On the Design of Temporal Compression Strategies for Energy Harvesting Sensor Networks

Davide Zordan, Tommaso Melodia, *Member, IEEE*, and Michele Rossi, *Senior Member, IEEE*

***Abstract*—Recent advances in energy harvesting devices and low-power embedded systems are enabling energetically self-sustainable wireless sensing systems able to sense, process, and wirelessly transmit environmental data. In such systems, energy resources need to be judiciously allocated to processing and transmission tasks to guarantee high-fidelity reconstruction of the phenomenon under observation while keeping the system operational over extended periods of time. Within this context, this paper addresses the problem of designing efficient policies to control the task of lossy data compression for wireless transmission over fading channels in the presence of a stochastic energy input process and a replenishable energy buffer. As a first contribution, the transmission and energy dynamics of a sensor node implementing practical lossy compression methods are modeled as a constrained Markov decision problem (CMDP). Then, an algorithm is designed to derive optimal compression/transmission policies through a Lagrangian relaxation approach combined with a dichotomic search for the Lagrangian multiplier, while also obtaining theoretical results on the optimal policy structure. Furthermore, a thorough numerical evaluation of optimal and heuristic policies is conducted under different scenarios. Finally, the impact of practical operating conditions, including perfect versus delayed channel state information and power control, is evaluated.**

***Index Terms*—Energy harvesting wireless sensor networks, optimal control, protocol design, temporal compression.**

I. Introduction

ENERGY management in Wireless Sensor Networks (WSNs) has played a central role in recent years. Previous work has covered efficient designs in terms of channel access [1], data compression [2] and compressive data gathering and aggregation [3], [4]. However, the common objective of this body of work has typically been that of maximizing the network lifetime through efficient use of the available energy, considering non-replenishable batteries. Recent advances in energy harvesting devices and systems have opened up new opportunities for energetically self-sustainable (or energy-neutral) designs. In this case, optimal policies no longer need to be energy frugal through the entire lifetime of the nodes, but should intelligently assess when energy is available, delivering high performance when the energy income is abundant and putting the nodes into energy saving modes when it is scarce [5].

In the last few years, a vast literature has emerged on Energy Harvesting Sensor Networks (EHWSNs), attracting the interest of theoretical researchers and implementors. Most existing approaches have focused on the design of energy-neutral transmission policies [6]–[10]. These papers commonly assume some form of knowledge (either deterministic or statistical) of the energy and data arrival processes, which is then leveraged to derive optimal power allocation strategies and transmission schedules to either maximize the throughput or to minimize the transmission completion time. In [8], a single sensor node transmitting over a fading channel affected by additive Gaussian noise is considered. Energy arrivals and channel states are modeled as Poisson processes and are assumed to be known causally. Optimal transmission policies are then evaluated offline for static as well as fading channels. Online policies are also designed through a dynamic programming formulation. References [6] and [7] extend and generalize the results of [8] by relaxing the assumptions on packet arrivals and infinite size of the energy buffer, respectively. The authors of [9] present structural results on the optimal transmission policies in the presence of a Markovian energy input process, and assess the performance gap with respect to lightweight heuristics. Other approaches attempt to derive online adaptive optimal transmission policies for multi user scenarios [10], [11] by maximizing the long-term transmission rate, whilst satisfying energy and power constraints. Reference [10] uses Lyapunov optimization to devise online scheduling algorithms that achieve close-to-optimal routing performance for EHWSNs with time varying channels. Reference [11] generalizes [10], proposing utility-based and online routing policies that approach optimality as the network size increases.

Fewer papers are available on optimal lossy compression techniques. In [2], we studied the tradeoffs between reconstruction fidelity and compression/transmission energy for practical lossy compression schemes showing that, depending on the specific hardware architecture, compression makes it possible to trade some reconstruction accuracy (data fidelity) for substantial gains in terms of energy expenditure. Although energy harvesting is not considered in [2], the energy-fidelity relationships that were empirically derived in that paper are instrumental to the present one.

We remark that in EHWSNs the energy buffer dynamics allow for additional tradeoffs. Specifically, when nodes are

Manuscript received February 24, 2014; revised August 27, 2014, March 10, 2015, and September 17, 2015; accepted September 30, 2015. Date of publication October 9, 2015; date of current version February 8, 2016. This work was supported in part by the EU FP7 SWAP project, GA 251557, in part by the MOSAICS project funded by the University of Padova under Grant CPDA094077, and in part by the U.S. Office of Naval Research under Grant N00014-11-1-0848. The associate editor coordinating the review of this paper and approving it for publication was Guoliang Xue.

D. Zordan and M. Rossi are with the Department of Information Engineering, University of Padova, Padova 35131, Italy (e-mail: zordanda@dei.unipd.it; rossi@dei.unipd.it).

T. Melodia is with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115 USA (e-mail: melodia@ece.neu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2015.2489200

running out of energy, compression may help in lowering their energy consumption thanks to a reduced transmission time. Conversely, when the energy income is abundant, good policies will encourage the transmission of uncompressed data, so as to maximize the reconstruction fidelity. Hence, optimal policies in this case depend on energy buffer state, energy consumption (as caused by processing and communication) and on the dynamics of the energy harvesting process.

Previous papers that address the problem of designing optimal compression policies for EHWSNs are [12], [13]. The authors of [12] introduced the concept of *energy-neutral source-channel coding*, studying the relevant tradeoffs involved when sensor nodes allocate energy over source acquisition, compression and transmission. There, optimal and heuristic policies were analytically characterized and computed through convex optimization and dynamic programming techniques. Among other facts, the authors of [12] have shown that optimal policies imply the disjoint control of source and channel coding when the size of data and energy buffers is sufficiently large, whereas joint designs are needed for small buffers. The authors of [13] generalized [12] by considering finite buffer sizes and proposing a heuristic scheme having a long-term average distortion that approaches that of the optimal policy. The authors of [14] obtained information theoretic bounds (capacity laws) for EHWSNs for a variety of settings, including the one that we consider in this paper. Paper [15] extends [12] to a multihop WSN scenario. There, the Lyapunov-optimization technique with penalty functions was used to concoct distributed algorithms that meet quality of service requirements such as queue stability and distortion minimization. Differently from [12]–[15] our approach uses empirical curves that accurately capture these tradeoffs for real lossy compression schemes.

Our paper bears similarities with [12], [13]. Nevertheless, our present work differs in terms of data compression, system models, optimization objectives and results. In fact, we deal with realistic models for all system blocks and this somehow requires to give away some of the level of abstraction that was used in previous articles and adopt a different modeling approach. In return, our pragmatic design allows to plug any temporal compression scheme, given the empirical relationship governing its energy, rate and representation accuracy. For the channel and energy harvesting, we account for Markov-modulated processes where the channel is Rayleigh faded and power controlled, whereas the energy arrival process resembles of that of a real sun harvester, which was validated in [16]. For the temporal compression of time series, we consider the Lightweight Temporal Compression (LTC) scheme [17], that has been shown to provide excellent tradeoffs in terms of distortion vs energy consumption and, in turn, is one of the algorithms of choice for resource constrained sensor network technology, see [2]. Our optimality objective is the minimization of the long-term average distortion subject to the energetic sustainability of the sensor. Moreover, besides providing efficient algorithms for the numerical solution of the resulting constrained Markov decision process and obtaining optimal policies, we theoretically unveil their structure and study the impact of perfect vs delayed Channel State Information (CSI) at the transmitter and of power control over a multipath channel.

The contributions of our work are summarized as follows:

- We model, through a Constrained Markov Decision Problem (CMDP), the transmission and energy dynamics of a sensor node implementing the LTC compression method. As we shall see below, our approach is rather general and can be promptly extended to other compression schemes as well, provided that the relationship governing their distortion-energy tradeoff is known;
- We design an algorithm to derive optimal compression/transmission policies, using a Lagrangian relaxation approach combined with a dichotomic search[1] for the Lagrangian multiplier. These dynamically select the compression level to maximize the long-term data reconstruction fidelity at the receiver, while meeting stability constraints on the energy buffer;
- We derive a set of theoretical results on the structure of the optimal policy; specifically, we demonstrate that the optimal policy is non-decreasing in each of the system state components (energy buffer, channel and energy source states) under realistic assumptions. Moreover, the policy has a threshold structure that allows for its efficient storage in memory arrays;
- We conduct a thorough numerical evaluation of optimal as well as heuristic policies for different scenarios and evaluate the impact of perfect vs delayed CSI and power control.

The reminder of this paper is organized as follows. In Section II, we give a high-level view of the problem under analysis and identify the various system blocks. In Section III-A, we present our CMDP formulation, whereas in Section III-B we detail a model for a power controlled Rayleigh faded channel. In Section IV we propose an algorithm to numerically find the optimal policies and discuss the theorems unveiling their structure. In Section V, we validate the main theorems, and analyze various performance tradeoffs, the role of CSI and power control. In Section VI, we report our concluding remarks.

II. System Model

We consider a power controlled transmission scenario where a sensor node communicates wirelessly with a data collector (the sink). The sender is powered by a rechargeable battery (energy buffer) that stores the energy harvested from a renewable source, such as a solar panel. The task of the sensor is to sense some physical signal every T_{sens} seconds and report its measurements to the sink. Thus, we consider applications with a constant sampling rate of $1/T_{\text{sens}}$ samples/s. Also, we assume that the sensor node, in each data reporting window of duration T_{rep}, stores a certain number of measurements N_{m} in its memory buffer, where $T_{\text{rep}} = N_{\text{m}} T_{\text{sens}}$. At the end of the window, these readings are compressed into one or more packets of data and sent to a data collector. Note that, unlike [12], [13] the data queue stability is not an issue in our framework as exactly N_{m} readings per window are processed and the memory is always flushed at the end of each window of data. This amounts to assuming that the source rate is smaller than the

[1]This algorithm halves the search interval at each iteration as in a divide and conquer technique.

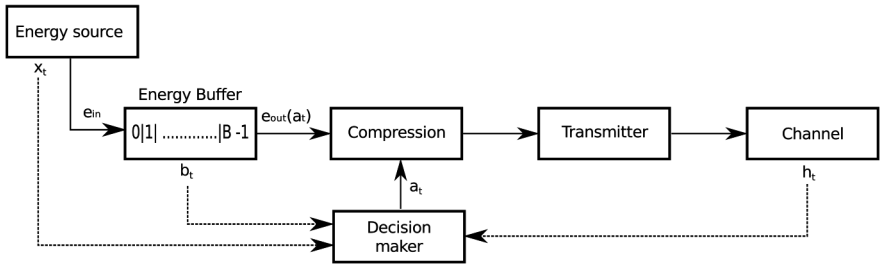


Fig. 1. Sensor Node diagram.

channel rate (i.e., that the source is slow-varying). This model is a simplification with respect to the analysis in [12], [13] but is accounted here to capture the complex energy-distortion trade-offs of practical lossy compression schemes through empirical curves, see Section III-A. Also, this assumption is not restrictive for typical WSN applications where, due to energy efficiency considerations, channel resources are sparingly used.

For the use of compression, note that most physical signals exhibit strong temporal correlation. Thus, it makes sense to compress the N_{m} samples within each reporting window through a *lossy* compression method, which allows trading accuracy in the data reconstruction at the sink for energy savings at the transmitter (fewer bits to transmit). Power control is instead exploited at the sender to keep the probability of successful reception around a predetermined target value, see Section III-B.

For the signal model, we consider stationary and ergodic time series $z(k)$ with given mean, variance and temporal correlation structure. To emulate the behavior of real WSN signals, a noise component is superimposed to $z(k)$, so as to mimic account for perturbations caused for example by the limited precision of the sensing hardware. The noise is modeled as a zero mean white Gaussian process with variance $\sigma^2_{\mathrm{noise}}$. Hence, the average Signal to Noise Ratio of the resulting signal is $SNR = E[z(k)^2]/\sigma^2_{\mathrm{noise}}$. We use the following correlation model: let $\rho_z(n)$, $n \geq 0$, be the autocorrelation function associated with $z(k)$. We define correlation length of $z(k)$ as the smallest value $n^\star$ for which $\rho_z(n)$ is smaller than a predetermined threshold ρ_{th}. Formally: $n^\star = \mathrm{argmin}_{n\geq 0}\{\rho_z(n) < \rho_{\mathrm{th}}\}$. More details on the selection of these parameters are given in Section V.

