = \rho_n\big(\sum_{k=1}^K \psi_{k,n} \log(1 + h_{k,n}P_{k,n})\big) - \big(\rho_n P_c + \sum_{k=1}^K P_{k,n} + \psi_{k,n}P_d\big)/W_j^{(q)}$ with $W_j^{(q)} = (T_s \sum_{\ell=j}^J \lambda_\ell^{(q)})^{-1}$ being the water level of the j-th epoch at the q-th iteration.

Note that $\max_{\mathbf{p}_n, \rho_n, \boldsymbol{\psi}_n \in \mathcal{F}} g_n(\mathbf{p}_n, \rho_n, \boldsymbol{\psi}_n)$ is still a non-convex problem due to the binary variables. However, after applying decomposition, it is feasible to perform an exhaustive search over ρ_n as there are only two possibilities either $\rho_n = 0$ or $\rho_n = 1$. Thus, we can solve two separated maximization problems and select the pointwise maximum of the two, i.e.,

$$\max_{\mathbf{p}_n, \rho_n, \boldsymbol{\psi}_n \in \mathcal{F}} g_n(\mathbf{p}_n, \rho_n, \boldsymbol{\psi}_n) = \tag{11}$$
$$\max\Big\{ \underbrace{\max_{\mathbf{p}_n, \boldsymbol{\psi}_n \in \mathcal{F}} g_n(\mathbf{p}_n, 0, \boldsymbol{\psi}_n)}_{\text{(SP 1)}}, \underbrace{\max_{\mathbf{p}_n, \boldsymbol{\psi}_n \in \mathcal{F}} g_n(\mathbf{p}_n, 1, \boldsymbol{\psi}_n)}_{\text{(SP 2)}} \Big\}.$$

$$\begin{aligned}\max_{\mathbf{P},\boldsymbol{\rho},\boldsymbol{\Psi}\in\mathcal{F}} \quad & \mathcal{L}(\mathbf{P},\boldsymbol{\rho},\boldsymbol{\Psi},\boldsymbol{\lambda}^{(q)}) = \max_{\mathbf{P},\boldsymbol{\rho},\boldsymbol{\Psi}\in\mathcal{F}} \mathcal{I}(\mathbf{P},\boldsymbol{\rho},\boldsymbol{\Psi}) - \sum_{\ell=1}^{J}\lambda_\ell^{(q)}\left(-\beta_\ell + T_s\sum_{j=1}^{\ell}\sum_{n\in\tau_j}\left(\rho_n P_c + \sum_{k=1}^{K} P_{k,n} + \psi_{k,n}P_d\right)\right) \\ = \quad & \sum_{\ell=1}^{J}\lambda_\ell^{(q)}\beta_\ell + \max_{\mathbf{P},\boldsymbol{\rho},\boldsymbol{\Psi}\in\mathcal{F}}\sum_{j=1}^{J}\sum_{n\in\tau_j}\left[\rho_n\Big(\sum_{k=1}^{K}\psi_{k,n}\mathsf{log}(1+h_{k,n}P_{k,n})\Big) - T_s\Big(\rho_n P_c + \sum_{k=1}^{K}P_{k,n} + \psi_{k,n}P_d\Big)\sum_{\ell=j}^{J}\lambda_\ell^{(q)}\right] \\ = \quad & \sum_{\ell=1}^{J}\lambda_\ell^{(q)}\beta_\ell + \sum_{j=1}^{J}\sum_{n\in\tau_j}\max_{\mathbf{p}_n,\rho_n,\boldsymbol{\psi}_n\in\mathcal{F}} g_n(\mathbf{p}_n,\rho_n,\boldsymbol{\psi}_n). \end{aligned} \tag{10}$$

Table III
OPTIMAL SOLUTION TO THE SUBPROBLEMS IN (11).

Subproblem	Maximum value	Optimal $\mathbf{p}_n$	Optimal $\boldsymbol{\psi}_n$
(SP 1)	0	$\mathbf{p}_n^{(\text{SP 1})} = \mathbf{0}$	$\boldsymbol{\psi}_n^{(\text{SP 1})} = \mathbf{0}$
(SP 2)	$-\frac{P_c}{W_j^{(q)}} + \sum_{k=1}^{A_n^{(q)}}\Big(\mathsf{log}(W_j^{(q)}h_{k,n}) -1 + \frac{1}{W_j^{(q)}h_{k,n}} - \frac{P_d}{W_j^{(q)}}\Big)$	$[\mathbf{p}_n^{(\text{SP 2})}]_{k\in[1,A_n^{(q)}]} = W_j^{(q)} - h_{k,n}^{-1}$ $[\mathbf{p}_n^{(\text{SP 2})}]_{k\in(A_n^{(q)},K]} = 0$	$[\boldsymbol{\psi}_n^{(\text{SP 2})}]_{k\in[1,A_n^{(q)}]} = 1$ $[\boldsymbol{\psi}_n^{(\text{SP 2})}]_{k\in(A_n^{(q)},K]} = 0$

Table IV
OPTIMAL SOLUTION TO THE SUBPROBLEMS IN (12).

Subproblem	Maximum value	Optimal $P_{k,n}$
(SP 2.1)	0	$P_{k,n}^{(\text{SP 2.2})} = 0$
(SP 2.2)	$\mathsf{log}(W_j^{(q)}h_{k,n}) - 1 + \frac{1}{W_j^{(q)}h_{k,n}} - \frac{P_d}{W_j^{(q)}}$	$P_{k,n}^{(\text{SP 2.2})} = W_j^{(q)} - h_{k,n}^{-1}$

These two problems are solved in the following lines and Table III summarizes the obtained results.

Solution to (SP 1): By observing the objective function of (SP 1), i.e.,

$$g_n(\mathbf{p}_n, 0, \boldsymbol{\psi}_n) = -\frac{\sum_{k=1}^{K}P_{k,n} + \psi_{k,n}P_d}{W_j^{(q)}}$$

and by noting that $W_j^{(q)}$ is positive, it is straight-forward to show that the optimal transmitted powers and stream indicator variables of the n-th channel access are $\mathbf{p}_n^{(\text{SP 1})} = \boldsymbol{\psi}_n^{(\text{SP 1})} = \mathbf{0}$ and the maximum value of the objective function is 0, as expected since $\rho_n = 0$.

Solution to (SP 2): To solve the second subproblem, which is nonconvex due to the stream indicator variables, we can again apply decomposition as follows:

$$\begin{aligned}&\max_{\mathbf{p}_n,\boldsymbol{\psi}_n\in\mathcal{F}} g_n(\mathbf{p}_n, 1, \boldsymbol{\psi}_n) = \\ &-\frac{P_c}{W_j^{(q)}} + \sum_{k=1}^{K}\max_{P_{k,n},\psi_{k,n}\in\mathcal{F}} g_{k,n}(P_{k,n},\psi_{k,n}),\end{aligned}$$

where $g_{k,n}(P_{k,n},\psi_{k,n}) = \psi_{k,n}\mathsf{log}(1+h_{k,n}P_{k,n}) - (P_{k,n} + \psi_{k,n}P_d)/W_j^{(q)}$. As before, after applying decomposition, we can perform an exhaustive search over $\psi_{k,n}$ since there are only two possibilities, i.e., either $\psi_{k,n} = 0$ or $\psi_{k,n} = 1$. Thus, $\max_{P_{k,n},\psi_{k,n}\in\mathcal{F}} g_{k,n}(P_{k,n},\psi_{k,n}) =$

$$\max\Big\{\underbrace{\max_{P_{k,n}\in\mathcal{F}} g_{k,n}(P_{k,n},0)}_{\text{(SP 2.1)}}, \underbrace{\max_{P_{k,n}\in\mathcal{F}} g_{k,n}(P_{k,n},1)}_{\text{(SP 2.2)}}\Big\}. \tag{12}$$

Now, both subproblems are convex and can be easily solved. Table IV summarizes the maximum achieved value and the optimal transmission power of each subproblem. Thus, the k-th stream is active if $\mathsf{log}(W_j^{(q)}h_{k,n}) - 1 + 1/(W_j^{(q)}h_{k,n}) - P_d/W_j^{(q)} > 0$. Solving the previous equation (set $M := 1$, $\hat{W} := \bar{W}_{k,n}$, $H_1 := h_{k,n}$ and $P := P_d$ in Appendix D), we obtain an equivalent condition for the k-th stream activation, i.e., $W_j^{(q)} > \bar{W}_{k,n}$, where $\bar{W}_{k,n}$ is the stream cutoff water level given in (9). Thus, after evaluating the condition $W_j^{(q)} > \bar{W}_{k,n}$, $\forall k$, we obtain the number of streams that are activated if the channel access is active, $A_n^{(q)}$.

