

Dynamic Compression-Transmission for Energy-Harvesting Multihop Networks With Correlated Sources

Cristiano Tapparello, *Member, IEEE*, Osvaldo Simeone, *Member, IEEE*, and Michele Rossi, *Senior Member, IEEE*

Abstract—Energy-harvesting wireless sensor networking is an emerging technology with applications to various fields such as environmental and structural health monitoring. A distinguishing feature of wireless sensors is the need to perform both source coding tasks, such as measurement and compression, and transmission tasks. It is known that the overall energy consumption for source coding is generally comparable to that of transmission, and that a joint design of the two classes of tasks can lead to relevant performance gains. Moreover, the efficiency of source coding in a sensor network can be potentially improved via distributed techniques by leveraging the fact that signals measured by different nodes are correlated. In this paper, a data-gathering protocol for multihop wireless sensor networks with energy-harvesting capabilities is studied whereby the sources measured by the sensors are correlated. Both the energy consumptions of source coding and transmission are modeled, and distributed source coding is assumed. The problem of dynamically and jointly optimizing the source coding and transmission strategies is formulated for time-varying channels and sources. The problem consists in the minimization of a cost function of the distortions in the source reconstructions at the sink under queue stability constraints. By adopting perturbation-based Lyapunov techniques, a close-to-optimal online scheme is proposed that has an explicit and controllable tradeoff between optimality gap and queue sizes. The role of side information available at the sink is also discussed under the assumption that acquiring the side information entails an energy cost.

Index Terms—Data gathering, distributed source coding, energy harvesting, Lyapunov optimization, wireless sensor networks.

I. INTRODUCTION

WIRELESS sensor networks have found applications in a large number of fields such as environmental sensing and structural health monitoring [1]. In such applications, the maintenance necessary to replace the batteries when depleted

Manuscript received March 15, 2012; revised October 16, 2012 and February 22, 2013; accepted September 13, 2013; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor D. Manjunath. Date of publication October 17, 2013; date of current version December 15, 2014. The work of C. Tapparello and M. Rossi was supported in part by the European Commission under the 7th Framework Programme through the “SWAP” project, G.A. No. 251557. The work of O. Simeone was supported in part by the US NSF under Grant CCF-0914899.

C. Tapparello and M. Rossi are with the Department of Information Engineering (DEI), University of Padova, 35131 Padova, Italy (e-mail: tapparello@dei.unipd.it).

O. Simeone is with the Center for Wireless Communications and Signal Processing Research (CWCSRP), New Jersey Institute of Technology (NJIT), Newark, NJ 07102 USA.

Digital Object Identifier 10.1109/TNET.2013.2283071

is often of prohibitive complexity, if not impossible. Therefore, sensors that harvest energy from the environment, e.g., in the form of solar, thermal, vibrational or radio energy [2], [3], have been proposed and are now commercially available.

Given the interest outlined above, the problem of designing optimal transmission protocols for energy-harvesting wireless sensor networks has recently received considerable attention. In the available body of work reviewed in Section I-B, the only source of energy expenditure is the energy used for transmission. This includes, e.g., the energy used by the power amplifiers. However, a distinguishing feature of sensor networks is that the sensors not only have to carry out transmission tasks, but also *sensing and source coding tasks*, such as compression. The source coding tasks entail a nonnegligible energy consumption. In fact, [4] demonstrates that the overall cost required for compression¹ is comparable to that needed for transmission, and that a joint design of the two tasks can lead to significant energy saving gains. Another distinguishing feature of sensor networks is that the efficiency of source coding can be improved via *distributed source coding techniques* (see, e.g., [5]) by leveraging the fact that sources measured by different sensors are generally correlated (see, e.g., [6]).

A. Contributions

In this paper, we focus on an energy-harvesting wireless sensor network and account for the energy costs of both source coding and transmission. Moreover, we assume that the sensors can perform distributed source coding to leverage the correlation of the sources measured at different sensors. A key motivation for enabling distributed source coding in energy-harvesting networks is that this enables sensors with correlated measurements to trade energy resources among them, to an extent determined by the amount of correlation. For instance, a sensor that is running low on energy can benefit from the energy potentially available at a nearby node if the latter has correlated measurements. This is because, through distributed source coding, the transmission requirements on the first sensor are eased by the transmission of correlated information from the nearby sensor.