Now, consider the diagram in Fig. 1. We assume that the sensor operates in slotted time $t = 0, 1, 2, \ldots$ and identify the following elements:

- **Energy Source**. This block models the energy source. The output of this block is e_{in}, i.e., the number of energy quanta that are harvested in a time slot (the energy inflow and the energy cost for compression and transmission have been quantized for computational efficiency, see also [9]). The source dynamics are tracked through a two-state Markov model, see [16], where x_t is the source state in the current slot t. In detail, $x_t = 0$ represents a low energy state (e.g., "night") and $x_t = 1$ represents a high energy state ("day"). Additional details on the source model are given in Section III.
- **Energy buffer**. The energy harvested by the Energy Source block is stored in the energy buffer, which could be a rechargeable battery, a supercapacitor or any other energy storage device. The state of the buffer is indicated as b_t. e_{out} energy quanta are drawn from the buffer in each time slot. Note that e_{out} depends on the compression ratio adopted by the sensor in the current time slot and on the channel state (through power control as discussed in detail in Section III-B).
- **Compression**. This block describes the compression mechanism, linking the energy consumption, the size of the output packet and the representation accuracy of the received data to the action taken by the sensor. In this paper, we consider the Lightweight Temporal Compression (LTC) method, whose compression, energy and error figures are evaluated in [2]. LTC compresses a time series by creating an approximate representation through line segments. The data to be sent by the transmitter are then the parameters of this representation, i.e., the end points of the line segments. Note that we consider the energy spent in the sensing process to be negligible with respect to that spent for compression and transmission [18]. This is in fact the case for most of the sensors involved in environmental monitoring application or wearable IoT devices. If the energy consumed for the sensing process is too high, one may adopt adaptive sampling techniques which are however not considered in the present work.
- **Transmitter**. This block accounts for the energy consumption associated with data transmission, which depends on the size of the packet to be transmitted (and hence on the compression ratio) and on the radio transceiver transmission power. The latter, in turn, depends on the channel state as will be discussed in Section III-B.
- **Channel**. The channel block models channel variations in time. The channel evolution is tracked through a Markov chain, where the channel state is indicated with h_t, see Section III-B.
- **Decision Maker**. This block takes as input the three components of the system state, namely b_t, x_t and h_t and makes a decision on the current compression ratio a_t, according to a certain quality criterion (see Section III).

III. Problem Formulation

In the following Section III-A, we formulate the joint data transmission and compression problem as a Constrained Markov Decision Process (CMDP). We will then use a

TABLE I
NOTATION AND SYMBOLS

SYMBOL	MEANING
$t = 0, 1, 2, \dots$	discrete time index
$\mathcal{S}$, $\mathcal{A}$	system state and action sets
$s_t = [b_t, x_t, h_t] \in \mathcal{S}$	system state in slot t
$b_t \in \mathcal{B} = \{0, 1, \dots, B\}$	energy buffer state in slot t
$x_t \in \mathcal{X} = \{0, 1\}$	energy source state in slot t
$h_t \in \mathcal{H} = \{1, \dots, H\}$	channel state in slot t
$p(s_{t+1}\|s_t, a_t)$	transition prob. from s_t to s_{t+1} for action a_t
$a_t = \{0, 1, \dots, n\} \in \mathcal{A}$	action taken (compression level) in slot t
$e_c(a_t)$	processing energy in slot t
$e_{tx}(a_t, h_t)$	transmission energy in slot t
$\mu_t : \mathcal{S} \to \mathcal{A}$	decision function $a_t \leftarrow \mu_t(s_t)$
$\pi = \{\mu_1, \mu_2, \dots\}$	compression policy
$v^\pi(s)$	expected average reward for initial state s and policy π

Lagrangian relaxation approach to deal with the energy buffer constraint, which assures that the node remains energetically self-sufficient, while delivering the best performance in terms of reconstruction fidelity at the data collector. In Section III-B, we present the adopted channel and power control model.

A. Constrained Markov Decision Process Formulation

The authors of [12], [13] use the following decision variables: the amount of energy to allocate to source and channel encoders and the distortion associated with the compressed data in each time slot t. The objective of their optimization is to keep the data queue stable, while assuring that the average distortion at the receiver remains above a certain minimum threshold. Here, we focus on practical compression schemes for which the queue stability is no longer an issue, in fact data is acquired at a constant rate and arrives in batches of fixed size at the compressor queue.[2] Further, we consider the sole compression ratio η_t as the decision variable and, given η_t, the amount of energy allocated for compression and transmission depend on the considered compression scheme (compression energy), power control and channel model (transmission energy). Thus, we obtain the compression policy that minimizes the average long-term distortion subject to the fact that the energy queue level shall never decrease below a predetermined threshold. We do so by modeling practical lossy compression schemes from the literature in terms of fidelity-energy tradeoffs, obtaining optimal policies for Markov fading channels and a Markov-modulated energy harvesting process in the presence of packet losses. Our pragmatic design methodology allows our framework to be quickly customized to different compression approaches or power control strategies.

The notation and symbols used in the following analysis are summarized in Table I. We consider a transmission system that operates in a time-slotted fashion, where $t = 0, 1, 2, \dots$ is the time index. The state space is denoted by $\mathcal{S}$ and the system evolves slot by slot according to the behavior of the energy source and of the channel, and according to the action taken by the considered node, i.e., the selected transmission/compression policy. We model the problem as a CMDP defined through the tuple $(\mathcal{S}, \mathcal{A}, P, r(\cdot), c(\cdot))$, where $\mathcal{S}$ represents the system state, $\mathcal{A}$ is the action state, P are the transition probabilities (representing the system dynamics), $r(\cdot)$ and $c(\cdot)$ represent a reward and a cost function, respectively. Both depend on the system state and on the action taken in the current time slot. Specifically, let $s \in \mathcal{S}$ be the system state at time t, with $s_t = s$. We refer to $a_t \in \mathcal{A}_s \subseteq \mathcal{A}$ as the action taken by the decision maker at time t, where $\mathcal{A}_s$ is a finite set that includes all feasible actions in state s.[3] The system evolution is Markovian with transition probabilities

$$p(j|i, a) = P[s_{t+1} = j | s_t = i, a_t = a] \ , \ i, j \in \mathcal{S} \ . \quad (1)$$

A decision rule is a function $\mu_t : \mathcal{S} \to \mathcal{A}$ that dictates the action to be taken when the system is in state s at decision epoch t. In our formulation, actions correspond to the degree of compression, as we further discuss below. A policy $\pi = (\mu_1, \mu_2, \dots)$ is a sequence of decision rules. Let Φ denote the set of all stationary policies $\pi = (\mu, \mu, \dots)$ where μ is a (possibly probabilistic) function of the current state s and does not depend on time t. The (finite) reward $r(s_t, a_t) \geq 0$ is the instantaneous payoff of taking action a_t in state s_t. The notion of reward is used to quantify the reconstruction fidelity of the compressed signal. For any admissible policy $\pi \in \Phi$, the expected average reward is

$$v^\pi(s) = \lim_{N \to \infty} \frac{1}{N} \mathbb{E}_s^\pi \left[\sum_{t=0}^{N-1} r(s_t, a_t) \middle| s_0 = s \right] . \quad (2)$$

Our goal is to compute the optimal policy $\pi^\star$ that maximizes the expected average reward (2)

$$v^{\pi^\star}(s) = \max_{\pi \in \Phi} v^\pi(s) \ , \ \forall s \in \mathcal{S} \quad (3)$$

subject to the global constraint

$$C^{\pi^\star}(s) = \lim_{N \to \infty} \frac{1}{N} \mathbb{E}_s^{\pi^\star} \left[\sum_{t=0}^{N-1} c(s_t, a_t) \middle| s_0 = s \right] \leq C_{\max} \quad (4)$$

Here, $c(s_t, a_t) \geq 0$ is the (known) finite cost incurred in slot t, which is related to the energy buffer state, whereas $C_{\max} \geq 0$ is a user defined parameter. Constraint $C_{\max}$ is termed *feasible* when there exists at least one policy $\pi \in \Phi$ that satisfies (4).

For pure policies the decision μ is deterministic in the current state. As we shall discuss in Section IV, the optimal policy, i.e., the policy that maximizes the long-term average reward (2), while satisfying (4) with equality, is in general a mixture of two pure policies [19].

Next, we specify each element of the tuple $(\mathcal{S}, \mathcal{A}, P, r(\cdot), c(\cdot))$ for the problem under analysis.

1) States: Each state $s \in \mathcal{S}$ consists of three components $s = [b, x, h]$, where $\mathcal{S} = \mathcal{B} \times \mathcal{X} \times \mathcal{H}$. $\mathcal{B}$ denotes the energy buffer state space, $\mathcal{X}$ denotes the state space of the energy

[2] Although this is a simplification with respect to a variable queue length, this allows us to model the relations between fidelity and energy consumption for practical and complex temporal compression algorithms, see Section II.

[3] At time t, an action a_t is termed infeasible if taking it would imply a negative energy buffer in the next time slot.

income process, and $\mathcal{H}$ denotes the channel state space. $b \in \mathcal{B}$ is the energy buffer state, $\mathcal{B} = \{0, 1, \ldots, B\}$ contains all possible energy buffer levels, where the available energy has been discretized into $B+1$ energy quanta. $x \in \mathcal{X} = \{0, 1\}$ is the energy source state (assumed to be observable). In detail, a certain amount of energy can be harvested in each time slot t. When $x = 0$, the energy source is in a "bad" state, where the amount of energy quanta harvested from the source in slot t is $e_{\text{in}} = 0$ w.p. 1. When $x = 1$, a number $e_{\text{in}} \in \mathcal{E}^{(1)} = \{1, \ldots, E\}$ of energy quanta is harvested in the current slot t, where e_{in} is distributed according to some mass distribution function (mdf). Hence, the evolution of the energy source state is represented through a two state Markov chain, where E represents the maximum number of energy quanta that can enter the energy buffer in a time slot.

The channel state $h \in \mathcal{H} = \{1, \ldots, H\}$ affects the reception probability at the receiver and, since we consider a power controlled transmission system, it also affects the power consumption at the transmitter. We assume h to be independent of x and b and we model it as a Markov chain. The procedure used to derive the channel state transition matrix for a Rayleigh channel model is detailed in Section III-B. Optimal policies are computed in the case where the transmitter has perfect CSI, where the CSI at the transmitter is obtained with a certain delay (delayed CSI) and where the transmitter has no CSI at all.

2) Actions: At each decision epoch t the transmitter (decision maker) observes the system state $s = s_t$ and chooses an action $a = a_t$ from the action set $\mathcal{A}_s$. In our model, a_t corresponds to the compression ratio η_t to be used by the transmitter in the current slot t. Here, η_t is defined as the ratio of the size of the compressed signal with respect to that of its uncompressed version. For the action set, we have $\mathcal{A}_s \subseteq \{0, 1, \ldots, n\}$. In particular, when $a_t = 0$ the transmitter will be silent in the current transmission slot t. For $0 < a_t < n$, the transmitter will compress the data stored in the memory buffer with a compression ratio $\eta_t = a_t/n$ and it will thus send the compressed data to the collector. When $a_t = n$, the transmitter will send the data without performing any compression ($\eta_t = 1$).

3) Transition Probabilities: Let t and $s_t = [b_t, x_t, h_t]$ represent the current time index and the system state in slot t, respectively. We indicate with a_t the action taken in the current slot t and we refer to $e_{\text{out}}(s_t, a_t)$ as the number of energy quanta that are taken from the buffer (i.e., the energy consumption) given that action a_t is chosen. Hence, the energy buffer evolves as:

$$\begin{aligned} b_{t+1} &= \max\{0, \min\{b_t + e_{\text{in}} - e_{\text{out}}(s_t, a_t), B\}\} \\ &\triangleq [b_t + e_{\text{in}} - e_{\text{out}}(s_t, a_t)]^{\dagger} . \end{aligned} \tag{5}$$

If s_t is the system state, action a_t is admissible only if $e_{\text{out}}(s_t, a_t) \leq b_t$. Further, $e_{\text{out}}(s_t, a_t)$ is given by the sum of two components: the energy consumption associated with the compression task $e_{\text{c}}(a_t)$ and that associated with the transmission task $e_{\text{tx}}(a_t, h_t)$.[4] In the present analysis, as a compression technique we select the Lightweight Temporal Compression (LTC) algorithm, since it strikes a good balance in terms of compression vs energy consumption. As shown in [2], for LTC the processing energy $e_{\text{c}}(a_t)$ is related to the compression ratio $\eta_t = a_t/n$ through the linear relationship:

$$e_{\text{c}}(a_t) = \begin{cases} 0 & a_t \in \{0, n\} \\ \left(\kappa \frac{a_t}{n} + \ell\right) N_b E_0 & 0 < a_t < n \ . \end{cases} \tag{6}$$

Note that when $a_t \in \{0, n\}$ the node does not perform any compression, thus we have $e_{\text{c}}(0) = 0$. In (6), N_b is the number of bits to be compressed, that is defined as $N_b = N_{\text{m}} n_b$, where n_b is the number of bits used to represent a single measurement and N_{m} represents the number of measurements, to be sent every T_{rep} seconds. E_0 is the energy consumption of the micro controller in one clock cycle, $\kappa = 16.1$ and $\ell = 105.4$ are two fitting coefficients, see [2]. An observation is in order. One might expect that the compression cost decreases with a_t, which is not the case for (6). This is caused by practical implementation details. In fact, as $\eta_t = a_t/n$ decreases (higher compression), LTC operates using a larger error tolerance, which results in a lower number of operations. For For further details see [2]. Note that the results that we discuss in this paper are rather general and can be readily extended to other compression approaches, i.e., by just replacing (6) with a appropriate function.