Now that both (SP 1) and (SP 2) are solved (Table III summarizes the obtained results), we can conclude that the n-th channel access is active if $-P_c/W_j^{(q)} + \sum_{k=1}^{A_n^{(q)}}\Big(\mathsf{log}(W_j^{(q)}h_{k,n}) - 1 + 1/(W_j^{(q)}h_{k,n}) - P_d/W_j^{(q)}\Big) > 0$ or, equivalently, if $W_j^{(q)} > \hat{W}_n(A_n^{(q)})$ (set $M := A_n^{(q)}$, $\hat{W} := \hat{W}_n(A_n^{(q)})$, $H_k := h_{k,n}$, and $P := A_n^{(q)}P_d + P_c$ in Appendix D to show this equivalence).[9]

In summary, the optimal primal variables at the q-th iteration of the subgradient can be obtained by checking the condition $W_j^{(q)} > \bar{W}_{k,n}$, $\forall k$, to obtain the number of streams that would become active if the channel access becomes active, $A_n^{(q)}$, and then checking the condition $W_j^{(q)} > \hat{W}_n(A_n^{(q)})$ to find out whether the channel access is active or not. If the channel access is active, the optimal primal variables at the q-th iteration of the subgradient are $\mathbf{p}_n^{(q)} = \mathbf{p}_n^{(SP2)}$, $\boldsymbol{\psi}_n^{(q)} = \boldsymbol{\psi}_n^{(SP2)}$, and $\rho_n^{(q)} = 1$. Otherwise, we have that $\mathbf{p}_n^{(q)} = \mathbf{0}$, $\boldsymbol{\psi}_n^{(q)} = \mathbf{0}$, and $\rho_n^{(q)} = 0$. However, this procedure might be quite inefficient when the number of streams is large

[9]Note that when $W_j^{(q)} = \bar{W}_{k,n}$ (or $W_j^{(q)} = \hat{W}_n(A_n^{(q)})$) it is equivalent to activate or not the stream (or channel access) since both achieve the same value of the objective function.

Algorithm 3 Maximization of the Lagrangian

Data: $\boldsymbol{\lambda}^{(q)}$, $\hat{W}_n(M_n^\star)$.
Compute $W_j^{(q)} = (T_s \sum_{\ell=j}^{J} \lambda_\ell^{(q)})^{-1}$, $\forall j = 1, \ldots, J$.
▷ Compute the water level in all the epochs given $\boldsymbol{\lambda}^{(q)}$.
for $n \in \tau_j$, $j := 1, \ldots, J$ **do** ▷ For all the channel accesses.
  if $W_j^{(q)} > \hat{W}_n(M_n^\star)$ **then** ▷ Check if the channel access is active.
    $A_n^{(q)} := M_n^\star$; ▷ The channel access is active.
    ▷ Count the number of active streams.
    for $k := M_n^\star + 1, \ldots, K$ **do**
      if $W_j^{(q)} > \bar{W}_{k,n}$ **then**
        $A_n^{(q)} := A_n^{(q)} + 1$;
        ▷ The k-th stream is active at the q-th iteration water level.
      end if
    end for ▷ End of counting.
    $\mathbf{p}_n^{(q)} = \mathbf{p}_n^{(SP2)}$, $\boldsymbol{\psi}_n^{(q)} = \boldsymbol{\psi}_n^{(SP2)}$, and $\rho_n^{(q)} = 1$;
  else
    $\mathbf{p}_n^{(q)} = \mathbf{0}$, $\boldsymbol{\psi}_n^{(q)} = \mathbf{0}$, and $\rho_n^{(q)} = 0$;
    ▷ The n-th channel access is turned off.
  end if
end for

($K \gg 1$) and it can be avoided by first checking whether the channel access is active. Note that for any value of $A_n^{(q)}$ the channel access is active if and only if $W_j^{(q)} > \hat{W}_n(M_n^\star)$.[10] In this context, the procedure in Algorithm 3 is equivalent to the proposed above, but more computationally efficient.

Remark 6. The resource allocation obtained by solving the dual problem is almost equal to the one obtained by means of the relaxed problem, which is given in Lemmas 1 and 2. The channel access or stream activation conditions obtained in this section only differ from the ones in Lemmas 1 and 2 when the water level is equal to the stream or channel access cutoff water levels. In this section we have seen that if $W_j = \hat{W}_n(M_n^\star)$ or $W_j = \bar{W}_{k,n}$ for $k \in (M_n^\star, K]$, it is indifferent to have the channel access active or inactive since both situations achieve the same value of the dual function, whereas, in the relaxed problem we obtained that a partial use of the channel access or the stream is optimal, which is not allowed in the problem considered in this section due to the binary feasible set of the indicator variables.

C. The Boxed Water-Flowing interpretation

In this section, we provide a graphical representation of the asymptotically optimal offline solution named the *Boxed Water-Flowing* interpretation, which is depicted in Figure 3. This interpretation follows directly from the concept of the cutoff water levels and it generalizes the DWF interpretation in [9] by considering the different sources of energy consumption at the transmitter. The interpretation is the following:

1) Each stream is represented with a water-porous vessel with base equal to T_s.[11] There are two types of boxes, namely, the *channel access box* and the *stream box*. At the n-th channel access, the *channel access box* with height $\hat{W}_n(M_n^\star)$ is shared among the streams $k = 1, \ldots, M_n^\star$. The remaining streams, i.e., $k > M_n^\star$, have their own stream box with height equal to $\bar{W}_{k,n}$.[12] A water *right-permeable* material is used to separate the different epochs.
2) Each box is filled by a solid substance up to a height equal to $h_{k,n}^{-1}$ and the boxes are closed by a lid. The cost (in terms of water) of opening the channel access box is $(P_c + M_n^\star P_d)T_s$, whereas, the cost of opening each stream box is $P_d T_s$.
3) The water level is progressively increased to all epochs at the same time by adding the necessary amount of water to each epoch. The maximum amount of water that can be externally added at some epoch is given by the epoch's harvested energy (depicted with the top-down arrows in Figure 3).[13] When some epoch runs out of water, it uses water that flows from previous epochs (if any is available) in order to continue increasing the water level simultaneously. When the water level reaches the lid of some box, check if there is enough available water (in the current and previous epochs) to pay the cost of opening the lid and to fill in the whole box with water: If there is enough water, remove the lid (which means that the amount of water associated with the lid opening cost is lost), let the water fill the box and go back to Step 3; otherwise, keep the lid in the box and go back to Step 3.
4) When all the available water has been poured, the optimal power allocation is found as the amount of water in each of the vessels divided by T_s or, equivalently, as the height of the water in each vessel, i.e., $P_{k,n} = (W_j - h_{k,n}^{-1})^+$.

Interestingly, by particularizing the *Boxed Water-Flowing* interpretation to the case in which there is no circuitry consumption ($P_c = 0$ and $P_d = 0$), the heights of the boxes reduce to its minimum possible value, i.e., $h_{k,n}^{-1}$, (set $P_c = 0$ and $P_d = 0$ in Lemmas 1 and 2) and the DWF graphical interpretation in [9] is recovered.

Remark 7. Having the graphical representation of the asymptotically optimal solution in mind, it is easy to understand why its performance is close to $\mathcal{I}^\star$. In the representation shown in Figure 3, all the streams are using the optimal resource allocation to (5), except the second stream of τ_1. After solving the integer relaxation problem, we would have obtained that a fractional use of this stream would be optimal. However, as this fractional use is not allowed in a discrete channel model, we do not know where to optimally allocate the small remaining energy in τ_1. In summary, the following arguments justify why the optimality gap of the *Boxed Water-Flowing* solution is small:

- As mentioned before, if the water level is different than

[10] By using the definition of $M_n^\star$ and Lemmas 3 and 4 in Appendix B, it is easy to show that: if $A_n^{(q)} > M_n^\star$, then $\hat{W}_n(M_n^\star) < \hat{W}_n(A_n^{(q)}) < \bar{W}_{A_n^{(q)},n} < W_j$; and if $A_n^{(q)} < M_n^\star$, then $\hat{W}_n(A_n^{(q)}) \geq \hat{W}_n(M_n^\star) \geq \bar{W}_{M_n^\star,n} \geq W_j > \bar{W}_{A_n^{(q)},n}$. Thus, from these inequalities, we can compare W_j directly with $\hat{W}_n(M_n^\star)$ to determine whether the channel access is active or not.

[11] The vessel boundaries are not depicted in Figure 3 for the sake of simplicity.

[12] Thus, the n-th channel access has $K - M_n^\star$ stream boxes and one channel access box.

[13] The amount of water corresponds to energy, whereas, the water level, i.e., the height of the water, corresponds to power.