We study the problem of dynamically and jointly optimizing the source coding and transmission strategies over time-varying channels and sources. The problem consists in the minimization of a cost function of the distortions in the source reconstructions at the sink under queue stability constraints. Our approach is

¹This reference considers transmission of Web data.

based on the Lyapunov optimization strategy with weight perturbation developed in [7]. It is emphasized that the application of this strategy to the problem at hand poses significant technical challenges due to the presence of correlated sources, distributed source coding, and sensing costs, which were not considered in previous work. As it will be discussed, this is primarily a consequence of the need to properly design the choice of the weight vector to be used in the weight perturbation approach: The optimal design of the weight vector is problem-dependent and thus cannot be obtained from [7]. We devise an efficient online algorithm that only takes actions based on the harvested energy; on the current state of channel, queues, and energy reserves; and also based on the statistical description of the source correlation. We prove that the proposed policy achieves an average network cost that can be made arbitrarily close to the optimal one with a controllable tradeoff between the sizes of the queues and batteries.

We also investigate the role of side information available at the sink under the assumption that acquiring the side information entails an energy cost. It is shown that properly allocating the available (harvested) energy to both the tasks of transmission and side information measurement has significant benefits both in terms of overall network cost function and queue sizes.

B. Prior Work

We start by introducing related prior work that assumes energy harvesting. The literature on this topic is quickly increasing in volume, but it mostly (with the exception of [8]) accounts only for the energy consumption of the transmission component and does not model the contribution of the source coding part. In this context, [9] and [10] studied the problem of maximizing the throughput or minimizing the completion time for a single link energy-harvesting system by focusing on both offline and online policies (see also [11] and [12]). A related work is also [13], which finds a power allocation policy that stabilizes the data queue whenever feasible. Still, for a point-to-point system, using large deviation tools, the effect of finite data queue length and battery size is studied in [14] in terms of scaling results as the battery and queue grow large.

We now consider work on multihop energy-harvesting networks. As mentioned above, all the works at hand only account for the energy used for transmission. Moreover, source correlations and distributed source coding are not accounted for. In [7], assuming independent and identically distributed (i.i.d.) channel states and energy-harvesting processes, a Lyapunov optimization technique with weight perturbation [15] is leveraged to obtain approximately optimal strategies in terms of a general function of the data rates under queue stability constraints. The proposed technique obtains an explicit tradeoff in terms of data queue length and battery size. An extension of this work that assumes more general arrival, channel state, and recharge processes along with finite batteries and queues is put forth in [16]. Also related are [17]–[19] that tackle similar problems, and [20] that solves the power allocation problem for a multi-commodity scenario with independent sources, finding the maximum average transmission rates for fixed channel states, by additionally considering fixed reception, sensing, and processing

energy consumptions, under average energy availability constraints. This model substantially differs from ours as we consider correlated sources and time-variant channels and we require that the energy availability constraint be verified in every time-slot.

We now discuss work that accounts for the energy tradeoffs related to source coding and transmission. These works (except [18]) do not model the additional constraints arising from energy harvesting. Moreover, they do not allow for distributed source coding. The joint design of source coding and transmission parameters is investigated through various algorithms, for either static scenarios in [21] and [22] or dynamic scenarios in [23] and [24]. Specifically, [23] and [24] studied the tradeoffs between energy used for compression, or more generally source coding, and transmission by assuming i.i.d. source and channel processes and arbitrarily large data buffer. Using Lyapunov optimization techniques, a policy with close-to-optimal power expenditure and an explicit tradeoff with the delay is derived for a given average distortion. The problem of optimal energy allocation between source coding and transmission for a point-to-point system was studied in [8].

Finally, distributed source coding techniques for multihop sensor networks have been studied in [25] and [26]. In [25], the problem of optimizing the transmission and compression strategy was tackled under distortion constraints in a centralized fashion, whereas [26] proposed a distributed algorithm that maximizes an aggregate utility measure defined in terms of the distortion levels of the sources. Both these works do not consider energy harvesting nor the energy consumption of the sensing process.

C. Paper Organization

The rest of the paper is organized as follows. In Section II, we present the system model and state the optimization problem. In Section III, we obtain a lower bound on the optimal network cost for the proposed problem. In Section IV, we present the proposed algorithm designed following the Lyapunov optimization framework and show how it can be implemented in a distributed fashion. Section V formalizes the main results of our paper and provide analytical insights into the performance of the proposed policy. Section IV proposes an extended version of the problem, where the sink node acts as a cluster head that is able to acquire correlated side information to improve the system performance. In Section VII, we prove the effectiveness of our analytical analysis and discuss the impact of the optimization parameters. Section VIII concludes the paper.