The energy consumption $e_{\text{tx}}(a_t, h_t)$ depends on the number of bits N_b to be transmitted, on the action a_t and on the channel state h_t:

$$e_{\text{tx}}(a_t, h_t) = \frac{a_t}{n} N_b E_{\text{tx}}(h_t) \ , \tag{7}$$

where $E_{\text{tx}}(h_t)$ is the energy consumption associated with the transmission of one bit, which depends on the specific radio technology and on the channel state h_t (through the power control mechanism).

The transition probability from state $s_t = [b_t, x_t, h_t]$ to state $s_{t+1} = [b_{t+1}, x_{t+1}, h_{t+1}]$ given that action a_t is selected is:

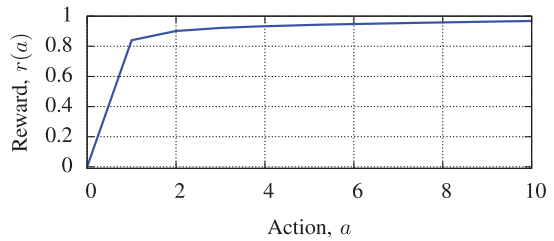
$$\begin{aligned} p(s_{t+1}|s_t, a_t) = \delta\left(b_{t+1} - [b_t + e_{\text{in}} - e_{\text{out}}(s_t, a_t)]^{\dagger}\right) \cdot \\ \cdot\, p_{e_{\text{in}}}(e_{\text{in}}|x_t) \cdot p_x(x_{t+1}|x_t) \cdot p_h(h_{t+1}|h_t) \ , \end{aligned} \tag{8}$$

where $\delta(\cdot)$ is the indicator function (equal to 1 if the argument is zero and null otherwise), $p_{e_{\text{in}}}(e_{\text{in}}|x_t)$ is the mdf of the input energy in state x_t, while $p_x(x_{t+1}|x_t)$ and $p_h(h_{t+1}|h_t)$ are respectively obtained from the transition probability matrices of the energy source and the channel. We remark that our framework can be promptly adapted to any Markov-modulated energy arrival and channel processes by respectively changing the transition probabilities $p_{e_{\text{in}}}(e_{\text{in}}|x_t)$ and $p_h(h_{t+1}|h_t)$.

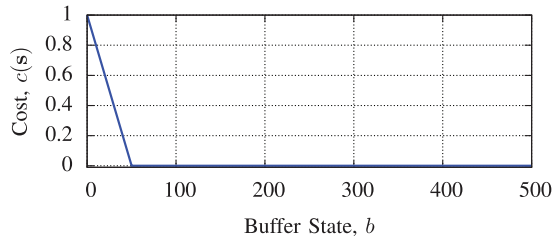
4) Reward Function: To maximize the reconstruction fidelity at the data collector, the reward function $r(a_t)$ is chosen to be a strictly increasing function of the selected compression ratio $\eta_t = a_t/n$. The reward is zero if the selected action is $a_t = 0$ and reaches one when $a_t = n$. We define $r(a_t)$ as

$$r(a_t) = \begin{cases} 0 & a_t = 0 \\ 1 - \left(\frac{p_1 (a_t/n)^2 + p_2 a_t/n + p_3}{a_t/n + q_1}\right) \sigma^2_{\text{noise}} & a_t \neq 0, \end{cases} \tag{9}$$

[4] In this work, the sampling rate is fixed. A non-negligible sensing cost can be accounted for by adding the related (constant) energy consumption e_{s} to the energy budget, i.e., $e_{\text{out}}(s_t, a_t) = e_{\text{c}}(a_t) + e_{\text{tx}}(a_t, h_t) + e_{\text{s}}$.



(a) Reward function $r(a)$ for $n = 10$.



(b) Cost function $c(b)$ for $b_{\text{th}} = 50$ and $B = 500$.

Fig. 2. Example of reward and cost functions.

where the polynomial function in (9) is the LTC rate distortion curve obtained in [2], which relates the error in the reconstruction process to the compression ratio η_t. Here, $p_1 = -0.870$, $p_2 = 1.469$, $p_3 = 0.191$ and $q_1 = -0.017$ are suitable constants, that in general depend on the temporal correlation exhibited by the signal, whereas σ^2_{noise} is the variance of the white noise that is superimposed to the sensed signal. Different compression schemes and correlations can be considered in this framework by substituting (9) with the appropriate rate-distortion function.

5) Cost Function: Since we want to prevent depletion of the energy buffer, we impose a positive cost for those states where the energy buffer level is below a pre-determined threshold b_{th} (i.e., a soft-constraint on the energy buffer). Also, the cost increases linearly when b decreases below b_{th}. We underline that b_{th} in general depends on application requirements (such as the minimum energy reserve to react to unexpected and important events) and on hardware constraints (a positive but too low charge may not be sufficient to guarantee the correct operation of the sensor). This extends the models usually considered in the related literature, see, e.g., [9], making the analysis of optimal policies more convoluted but, at the same time, allowing for more practical designs.

Hence, at time t, the cost is defined as:

$$c(b_t) = \begin{cases} 0 & b_t \geq b_{\text{th}} \\ \dfrac{b_{\text{th}} - b}{b_{\text{th}}} & b_t < b_{\text{th}} \end{cases} . \tag{10}$$

Note that in our case $r(s_t, a_t)$ only depends on action a_t and cost $c(s_t, a_t)$ only depends on the state component b_t, see Fig. 2. These facts will be used in Section IV-B to characterize the structure of the optimal policies.

B. Channel Model

Next, we present an H-state Markovian model for a power controlled transmission link in the presence of Rayleigh fading. With respect to standard modeling approaches [20], [21], here we partition the Signal to Noise Ratio (SNR) space according to power control via channel inversion, so that a minimum Quality of Service (QoS) level is granted for each channel state, whenever possible. Specifically, we consider that the node adapts its transmission power according to H (radio specific) power levels $P_{\text{tx}}[i]$, with $i \in \mathcal{H} = \{1, \ldots, H\}$ and $P_{\text{tx}}[1] > P_{\text{tx}}[2] > \cdots > P_{\text{tx}}[H]$. A specific power level is assigned to each channel state to assure that the packet error probability P_{pkt} remains smaller than or equal to a given target QoS parameter ζ, i.e., $P_{\text{pkt}} \leq \zeta$ for states $i > 1$. This amounts to identifying $H + 1$ thresholds for the fading power. For a certain SNR γ and compression level a, the packet error rate is:

$$P_{\text{pkt}}(\gamma, a) = 1 - (1 - P_{\text{bit}}(\gamma))^{L(a)} , \tag{11}$$

where $L(a)$ is the packet size expressed in bits and $P_{\text{bit}}(\gamma)$ is the bit error rate formula. Note that the packet size depends on the level of compression $a \in \{0, 1, \ldots, n\}$ in the current time slot, i.e., $L(a) = aL_{\max}/n$, where $L_{\max}$ is the maximum packet size expressed in bits.[5] Assuming a $\pi/4$-DQPSK modulation, the bit error rate is approximated as:[6]

$$P_{\text{bit}}(\gamma) = \frac{4}{3}\text{erfc}\left(\sqrt{\gamma}\right) , \tag{12}$$

where "erfc" is the complementary error function. From (11) and (12), assuming $L(a) = L_{\max}$ (i.e., $a = n$), the SNR threshold γ_{th} corresponding to the target packet error probability ζ is:

$$\gamma_{\text{th}} = \left(\text{erfc}^{-1}\left(\frac{3}{4}\left(1 - (1 - \zeta)^{1/L_{\max}}\right)\right)\right)^2 . \tag{13}$$

Thus, the QoS requirement $P_{\text{pkt}}(\gamma, a) \leq \zeta$, with $a \in \{1, \ldots, n\}$ corresponds to requiring that $\gamma \geq \gamma_{\text{th}}$. For state $i \in \mathcal{H}$, the SNR γ at the receiver in slot t is $\gamma(t) = \gamma_0[i]\alpha(t)$, where $\gamma_0[i] = P^0_{\text{rx}}[i]/(N_0 B)$ is the average SNR, which depends on the path loss and on the noise power $N_0 B$ (N_0 is the power spectral density and B is the transmission bandwidth), whereas $\alpha(t)$ is the fading power. The average received power $P^0_{\text{rx}}[i]$ for transmission level i is obtained as: $\left[P^0_{\text{rx}}[i]\right]_{\text{dB}} = [P_{\text{tx}}[i]]_{\text{dB}} + [G]_{\text{dB}} - [A]_{\text{dB}} - [\mathcal{PL}]_{\text{dB}}$, where $[G]_{\text{dB}}$ and $[A]_{\text{dB}}$ respectively represent the total antenna gains (transmitter and receiver) and the total attenuation losses (transmitter and receiver), whereas $[\mathcal{PL}]_{\text{dB}}$ is the path loss attenuation, expressed in dB, that also depends on the distance d between the transmitter and the receiver, see, e.g., Chapter 2 of [22]. The fading thresholds $\alpha[0] \leq \alpha[1] \leq \cdots \leq \alpha[H]$ are evaluated as follows. In state $i \in \{2, \ldots, H\}$, the transmission power is $P_{\text{tx}}[i]$ and the fading is distributed in $[\alpha[i-1], \alpha[i]]$. From the condition $\gamma \geq \gamma_{\text{th}}$, for state i we must have that:

$$[\alpha[i-1]]_{\text{dB}} \geq [\gamma_{\text{th}}]_{\text{dB}} - [\gamma_0[i]]_{\text{dB}} , \tag{14}$$

[5]Taking the IEEE 802.15.4 standard as a reference, in this paper we consider $L_{\max} = 127$ bytes.

[6]Here, we do not consider coding as current standards for resource constrained sensor nodes, e.g., IEEE 802.15.4, do not use it. However, the impact of techniques such as coding and direct sequence spread spectrum can be assessed by modifying (12) and does not require any other change in the analysis.

using equality in the previous equation returns the lower edge of the interval $[\alpha[i-1], \alpha[i]]$. Moreover, we have $\alpha[0] = 0$ and $\alpha[H] = +\infty$. Note that 1 is the only channel state for which the QoS requirement $P_{\text{pkt}} \leq \zeta$ cannot be met. Because of this, for all states $s = [b, x, h]$ with $h = 1$, we set $\mathcal{A}_s = \{0\}$ to avoid degenerative situations where it is convenient to transmit in this state, as the reward does not deteriorate. Thus, the SNR interval associated with state $i \in \{1, \ldots, H\}$ is $[\alpha[i-1], \alpha[i]]$ and power $P[i]$ is used when the channel is in that state. To fully characterize the Markov channel, the transition probabilities from state i to state j, $p_h(j|i)$, are derived as in [20], [21] through a double integration over the respective SNR intervals, normalized to the steady state probability associated with state i.

IV. Optimal Policies

The unconstrained problem (3) can be solved by first writing the Bellman's optimality equation

$$v(s) = \max_{a \in \mathcal{A}_s} \left\{ r(s,a) + \sum_{s' \in \mathcal{S}} p\left(s'|s,a\right) v(s') \right\} , \tag{15}$$

where the corresponding optimal policy is given by

$$a^\star(s) = \operatorname*{argmax}_{a \in \mathcal{A}_s} \left\{ r(s,a) + \sum_{s' \in \mathcal{S}} p\left(s'|s,a\right) v(s') \right\} . \tag{16}$$

The optimal average reward $v(s)$ can be obtained through, e.g., the Value Iteration algorithm [23], whereas $a^\star(s)$ indicates the mapping from s to the optimal action, i.e., $\mu : \mathcal{S} \to \mathcal{A}$. When the constraint (4) is added to the problem, a Lagrangian relaxation approach can be used to convert the constrained formulation (see (3) and (4)) into an equivalent unconstrained MDP. This is achieved by defining a new reward function, $r(s,a;\beta)$, through the Lagrangian multiplier $\beta > 0$

$$r(s,a;\beta) = r(s,a) - \beta c(s,a) . \tag{17}$$

The new optimality equation is thus

$$v_\beta(s) = \max_{a \in \mathcal{A}_s} \left\{ r(s,a;\beta) + \sum_{s' \in \mathcal{S}} p\left(s'|s,a\right) v_\beta(s') \right\} , \tag{18}$$

that for any fixed value of β can be solved via Value Iteration. Theorem 12.7 of [19] proves the existence of an optimal policy and also the fact that the Lagrangian formulation (18) can be solved through a search on the parameter β as we now explain.