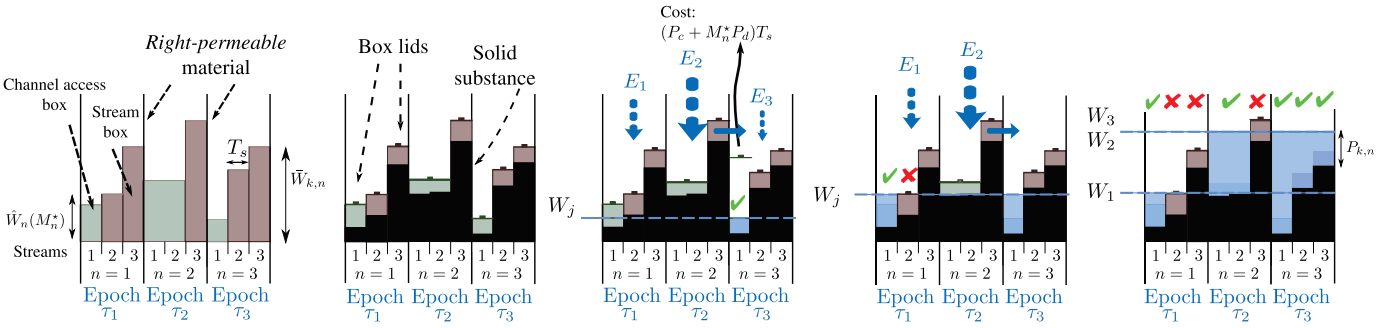

Figure 3. The *Boxed Water-Flowing* interpretation. For graphical simplicity, only one channel access is contained in each epoch. Figures (a) and (b) depict steps 1 and 2 of the explanation, respectively. Figures (c) to (e) depict step 3 where the water level is progressively increased and different situations occur: in (c), the box is opened because, by using water from τ_2 and τ_3, there is enough water to fill the box and pay the opening cost of the channel access box; in (d), when the water level reaches the stream box ($n = 1, k = 2$) the remaining water in the first epoch is not enough to pay the opening cost and fill in the whole box with water (see Remark 7 for a discussion on what happens with this remaining water); finally, (e) depicts the obtained resource allocation once all the available water has been poured.

the cutoff water level of all the boxes ($\tilde{\rho}_n^\star \in \{0,1\}$ and $\tilde{\psi}_{k,n}^\star \in \{0,1\}\ \forall k,n$), then the *Boxed Water-Flowing* solution is optimal and the optimality gap is zero.

- Otherwise, when some stream or slot is partially used ($\tilde{\rho}_n^\star \in (0,1)$ or $\tilde{\psi}_{k,n}^\star \in (0,1)$), the remaining energy in the epoch is very small and can be allocated in any of the active channels without having a relevant impact on the total mutual information.
  - Within each epoch, the water level can only be equal to one specific box height, which might be present in different channels as far as they have the same channel gain. This means that all the channels with different box heights (or different channel gains) are using the optimal resource allocation to (5).
  - If this specific box height is present in several channel accesses or streams of that same epoch, then the time sharing argument can be used to allocate the little remaining energy in the epoch, i.e., a fraction $\tilde{\rho}_n^\star$ ($\tilde{\psi}_{k,n}^\star$) of channel access (stream) boxes are opened and the remaining ones are kept closed.

Due to this, we can expect a small optimality gap as it is confirmed by the conducted experimental results presented in Section V.

IV. Online Resource Allocation

Up to now, we have assumed that the transmitter has non-causal knowledge of both the channel state and the energy harvesting process, which is only a realistic assumption under very specific scenarios, e.g., when the channel is static and the energy source is controllable. In this section, we develop an online algorithm, which does not require future knowledge of neither the energy arrivals nor the channel state, that is based on the structure of the *Boxed Water-Flowing*, the asymptotically optimal offline resource allocation that we derived in the previous section.

Let F_w be the *flowing window* that is an input parameter of the online algorithm that refers to the number of channel accesses in which the water is allowed to flow, which can be obtained by a previous training under the considered (or measured) energy harvesting profile, and let an *event*, s_t, denote the time index of a channel access in which either a change in the channel state is produced or an energy packet is harvested (or both events take place at the same time), i.e., $s_t = \cup_{k=1}^{K}\{n|h_{k,n-1} \neq h_{k,n}\} \cup \{n|n = e_j, j = 1,\dots,J\}$, $t = 1,\dots,T$, where $T \in [J, N]$. In this context, the proposed online algorithm proceeds as follows: (1.) The initial energy in the battery, E_1, is allocated to the different streams of the first F_w channel accesses according to the *Boxed Water-Flowing* where the channel is expected to be static and equal to the observed channel at the first channel use, i.e., we assume as if $h_{k,n} = h_{k,1}, \forall n \in [1, F_w]$, $\forall k$. (2.) When the transmitter detects an event, it updates the allocated power of the channel accesses $n \in [s_t, \min\{s_t + F_w - 1, N\}]$ by using the *Boxed Water-Flowing* with the remaining energy in the battery and with the energy of the harvested packet (if the event is an energy arrival), i.e., $\sum_{j|e_j \leq s_t} E_j - T_s \sum_{n=1}^{s_t-1}(P_c\rho_n + \sum_k (P_{k,n} + P_d\psi_{k,n}))$, and by assuming that the channel remains constant during the flowing window, i.e., $h_{k,n} = h_{k,s_t}$, $\forall n \in [s_t, \min\{s_t + F_w - 1, N\}]$, $\forall k$.[14] Step (2.) is repeated until the N-th channel access is reached. A natural requirement of WEHNs is that they operate perpetually. Note that the proposed online algorithm can operate in an infinite time window, i.e., $N \to \infty$, where the algorithm continuously remains in Step (2.). The proposed online algorithm satisfies the ECCs and, as pointed out, does not require future information of neither the channel state nor the energy arrivals.

The mutual information that can be achieved by any online algorithm is inherently limited by the partial knowledge of the harvested energy and channel state. By using sophisticated prediction algorithms, the transmitter can have an estimation of the future harvested energy, which can be used to design online algorithms that perform close to the optimal offline algorithm. However, the use of these prediction algorithms has

[14]Note that the transmitter may stay silent in some channel accesses if the difference between two consecutive incoming energy packets is greater than the flowing window, $e_j - e_{j-1} > F_w$. This situation rarely takes place in practice since, in most common situations, F_w is several times the mean number of channel accesses per epoch. For example, in the simulated framework presented in Section V, we have obtained that F_w is 5 times the mean number of channel accesses per epoch.

two major drawbacks: (i) currently, there is a lack of models and prediction algorithms of the energy harvesting process; (ii) the computational complexity required for the prediction has to be as low as possible since the energy spent in the computation of the prediction cannot be used for transmission, which directly affects the achievable mutual information. We believe that our proposed online algorithm correctly balances these two points since it is a low-complexity online algorithm (the estimation of F_w can be done during the node deployment when the node is not limited by the harvested energy) that achieves a remarkably high fraction of the mutual information as shown by numerical simulation in the next section.

V. Simulation Results

By numerical simulation, in this section we evaluate the performance of the different solutions presented in the previous sections. We have considered a total of $N = 100$ channel accesses in which symbols are transmitted through $K = 8$ parallel streams. The channel access duration is $T_s = 10$ ms. The power consumptions associated with the channel access and stream activation are $P_c = 100$ mW and $P_d = 10$ mW, respectively [1]. A Rayleigh fading channel has been considered where the channel power gain satisfies $\mathbb{E}\{h_{k,n}\} = 1$. The energy harvesting process is modeled as a compound Poisson process as done in [9], where the packet arrival instants follow a Poisson distribution with rate $1/5$ and the energy in the packets is drawn from a uniform distribution and normalized by the total harvested energy that varies along the x-axis of Figures 4-8.

In the setup above, Figure 4 shows the achieved mutual information with the different presented resource allocation strategies: $\tilde{\mathcal{I}}^\star$ is the upper bound obtained in Section III-A by relaxing the feasible set of the stream and channel access indicator variables to the integer interval $[0, 1]$; $\hat{\mathcal{I}}$ is the mutual information achieved by the feasible resource allocation $\{\hat{\mathbf{P}}, \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\Psi}}\}$ that is obtained by projecting ρ_n and $\psi_{k,n}$ into the set $\{0, 1\}$ as explained in Section III-A, where we have used $\gamma = 0.5$ as it provides a good performance; *Duality* shows the mutual information achieved by solving the dual problem as explained in Section III-B; and *Online* depicts the mutual information achieved by the online algorithm presented in Section IV. Additionally, to assess the impact of energy harvesting versus traditional non-harvesting nodes, we have evaluated the performance of a virtual non-harvesting node in which the battery of the node is replaced by a new battery containing E_j Joules at the channel access e_j. Although this battery replacement is not feasible in practice, it allows us to fairly compare the performance of the energy harvesting node and the virtual battery operated node since both nodes have the same energy levels. For the non-harvesting node, we have designed a resource allocation strategy, named *Epoch by Epoch* (*EbE*), that uses the *Boxed Water-Flowing* in each epoch independently, i.e., water is not allowed to flow between epochs (due to the virtual battery replacement). Finally, we also compare our strategies with *DWF & PP* that uses the DWF in (1) with an additional post processing stage that scales the transmission powers to guarantee that the ECCs containing the circuitry power consumption are satisfied.

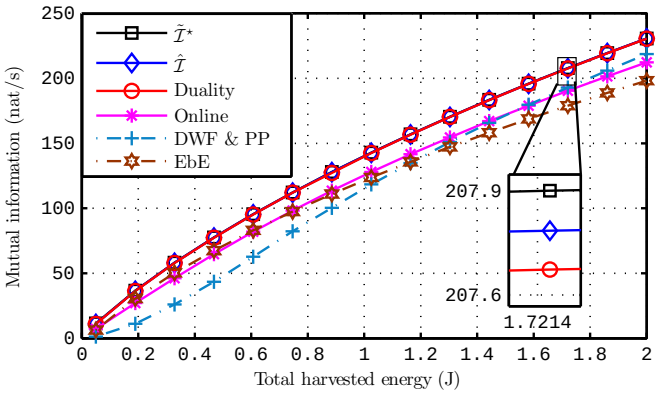

Figure 4. Achieved mutual information versus total harvested energy for the different algorithms.