II. SYSTEM MODEL

We consider a wireless network modeled by a directed graph $\mathcal{G} = (\mathcal{N} \cup \{d\}, \mathcal{L})$, where $\mathcal{N} = \{1, 2, \dots, N\}$ is the set of nodes in the network, d is the destination (or sink), and $\mathcal{L} \subset \{(n, m): n, m \in \mathcal{N} \cup \{d\}, n \neq m\}$ represents the set of communication links; see Fig. 1 for an illustration. We define l_{\max} as the maximum number of transmission links that any node $n \in \mathcal{N} \cup \{d\}$ can have. As discussed in the following, we allow for fairly general interference models. We will consider a more general model in Section VI in which the sink acts as a cluster

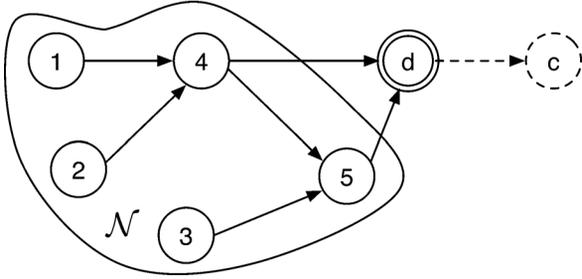


Fig. 1. Set \mathcal{N} of energy-harvesting nodes communicates correlated sources to a destination d . For the more general model of Section VI, the destination d acts as a cluster head and communicates to a network collector node (shown in dashed lines). In this latter model, the node d can collect side information correlated to the sources measured by the nodes.

head for the set of nodes \mathcal{N} and reports to a collector node c (see Fig. 1).

We make the assumption that the network operates in slotted time, where during each time-slot every node is able to perform different tasks, as we detail next.

A. Sensor Model

Each wireless sensor node n is modeled according to the diagram of Fig. 2. Specifically, in every time-slot t , an amount of energy $H_n(t)$ is harvested from the environment. A portion $\tilde{H}(t) \leq H_n(t)$ of that energy is stored in a local energy buffer, depending on its capacity and current state. In the considered time-slot, node n receives an amount of exogenous data at an overall rate $\mu_{*,n}(t)$ from its neighboring nodes, and also generates endogenous data at rate $R_n(t)$ through the acquisition and compression of a spatially correlated signal. The sum of endogenous and exogenous data is thus stored into a local data queue, and data are transmitted from this queue at an overall rate $\mu_{n,*}(t)$. Note that the node can control the amount of compression (i.e., the rate $R_n(t)$) and the power used for the transmission in the current slot. Also, compression and transmission activities must be both accomplished using the energy available in the energy buffer. The choices at different nodes are coupled by the spatial correlation of the measured sources, as will be discussed. The transmission model is detailed in Section II-B, the data acquisition and compression process is modeled in Section II-C, whereas energy model and queuing dynamics are specified in Sections II-D and II-E, respectively.

We remark that, in the existing literature, the designs of optimal transmission scheduling for harvesting networks [7] and of dynamic compression and transmission [23] have been investigated separately. Moreover, unlike [23], we allow for distributed source coding at the sensors, thus leveraging the spatial correlation of the sensors' measurements in order to improve the energy efficiency of the network. The goal of the analysis is the design of dynamic distributed compression and transmission strategies that are able to maximize the fidelity of the reconstructions at the destination, while maintaining the network is energetically self-sufficient and ensuring queue stability. We emphasize that the need to properly balance the use of the available harvested energy across the nodes for distributed source coding and transmission creates significant technical challenges as compared to the previous work [7], [23]. This is because the

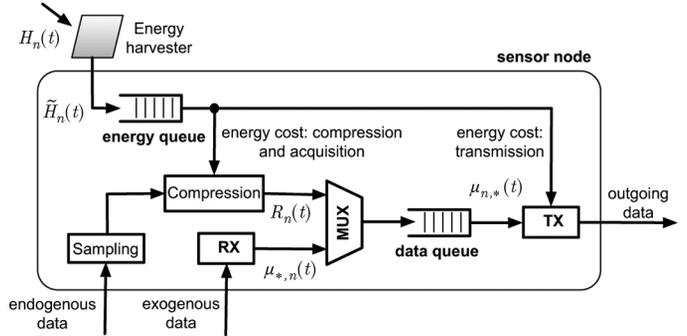


Fig. 2. Diagram of a sensor node.

decisions made at the different nodes become coupled by the constraints imposed by distributed source coding, as discussed in Section II-C