In detail, we need to find the optimal Lagrangian multiplier $\beta^\star$ for which constraint (4) is satisfied with equality. However, in CMDP with finite state and action sets, we have a finite number of feasible policies; a single pure policy that satisfies the constraint with equality, with $\beta^\star$ as its associated multiplier, may not exist. Instead, the optimal policy for this class of CMDP is a mixture of two policies, a first policy π^- that satisfies the constraint (being as close as possible from below), and a second one, π^+, that does not satisfy the constraint (but again having a cost that is the closest to $C_{\max}$ from above). By mixing these two policies through an appropriately chosen weight parameter $q \in [0,1]$, the constraint can be satisfied with equality [24]. In our paper, as in [25], [26], we apply a Q-learning algorithm to determine the appropriate Lagrangian multiplier for a given $C_{\max}$. This algorithm finds the two policies π^- and π^+ from which the optimal policy π^* is obtained. Specifically, π^- and π^+ are two pure and stationary policies [19], whose actions differ in a single state. To obtain these policies, we iteratively solve the constrained problem for two values of β, namely β^- and β^+, leading to policies π^- and π^+, respectively. The values of β^- and β^+ are iteratively updated to get closer to $\beta^\star$, as discussed in detail in Algorithm 2 of the next section. We denote the costs of the two policies π^- and π^+ by $C^- \leq C_{\max}$ and $C^+ \geq C_{\max}$, respectively. When the algorithm stops, q (the weight parameter of the mixed policy) is computed by solving the following equalion:

$$C_{\max} = qC^- + (1-q)C^+ . \tag{19}$$

Hence, the optimal policy $\pi^\star$ in each state $s \in \mathcal{S}$ is given by:

$$\pi^\star(s) = \begin{cases} \pi^-(s) \text{ w.p. } q \\ \pi^+(s) \text{ w.p. } 1-q . \end{cases} \tag{20}$$

Note that when $C_{\max} = 0$ the probability that the buffer is emptied below b_{th} is zero (buffer outage), whereas an increasing $C_{\max}$ corresponds to higher buffer outage probabilities.

A. Numerical Solver

In this section we outline the algorithms used to solve the CMDP. Algorithms 1 is used to solve the unconstrained MDP obtained through Lagrangian relaxation for a given value of β. Here, we use (18) as an update rule to iteratively compute the optimal average reward $v_\beta(s)$.The convergence $v_\beta^n(s) \to v_\beta(s)$ is assured since state ($\mathcal{S}$) and action ($\mathcal{A}_s$) sets are finite. Also, the process is unichain, which derives from the fact that we have a single recurrent class for any possible policy. For unichain CMPD with bounded cost and reward functions and finite state and action sets, the limits in (2) and (4) exist and are well defined and this ensures that Value Iteration converges to the optimal solution [23]. Algorithm 2 implements a dichotomic search over the β parameter, as an outer loop. Note that this search strategy is effective because, as proven, e.g., in Lemmas 3.1 and 3.2 of [24], the optimal Lagrangian reward $v_\beta(s)$ is a uniformly absolutely continuous, monotone and non-increasing function of β. Previous algorithms implemented similar searches by starting with a small (large) initial value of β for which the cost constraint is not (is) satisfied and increase (decrease) it using suitable step sizes (that usually get smaller as the algorithm gets closer to the optimal solution, see, e.g., Eq. (3.9) of [25]). For each β, $v_\beta(s)$ is found through the Value Iteration of Algorithm 1 (as we do here), and the associated optimal policy π_β is obtained. The algorithm stops when the average cost gets sufficiently close to the chosen constraint $C_{\max}$, see also Algorithm 2 of [26]. The design of the (decreasing) step size is however critical. Our algorithm is instead more robust as it does not require any heuristic for the adaptation of the step size, which is automatically updated by the dichotomic

Algorithm 1. Value Iteration for a fixed value of β. The span seminorm operator $\mathrm{sp}(v)$ [27] is defined as: $\mathrm{sp}(v) \triangleq \max_{s\in\mathcal{S}} v(s) - \min_{s\in\mathcal{S}} v(s)$

Value Iteration Algorithm
select $v_\beta^0 \in V$, $\epsilon > 0$ and set $n = 0$;
repeat
  $v_\beta^{n+1}(s) = \max_{a\in\mathcal{A}_s} \left\{ r(s,a;\beta) + \sum_{s'\in\mathcal{S}} p(s'|s,a) v_\beta^n(j) \right\}$;
  $n = n + 1$;
until $\mathrm{sp}\left(v_\beta^{n+1} - v_\beta^n\right) < \epsilon$;
$a_\beta^\star(s) = \operatorname{argmax}_{a\in\mathcal{A}_s} \left\{ r(s,a;\beta) + \sum_{s'\in\mathcal{S}} p(s'|s,a) v_\beta^n(j) \right\}$;

Algorithm 2. Dichotomic Algorithm for the Lagrangian Multiplier Update

Lagrangian multiplier update
set $n = 0$, $\beta^- = 0$, β^+, $\beta^0 = \beta^-$;
repeat
  $\beta = \beta^n$;
  compute $\pi_\beta^\star(s)$ via Algorithm 1;
  compute the stationary distribution $\rho(s)$ induced by $\pi_\beta^\star(s)$;
  if $\sum_{s\in\mathcal{S}} \rho(s) c(s, a_\beta^\star(s)) > 0$ **then**
    $\beta^{n+1} = \dfrac{\beta^n + \beta^+}{2}$;
    $\beta^- = \beta^n$;
  else
    $\beta^{n+1} = \dfrac{\beta^n + \beta^-}{2}$;
    $\beta^+ = \beta^n$;
  end
  $n = n + 1$
until $|\beta^{n+1} - \beta^n| < \epsilon$;

adaptation rule $\beta^{n+1} \leftarrow f(\beta^n)$, where $f(\cdot)$ is a suitable function (see Algorithm 2). Also, provided that the initial value for β^+ is selected so that the average cost of the optimal policy $\pi_{\beta^+}^\star$ is larger than $C_{\max}$, Algorithm 2 is guaranteed to converge and this descends from the monotonic behavior of $v_\beta(s)$. Thus, it will always find a pair (β^+, β^-) that verifies the above properties.

B. Structure of the Optimal Policies

In this section, we present some results on the structure of the optimal policies arising from the optimization technique described in Section IV. For improved clarity, the proof of the theorems is given in the Appendix. Moreover, the properties that we discuss here are further elaborated in Section V where we present some numerical results.

Before delving into the description of the main theorems, in the following we introduce some useful definitions.

Definition 1 (Supermodularity): A function $f : \mathcal{X} \times \mathcal{A} \times \mathcal{L} \to \mathbb{R}$ is supermodular in $(x, a) \in \mathcal{X} \times \mathcal{A}$ for a fixed parameter $\beta \in \mathcal{L}$, if for all $x' \geq x$ and $a' \geq a$

$$f(x', a'; \beta) - f(x', a; \beta) \geq f(x, a'; \beta) - f(x, a; \beta) \ . \quad (21)$$

In words, supermodularity means that $f(x, a; \beta)$ has non-decreasing differences in (x, a), i.e., the difference of $f(\cdot)$ computed in $a' \geq a$ and a is larger for increasing values of the state x.

Definition 2 (First-order Stochastic Dominance): Let X_1 and X_2 be random variable with the same support $\mathcal{X}$. X_1 first-order dominates X_2, or $X_1 \succeq X_2$ if

$$F_{X_1}(x) \leq F_{X_2}(x) \quad (22)$$

for all $x \in \mathcal{X}$, where F_{X_i} is the cumulative distribution function (cdf) of X_i.

Definition 3 (Stochastically Increasing Family): Let $\{X_\theta\}_{\theta\in\mathbb{R}}$ be a family of random variables on the same support $\mathcal{X}$. $\{X_\theta\}$ is stochastically increasing if

$$X_{\theta'} \succeq X_\theta \quad (23)$$

whenever $\theta' \geq \theta$.

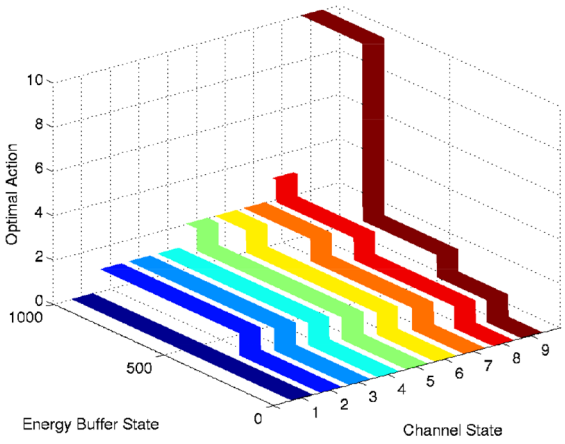
We now present some results on the structure of the optimal policy. Specifically, we prove that given some properties for the reward function and for the transition probabilities of the channel and the source states, the optimal policy has a three-dimensional threshold structure and is monotonically increasing in each system state component. This means that the policy has a staircase shape for each of the state variables b (energy buffer), x (energy influx) and h (channel state). In light of this, optimal policies can be efficiently stored in the node memory in the form of a lookup table with a few entries. The thresholds are the breaking points where the policy changes (see Section V for a graphical example) and are obtained through Algorithm 2. In each decision epoch then, the node just needs to evaluate its state and accordingly select the proper action from the lookup table.

The three theorems that follow prove that optimal policies (i.e., the corresponding optimal actions) are non-decreasing in the energy buffer state b, in the channel state h and in the energy source state x. This means that optimal policies are jointly non-increasing in the three state components and, in turn, have a staircase structure due to the quantized nature of states and actions. This is exemplified in Fig. 3 (discussed in Section V) and proves that optimal policies have a threshold structure.

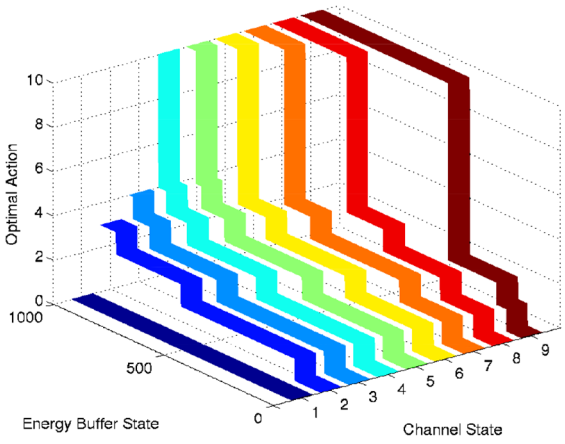
Theorem 1: Let the instantaneous Lagrangian reward function $r(s, a; \beta)$ be supermodular in the pair (b, a), concave and non-decreasing in b. Then, the optimal policy $\pi_\beta^\star$ is a monotone non-decreasing function of the buffer state b.

Theorem 2: Let the assumptions of Theorem 1 hold, let $r(s, a; \beta)$ be supermodular in the pair (h, a), and let $p_h(h'|h)$ be stochastically increasing in h. Then, the optimal policy $\pi_\beta^\star$ is a monotone non-decreasing function of the channel state h.

Theorem 3: Let the assumptions of Theorem 1 hold, let $r(s, a; \beta)$ be supermodular in the pair (x, a), and let $p_x(x'|x)$ be stochastically increasing in x. Then, the optimal policy $\pi_\beta^\star$ is a monotone non-decreasing function of the energy source state x.



(a) Night



(b) Day

Fig. 3. Structure of the optimal policy.

Discussion: We recall that, as discussed in Section III-A, for the considered scenario the reward $r(s,a)$ only depends on action a and the cost $c(s,a)$ only depends on the state component b. Given this, for the sake of clarity, and with a slight abuse of notation, reward and cost functions will be respectively written as $r(a)$ and $c(b)$. Also, since state s is defined through the tuple $[b,x,h]$, when checking the supermodularity considering one of such variables, the remaining two are kept fixed: the above theorems consider one variable at a time. For Theorem 1, the supermodularity of $r(s,a;\beta)$ in the pair (b,a) is a relatively weak assumption, in fact, if the reward function $r(a)$ is an increasing function of a and the cost function $c(b)$ is a decreasing function of b the assumption is satisfied. These are rather obvious choices for a model that has to optimize the representation accuracy of transmitted data ($r(a)$) giving a penalty for low energy buffer states ($c(b)$). For Theorem 2, the supermodularity of $r(s,a;\beta)$ in the pair (h,a) in our case is assured by the assumption of Theorem 1 and by the fact that $e_{\text{out}}(h,a)$ is monotonically increasing in h. This means that the energy consumption for a transmission when the channel is in a bad state is higher than that for a transmission over a good channel state. The assumption on $p_h(h'|h)$ to be stochastically increasing in h means that the probability of going to a state $h' \in \mathcal{H}$ is higher from a state $h_1 \in \mathcal{H}$ that is closer to h' than from any other state $h_2 \in \mathcal{H}$ such that $|h_1 - h'| < |h_2 - h'|$. This is true if the channel state transitions are assumed to be correlated, as in our case, but also if the channel states are assumed to be equally likely. For Theorem 3 the supermodularity of $r(s,a;\beta)$ in the pair (x,a) is assured by the assumption of Theorem 1 and by the fact that $e_{\text{out}}(h,a)$ is monotonically increasing in a. This means that the energy cost of transmitting a packet grows with the number of bits therein, which is realistic. Moreover, the assumption on $p_x(x'|x)$ to be stochastically increasing in x means that, for the energy source, for two consecutive time slots the probability of changing state is no higher than that of remaining in the current one. This is true in our case, were the two energy source states represent night and day, but it would also be true in the limiting case of equally-likely transition probabilities (i.e., when the process is iid).