In the magnified plot in Figure 4, one can observe that the optimality gap is almost zero since the difference between the upper bound, $\tilde{\mathcal{I}}^\star$, and the strategies $\hat{\mathcal{I}}$ and *Duality* is almost zero (remember that $\mathcal{I}^\star - \hat{\mathcal{I}} \leq \tilde{\mathcal{I}}^\star - \hat{\mathcal{I}}$ and the same applies for *Duality*). As expected, the proposed online algorithm has performance loss in comparison to the optimal offline solution as it has no knowledge of the future channel state and energy arrivals. This performance loss is evaluated in Figure 5 both in absolute (left y-axis) and relative (right y-axis) terms. It is observed that when the harvested energy is above 1 J, the relative performance loss is around 10%. The *Boxed Water-flowing* solutions ($\mathcal{I}^\star$ and *Duality*) also outperform the *EbE* scheme (for the battery operated node) and the *DWF & PP*. It is observed that the performance of the *EbE* scheme drops down for high energy levels because water is not allowed to flow across epochs. In opposition, *DWF & PP* has a very poor performance for low energy levels since the circuitry power consumption has not been accounted for in the optimization and plays an important role specially for low levels of harvested energy.

Figure 6 shows the percentage of the total harvested energy that is expended in the circuitry, i.e., $T_s \sum_{n=1}^{N} \left(P_c \rho_n + \sum_{k=1}^{K} P_d \psi_{k,n} \right) 100/\beta_J$. It is observed that when the harvested energy is low, the amount of energy spent in the circuitry components is a relatively high fraction of the total harvested energy. Additionally, when the harvested energy is high the different strategies show a similar percentage of circuitry energy consumption. However, the *Boxed Water-Flowing* strategies ($\mathcal{I}^\star$ and *Duality*) achieve a higher mutual information as seen in Figure 4. This is because the *Boxed Water-Flowing* solutions are able to activate the channel accesses and streams that contribute the most to the mutual information.

The computational complexities of the different strategies are compared in Figure 7 in terms of measured execution time versus harvested energy.[15] Observe that *Duality* requires a much lower execution time than $\hat{\mathcal{I}}$. However, to obtain such a good performance, the step size to update the dual variable must be carefully selected depending on the energy harvesting profile. In some manner, $\hat{\mathcal{I}}$ is more robust to variations of the

[15]Note that the execution time is approximately proportional to the algorithmic computational complexity.

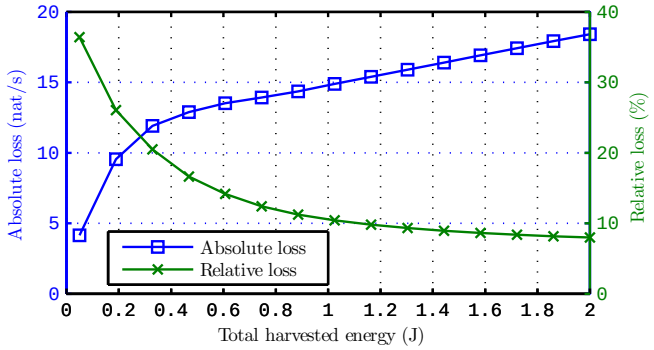

Figure 5. Performance loss of the proposed online algorithm versus the

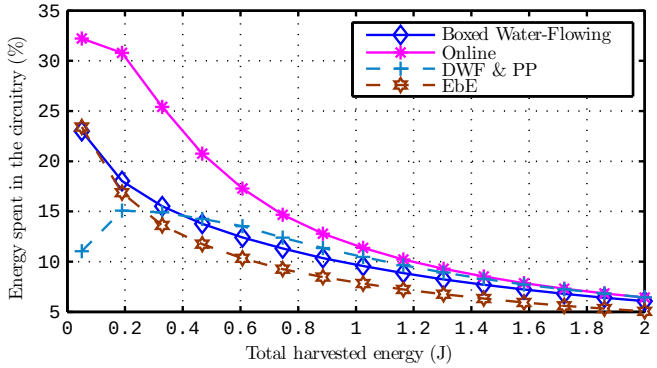

Figure 6. Percentage of the total harvested energy expended in the circuitry.

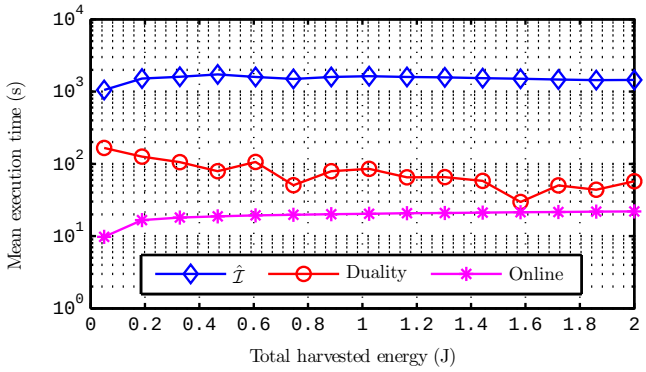

Figure 7. Mean execution time versus total harvested energy.

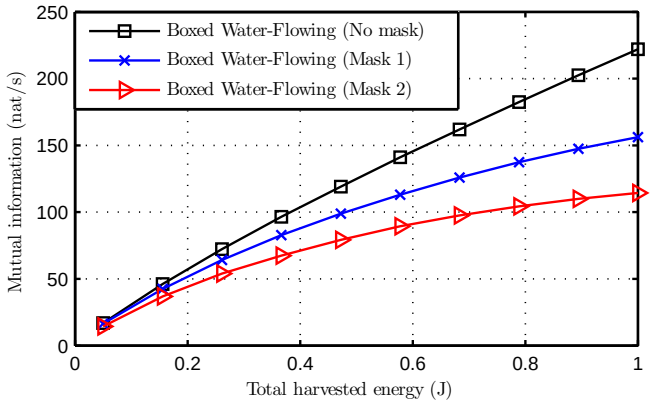

Figure 8. Mutual information of the *Boxed Water-Flowing* solution given different mask constraints.

energy harvesting profile; however, at a cost of having a higher computational complexity. Moreover, the complexity of the proposed *Online* solution is remarkably low, which makes it a good candidate solution to be implemented in wireless devices.

As pointed out in Remark 2, the *Boxed Water-Flowing* algorithms in Algorithms 1 and 2, can be trivially extended to include mask constraints. Figure 8 evaluates the impact of two different mask constraints on the achieved mutual information for $K = 80$ parallel streams where $P_d = 1$ mW (the remaining system parameters are the ones mentioned above). We have considered two different mask constraints: the first mask, *Mask 1*, limits the transmission power in each stream as $P_{k,n} \leq 25$ mW, $\forall k, n$; in the second mask, *Mask 2*, the transmission power of the external streams is further limited to avoid interferences to other possible transmissions, i.e., $P_{k,n} \leq 5$ mW, $k \in [1, 20] \cup [61, 80], \forall n$, and $P_{k,n} \leq 25$ mW, $\forall k \in [21, 60], \forall n$. As expected, at low energy levels, the mask constraint does not have a significant impact on the achieved mutual information because the transmission power in the different subchannels is low; however, when the harvested energy increases, the mask constraint limits the transmission power in the different subchannels and, as a result, the mutual information is reduced.

VI. Conclusions

In this paper, we have studied the resource allocation for a WEHN that maximizes the mutual information along N independent channel accesses in which symbols are sent through K parallel streams. The main contribution with respect to previous works is that we not only account for the transmission radiated power but we also consider the remaining sources of energy consumption at the transmitter. First, we have studied the offline maximization problem (where the transmitter has full knowledge of the harvested energy and channel state) and we have shown that it is not a convex optimization problem. Due to this lack of convexity, we have proposed and studied two different problems (the integer relaxation and the dual problem) from which we have obtained two suboptimal solutions of the offline maximization problem that asymptotically tend to the optimal solution when the number of channel accesses or streams per epoch is large. From these two problems, we have obtained a common condition for the activation of the channel access and streams, i.e., if the epoch water level is greater than the corresponding cutoff water level. Based on the cutoff water level concept, we have devised the *Boxed Water-Flowing*, a novel graphical representation of the asymptotically optimal offline resource allocation. Additionally, we have proposed a practical online algorithm that does not require knowledge of the future energy arrivals nor the channel state. From the simulation results, we have confirmed that the *Boxed Water-Flowing* resource allocation is the asymptotically optimal offline resource allocation and that the performance loss of the proposed *online* solution is very small. Moreover, we have evaluated the computational complexity of the different resource allocation strategies and obtained that the online solution is the one that requires the lowest execution time.