B. Transmission Model

The transmission model follows the framework of, e.g., [27]. According to this model, at every time-slot $t = 1, 2, \dots$, each node $n \in \mathcal{N}$ allocates power $P_{n,m}(t) \geq 0$ to each outgoing link $(n, m) \in \mathcal{L}$ for data transmission. In what follows, we refer to the number of *channel uses* (or transmission symbols) per time-slot as the baud rate multiplied by the slot duration. At the generic time-slot t , we define $\mathbf{P}(t) = \{P_{n,m}(t)\}$, with $(n, m) \in \mathcal{L}$, as the power allocation matrix and the total transmission power of node n , that is

$$P_n(t) = \sum_{m:(n,m) \in \mathcal{L}} P_{n,m}(t) \quad (1)$$

which is assumed to satisfy the constraint $P_n(t) \leq P_{\max}$, for some $P_{\max} < \infty$. The transmission rate $\mu_{n,m}(t)$ on link (n, m) depends on the power allocation matrix $\mathbf{P}(t)$ and on the current *channel state* $\mathbf{S}(t) = \{S_{n,m}(t)\}$ with $(n, m) \in \mathcal{L}$. The latter accounts, for instance, for the current fading channels or for the connectivity conditions on the network links. We assume that $\mathbf{S}(t)$ takes values in some finite set $\mathcal{S} = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_M\}$, is constant within a time-slot, but is i.i.d. across time-slots. We use $\rho_{\mathbf{s}_i} = \Pr[\mathbf{s}(t) = \mathbf{s}_i]$ for $i = 1, \dots, M$. We write

$$\mu_{n,m}(t) = C_{n,m}(\mathbf{P}(t), \mathbf{S}(t)) \quad (2)$$

where $C_{n,m}(\mathbf{P}(t), \mathbf{S}(t))$ is the capacity-power curve for link (n, m) expressed in terms of bits per channel use (transmission symbol). The latter depends on the specific network transmission strategy, which includes the modulation and coding/decoding schemes used on all links. We assume that function $C_{n,m}(\mathbf{P}(t), \mathbf{S}(t))$ is continuous in $\mathbf{P}(t)$ and nondecreasing in $P_{n,m}(t)$. An example of the function $C_{n,m}(\mathbf{P}(t), \mathbf{S}(t))$ is the Shannon capacity obtained by treating interference as noise at the receivers, namely

$$C_{n,m}(\mathbf{P}(t), \mathbf{S}(t)) \propto \log \left(1 + \frac{P_{n,m}(t) S_{n,m}(t)}{N_0 + \sum_{l \in \mathcal{N} \setminus \{n\}} P_l(t) S_{l,m}(t)} \right) \quad (3)$$

where $S_{n,m}(t)$ represents the channel power gain on link (n, m) and N_0 is the noise spectral density. We assume that there ex-

ists some finite constant μ_{\max} such that $\mu_{n,m}(t) \leq \mu_{\max}$ for all t , any power allocation vector $\mathbf{P}(t)$, and channel state $\mathbf{S}(t)$. Moreover, following [7], we assume that the function $C_{n,m}(\mathbf{P}(t), \mathbf{S}(t))$ satisfies the following properties.

Property 1: For any power allocation matrix $\mathbf{P}(t)$, we have

$$C_{n,m}(\mathbf{P}(t), \mathbf{S}(t)) \leq \xi P_{n,m}(t) \quad (4)$$

for some finite constant $\xi > 0$;

Property 2: For any power allocation matrix $\mathbf{P}(t)$ and matrix $\mathbf{P}'(t)$ obtained by $\mathbf{P}(t)$ by setting the entry $P_{n,m}(t)$ to zero for a given (n, m) pair, we have

$$C_{a,b}(\mathbf{P}(t), \mathbf{S}(t)) \leq C_{a,b}(\mathbf{P}'(t), \mathbf{S}(t)) \quad (5)$$

for all $(a, b) \in \mathcal{L}$, with $(a, b) \neq (n, m)$.

Note that both properties are satisfied by typical choices of function $C_{n,m}(\mathbf{P}(t), \mathbf{S}(t))$ such as (3). In fact, Property 1 is satisfied if function $C_{n,m}(\mathbf{P}(t), \mathbf{S}(t))$ is concave with respect to $P_{n,m}(t)$, while Property 2 states that interference due to power spent on other links cannot be beneficial.² Finally, we define the total outgoing transmission rate $\mu_{n,*}(t)$ from a node $n \in \mathcal{N}$ at time t as

$$\mu_{n,*}(t) = \sum_{m:(n,m) \in \mathcal{L}} \mu_{n,m}(t) \quad (6)$$

and the total incoming transmission rate $\mu_{*,n}(t)$ at a node $n \in \mathcal{N}$ as

$$\mu_{*,n}(t) = \sum_{m:(m,n) \in \mathcal{L}} \mu_{m,n}(t). \quad (7)$$

Note that, in general, the way in which transmission and reception activities are scheduled within a time-slot depends on the specific channel access and physical-layer technology in use. In this paper, these aspects are modeled through a suitable capacity-power curve—see (3)—thus abstracting the operation of specific transmission protocols, as it is standard practice in this type of analysis.