V. Results

In this section, we discuss the performance of the optimal policies obtained as discussed in detail in Section III. We first look at their structure, validating what predicted by Theorems 1, 2 and 3. After that, we present a thorough performance evaluation for different energy budget scenarios, highlighting the relationships among the system parameters and the average reconstruction fidelity at the sink. Moreover, we show how the performance scales when the assumption of perfect CSI is removed, i.e., when the channel state information is retrieved with a certain delay (delayed CSI). Finally, we describe two heuristic policies, comparing them against the optimal solution and investigating the impact of power control. In the results that we discuss in this paper, we use $C_{\max} = 0$ so that $\pi^\star$ always avoids actions that lead to buffer outage events. Since the cost in our model is non negative and $C_{\max} = 0$, the weight parameter q is always equal to 0 and we always take π^- as the optimal policy. For the power control we have set a target packet error probability of $\zeta = 0.001$. Also, $N_{\text{m}} = 60$ measurements with $n_b = 16$ bits and $T_{\text{sens}} = 30$ s, which gives $T_{\text{rep}} = N_{\text{m}} T_{\text{sens}} = 30$ minutes. For the following plots, we use environmental signals from [28], which are collected from Global Historical Climatology Network and Legates and Willmott's meteorological stations of air temperature and precipitation. These, for the considered sampling time T_{sens} and threshold $\rho_{\text{th}} = 0.05$ have a correlation length $n^\star$ ranging between 100 and 500 samples. As shown in [2], the compression performance of LTC is nearly constant within that interval.

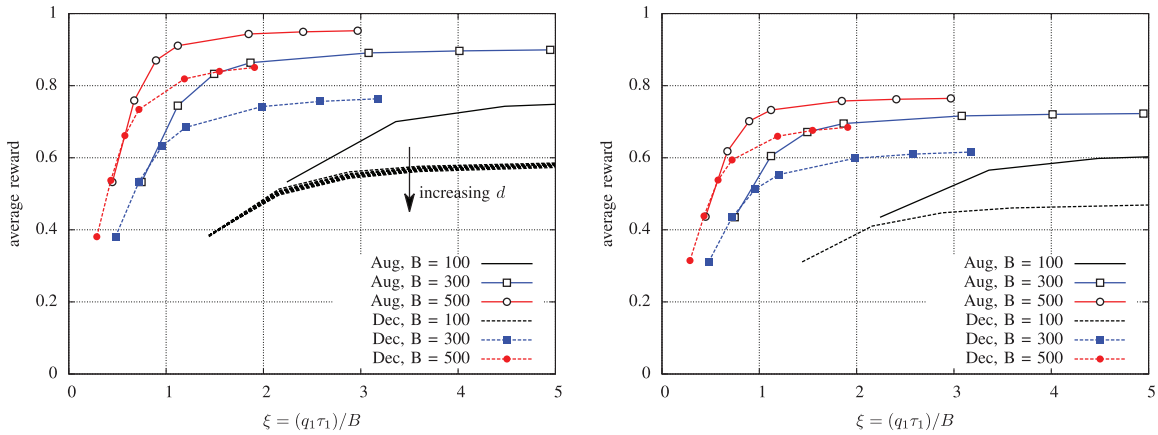
A. Optimal Policies: Structural Results

Theorem 1 states that the optimal policy is non-decreasing in the energy buffer state b. For this result to hold, we need the Lagrangian reward function $r(s,a;\beta)$ to be:

A1) supermodular in the pair (b,a);

A2) concave and non decreasing in b.

From (17), we have that the reward function only depends on a, whereas from (10) it descends that the cost function only depends on b. From these facts and $\beta > 0$, it follows that the Lagrangian reward function $r(s,a;\beta)$ satisfies A1 and A2.



(a) Perfect CSI. (b) Delayed CSI ($T = 3$ time slots).

Fig. 4. Average reconstruction fidelity at the sink vs ξ (see (24) and (25)). $\tau_1 \in \{10, 14\}$ hours, $B \in \{100, 300, 500\}$ energy quanta.

Theorem 2 states that the optimal policy is also non-decreasing in the channel state component of the system state, provided that:

A3) the family of transition probabilities $p_h(h'|h)$ is stochastically increasing in h;

A4) $r(s, a; \beta)$ is supermodular in the pair (h, a),

With our channel modeling technique of Section III-B assumption A3 is verified. Besides that, since the reward only depends on s through its buffer component b, condition A4 is also met.

In order for Theorem 3 to hold, we need the following assumptions to be verified:

A5) $p_x(x'|x)$ has to be stochastically increasing in x;

A6) the energy income in states x_i has to take values in disjoint ordered sets;

A7) $r(s, a; \beta)$ must be supermodular in the pair (x, a).

Here, we model the energy source through a Markov chain with transition probabilities that verify A5. Moreover, we deterministically set the energy income in state $x = 0$ to 0, while the energy income in state $x = 1$ is distributed according to a truncated Gaussian r.v. that takes values in $[1, E]$: hence, the assumption A6 is also verified. Finally, A7 is also met based on arguments similar to those discussed for Theorem 1 above.

In Fig. 3, we show an example policy for $B = 1000$ energy quanta, where the action space is quantized according to $\mathcal{A}_s = \{0, 1, \ldots, n\}$ with $n = 10$ and there are nine channel states ($H = 9$). From this plot we see that the results of the theorems hold. In detail, the optimal policy is monotonically non decreasing in the buffer state (i.e., along each stripe in the figures), in the channel state (across different stripes in the same figure) and in the source state (see any two stripes, for the same channel state, in Figs. 3a and 3b).

B. Optimal Policies: Performance Evaluation

In this section, we discuss the performance of the optimal policies that have been numerically obtained utilizing the algorithms of Section IV-A. Different scenarios have been considered, varying the energy buffer size, the average energy income in state $x = 1$, the transition probabilities for the energy source and the distance d between the sensor node and the sink. These settings have been summarized introducing the new variable ξ, which describes the average energy income for each set of parameters. Specifically, ξ is defined as $\xi = q_1\tau_1/B$, where q_1 is the average energy income in the "good" state $x = 1$, τ_1 represents the average time the energy source model stays in $x = 1$ before moving to $x = 0$ and B is the energy buffer size. We start by investigating the average reward (reconstruction fidelity) that is earned at the sink by the optimal policy, which is evaluated as

$$R \stackrel{\text{def}}{=} \sum_{s \in \mathcal{S}} \rho(s) r(s, a^\star(s)) P_{\text{suc}}(s, a^\star(s)) \,, \tag{24}$$

where $\rho(s)$ is the steady state probability distribution induced by the optimal policy $\pi^\star$, $r(s, a^\star(s))$ is the immediate reward accrued by $\pi^\star$ in state $s = [b, x, h]$, see (9), and $P_{\text{suc}}(s, a^\star(s)) = E[(1 - P_{\text{pkt}}(\gamma, a^\star(s)))]$ (see (11)) is the average probability that a packet compressed according to the optimal action $a^\star(s)$ is successfully received at the sink, when the channel is in state $h \in \mathcal{H}$ (the expectation is taken for the SNR γ in the interval $[\gamma_0[h]\alpha[h-1], \gamma_0[h]\alpha[h]]$, which corresponds to the SNR range associated with state h, see Section III-B). Note that compressing the packet provides an additional gain in terms of $P_{\text{suc}}(s, a^\star(s))$, as smaller packet sizes result in smaller packet error probabilities. Hence, the average reconstruction fidelity at the sink, evaluated through (24), benefits from a higher degree of compression when the SNR is low. Conversely, for high SNR values, a higher fidelity is obtained for $a = 1$ as in this case the dominating term in (24) is the immediate reward $r(s, a^\star(s))$.

Figure 4 shows the results for the average reward R for different values of ξ. In particular, we considered two different scenarios for the average duration τ_1 of the "good" energy state $x = 1$, picking $\tau_1 = 14$ hours and $\tau_1 = 10$ hours. These parameters are taken from [16], where a Markovian model for the statistical description of the energy harvested by outdoor micro-solar panels for WSN applications is presented. According to the results in this paper, $\tau_1 = 14$ and $\tau_1 = 10$ respectively correspond to the average duration of sunlight in a day for the city of Los Angeles in the months of August and December. For the energy buffer size, we considered three values, i.e., $B \in$

{100, 300, 500} energy quanta. Finally, for each pair (τ_1, B), the average energy income in a time slot q_1 is varied in the set {8, 12, 16, 20, 33, 43, 53} energy quanta. Also, in Figs. 4a and b we plot R for the two cases of perfect and delayed-CSI at the transmitter, respectively.

From Fig. 4, we see that for any given pair (τ_1, B), the average reward R increases with ξ. This is due to the higher average energy income q_1, which allows the system to reach the end of a "good" period $x = 1$ (i.e., day) with a higher residual level of energy in the energy buffer which, in turn, permits to additionally transmit a certain amount of data during the "bad" state (night). R is also monotonically increasing in B and τ_1. Again, this happens because a larger energy buffer (higher B) or a longer duration of the "good" state (higher τ_1) both result in higher energy availability.

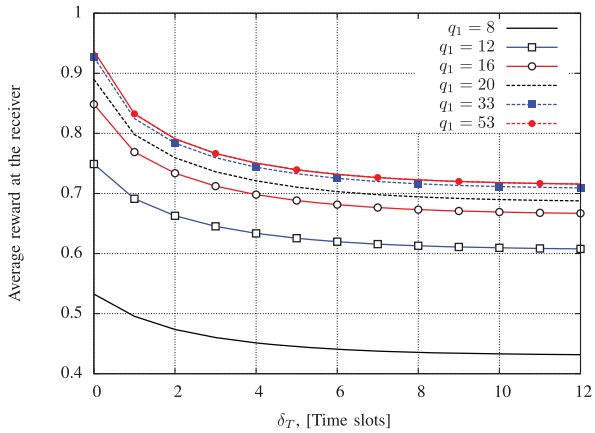
In Fig. 4a we also plot results for different values of d, i.e., the distance between the transmitting node and the sink. We only show these results for one scenario (December, $B = 100$ energy quanta) - other scenarios show a similar behavior. d only affects the channel transition probabilities $p_h(h'|h)$, with higher values of d leading to worse performance, since on average the channel is more likely to be in a SNR region where the required energy consumption (for the given quality parameter ζ) is higher.

In Fig. 4b we show the average reconstruction fidelity at the sink, $R(T)$, when the CSI at the transmitter is known subject to a certain delay $T \geq 1$ (expressed in time slots):

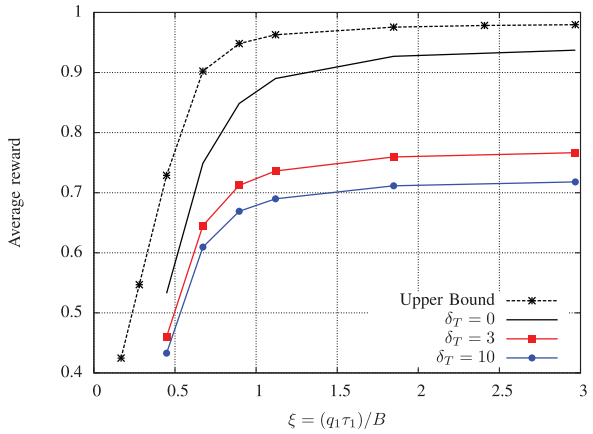
$$R(T) \stackrel{\text{def}}{=} \sum_{s \in \mathcal{S}} \rho(s) r(s, a^\star) \left[\sum_{h' \in \mathcal{H}} p_h^T(h'|h) P_{\text{suc}}(s', a^\star(s)) \right], \tag{25}$$

where $s = [b, x, h]$, $s' = [b, x, h']$ and $p_h^T(h'|h)$ corresponds to the probability that the channel is in state h' in slot $t + T$, given that the channel state is h in slot t, for any given $t \geq 0$. If $\mathbf{H}$ is the channel transition probability matrix, see Section III-B, $p_h^T(h'|h)$ is the entry in position (h, h') of the T-step channel transition probability matrix $\mathbf{H}^T$. This probability is used to track a delayed representation of the channel behavior, where T is the delay in the acquisition of the CSI. Note that for $T = 0$, $\mathbf{H}^0$ is the identity matrix and (25) reduces to (24). As expected, from Figs. 4a and 4b we see that adding some uncertainty in the channel state at the transmitter results in an overall reduction of the system performance. This is especially detrimental when the transmitter thinks that the channel in the current slot is good (high SNR) and, in turn, sends its packet uncompressed and using power level. In fact, if the actual channel state is instead rather bad (low SNR), this behavior is exactly the opposite of what the transmitter should do. As a result, the packet is lost with high probability and this corresponds to a waste of energy and also to a loss of reward, as no signal is recovered at the sink for this time slot.

In Fig. 5a we show the average reward $R(T)$ as a function of T in the August scenario, with $B = 500$ and varying q_1. We observe that as the value of T increases, the performance decreases until reaching a minimum value (around $T \simeq 10$ time slots) and this occurs for all the considered values of q_1.



(a) $R(T)$ as a function of the CSI delay T.



(b) $R(T)$ as a function of ξ for $T \in \{0, 3, 10\}$.