APPENDIX

A. Proofs of Lemmas 1 and 2

1) Proof of Lemma 1: We know that the $M_n^\star$ streams that contribute to the channel access activation are the ones with the best channel gains, i.e., $h_{1,n}, \ldots, h_{M_n^\star,n}$, because these streams are the ones that contribute the most to the objective function. Given $M_n^\star$, the following two conditions must be satisfied at the cutoff water level, i.e., when $W_j = \hat{W}_n(M_n^\star)$:
(C1) $\forall k \leq M_n^\star$, the following relations must be fulfilled: (i) $P_{k,n} > 0$; (ii) $\psi_{k,n} \in (0,1)$; (iii) $\hat{\eta}_{k,n} = 0$; (iv) $\check{\eta}_{k,n} = 0$; and (v) $\xi_{k,n} = 0$.[16]
(C2) $\forall k > M_n^\star$, the following conditions must be satisfied: (i) $P_{k,n} = 0$; (ii) $\psi_{k,n} = 0$; (iii) $\check{\eta}_{k,n} = 0$; and (iv) $\mu_{k,n} = 0$.[17]

Going back to (7), the radiated power of the streams $k \leq M_n^\star$ at the channel access cutoff water level is $P_{k,n} = \psi_{k,n}\left(\hat{W}_n(M_n^\star) - h_{k,n}^{-1}\right)$, where we have used that $W_j = \hat{W}_n(M_n^\star) = (T_s \sum_{\ell=j}^{J} \lambda_\ell)^{-1}$. Plugging this into the KKT condition in (6b), we have that all the streams that contribute to the n-th channel access activation, i.e., $k \leq M_n^\star$, must satisfy:

$$\frac{\partial \tilde{\mathcal{L}}}{\partial \psi_{k,n}} = \log(h_{k,n}\hat{W}_n(M_n^\star)) - 1 + \frac{1}{\hat{W}_n(M_n^\star)h_{k,n}} - \frac{P_d}{\hat{W}_n(M_n^\star)} - \mu_{k,n} = 0, \quad \forall k \leq M_n^\star. \tag{13}$$

Note that we cannot isolate $\hat{W}_n(M_n^\star)$ in the previous equation due to the dependence on the Lagrange multiplier $\mu_{k,n}$. To get rid of this dependence, we can use the KKT condition in (6c) evaluated at the cutoff water level, i.e.,

$$\sum_{k=1}^{K} \mu_{k,n} = \sum_{k=1}^{M_n^\star} \mu_{k,n} = \frac{P_c}{\hat{W}_n(M_n^\star)}, \tag{14}$$

where we have used that $\mu_{k,n} = 0$,$\forall k > M_n^\star$, which follows from (C2).

With this, we can obtain an equation that does not depend on $\mu_{k,n}$ by adding the equations $\partial\tilde{\mathcal{L}}/\partial\psi_{k,n} = 0$, $\forall k \leq M_n^\star$. Thus, the n-th channel access cutoff water level, $\hat{W}_n(M_n^\star)$, is obtained by solving $\sum_{k=1}^{M_n^\star} \partial\tilde{\mathcal{L}}/\partial\psi_{k,n} = 0$, which is performed in Appendix D (set $M := M_n^\star$, $\hat{W} := \hat{W}_n(M_n^\star)$, $H_k := h_{k,n}$, and $P := M_n^\star P_d + P_c$), and is the one given in (8).

Up to now, we have shown that if $W_j = \hat{W}_n(M_n^\star)$, the channel access is "partially on", i.e., $\rho_n \in (0,1)$, $\psi_{k,n} \in (0,1), \forall k \leq M_n^\star$ and $\psi_{k,n} = 0, \forall k > M_n^\star$. Note that if $W_j = \hat{W}_n(M_n^\star) + \epsilon$, with $\epsilon > 0$, then we have that $\log(h_{k,n}W_j) - 1 + 1/(W_j h_{k,n}) - P_d/W_j - \mu_{k,n} > 0$. Thus, in order to satisfy (6b), we must have $\check{\eta}_{k,n} > 0$, $\forall k \leq M_n^\star$. Then, from the slackness condition (6g), we know that $\psi_{k,n} = 1$, $\forall k \leq M_n^\star$, and hence $\rho_n = 1$. A similar approach can be used to show that if $W_j < \hat{W}_n(M_n^\star)$, then $\psi_{k,n} = 0$, $\forall k$, and hence $\rho_n = 0$. ∎

[16]Where (i) and (ii) follow from the fact that the stream must contribute to the channel access activation and (iii), (iv), and (v) follow from the slackness conditions (6f), (6g), and (6e), respectively.

[17]Where (i) and (ii) follow from the fact that the stream must not contribute to the channel access activation and (iii) and (iv) follow from the slackness conditions (6g) and (6h).

2) Proof of Lemma 2: Now, we derive the expression of the k-th *stream cutoff water level*, $\bar{W}_{k,n}$, i.e., the water level at which the k-th stream becomes partially active, where now $k > M_n^\star$. Similarly as in the proof of Lemma 1, in the k-th stream cutoff water level, we must have that $P_{k,n} > 0$, $\psi_{k,n} \in (0,1)$, $\rho_n = 1$, and from the slackness conditions we know that $\check{\eta}_{k,n} = 0$, $\hat{\eta}_{k,n} = 0$, $\xi_{k,n} = 0$, $\mu_{k,n} = 0$. The k-th stream cutoff water level is obtained by solving the equation obtained from the KKT condition in (6b) for $\bar{W}_{k,n}$ (set $M := 1$, $\hat{W} := \bar{W}_{k,n}$, $H_1 := h_{k,n}$ and $P := P_d$ in Appendix D) and is the one given in (9). Following the same procedure as in the last paragraph of the proof of Lemma 1, it can be shown that if $W_j > \bar{W}_{k,n}$, $k \in (M_n^\star, K]$, then $\psi_{k,n} = 1$. Alternatively, if $W_j < \bar{W}_{k,n}$, $k \in (M_n^\star, K]$, then $\psi_{k,n} = 0$. ∎

B. Proof of Proposition 1

In this appendix, we prove that $M_n^\star$ can be found by performing a forward search over M_n and that it is the smallest M_n that satisfies the condition $\hat{W}_n(M_n) < \bar{W}_{M_n+1,n}$. To prove Proposition 1, we need to make use of the following two lemmas:

Lemma 3. *If $\hat{W}_n(M_n) = \bar{W}_{M_n+1,n}$, then $\hat{W}_n(M_n+1) = \hat{W}_n(M_n)$.*

Proof: See Appendix B-1. ∎

Lemma 4. *If $\hat{W}_n(M_n) \geq \bar{W}_{M_n+1,n}$, then: (a) $\hat{W}_n(M_n+1) \geq \bar{W}_{M_n+1,n}$; (b) the function $\hat{W}_n(M_n+1)$ is monotonically decreasing with $h_{M_n+1,n}$.*

Proof: See Appendixes B-2 and B-3. ∎

Observe that $\bar{W}_{M_n+1,n}$ is a function of $h_{M_n+1,n}$ in opposition to $\hat{W}_n(M_n)$ that does not depend on $h_{M_n+1,n}$. In this context, let $\tilde{h}_{M_n+1,n}$ be the specific value of the channel gain of the stream $k = M_n+1$ that satisfies $\bar{W}_{M_n+1,n} = \hat{W}_n(M_n)$.

We first show that any M_n smaller than $M_n^\star$ is suboptimal. We know that $\hat{W}_n(M_n) \geq \bar{W}_{M_n+1,n}$, $\forall M_n < M_n^\star$, since the condition $\hat{W}_n(M_n) < \bar{W}_{M_n+1,n}$ is not satisfied until $M_n := M_n^\star$. Since $\bar{W}_{M_n+1,n}$ is decreasing with the channel gain (see Remark 4), the condition $\hat{W}_n(M_n) \geq \bar{W}_{M_n+1,n}$ implies $h_{M_n+1,n} \geq \tilde{h}_{M_n+1,n}$. From Lemma 3, we have that at $\tilde{h}_{M_n+1,n}$, $\hat{W}_n(M_n+1) = \hat{W}_n(M_n)$ and from Lemma 4.b if $h_{M_n+1,n} \geq \tilde{h}_{M_n+1,n}$, then $\hat{W}_n(M_n+1) \leq \hat{W}_n(M_n)$, $\forall M_n < M_n^\star$ or, equivalently, $\hat{W}_n(1) \geq \hat{W}_n(2) \geq \cdots \geq \hat{W}_n(M_n^\star)$. Therefore, any M_n in [1, $M_n^\star$) is suboptimal.

Now, we prove the suboptimality of any M_n greater than $M_n^\star$. From the structure of the forward search, the following relationship is satisfied $\hat{W}_n(M_n) < \bar{W}_{M_n+1,n}$, $\forall M_n \in [M_n^\star, K]$. Thus, the streams $k > M_n^\star$ cannot contribute to the channel access activation since these streams are not active at the cutoff water level.