C. Data Acquisition, Compression, and Distortion Model

At each time-slot, each node of the network is able to sense the environment and to acquire spatially correlated measurements. The measurements are then routed through the network to be gathered by a sink node, as illustrated in Fig. 1. Before transmission, the acquired data is compressed via adaptive lossy source coding by leveraging the spatial correlation of the measurements. Specifically, we define the *source state* at time t as the spatial correlation matrix describing the signal within this time-slot, which is referred to as $\mathbf{O}(t) = \{O_{n,m}(t)\}$ with $n, m \in \mathcal{N}$. We assume that $\mathbf{O}(t)$ takes values in some finite set $\mathcal{O} = \{\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_L\}$, remains constant within a time-slot, but is i.i.d. across time-slots. Additionally, we define the probability $\rho_{\mathbf{o}_i} = \Pr[\mathbf{O}(t) = \mathbf{o}_i]$. We remark that in this paper we propose optimal algorithms considering the evolution of $\mathbf{O}(t)$ to be known (causally) by all nodes. As an example, $\mathbf{O}(t)$ can represent a particular phase of an industrial process monitored by the wireless sensor network. The current phase is thus known to all nodes. In the case where $\mathbf{O}(t)$ is unknown, suitable solutions

²This may not be the case if sophisticated physical-layer techniques are used, such as successive interference cancellation (see, e.g., [5]).

must be devised for the distribution of information about $\mathbf{O}(t)$ to the nodes. We leave the study of these algorithms for future work. Each node $n \in \mathcal{N}$ compresses the measured source with rate $R_n(t) \leq R_{\max}$ bits per source symbol and targets a reproduction distortion at the sink of $D_{\min} \leq D_n(t) \leq D_{\max}$, with $0 < R_{\max}, D_{\min} \leq D_{\max} < \infty$. Note that imposing a strictly positive lower bound on $D_n(t)$ is without loss of generality because the rate $R_n(t)$ is upper-bounded by a finite constant and therefore the distortion $D_n(t)$ cannot in general be made arbitrarily small (see, e.g., [5]). The distortion is measured according to some fidelity criterion such as mean square error (MSE). We define the rate vector as $\mathbf{R}(t) = (R_1(t), \dots, R_N(t))$ and the distortion vector as $\mathbf{D}(t) = (D_1(t), \dots, D_N(t))$. Due to the spatial correlation of the measurements, *distributed source coding techniques* can be leveraged. Thanks to these techniques, the rates of different users can be traded without affecting the achievable distortions, to an extent that depends on the amount of spatial correlation [5]. The adoption of distributed source coding entails that, given certain distortion levels $\mathbf{D}(t)$, the rates $\mathbf{R}(t)$ can be selected arbitrarily as long as they satisfy appropriate joint constraints. Under such constraints, a sink receiving data at the specified rates is able to recover all sources at the given distortion levels.³

To elaborate on this point, consider the following conditions on the rates $R_n(t)$ and distortions $D_n(t)$ for $n \in \mathcal{N}$

$$\sum_{n \in \mathcal{X}} R_n(t) \geq g(\mathcal{X}, \mathbf{O}(t)) - \log \left((2\pi e)^{|\mathcal{X}|} \prod_{n \in \mathcal{X}} D_n(t) \right) \quad (8)$$

for all $\mathcal{X} \subseteq \mathcal{N}$, where $g(\mathcal{X}, \mathbf{O}(t))$ denotes the joint conditional differential entropy of the sources measured by the nodes in the subset \mathcal{X} , where conditioning is with respect to the sources measured by the nodes in the complement $\mathcal{N} \setminus \mathcal{X}$.

The case of jointly Gaussian sources is particularly significant since, for a given covariance matrix, jointly Gaussian sources require the highest rates to meet a given set of distortion constraints [28]. Thus, the results obtained by considering the Gaussian source model provide a worst-case analysis for sources with a given correlation matrix. We finally remark that the problem of characterizing the set of all achievable rate-distortion pairs in general is still open [5].

For jointly Gaussian sources with zero mean and correlation matrix $\mathbf{O}(t)$, the conditional entropy in (8) is given by

$$g(\mathcal{X}, \mathbf{O}(t)) = \frac{1}{2} \log \left(\frac{\det \mathbf{O}(t)}{\det \mathbf{O}(t)|_{\mathcal{N} \setminus \mathcal{X}}} \right) \quad (9)$$

where $\mathbf{O}(t)|_{\mathcal{N} \setminus \mathcal{X}}$ represents the correlation submatrix of the sources measured by nodes in $\mathcal{N} \setminus \mathcal{X}$. If the rates satisfy conditions (8), it is known [29] that for any well-behaved joint (analog) source distribution, the sink is able to recover all the sources within MSE levels $D_n(t)$, for all $n \in \mathcal{N}$. Moreover, the rate-distortion region (8) exhausts all possible achievable rate-distortion pairs if the distortion levels are small enough [29].