Fig. 5. Average reconstruction fidelity at the sink $R(T)$: impact of delayed CSI.

This is because, with increasing T, $\mathbf{H}^T$ converges towards the stationary distribution of the channel.

Figure 5b shows the average reward $R(T)$ in the same scenario for different values of T. In this plot, the upper bound from [13] is also shown for comparison. This bound has been obtained adapting the analysis in [13], accounting for the rate-distortion curve of LTC and removing the data queue stability constraint, as per our application model (see Section II). Also, the Shannon capacity formula has been considered to map the SNR into the corresponding transmission rate, while transmission power and compression level were jointly modulated. In contrast to this, in our model the transmission power is independently adapted, according to the channel inversion criterion of Section III-B, channel errors are accounted through (12) and the optimization is solely carried out over the compression action. For $T \in \{0, 3, 10\}$, the performance gap $R(0) - R(T)$, with $T \geq 1$, is an increasing function of ξ. This is because, as discussed above, an increasing T corresponds to a higher transmission error probability. Moreover, for higher value of ξ, the optimal action in most of the system states corresponds to transmitting the data packets uncompressed and transmission errors when ξ is higher have a higher impact on the performance. In fact, errors on uncompressed packets are more detrimental (both in terms of energy expenditure and loss of accuracy) and

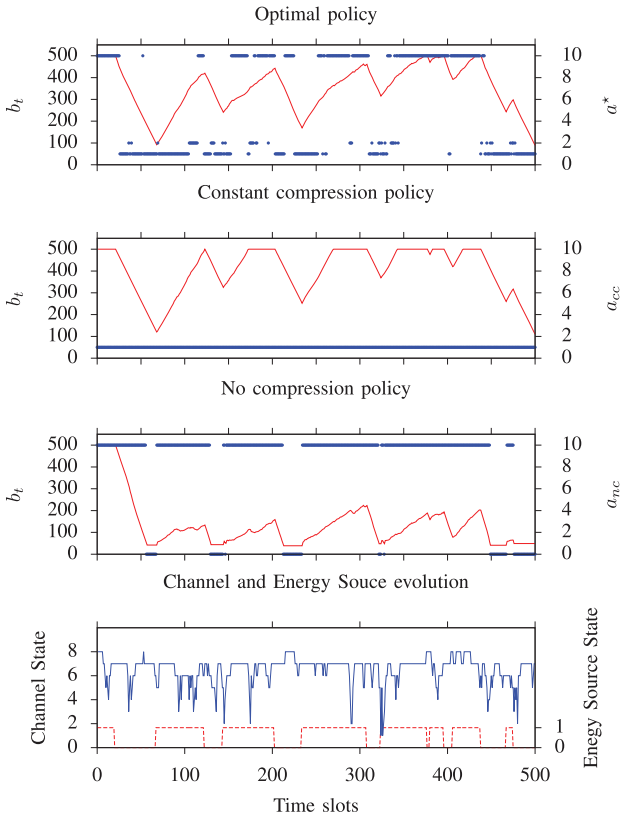


Fig. 6. Performance comparison between optimal and heuristic policies: temporal evolution example.

when T and ξ are large these error events are more likely. As expected, the upper bound achieves higher rewards across the entire range of q_1.

C. Optimal vs Heuristics Policies

In this section, we compare the performance of the optimal policy against that of two heuristic policies:

- Constant Compression Policy (CC): in any given time slot, CC transmits a compressed data packet according to a constant compression level $a_{\text{CC}} \in \{1, \ldots, n\}$ for each state s where $b > b_{\text{th}}$, whereas no transmission is performed otherwise. With this policy, the reward and the energy consumption per transmitted packet are constant, and can be tuned according to the selected compression level. In the results that follow, we have set $a_{\text{CC}} = 1$, corresponding to the smallest compression ratio $a_{\text{CC}}/n = 0.1$.
- No Compression Policy (NC): in any given time slot, NC transmits a data packet uncompressed ($a_{\text{NC}} = 0$) if $b > b_{\text{th}}$, whereas no transmission is performed otherwise. Note that an uncompressed packet implies the highest reward, but at the same time the energy consumption associated with the transmission of a full packet is the highest.

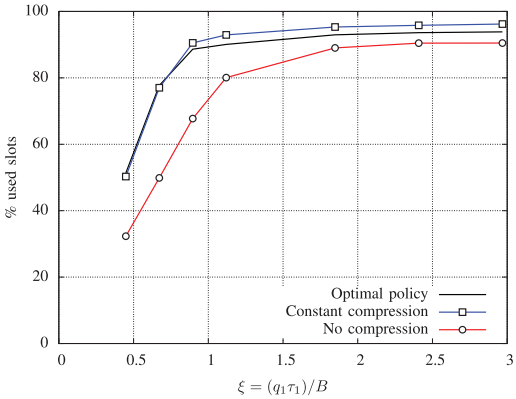
In Fig. 6, we show the temporal evolution for the first 500 time slots for CC, NC and the optimal policy, along with the corresponding evolution of channel and source states. The first three graphs show the energy buffer state evolution (red solid line) and the action taken in each time slot (blue dots) for the optimal policy, CC and NC, respectively. The last graph shows the channel state (blue solid line) and the energy source state (red dashed line). It is interesting to note how the optimal policy manages to keep the energy buffer in a state that is neither fully charged nor below the threshold $b_{\text{th}} = 50$, modulating the action that the system takes in each state. In this way, the buffer constraint is met, the average fidelity is maximized and no input energy is wasted. This is not the case for the other two policies. In particular, CC loses some efficiency as it wastes some harvested energy when the energy buffer is full, while NC drains out the energy buffer too fast and, in turn, forces the node to remain idle more often.

Fig. 7 shows the percentage of used transmission slots and the average reward per slot R. We consider the same scenario of Fig. 5a, with perfect CSI (i.e., $T = 0$). For all the policies the percentage of used slots increases with ξ. CC achieves the best performance since its energy consumption is always that of $a_{cc} = 1$, which is the least energy demanding action. NC, on the other hand, always uses the most expensive action and thus it more often drains out the energy buffer. Thus, the percentage of used slots for this NC is smaller. The optimal policy modulates the behavior of the node depending on the system state, the percentage of used slots is in between that of CC and NC, but as Fig. 7b shows, its average reward is always higher.

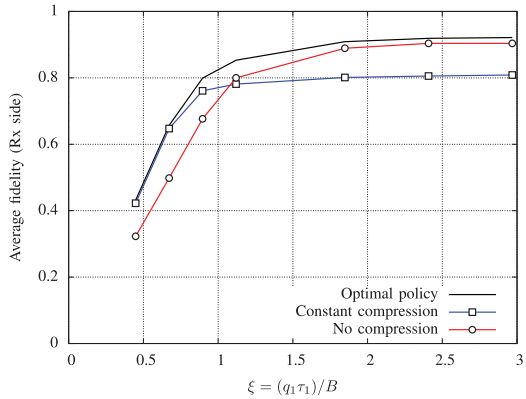
D. Impact of Power Control

In this section, we evaluate the impact of power control at the transmitter. This means that the system state is only composed by the energy buffer state b and the energy source state x. Thus, we obtain the optimal policy for this new system state and we investigate its performance in the presence of a Rayleigh faded channel, see Section III-B, and a fixed transmission power. The probability of a successful transmission is still a function of the channel state, but with a fixed transmission power we can no longer assure that the error probability will be bounded. In fact, setting a low transmission power will lead to small energy consumption, but the only successful transmissions will occur when the channel state is good (i.e., high SNR). On the other hand, with a high transmission power almost all the transmitted packets will be correctly received, but the energy consumption will be very high and the energy buffer will be drained quickly, forcing the transmitter to stop and wait for incoming energy in order to satisfy the constraint on the energy buffer level.

Fig. 8 shows the performance of the optimal policy ("Optimal policy") along with the performance of CC and NC when power control is not applied. We set the transmission power to the maximum level for this graph. For comparison, we also show the performance of the optimal policy when power control is applied ("Optimal policy-PC"). The performance without power control is lower for all policies, both in terms of transmission activity (% of used slots) and fidelity R. The optimal policy without power control shows a transition around $\xi = 1$. This happens because, in the considered example, when $\xi < 1$ the average amount of energy harvested during a "good" period is no longer sufficient to fully replenish the energy buffer. Thus,

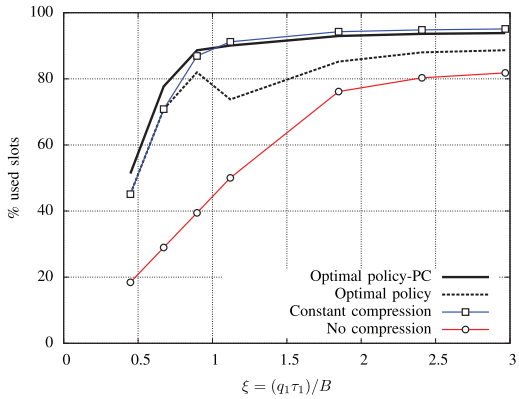


(a) Percentage of used transmission slots.

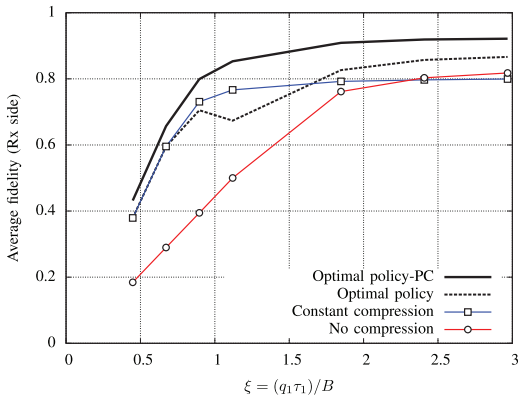


(b) Average reconstruction fidelity at the sink R.

Fig. 7. Performance comparison between optimal and heuristic policies.



(a) Percentage of used slots.



(b) Average reconstruction fidelity at the sink R.

Fig. 8. Impact of power control.

the behavior of the optimal policy tends to mimic that of CC, favoring the transmission of many compressed packets with low average reward over the transmission of a smaller number of high rewarding (uncompressed) data packets. On the other hand, when $\xi > 1$, the input energy is more abundant, and the behavior of the optimal policy is similar to NC.

VI. Conclusions

We studied the problem of of designing efficient policies to control the process of data compression for wireless transmission over fading channels in the presence of a stochastic energy input process and a replenishable energy buffer. We first modeled the transmission and energy dynamics of a sensor node implementing practical lossy compression methods as a constrained Markov decision problem. With an optimality criterion identified as maximizing the long-term data reconstruction fidelity at the receiver (while meeting stability constraints on the energy buffer), we derived optimal compression/transmission policies through a Lagrangian relaxation approach combined with a dichotomic search for the Lagrangian multiplier.

We finally characterized theoretically and numerically the derived optimal policies. We proved that, under realistic assumptions, the optimal policy is non-decreasing in each of the system state components, i.e., energy buffer state, channel state, and energy source state. We also presented a thorough performance evaluation of the optimal policy and of a set of heuristic policies under different energy budget scenarios, highlighting the relationships among system parameters and reconstruction fidelity at the sink. Future work will address modeling of the coupled dynamics of multiple energy-harvesting sensors interacting through a shared wireless channels in a multi-user scenario.

Appendix

In this Section we give the proof of Theorem 1, 2 and 3. In what follows, we omit the dependence on the time index t. For improved readability and with a slight abuse of notation, with a_i, b_i, h_i and x_i we respectively indicate elements of action, energy buffer, channel and energy source sets.

Lemma 1: Under the assumptions of Theorem 1, the expected average reward function $v_\beta(\mathbf{s})$ is concave and monotonically non-decreasing in the energy buffer state b, for any channel state h, and energy source state x.