Finally, we must show that the streams $k \leq M_n^\star$ are active in the channel access cutoff water level, i.e., $\bar{W}_{k,n} \leq \hat{W}_n(M_n^\star)$, which is verified as proved in Lemma 4.a. ∎

$$\frac{(\prod_{k=1}^{M_n} h_{k,n}^{\frac{1}{M_n}})\bar{W}_{M_n+1,n}}{e}\mathsf{log}\left(\frac{(\prod_{k=1}^{M_n} h_{k,n}^{\frac{1}{M_n}})\bar{W}_{M_n+1,n}}{e}\right)=\frac{\prod_{k=1}^{M_n} h_{k,n}^{\frac{1}{M_n}}}{eM_n}(P_c+M_nP_d-\sum_{k=1}^{M_n}h_{k,n}^{-1})\Rightarrow \tag{16}$$

$$\tilde{X}_{M_n+1}=\frac{\prod_{k=1}^{M_n+1} h_{k,n}^{\frac{1}{M_n+1}}}{e(M_n+1)}\left[M_n\bar{W}_{M_n+1,n}\mathsf{log}\left(\frac{(\prod_{k=1}^{M_n} h_{k,n}^{\frac{1}{M_n}})\bar{W}_{M_n+1,n}}{e}\right)+P_d-h_{M_n+1,n}^{-1}\right]\Rightarrow \tag{17}$$

$$\tilde{X}_{M_n+1}=\frac{(\prod_{k=1}^{M_n+1} h_{k,n}^{\frac{1}{M_n+1}})M_n\bar{W}_{M_n+1,n}}{e(M_n+1)}\left[\mathsf{log}\left(\frac{(\prod_{k=1}^{M_n} h_{k,n}^{\frac{1}{M_n}})\bar{W}_{M_n+1,n}}{e}\right)+\frac{\mathcal{W}_0(\frac{P_dh_{M_n+1,n}{}^{-1}}{e})}{M_n}\right]\Rightarrow \tag{18}$$

$$\tilde{X}_{M_n+1}=\frac{(\prod_{k=1}^{M_n+1} h_{k,n}^{\frac{1}{M_n+1}})\bar{W}_{M_n+1,n}}{e(M_n+1)}\mathsf{log}\left(\frac{(\prod_{k=1}^{M_n} h_{k,n})\bar{W}_{M_n+1,n}^{M_n}}{e^{M_n}}e^{\mathcal{W}_0(\frac{P_dh_{M_n+1,n}{}^{-1}}{e})}\right)\Rightarrow \tag{19}$$

$$\tilde{X}_{M_n+1}=\frac{(\prod_{k=1}^{M_n+1} h_{k,n}^{\frac{1}{M_n+1}})\bar{W}_{M_n+1,n}}{e}\mathsf{log}\left(\frac{(\prod_{k=1}^{M_n+1} h_{k,n}^{\frac{1}{M_n+1}})\bar{W}_{M_n+1,n}}{e}\right) \tag{20}$$

---

1) Proof of Lemma 3: Let the channel access cutoff water level for a given M_n be expressed as

$$\hat{W}_n(M_n)=\frac{X_{M_n}}{\frac{1}{e}(\prod_{k=1}^{M_n} h_{k,n}^{\frac{1}{M_n}})\mathcal{W}_0(X_{M_n})},$$

where X_{M_n} is the argument of the Lambert function, i.e.,

$$X_{M_n}=\frac{\prod_{k=1}^{M_n} h_{k,n}^{\frac{1}{M_n}}}{eM_n}(P_c+M_nP_d-\sum_{k=1}^{M_n}h_{k,n}^{-1}). \tag{15}$$

In (16)-(20), we impose the condition $\bar{W}_{M_n+1,n}=\hat{W}_n(M_n)$, and, after some algebra, we obtain the argument of the Lambert function in $\hat{W}_n(M_n+1)$, i.e., $\tilde{X}_{M_n+1}$, where the tilde denotes that it is the argument that satisfies $\bar{W}_{M_n+1,n}=\hat{W}_n(M_n)$. To obtain (17), we have multiplied both sides in (16) by

$$\frac{M_n\prod_{k=1}^{M_n+1} h_{k,n}^{\frac{1}{M_n+1}}}{\prod_{k=1}^{M_n} h_{k,n}^{\frac{1}{M_n}}}$$

and used the definition of $\tilde{X}_{M_n+1}$, which follows from (15). In (20), we have used that $e^{\mathcal{W}(z)}=z/\mathcal{W}(z)$, which directly follows from the definition of the Lambert function. The cutoff water level for M_n+1 active streams is

$\hat{W}_n(M_n+1)=\frac{\tilde{X}_{M_n+1}}{\frac{1}{e}(\prod_{k=1}^{M_n+1} h_{k,n}^{\frac{1}{M_n+1}})\mathcal{W}_0(\tilde{X}_{M_n+1})}$. Thus, $\tilde{X}_{M_n+1}=\frac{\prod_{k=1}^{M_n+1} h_{k,n}^{\frac{1}{M_n+1}}}{e}\hat{W}_n(M_n+1)\mathsf{log}\left(\frac{\prod_{k=1}^{M_n+1} h_{k,n}^{\frac{1}{M_n+1}}}{e}\hat{W}_n(M_n+1)\right)$.

By comparing this expression with (20), we have that $\hat{W}_n(M_n+1)$ must be equal to $\bar{W}_{M_n+1,n}$ and, thus, we have $\hat{W}_n(M_n+1)=\hat{W}_n(M_n)$. ∎

2) Proof of Lemma 4.a : Following similar steps as in (16)-(20), it is easy to show that if $\hat{W}_n(M_n)\geq\bar{W}_{M_n+1,n}$, then $X_{M_n+1}\geq\tilde{X}_{M_n+1}$. From where, it follows

$$\frac{X_{M_n+1}}{\mathcal{W}_0(X_{M_n+1})}\geq\frac{(\prod_{k=1}^{M_n+1} h_{k,n}^{\frac{1}{M_n+1}})\bar{W}_{M_n+1,n}}{e}$$

and, therefore, we also have $\hat{W}_n(M_n+1)\geq\bar{W}_{M_n+1,n}$. ∎

3) Proof of Lemma 4.b: In the following lines, we demonstrate that $\partial\hat{W}_n(M_n+1)/\partial h_{M_n+1,n}\leq 0$ for $\hat{W}_n(M_n)\geq\bar{W}_{M_n+1,n}$ or, equivalently, for $X_{M_n+1}\geq\tilde{X}_{M_n+1}$. Thus, in (21)-(24) we analyze the sign of the derivative. In (22), we have used that

$$\frac{\partial X_{M_n+1}}{\partial h_{M_n+1,n}}=\frac{X_{M_n+1}}{(M_n+1)h_{M_n+1,n}}+\frac{\prod_{k=1}^{M_n+1} h_{k,n}^{\frac{1}{M_n+1}}}{e(M_n+1)h_{M_n+1,n}^2}.$$

In (23), we have defined

$$m\triangleq eh_{M_n+1,n}\left(\prod_{k=1}^{M_n+1} h_{k,n}^{\frac{1}{M_n+1}}\right)^{-1}$$

and evaluated the derivative of the Lambert function, i.e.,

$$\frac{\mathrm{d}\mathcal{W}_0(X_{M_n+1})}{\mathrm{d}X_{M_n+1}}=\frac{\mathcal{W}_0(X_{M_n+1})}{X_{M_n+1}(1+\mathcal{W}_0(X_{M_n+1}))}.$$

From (24), we see that the sign of the derivative depends on the product of the Lambert function (which is positive for $X_{M_n+1}>0$ and negative for $X_{M_n+1}<0$) and the difference between the Lambert function and a line with slope m. To demonstrate that $\hat{W}_n(M_n+1)$ is monotonically decreasing with $h_{M_n+1,n}$ for $X_{M_n+1}\geq\tilde{X}_{M_n+1}$, we must show that

$$\mathcal{W}_0(X_{M_n+1})\overset{(a)}{<}\bar{m}X_{M_n+1}\overset{(b)}{\leq}mX_{M_n+1},\quad\forall X_{M_n+1}>\max\{0,\tilde{X}_{M_n+1}\}, \tag{25}$$

$$\mathcal{W}_0(X_{M_n+1})\overset{(a)}{>}\bar{m}X_{M_n+1}\overset{(b)}{\geq}mX_{M_n+1},\quad\forall\tilde{X}_{M_n+1}<X_{M_n+1}<0, \tag{26}$$

where $\bar{m}$ is an arbitrary constant. Observe that when $\bar{m}\leq m$, inequalities (b) in (25) and (26) are satisfied. In the following lines, we propose a specific $\bar{m}$ that allows us to prove inequalities in (a) and thus to demonstrate that $\hat{W}_n(M_n+1)$ is monotonically decreasing with $h_{M_n+1,n}$.