As an example, the rate region for $\mathcal{N} = \{1, 2\}$ is sketched in Fig. 3. The rates $R_1(t)$ and $R_2(t)$ at which the two source

³Note that we do not penalize for source coding delay because our concern is the average distortion perceived at the sink across time.

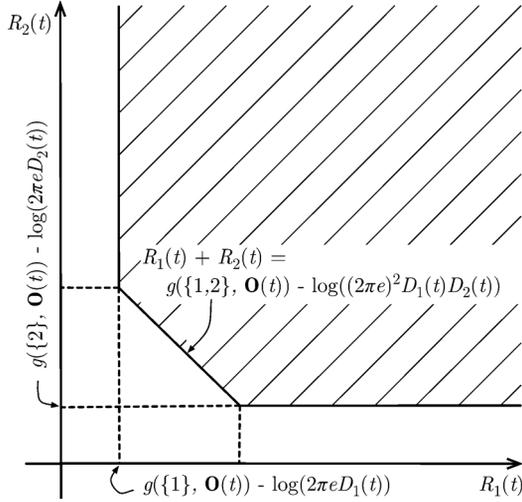


Fig. 3. Illustration of the rate region (8) for correlated sources and $\mathcal{N} = \{1, 2\}$. For all rate pairs $(R_1(t), R_2(t))$, there exists a coding scheme that enables the sink to recover the two sources with distributed distortion (MSE) levels $D_1(t)$ and $D_2(t)$, respectively.

sequences are acquired and compressed at the two nodes can be traded with one another without affecting the distortions of the reconstructions at the sink, as long as they remain in the shown rate region (8)

We account for the cost of source acquisition and compression by defining a function $P_n^c(R_n(t))$ that provides the power spent for compressing the acquired data at a particular rate $R_n(t)$. In this paper, we assume a linear relationship between the rate $R_n(t)$ and $P_n^c(R_n(t))$, as follows:

$$P_n^c(R_n(t)) = \alpha_n R_n(t) \quad (10)$$

for some coefficient $\alpha_n \geq 0$. This can be justified through the following argument. A vector quantizer of rate R , which is required at each node in order to achieve the rate-distortion performance (8), has a complexity per symbol of 2^{qR} (q is the block length). This is because optimal vector quantization requires to compare the source vector to all the 2^{qR} reconstruction vectors belonging to the quantization codebook. However, in practice, various schemes have reported smaller complexity orders (see, e.g., [30] and [31]). Thus, our assumption can be seen as a small rate approximation of the exponential order and as a pragmatic way to capture more efficient algorithms. Furthermore, recent results point to the possibility for polynomial close-to-optimal distributed compression strategies [32]. Finally, we remark that the destination d is assumed not to have sensing capabilities, and thus is not able to acquire any measurements. We will treat the extension to this setting in Section VI

D. Energy Model

Every node in the network is assumed to be powered via energy harvesting. The harvested energy is stored in an energy storage device, or battery, which is modeled as an energy queue, as in, e.g., [7]. The energy queue size $E_n(t)$ at a node $n \in \mathcal{N}$ measures the amount of energy left in the battery of a node n at the beginning of time-slot t . For convenience, we normalize the available energy to the number of channel uses (transmission symbols) per slot. Without loss of generality, we assume

unitary slot duration so that the amount of power consumed for transmission and acquisition/compression is equivalent to the energy spent in a time-slot. Therefore, at any time-slot t , the overall energy used at a node $n \in \mathcal{N}$ must satisfy the availability constraint

$$P_n(t) + P_n^c(R_n(t)) \leq E_n(t). \quad (11)$$

That is, the total consumed energy due to transmission and acquisition/compression must not exceed the energy available at the node.

We denote by $H_n(t) \leq H_{\max}$ the amount of energy harvestable by node n at time-slot t , and we define the vector $\mathbf{H}(t) = (H_1(t), \dots, H_N(t))$ as the *energy-harvesting state*. We assume that $\mathbf{H}(t)$ takes value in a finite set $\mathcal{H} = \{\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_P\}$ and is constant for the duration of a time-slot, but i.i.d. over time-slots. Finally, we define the probability $\rho_{\mathbf{h}_i} = \Pr[\mathbf{H}(t) = \mathbf{h}_i]$, which accounts for possible spatial correlation of the harvestable energy across different nodes.