Proof: Proceeding by induction, since the value iteration algorithm converges for any $v^0_\beta([b,x,h])$, let us choose $v^0_\beta([b,x,h])$ concave non-decreasing function of the buffer

state b. Now, assume that $v_\beta^m([b,x,h])$ is concave and non-decreasing in b. We have to prove that $v_\beta^{m+1}([b,x,h])$ is also concave and non-decreasing in b. We recall that $v_\beta^{m+1}([b,x,h])$ is defined as:

$$v_\beta^{m+1}([b,x,h]) = \max_a Q_\beta^m([b,x,h],a) \tag{26}$$

If v_β^m is concave in b it can be shown (see proof of Theorem 1) that $Q_\beta^m([b,x,h],a)$ is supermodular in the pair (b,a). Therefore,

$$\begin{aligned} &Q_\beta^m([b',x,h],a') - Q_\beta^m([b,x,h],a') \geq \\ &\quad \geq Q_\beta^m([b',x,h],a) - Q_\beta^m([b,x,h],a) \end{aligned} \tag{27}$$

for some $a' \geq a$ and $b' \geq b$. Using the concavity of $Q_\beta^m([b,x,h],a)$ in b, we can write:

$$\begin{aligned} &Q_\beta^m([b',x,h],a') - Q_\beta^m([b,x,h],a') \geq \\ &\quad \geq Q_\beta^m([b'+p,x,h],a) - Q_\beta^m([b+p,x,h],a) \end{aligned} \tag{28}$$

for some $p \geq 0$. Now substituting $b' = \bar{b}$ and $b = \bar{b} - p$, we obtain

$$\begin{aligned} &Q_\beta^m([\bar{b},x,h],a') - Q_\beta^m([\bar{b}-p,x,h],a') \geq \\ &\quad \geq Q_\beta^m([\bar{b}+p,x,h],a) - Q_\beta^m([\bar{b},x,h],a)\ , \end{aligned} \tag{29}$$

and rearranging the terms

$$\begin{aligned} &Q_\beta^m([\bar{b}+p,x,h],a) - Q_\beta^m([\bar{b},x,h],a') \leq \\ &\quad \leq Q_\beta^m([\bar{b},x,h],a) - Q_\beta^m([\bar{b}-p,x,h],a')\ . \end{aligned} \tag{30}$$

Taking $a' = a = \text{argmax}_a\, Q_\beta^m([\bar{b},x,h],a)$ and using (26), we get

$$\begin{aligned} &v_\beta^{m+1}([\bar{b}+p,x,h]) - v_\beta^{m+1}([\bar{b},x,h]) \leq \\ &\quad \leq v_\beta^{m+1}([\bar{b},x,h]) - v_\beta^{m+1}([\bar{b}-p,x,h])\ , \end{aligned} \tag{31}$$

which proves concavity of v_β^{m+1} in b. ■

A. Proof of Theorem 1

Proof: Since

$$\pi_\beta^\star([b,x,h]) = \underset{a}{\text{argmax}} \left\{ Q_\beta([b,x,h],a) \right\}\ , \tag{32}$$

to prove that the optimal policy is non-decreasing in the buffer state b, we have to prove that $Q_\beta([b,x,h],a)$ is supermodular in the pair (b,a). Now, $Q_\beta([b,x,h],a)$ is defined as:

$$Q_\beta([b,x,h],a) = r(s,a;\beta) + \sum_{s' \in \mathcal{S}} p(s'|s,a) v_\beta(s'), \tag{33}$$

where the first term in the right hand side of the equation, i.e., $r(s,a;\beta)$, is supermodular by assumption. Let us rewrite the second term as:

$$\begin{aligned} Q_1([b,x,h],a) &= \sum_{s' \in \mathcal{S}} p(s'|s,a) v_\beta(s') \\ &= \sum_{h' \in \mathcal{H}} \sum_{x' \in \mathcal{X}} \sum_{e_{\text{in}}=0}^{E} p_h(h'|h)\, p_x(x'|x) \cdot \\ &\quad \cdot p_{e_{\text{in}}}(e_{\text{in}}|x) v_\beta([b + e_{\text{in}} - e_{\text{out}}(h,a), x', h'])\ , \end{aligned} \tag{34}$$

where we used the independence of the channel evolution and of the source state evolution. We need to prove that the sum in (35) is supermodular in (b,a) for any $h \in \mathcal{H}$ and $x \in \mathcal{X}$. Using Lemma 1 we have that $v_\beta([b,x,h])$ is concave in b, thus it can be shown that $v_\beta([b + e_{\text{in}} - e_{\text{out}}(h,a), x', h'])$ is supermodular in (b,a) for any $h' \in \mathcal{H}$ and $x' \in \mathcal{X}$. In fact, since $v_\beta([b,x,h])$ is concave in b, it holds that:

$$\begin{aligned} &v_\beta([b_1,x,h]) + v_\beta([b_2,x,h]) \leq v_\beta([\alpha b_1 + (1-\alpha) b_2, x, h]) + \\ &\quad + v_\beta([(1-\alpha) b_1 + \alpha b_2, x, h]) \end{aligned} \tag{36}$$

for $0 \leq \alpha \leq 1$. Substituting $b_1 = b + e_{\text{in}} - e_{\text{out}}(h,a')$, $b_2 = b' + e_{\text{in}} - e_{\text{out}}(h,a)$ and $\alpha = (e_{\text{out}}(h,a') - e_{\text{out}}(h,a))/(e_{\text{out}}(h,a') - e_{\text{out}}(h,a) + b' - b)$, and rearranging the terms we get:

$$\begin{aligned} &v_\beta([b' + e_{\text{in}} - e_{\text{out}}(h,a'), x, h]) - \\ &\qquad - v_\beta([b' + e_{\text{in}} - e_{\text{out}}(h,a), x, h]) \geq \\ &\geq v_\beta([b + e_{\text{in}} - e_{\text{out}}(h,a'), x, h]) - \\ &\qquad - v_\beta([b + e_{\text{in}} - e_{\text{out}}(h,a), x, h])\ , \end{aligned} \tag{37}$$

that for $a' \geq a$ and $b' \geq b$ proves the supermodularity of v_β in (b,a) for any channel state h and source state x. Furthermore, positive weighted sum of supermodular function is also supermodular, hence $Q_\beta([b,x,h],a)$ is supermodular in (b,a) and the monotonic structure of the optimal policy in the energy state b is proven. ■

B. Proof of Theorem 2

Proof: Since

$$\pi_\beta^\star([b,x,h]) = \underset{a}{\text{argmax}} \left\{ Q_\beta([b,x,h],a) \right\}\ , \tag{38}$$

to prove that the optimal policy is non-decreasing in the channel state h, we have to prove that $Q_\beta([b,x,h],a)$ is supermodular in (h,a). Now, $Q_\beta([b,x,h],a)$ is defined as

$$Q_\beta([b,x,h],a) = r(s,a;\beta) + \sum_{s' \in \mathcal{S}} p(s'|s,a) v_\beta(s'), \tag{39}$$

where the first term in the right hand side of the equation, i.e., $r(s,a;\beta)$, is supermodular in (h,a) since it does not depend on h and it is monotonically non-decreasing in a. Let us rewrite the second term as

$$\begin{aligned} Q_1([b,x,h],a) &= \sum_{s' \in \mathcal{S}} p(s'|s,a) v_\beta(s') \\ &= \sum_{h' \in \mathcal{H}} \sum_{x' \in \mathcal{X}} \sum_{e_{\text{in}}=0}^{E} p_h(h'|h)\, p_x(x'|x) \cdot \\ &\quad \cdot p_{e_{\text{in}}}(e_{\text{in}}|x) v_\beta([b + e_{\text{in}} - e_{\text{out}}(h,a), x', h'])\ , \end{aligned} \tag{40}$$

(41)

where we used the independence of the channel evolution and of the source state evolution. We need to prove that the sum in (41) is supermodular in (h, a) for any $b \in \mathcal{B}$ and $x \in \mathcal{X}$, that is

$$\begin{aligned}
&\sum_{x' \in \mathcal{X}} p_x(x'|x) \sum_{e_{\text{in}}=0}^{E} p_{e_{\text{in}}}(e_{\text{in}}|x) \sum_{h' \in \mathcal{H}} p_h(h'|h_1) \cdot \\
&\quad \cdot v_\beta([b + e_{\text{in}} - e_{\text{out}}(h_1, a_1), x', h']) - \\
&\quad - \sum_{x' \in \mathcal{X}} p_x(x'|x) \sum_{e_{\text{in}}=0}^{E} p_{e_{\text{in}}}(e_{\text{in}}|x) \sum_{h' \in \mathcal{H}} p_h(h'|h_1) \cdot \\
&\quad \cdot v_\beta([b + e_{\text{in}} - e_{\text{out}}(h_1, a_2), x', h']) \geq \\
&\geq \sum_{x' \in \mathcal{X}} p_x(x'|x) \sum_{e_{\text{in}}=0}^{E} p_{e_{\text{in}}}(e_{\text{in}}|x) \sum_{h' \in \mathcal{H}} p_h(h'|h_2) \cdot \\
&\quad \cdot v_\beta([b + e_{\text{in}} - e_{\text{out}}(h_2, a_1), x', h']) - \\
&\quad - \sum_{x' \in \mathcal{X}} p_x(x'|x) \sum_{e_{\text{in}}=0}^{E} p_{e_{\text{in}}}(e_{\text{in}}|x) \sum_{h' \in \mathcal{H}} p_h(h'|h_2) \cdot \\
&\quad \cdot v_\beta([b + e_{\text{in}} - e_{\text{out}}(h_2, a_2), x', h'])
\end{aligned} \tag{42}$$

for $h_1 \geq h_2, a_1 \geq a_2$, for any $x \in \mathcal{X}$ and any $b \in \mathcal{B}$. Getting rid of the constant terms, (42) can be rewritten as:

$$\begin{aligned}
&\sum_{h' \in \mathcal{H}} p_h(h'|h_1)(v_\beta([b + e_{\text{in}} - e_{\text{out}}(h_1, a_1), x', h']) - \\
&\quad - v_\beta([b + e_{\text{in}} - e_{\text{out}}(h_1, a_2), x', h'])) \geq \\
&\geq \sum_{h' \in \mathcal{H}} p_h(h'|h_2)(v_\beta([b + e_{\text{in}} - e_{\text{out}}(h_2, a_1), x', h']) - \\
&\quad - v_\beta([b + e_{\text{in}} - e_{\text{out}}(h_2, a_2), x', h'])) \, .
\end{aligned} \tag{43}$$

Since $p_h(h'|h)$ is stochastically increasing in h, a sufficient condition for (43) to hold (see Lemma 4.7.2 in [27]) is that

$$\begin{aligned}
&v_\beta([b + e_{\text{in}} - e_{\text{out}}(h_1, a_1), x', h']) - \\
&\quad - v_\beta([b + e_{\text{in}} - e_{\text{out}}(h_1, a_2), x', h']) \geq \\
&\geq v_\beta([b + e_{\text{in}} - e_{\text{out}}(h_2, a_1), x', h']) - \\
&\quad - v_\beta([b + e_{\text{in}} - e_{\text{out}}(h_2, a_2), x', h'])
\end{aligned} \tag{44}$$

i.e., v_β is supermodular in (h, a). Since b, x', h', e_{in} are fixed, and $e_{\text{out}}(h, a)$ is monotonically decreasing in h, (44) can be rewritten (with some abuse of notation) as

$$\begin{aligned}
&v_\beta([(b', a'), x', h']) - v_\beta([(b', a), x', h']) \geq \\
&\quad \geq v_\beta([(b, a'), x', h']) - v_\beta([(b, a), x', h'])
\end{aligned} \tag{45}$$

where $b' \geq b$ and $a' \geq a$. But this is the condition for the supermodularity of v_β in (b, a) that holds for Theorem 1. ■

C. Proof of Theorem 3

Proof: Since

$$\pi^\star_\beta([b, x, h]) = \underset{a}{\text{argmax}} \left\{ Q_\beta([b, x, h], a) \right\} \, , \tag{46}$$

to prove that the optimal policy is non-decreasing in the source state x, we have to prove that $Q_\beta([b, x, h], a)$ is supermodular in (x, a). Now, $Q_\beta([b, x, h], a)$ is defined as

$$Q_\beta([b, x, h], a) = r(s, a; \beta) + \sum_{s' \in \mathcal{S}} p(s'|s, a) v_\beta(s'), \tag{47}$$

where the first term in the right hand side of the equation, i.e., $r(s, a; \beta)$, is supermodular in (x, a) since it does not depend on x and it is monotonically non-decreasing in a. Let us rewrite the second term as

$$\begin{aligned}
Q_1([b, x, h], a) &= \sum_{s' \in \mathcal{S}} p(s'|s, a) v_\beta(s') \qquad (48) \\
&= \sum_{h' \in \mathcal{H}} \sum_{x' \in \mathcal{X}} \sum_{e_{\text{in}}=0}^{E} p_h(h'|h) p_x(x'|x) \cdot \\
&\quad \cdot p_{e_{\text{in}}}(e_{\text{in}}|x) \, v_\beta\left(\left[b + e_{\text{in}} - e_{\text{out}}(h, a), x', h' \right] \right), \qquad (49)
\end{aligned}$$

where we used the independence of the channel evolution and of the source state evolution. We need to prove that the sum in (49) is supermodular in (x, a) for any $b \in \mathcal{B}$ and $h \in \mathcal{H}$, that is

$$\begin{aligned}
&\sum_{h' \in \mathcal{H}} p_h(h'|h) \sum_{x' \in \mathcal{X}} p_x(x'|x_1) \sum_{e_{\text{in}}=0}^{E} p_{e_{\text{in}}}(e_{\text{in}}|x_1) \cdot \\
&\quad \cdot v_\beta([b + e_{\text{in}} - e_{\text{out}}(h, a_1), x', h']) - \\
&\quad - \sum_{h' \in \mathcal{H}} p_h(h'|h) \sum_{x' \in \mathcal{X}} p_x(x'|x_1) \sum_{e_{\text{in}}=0}^{E} p_{e_{\text{in}}}(e_{\text{in}}|x_1) \cdot \\
&\quad \cdot v_\beta([b + e_{\text{in}} - e_{\text{out}}(h, a_2), x', h']) \geq \\
&\geq \sum_{h' \in \mathcal{H}} p_h(h'|h) \sum_{x' \in \mathcal{X}} p_x(x'|x_2) \sum_{e_{\text{in}}=0}^{E} p_{e_{\text{in}}}(e_{\text{in}}|x_2) \cdot \\
&\quad \cdot v_\beta([b + e_{\text{in}} - e_{\text{out}}(h, a_1), x', h']) - \\
&\quad - \sum_{h' \in \mathcal{H}} p_h(h'|h) \sum_{x' \in \mathcal{X}} p_x(x'|x_2) \sum_{e_{\text{in}}=0}^{E} p_{e_{\text{in}}}(e_{\text{in}}|x_2) \cdot \\
&\quad \cdot v_\beta([b + e_{\text{in}} - e_{\text{out}}(h, a_2), x', h'])
\end{aligned} \tag{50}$$

for $x_1 \geq x_2, a_1 \geq a_2$, for any $h \in \mathcal{H}$ and any $b \in \mathcal{B}$. Getting rid of the constant terms, (50) can be rewritten as

$$\begin{aligned}
&\sum_{x' \in \mathcal{X}} p_x(x'|x_1) \sum_{e_{\text{in}}=0}^{E} p_{e_{\text{in}}}(e_{\text{in}}|x_1) \cdot \\
&\quad \cdot (v_\beta([b + e_{\text{in}} - e_{\text{out}}(h, a_1), x', h']) - \\
&\quad - v_\beta([b + e_{\text{in}} - e_{\text{out}}(h, a_2), x', h'])) \geq \\
&\geq \sum_{x' \in \mathcal{X}} p_x(x'|x_2) \sum_{e_{\text{in}}=0}^{E} p_{e_{\text{in}}}(e_{\text{in}}|x_2) \cdot \\
&\quad \cdot (v_\beta([b + e_{\text{in}} - e_{\text{out}}(h, a_1), x', h']) - \\
&\quad - v_\beta([b + e_{\text{in}} - e_{\text{out}}(h_2, a_2), x', h'])) \, .
\end{aligned} \tag{51}$$