Note that, by replacing m in (20), we obtain

$$\tilde{X}_{M_n+1}=\frac{\bar{W}_{M_n+1,n}h_{M_n+1,n}}{m}\mathsf{log}\left(\frac{\bar{W}_{M_n+1,n}h_{M_n+1,n}}{m}\right).$$

From this relation, we can express m as a function of $\tilde{X}_{M_n+1}$, i.e., $m=\bar{W}_{M_n+1,n}h_{M_n+1,n}\mathcal{W}_0(\tilde{X}_{M_n+1})/\tilde{X}_{M_n+1}$. Note that $\bar{W}_{M_n+1,n}h_{M_n+1,n}\geq 1$ because $\bar{W}_{M_n+1,n}\in[h_{M_n+1,n}^{-1},\infty)$ as pointed out in Remark 4. Thus, $m\geq\mathcal{W}_0(\tilde{X}_{M_n+1})/\tilde{X}_{M_n+1}$. Let $\bar{m}$ denote the minimum slope, i.e.,

$$\mathsf{sign}\left[\frac{h_{M_n+1,n}^{-2}}{M_n+1}\mathcal{W}_0(X_{M_n+1})-\frac{1}{M_n+1}\left((M_n+1)P_d+P_c-\sum_{k=1}^{M_n+1}h_{k,n}^{-1}\right)\frac{\mathrm{d}\mathcal{W}_0(X_{M_n+1})}{\mathrm{d}X_{M_n+1}}\frac{\partial X_{M_n+1}}{\partial h_{M_n+1,n}}\right] \tag{21}$$

$$= \mathsf{sign}\left[\mathcal{W}_0(X_{M_n+1})-X_{M_n+1}\left(1+\frac{eh_{M_n+1,n}X_{M_n+1}}{\prod_{k=1}^{M_n+1}h_{k,n}^{\frac{1}{M_n+1}}}\right)\frac{\mathrm{d}\mathcal{W}_0(X_{M_n+1})}{\mathrm{d}X_{M_n+1}}\right] \tag{22}$$

$$= \mathsf{sign}\left[\mathcal{W}_0(X_{M_n+1})-X_{M_n+1}\left(1+mX_{M_n+1}\right)\frac{\mathcal{W}_0(X_{M_n+1})}{X_{M_n+1}(1+\mathcal{W}_0(X_{M_n+1}))}\right] \tag{23}$$

$$= \mathsf{sign}\left[\mathcal{W}_0(X_{M_n+1})\left(\mathcal{W}_0(X_{M_n+1})-mX_{M_n+1}\right)\right] \tag{24}$$

$$\hat{W}\left(\prod_{k=1}^{M}H_k^{\frac{1}{M}}\right)\left[\mathsf{log}\left(\hat{W}\prod_{k=1}^{M}H_k^{\frac{1}{M}}\right)-\mathsf{log}e\right]+\frac{1}{M}\left(\prod_{k=1}^{M}H_k^{\frac{1}{M}}\right)\left(\sum_{k=1}^{M}H_k^{-1}\right)-\frac{P}{M}\left(\prod_{k=1}^{M}H_k^{\frac{1}{M}}\right)=0\Rightarrow \tag{27}$$

$$\frac{\hat{W}\prod_{k=1}^{M}H_k^{\frac{1}{M}}}{e}\mathsf{log}\left(\frac{\hat{W}\prod_{k=1}^{M}H_k^{\frac{1}{M}}}{e}\right)=\frac{\prod_{k=1}^{M}H_k^{\frac{1}{M}}}{Me}(P-\sum_{k=1}^{M}H_k^{-1})\Rightarrow\hat{W}=\frac{P-\sum_{k=1}^{M}H_k^{-1}}{M\mathcal{W}_0\left(\frac{\prod_{k=1}^{M}H_k^{\frac{1}{M}}}{Me}(P-\sum_{k=1}^{M}H_k^{-1})\right)} \tag{28}$$

$\bar{m}=\mathcal{W}_0(\tilde{X}_{M_n+1})/\tilde{X}_{M_n+1}$. Since $\bar{m}\leq m$ the inequalities (b) in (25) and (26) are satisfied. Note that the Lambert function, $\mathcal{W}_0(X_{M_n+1})$, and the line $\bar{m}X_{M_n+1}$ cross both at the origin ($X_{M_n+1}=0$) and at the point $X_{M_n+1}=\tilde{X}_{M_n+1}$, i.e., $\mathcal{W}_0(\tilde{X}_{M_n+1})=\bar{m}\tilde{X}_{M_n+1}$. Finally, by using the concavity of the positive branch of the Lambert function [18] and these two crossing points, it is straight forward to show that the inequalities (a) in (25) and (26) are satisfied. ■

C. Proof of Proposition 2

Let $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ be two different energy harvesting profiles, where the energy packet arrival instants are the same but the amount of energy in the packets is different, and let $\left\{\mathbf{P}^{\star}_{\boldsymbol{\beta}_1},\boldsymbol{\rho}^{\star}_{\boldsymbol{\beta}_1},\boldsymbol{\Psi}^{\star}_{\boldsymbol{\beta}_1}\right\}$ and $\left\{\mathbf{P}^{\star}_{\boldsymbol{\beta}_2},\boldsymbol{\rho}^{\star}_{\boldsymbol{\beta}_2},\boldsymbol{\Psi}^{\star}_{\boldsymbol{\beta}_2}\right\}$ be the associated optimal solutions to (4), respectively. Therefore, showing that the time-sharing condition is fulfilled is equivalent to demonstrating that

$$\begin{aligned}&I(\mathbf{P}^{\star}_{\theta\boldsymbol{\beta}_1+(1-\theta)\boldsymbol{\beta}_2},\boldsymbol{\rho}^{\star}_{\theta\boldsymbol{\beta}_1+(1-\theta)\boldsymbol{\beta}_2},\boldsymbol{\Psi}^{\star}_{\theta\boldsymbol{\beta}_1+(1-\theta)\boldsymbol{\beta}_2})\geq\\ &\quad I(\mathbf{P}_{\theta\boldsymbol{\beta}_1+(1-\theta)\boldsymbol{\beta}_2},\boldsymbol{\rho}_{\theta\boldsymbol{\beta}_1+(1-\theta)\boldsymbol{\beta}_2},\boldsymbol{\Psi}_{\theta\boldsymbol{\beta}_1+(1-\theta)\boldsymbol{\beta}_2})\geq\\ &\quad\quad \theta I(\mathbf{P}^{\star}_{\boldsymbol{\beta}_1},\boldsymbol{\rho}^{\star}_{\boldsymbol{\beta}_1},\boldsymbol{\Psi}^{\star}_{\boldsymbol{\beta}_1})+(1-\theta)I(\mathbf{P}^{\star}_{\boldsymbol{\beta}_2},\boldsymbol{\rho}^{\star}_{\boldsymbol{\beta}_2},\boldsymbol{\Psi}^{\star}_{\boldsymbol{\beta}_2}),\end{aligned}$$

where $\theta\in[0,1]$, $\left\{\mathbf{P}^{\star}_{\theta\boldsymbol{\beta}_1+(1-\theta)\boldsymbol{\beta}_2},\boldsymbol{\rho}^{\star}_{\theta\boldsymbol{\beta}_1+(1-\theta)\boldsymbol{\beta}_2},\boldsymbol{\Psi}^{\star}_{\theta\boldsymbol{\beta}_1+(1-\theta)\boldsymbol{\beta}_2}\right\}$ is the optimal resource allocation for an energy harvesting profile equal to $\theta\boldsymbol{\beta}_1+(1-\theta)\boldsymbol{\beta}_2$ and $\left\{\mathbf{P}_{\theta\boldsymbol{\beta}_1+(1-\theta)\boldsymbol{\beta}_2},\boldsymbol{\rho}_{\theta\boldsymbol{\beta}_1+(1-\theta)\boldsymbol{\beta}_2},\boldsymbol{\Psi}_{\theta\boldsymbol{\beta}_1+(1-\theta)\boldsymbol{\beta}_2}\right\}$ is any feasible resource allocation.

In the following lines, we construct a feasible resource allocation

$$\left\{\mathbf{P}_{\theta\boldsymbol{\beta}_1+(1-\theta)\boldsymbol{\beta}_2},\boldsymbol{\rho}_{\theta\boldsymbol{\beta}_1+(1-\theta)\boldsymbol{\beta}_2},\boldsymbol{\Psi}_{\theta\boldsymbol{\beta}_1+(1-\theta)\boldsymbol{\beta}_2}\right\}$$