The energy harvested at time t is assumed to be available for use at time $t + 1$. Moreover, each node $n \in \mathcal{N}$ can decide how much of the harvestable energy $H_n(t)$ to store in the battery at time-slot t , and we denote the harvesting decision by $\tilde{H}_n(t)$, with $0 \leq \tilde{H}_n(t) \leq H_n(t)$. We define the harvesting decision vector as $\tilde{\mathbf{H}}(t) = (\tilde{H}_1(t), \dots, \tilde{H}_N(t))$. Variable $\tilde{H}_n(t)$ is introduced, following [7], to address the issue of assessing the needs of the system in terms of capacities of the energy storage devices. In fact, as in [7], we do not make any assumption about the battery maximum size. However, it will be proved later that performance arbitrarily close to the optimal attainable with no limitations on the battery capacity can be achieved with finite-capacity batteries.

Remark 2.1: Note that the energy model does not account for contributions to the energy consumption that are constant and depend only on whether a node is active or not, such as the energy used in the transmit and receive circuitry [13], [33]. However, these can be incorporated into the model at the expense of a more complicated notation and by properly adjusting the performance analysis. We will not further elaborate on this aspect here.

E. Queueing Dynamics

We now detail the dynamics of the network queues. We define $\mathbf{E}(t) = (E_1(t), \dots, E_N(t))$ to be the vector of the energy queue sizes of all nodes at time t . From the discussion above, for each node $n \in \mathcal{N}$, $E_n(t)$ evolves as

$$E_n(t+1) = E_n(t) - P_n(t) - P_n^c(R_n(t)) + \tilde{H}_n(t) \quad (12)$$

since at each time-slot t , the energy $P_n(t) + P_n^c(R_n(t))$ is consumed, while energy $\tilde{H}_n(t)$ is harvested. We assume $E_n(0) \geq 0$ for all $n \in \mathcal{N}$.

We also define the vector $\mathbf{U}(t) = (U_1(t), \dots, U_N(t))$, for each time-slot t , to be the network data queue backlog, where $U_n(t)$ represents the amount of data queued at node n , which is normalized on the number of channel uses per time-slot for convenience of notation, which is expressed in terms of bits over channel uses per slot. Denote as b the ratio between the number

of channel uses per slot and the number of source samples per slot. Since b typically accounts for the ratio of the channel and source bandwidth, it is conventionally referred to as *bandwidth ratio* [5]. We assume that each queue $U_n(t)$ evolves according to the following dynamics:

$$U_n(t+1) \leq \max\{U_n(t) - \mu_{n,*}(t), 0\} + \mu_{*,n}(t) + \frac{R_n(t)}{b} \quad (13)$$

since at any time-slot t , each node $n \in \mathcal{N}$ can transmit, and thus remove from its data queue, at most $\mu_{n,*}(t)$ bits per channel use, while it can receive at most $\mu_{*,n}(t)$ bits per channel use due to transmissions from other nodes and $R_n(t)/b$ bits per channel use due to data acquisition/compression. We assume that $U_n(0) = 0$ for all $n \in \mathcal{N}$. Following standard definitions [34], we say that the network is stable if the following condition holds true:

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{n \in \mathcal{N}} \mathbb{E}[U_n(t)] < \infty. \quad (14)$$

Notice that the network stability condition (14) implies that the data queue $U_n(t)$ of each node $n \in \mathcal{N}$ is stable in the sense that $\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}[U_n(t)] < \infty$.

F. Optimization Problem

Define $\Theta(t) = (\mathbf{S}(t), \mathbf{O}(t), \mathbf{H}(t), \mathbf{U}(t), \mathbf{E}(t))$ as the state of the network at time-slot t . A (past-dependent) policy $\pi = \{\pi(t): t = 1, 2, \dots\}$ is a collection of mappings between the past and current states $\{\Theta(\tau): \tau = 1, \dots, t\}$ and the current decision $(\mathbf{R}(t), \mathbf{D}(t), \tilde{\mathbf{H}}(t), \mathbf{P}(t))$ on rates $\mathbf{R}(t)$, distortion levels $\mathbf{D}(t)$, harvested energy $\tilde{\mathbf{H}}(t)$, and transmission powers $\mathbf{P}(t)$. Moreover, for each node $n \in \mathcal{N}$, let $f_n(D_n(t))$ denote the cost incurred by node n when its corresponding distortion is $D_n(t)$. We assume that each function $f_n(D_n(t))$ is convex, finite, and nondecreasing in the interval $[D_{\min}, D_{\max}]$. Our objective is to solve the following optimization problem:

$$\underset{\pi}{\text{minimize}} F_0^\pi = \sum_{n \in \mathcal{N}} F_n^\pi \quad (15)$$

where

$$F_n^\pi = \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}[f_n(D_n(t))] \quad (16)$$

subject to the rate-distortion constraints (8), the energy availability constraint (11), and network stability constraint (14). Note that (16) is the per-slot average cost for node n .