Since $p_x(x'|x)$ is stochastically increasing in x, a sufficient condition for (51) to hold (see Lemma 4.7.2 in [27]) is that

$$\begin{aligned}
&\sum_{e_{\text{in}}=0}^{E} p_{e_{\text{in}}}(e_{\text{in}}|x_1)(v_\beta([b+e_{\text{in}}-e_{\text{out}}(h,a_1),x',h'])-\\
&\quad -v_\beta([b+e_{\text{in}}-e_{\text{out}}(h,a_2),x',h'])) \geq\\
&\geq \sum_{e_{\text{in}}=0}^{E} p_{e_{\text{in}}}(e_{\text{in}}|x_2)(v_\beta([b+e_{\text{in}}-e_{\text{out}}(h,a_1),x',h'])-\\
&\quad -v_\beta([b+e_{\text{in}}-e_{\text{out}}(h_2,a_2),x',h'])) \,. \qquad (52)
\end{aligned}$$

for each $x_1 \geq x_2$. If the energy income from states x_i takes values in disjoint ordered sets, i.e., $e_{\text{in}}^{x_i} \in \mathcal{E}^{x_i} = \{E_{\min}^{x_i}, E_{\min}^{x_i}+1, \ldots, E_{\max}^{x_i}\}$, $\mathcal{E}^{x_i} \cap \mathcal{E}^{x_j} = \emptyset$ for $i \neq j$, $\cup_i \mathcal{E}^{x_i} = \{0,1,\ldots,E\}$, $E_{\max}^{x_i} < E_{\min}^{x_{i+1}}$, and since the sums in (52) are convex combinations, we can write

$$\begin{aligned}
&\sum_{e_{\text{in}}=0}^{E} p_{e_{\text{in}}}(e_{\text{in}}|x_1)(v_\beta([b+e_{\text{in}}-e_{\text{out}}(h,a_1),x',h'])-\\
&\quad -v_\beta([b+e_{\text{in}}-e_{\text{out}}(h,a_2),x',h'])) \geq\\
&\geq v_\beta([b+E_{\min}^{x_1}-e_{\text{out}}(h,a_1),x',h'])-\\
&\quad -v_\beta([b+E_{\min}^{x_1}-e_{\text{out}}(h,a_2),x',h']) \geq\\
&\geq v_\beta([b+E_{\max}^{x_2}-e_{\text{out}}(h,a_1),x',h'])-\\
&\quad -v_\beta([b+E_{\max}^{x_2}-e_{\text{out}}(h,a_2),x',h']) \geq\\
&\geq \sum_{e_{\text{in}}=0}^{E} p_{e_{\text{in}}}(e_{\text{in}}|x_2)(v_\beta([b+e_{\text{in}}-e_{\text{out}}(h,a_1),x',h'])-\\
&\quad -v_\beta([b+e_{\text{in}}-e_{\text{out}}(h_2,a_2),x',h'])) \,. \qquad (53)
\end{aligned}$$

Thus, (52) holds if the innermost inequality in (53) holds. Rearranging the term, we can write

$$\begin{aligned}
&v_\beta([b+E_{\min}^{x_1}-e_{\text{out}}(h,a_1),x',h'])-\\
&\quad -v_\beta([b+E_{\max}^{x_2}-e_{\text{out}}(h,a_1),x',h']) \geq\\
&\geq v_\beta([b+E_{\min}^{x_1}-e_{\text{out}}(h,a_2),x',h'])-\\
&\quad -v_\beta([b+E_{\max}^{x_2}-e_{\text{out}}(h,a_2),x',h']) \,, \qquad (54)
\end{aligned}$$

which holds since e_{out} is monotonically increasing in a and v_β is concave and increasing in b because of Lemma 1. ■

REFERENCES

[1] T. Wang, W. Heinzelman, and A. Seyedi, "Link energy minimization for wireless sensor networks," *Elsevier Ad Hoc Netw.*, vol. 10, no. 3, pp. 569–585, 2012.

[2] D. Zordan, B. Martinez, I. Vilajosana, and M. Rossi, "On the performance of lossy compression schemes for energy constrained sensor networking," *ACM Trans. Sens. Netw.*, vol. 11, no. 1, pp. 15:1–15:34, Aug. 2014.

[3] L. Xiang, J. Luo, and A. Vasilakos, "Compressed data aggregation for energy efficient wireless sensor networks," in *Proc. IEEE Commun. Soc. Conf. Sens. Mesh Ad Hoc Commun. Netw. (SECON)*, Salt Lake City, UT, USA, Jun. 2011, pp. 46–54.

[4] G. Quer, R. Masiero, G. Pillonetto, M. Rossi, and M. Zorzi, "Sensing, compression and recovery for WSNs: Sparse signal modeling and monitoring framework," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3447–3461, Oct. 2012.

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

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

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

[8] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, "Transmission with energy harvesting nodes in fading wireless channels: Optimal policies," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1732–1743, Sep. 2011.

[9] N. Michelusi, K. Stamatiou, and M. Zorzi, "Transmission policies for energy harvesting sensors with time-correlated energy supply," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2988–3001, Jul. 2013.

[10] M. Gatzianas, L. Georgiadis, and L. Tassiulas, "Control of wireless networks with rechargeable batteries," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 581–593, Feb. 2010.

[11] L. Huang and M. J. Neely, "Utility optimal scheduling in energy-harvesting networks," *IEEE/ACM Trans. Netw.*, vol. 21, no. 4, pp. 1117–1130, Aug. 2013.

[12] P. Castiglione, O. Simeone, E. Erkip, and T. Zemen, "Energy management policies for energy-neutral source-channel coding," *IEEE Trans. Commun.*, vol. 60, no. 9, pp. 2668–2678, Sep. 2012.

[13] P. Castiglione and G. Matz, "Energy-neutral source-channel coding with battery and memory size constraints," *IEEE Trans. Commun.*, vol. 62, no. 4, pp. 1373–1381, Apr. 2014.

[14] R. Rajesh, V. Sharma, and P. Viswanath, "Capacity of Gaussian channels with energy harvesting and processing cost," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2563–2575, May 2014.

[15] C. Tapparello, O. Simeone, and M. Rossi, "Dynamic compression-transmission for energy-harvesting multihop networks with correlated sources," *IEEE/ACM Trans. Netw.*, vol. 22, no. 6, pp. 1729–1741, Dec. 2014.

[16] M. Miozzo, D. Zordan, P. Dini, and M. Rossi, "SolarStat: Modeling photovoltaic sources through stochastic Markov processes," in *Proc. IEEE Energy Conf. (ENERGYCON)*, Dubrovnik, Croatia, May 2014, pp. 688–695.

[17] T. Schoellhammer, B. Greenstein, E. Osterweil, M. Wimbrow, and D. Estrin, "Lightweight temporal compression of microclimate datasets [wireless sensor networks]," in *Proc. Annu. IEEE Int. Conf. Local Comput. Netw.*, Tampa, FL, USA, Nov. 2004, pp. 516–524.

[18] M. N. Halgamuge, M. Zukerman, K. Ramamohanarao, and H. L. Vu, "An estimation of sensor energy consumption," *Prog. Electromagn. Res. B*, vol. 12, pp. 259–295, 2009.

[19] E. Altman, *Constrained Markov Decision Processes: Stochastic Modeling*. London, U.K.: Chapman and Hall/CRC, 1999.

[20] C. Tan and N. Beaulieu, "On first-order Markov modeling for the Rayleigh fading channel," *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 2032–2040, Dec. 2000.

[21] M. Rossi, L. Badia, and M. Zorzi, "SR ARQ delay statistics on N-state Markov channels with non-instantaneous feedback," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1526–1536, Jun. 2006.

[22] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[23] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 4th ed., vol. 2, Belmont, MA, USA: Athena Scientific, 2012.

[24] F. J. Beutlerand and K. W. Ross, "Optimal policies for controlled Markov chains with a constraint," *J. Math. Anal. Appl.*, vol. 112, no. 1, pp. 236–252, Nov. 1985.

[25] D. Djonin and V. Krishnamurthy, "Structural results on optimal transmission scheduling over dynamical fading channels: A constrained Markov decision process approach," in *Wireless Communications*, vol. 143, New York, NY, USA: Springer, 2007, pp. 75–98.

[26] C. Sun, E. Stevens-Navarro, V. Shah-Mansouri, and V. W. Wong, "A constrained MDP-based vertical handoff decision algorithm for 4G heterogeneous wireless networks," *Wireless Netw.*, vol. 17, no. 4, pp. 1063–1081, May 2011.

[27] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Hoboken, NJ, USA: Wiley, 2005.

[28] C. J. Willmott and K. Matsuura. *Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series (1950–1999)* [Online]. Available: http://climate.geog.udel.edu/~climate/html_pages/README.ghcn_ts2.html, July 2001.

Davide Zordan received the M.Sc. degree in telecommunications engineering, and the Ph.D. degree from the University of Padova, Padova, Italy, in 2010 and 2014, respectively. He is currently a Postdoctoral Researcher with the Department of Information Engineering, University of Padova. His research interests include stochastic modeling and optimization, protocol design and performance evaluation for wireless networks, in-network processing techniques (including compressive sensing), energy efficient protocols, and energy harvesting techniques for WSNs and wearable IoT devices.

Tommaso Melodia (M'07) received the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2007. He is an Associate Professor with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, USA. His research was supported by the US National Science Foundation, Air Force Research Laboratory, and the Office of Naval Research, among others. His research interests include modeling, optimization, and experimental evaluation of networked communication systems, with applications to ultrasonic intra-body networks, cognitive and cooperative networks, multimedia sensor networks, and underwater networks. He has coauthored a paper that was recognized as the ISI Fast Breaking Paper in the field of computer science for February 2009. He was the Technical Program Committee Vice Chair for the IEEE Globecom 2013 and the Technical Program Committee Vice Chair for information systems for the IEEE INFOCOM 2013, and is an Area Chair for the IEEE INFOCOM 2016. He serves on the editorial boards of the IEEE TRANSACTIONS ON MOBILE COMPUTING, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON MULTIMEDIA, and *Computer Networks*. He was the recipient of the U.S. National Science Foundation CAREER Award and the ACM WUWNet 2013 Best Paper Award.

Michele Rossi (S'02–M'03–SM'13) is an Assistant Professor with the Department of Information Engineering, University of Padova, Padova, Italy. His research interests include centered around wireless sensor networks (WSNs), Internet of Things (networking protocol design), telecommunication and optimization aspects of smart power grids, signal processing for wearable devices (compression, activity, and motion analysis), and energy harvesting mobile networks. In the last few years, he has been actively involved in several EU projects, including FP7 SWAP (energy harvesting WSNs) and IOT-A (architectures and protocols for the Internet of Things). His current research is supported by the SAMSUNG Advanced Institute of Technology through a GRO SAMSUNG Research Award on Signal Processing and Compression for Wearable Devices and by the European Commission through the H2020 MSCA-ITN project SCAVENGE on Energy Harvesting Cellular Networks. He has been on the technical program committees of more than 80 among international conferences and workshops. He was the recipient of four best paper awards from the IEEE. He currently serves on the Editorial Board of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.