that satisfies the time sharing condition in an epoch by epoch basis. In this context, since the procedure is the same for all the epochs, we just explain how to obtain the resource allocation of the streams contained in a generic epoch τ_j. Let $x=1,\ldots,\mathcal{X}$ be an index used to indicate the different channel realizations observed within the streams in τ_j. Thus, $\mathcal{X}$ is the number of different channel realizations within the epoch. Let the set $\mathbf{s}_x$ contain all the streams in τ_j that have channel gain equal to $\bar{h}_x$, i.e., $\mathbf{s}_x=\{\{k,n\}|h_{k,n}=\bar{h}_x,n\in\tau_j\}$. K_x denotes the cardinality of $\mathbf{s}_x$, which is a large number since, in the proposition statement, we have considered that within each epoch every channel realization is observed a sufficiently large number of times. In the following lines, we explain how to construct the resource allocation of the channel accesses in $\mathbf{s}_x$. Since K_x is large, θ can be approximated as $\theta\approxeq\bar{N}_x/K_x$, where $\bar{N}_x$ is an integer in the interval $[0,K_x]$. Due to the nature of DWF, the power allocated by the optimal solution given some energy harvesting profile is equal for all the streams in $\mathbf{s}_x$ because when a certain channel access is active the transmission power in (1) only depends on the epoch water level and on the channel gain. We construct the resource allocation in the streams in $\mathbf{s}_x$ by assigning the resource allocation in $\{[\mathbf{P}^{\star}_{\boldsymbol{\beta}_1}]_{\{k',n'\}},[\boldsymbol{\rho}^{\star}_{\boldsymbol{\beta}_1}]_{\{k',n'\}},[\boldsymbol{\Psi}^{\star}_{\boldsymbol{\beta}_1}]_{\{k',n'\}}\}$ to $\bar{N}_x$ streams of $\mathbf{s}_x$ and the resource allocation in $\{[\mathbf{P}^{\star}_{\boldsymbol{\beta}_2}]_{\{k',n'\}},[\boldsymbol{\rho}^{\star}_{\boldsymbol{\beta}_2}]_{\{k',n'\}},[\boldsymbol{\Psi}^{\star}_{\boldsymbol{\beta}_2}]_{\{k',n'\}}\}$ to the remaining $K_x-\bar{N}_x$ streams, where $\{k',n'\}$ is any stream contained in $\mathbf{s}_x$. This procedure is repeated for the different channel realizations to obtain the resource allocation in all the streams of τ_j, then the total power consumption in τ_j is equivalent to θ times the power consumption in τ_j given by $\left\{\mathbf{P}^{\star}_{\boldsymbol{\beta}_1},\boldsymbol{\rho}^{\star}_{\boldsymbol{\beta}_1},\boldsymbol{\Psi}^{\star}_{\boldsymbol{\beta}_1}\right\}$ and $(1-\theta)$ times the power consumption in τ_j given by $\left\{\mathbf{P}^{\star}_{\boldsymbol{\beta}_2},\boldsymbol{\rho}^{\star}_{\boldsymbol{\beta}_2},\boldsymbol{\Psi}^{\star}_{\boldsymbol{\beta}_2}\right\}$. After repeating this process for all the epochs, the constructed resource allocation $\left\{\mathbf{P}_{\theta\boldsymbol{\beta}_1+(1-\theta)\boldsymbol{\beta}_2},\boldsymbol{\rho}_{\theta\boldsymbol{\beta}_1+(1-\theta)\boldsymbol{\beta}_2},\boldsymbol{\Psi}_{\theta\boldsymbol{\beta}_1+(1-\theta)\boldsymbol{\beta}_2}\right\}$ is a feasible solution as the ECCs are satisfied and the obtained mutual information is $\theta I(\mathbf{P}^{\star}_{\boldsymbol{\beta}_1},\boldsymbol{\rho}^{\star}_{\boldsymbol{\beta}_1},\boldsymbol{\Psi}^{\star}_{\boldsymbol{\beta}_1})+(1-\theta)I(\mathbf{P}^{\star}_{\boldsymbol{\beta}_2},\boldsymbol{\rho}^{\star}_{\boldsymbol{\beta}_2},\boldsymbol{\Psi}^{\star}_{\boldsymbol{\beta}_2})$. Therefore, we have shown that the *time-sharing condition* is satisfied. ■

D. Derivation of the cutoff water level

In this appendix, we use a generic notation that serves us to derive both $\hat{W}_n(M_n)$ and $\bar{W}_{k,n}$. Thus, $\hat{W}$ denotes the cutoff water level for the case of considering M active streams at the cutoff region with gains H_k, $k\in[1,M]$, and when the circuitry power consumption is P. Specifically, $\hat{W}$ is obtained by

solving the equation $\sum_{k=1}^{M}\left[\log(\hat{W}H_k)-1+1/(\hat{W}H_k)\right]-P/\hat{W}=0$ as done in (27)-(28). In (27), we have multiplied by $\hat{W}\left(\prod_{k=1}^{M}H_k^{\frac{1}{M}}\right)/M$. In (28), we have used that $b\log b=a\Leftrightarrow b=a/\mathcal{W}(a)$, which follows from the definition of the Lambert function [18]. Moreover, since $\hat{W}>H_k^{-1}$, $\forall k=[1,M]$, so that the streams are active in the cutoff region, the term b is always greater than e^{-1} and, thus, the positive branch of the Lambert function, which is denoted by $\mathcal{W}_0(\cdot)$, is used. ■

REFERENCES

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

[2] O. Orhan, D. Gunduz, and E. Erkip, "Throughput maximization for an energy harvesting communication system with processing cost," in *Proceedings of the IEEE Information Theory Workshop*, Sep. 2012, pp. 84–88.

[3] M. Gregori and M. Payaró, "Throughput maximization for a wireless energy harvesting node considering the circuitry power consumption," in *Proceedings of the IEEE Vehicular Technology Conference*, Sep. 2012, pp. 1–5.

[4] M. Gregori, A. Pascual-Iserte, and M. Payaró, "Mutual information maximization for a wireless energy harvesting node considering the circuitry power consumption," in *Proceedings of the IEEE Wireless Communications and Networking Conference*, Apr. 2013, pp. 4238–4243.

[5] T. M. Cover and J. A. Thomas, *Elements of information theory*. New York, NY, USA: Wiley-Interscience, 1991.

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

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

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

[9] 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, Sep. 2011.

[10] M. Gregori and M. Payaró, "Energy-efficient transmission for wireless energy harvesting nodes," *IEEE Trans. on Wireless Communications*, vol. 12, no. 3, pp. 1244–1254, Mar. 2013.

[11] ——, "On the precoder design of a wireless energy harvesting node in linear vector Gaussian channels with arbitrary input distribution," *IEEE Trans. on Communications*, vol. 61, no. 5, pp. 1868–1879, May 2013.

[12] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. on Wireless Communications*, vol. 4, no. 5, pp. 2349–2360, Sep. 2005.

[13] ——, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 6, pp. 1089 – 1098, Aug. 2004.

[14] C. Isheden, Z. Chong, E. Jorswieck, and G. Fettweis, "Framework for link-level energy efficiency optimization with informed transmitter," *IEEE Trans. on Wireless Communications*, vol. 11, no. 8, pp. 2946–2957, Aug. 2012.

[15] P. Youssef-Massaad, M. Medard, and L. Zheng, "Impact of processing energy on the capacity of wireless channels," in *Proceedings of International Symposium on Information Theory and its Applications*, Oct. 2004.

[16] P. Youssef-Massaad, L. Zheng, and M. Medard, "Bursty transmission and glue pouring: on wireless channels with overhead costs," *IEEE Trans. on Wireless Communications*, vol. 7, no. 12, pp. 5188 –5194, Dec. 2008.

[17] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge Univ Pr, 2004.

[18] R. Corless, G. Gonnet, D. Hare, D. Jeffrey, and D. Knuth, "On the LambertW function," *Advances in Computational mathematics*, vol. 5, no. 1, pp. 329–359, 1996.

[19] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. on Communications*, vol. 54, no. 7, pp. 1310–1322, Jul. 2006.

[20] S. Boyd, L. Xiao, and A. Mutapcic, "Subgradient methods," *lecture notes of EE392o, Stanford University, Autumn Quarter*, vol. 2004, 2003. [Online]. Available: http://www.stanford.edu/class/ee364b/lectures/subgrad_method_notes.pdf

Maria Gregori was born in Barcelona, Spain, in 1985. She received the Engineering Degree in Telecommunications and the European Master of Research on Information and Communication Technologies (MERIT) from the Universitat Politècnica de Catalunya (UPC), Barcelona, in 2009 and 2011, respectively. During the course 2008-2009, she pursued her bachelor thesis on NanoNetworks at the Georgia Institute of Technology with a fellowship from the Vodafone Foundation. From November 2009 to September 2010, she worked as research assistant on dynamic binary optimization for Intel-UPC. Since September 2010, she is a Ph.D. student at the Centre Tecnològic de Telecomunicacions de Catalunya (CTTC). She received predoctoral grants from the Generalitat de Catalunya and from the CTTC fellowship program. From September 2013 to March 2014, she held a research appointment at the Hong Kong University of Science and Technology. In 2013, she received the NEWCOM# Best young researcher's paper award and she was recognized as an exemplary reviewer of the IEEE Wireless Communications Letters. Her primary research interests are in the fields of energy harvesting for wireless communications, energy-aware communication system design, information theory, and NanoNetworks.

Miquel Payaró (S'03-M'08-SM'13) was born in Barcelona, Spain, in 1979. He received the degree in Electrical Engineering and the Ph.D. degrees from the Universitat Politècnica de Catalunya (UPC), Barcelona, in 2002 and 2007, respectively. He received a predoctoral grant from the Generalitat de Catalunya and from the Centre Tecnològic de Telecomunicacions de Catalunya (CTTC) predoctoral fellowships program. From October 2002 to February 2007, he was involved in several research projects in the field of wireless communications, such as MARQUIS included in the European program EUREKA-MEDEA+ or NEWCOM, funded by the Information Society Technologies program of the European Commission (EC). From September 2005 to February 2006 he held a visiting research appointment at the University of New South Wales (UNSW), Sydney, Australia. From February 2007 to December 2008 he held a postdoctoral position with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong. From January 2009, he is a Research Associate with the CTTC where, from January 2011 to March 2013, he led the Engineering Unit. In April 2013 he was appointed the Head of the Communications Technologies Division. In his second involvement with CTTC, he has participated in several research contracts with the industry as well as in research projects, such as BuNGee or BeFEMTO from the Seventh Framework Program of the EC. Currently, he is involved in Newcom#, where he is leading the experimental activities related to radio interfaces for 5G systems.