III. LOWER BOUND

In this section, we obtain a lower bound on the optimal network cost F_0^* of problem (15). This result will be used in Section V to obtain analytical performance guarantees on our online optimization policy, presented in Section IV. The lower bound is expressed in terms of an optimization problem over parameters $\mathbf{R}^{(\mathbf{o}_i)} = [R_1^{(\mathbf{o}_i)}, \dots, R_N^{(\mathbf{o}_i)}]$ and $\mathbf{D}^{(\mathbf{o}_i)} = [D_1^{(\mathbf{o}_i)}, \dots, D_N^{(\mathbf{o}_i)}]$ for all $\mathbf{o}_i \in \mathcal{O}$, $\mathbf{P}^{(\mathbf{s}_j)}$ with entries $P_{n,m}^{(\mathbf{s}_j)}$ for each $(n, m) \in \mathcal{L}$ and for all $\mathbf{s}_j \in \mathcal{S}$, and $\tilde{\mathbf{H}}^{(\mathbf{h}_k)} = [\tilde{H}_1^{(\mathbf{h}_k)}, \dots, \tilde{H}_N^{(\mathbf{h}_k)}]$ for all $\mathbf{h}_k \in \mathcal{H}$. The proof is based on relaxing the stability constraint (14) by imposing the necessary condition that the average arrival rate at each data queue be smaller than or equal to the average departure rate, and by also relaxing the energy availability constraint (11) by requiring it to be satisfied only on average. Finally, Lagrange relaxation is used on the resulting problem.

Theorem 3.1: The optimal network cost F_0^* satisfies the following inequality:

$$VF_0^* \geq d(\boldsymbol{\lambda}, \mathbf{v}, \boldsymbol{\chi}) \quad (17)$$

for all $\boldsymbol{\lambda} \in \mathbb{R}_+^{L(2^N-1)}$, $\mathbf{v} \in \mathbb{R}_+^N$, $\boldsymbol{\chi} \in \mathbb{R}^N$, where $V > 0$ is a constant and $d(\boldsymbol{\lambda}, \mathbf{v}, \boldsymbol{\chi})$ is given by

$$d(\boldsymbol{\lambda}, \mathbf{v}, \boldsymbol{\chi}) = \sum_{\mathbf{o}_i \in \mathcal{O}} \rho_{\mathbf{o}_i} \sum_{\mathbf{s}_j \in \mathcal{S}} \rho_{\mathbf{s}_j} \sum_{\mathbf{h}_k \in \mathcal{H}} \rho_{\mathbf{h}_k} d_{\mathbf{o}_i, \mathbf{s}_j, \mathbf{h}_k}(\boldsymbol{\lambda}^{(\mathbf{o}_i)}, \mathbf{v}, \boldsymbol{\chi}) \quad (18)$$

with $d_{\mathbf{o}_i, \mathbf{s}_j, \mathbf{h}_k}(\boldsymbol{\lambda}^{(\mathbf{o}_i)}, \mathbf{v}, \boldsymbol{\chi})$ defined in (19) at the bottom of the page, where the infimum is taken under constraints

$$0 \leq R_n^{(\mathbf{o}_i)} \leq R_{\max}, \quad D_{\min} \leq D_n^{(\mathbf{o}_i)} \leq D_{\max}, \quad \text{for all } n \in \mathcal{N}, \mathbf{o}_i \in \mathcal{O} \quad (20)$$

$$0 \leq P_n^{(\mathbf{s}_j)} \leq P_{\max}, \quad \text{for all } n \in \mathcal{N}, \mathbf{s}_j \in \mathcal{S} \quad (21)$$

$$0 \leq \tilde{H}_n^{(\mathbf{h}_k)} \leq h_{k,n}, \quad \text{for all } n \in \mathcal{N}, \mathbf{h}_k \in \mathcal{H}. \quad (22)$$

Proof: The proof follows from a similar approach as in [15, Theorem 1]. The complete derivation can be found in our technical report [35]. ■

$$d_{\mathbf{o}_i, \mathbf{s}_j, \mathbf{h}_k}(\boldsymbol{\lambda}^{(\mathbf{o}_i)}, \mathbf{v}, \boldsymbol{\chi}) = \inf_{\mathbf{R}^{(\mathbf{o}_i)}, \mathbf{D}^{(\mathbf{o}_i)}, \mathbf{P}^{(\mathbf{s}_j)}, \tilde{\mathbf{H}}^{(\mathbf{h}_k)}} \left\{ \sum_{n \in \mathcal{N}} V f_n(D_n^{(\mathbf{o}_i)}) + \sum_{m=1}^{2^N-1} \lambda_m^{(\mathbf{o}_i)} \left[g(\mathcal{X}_m, \mathbf{o}_i) - \log \left((2\pi e)^{|\mathcal{X}_m|} \prod_{n \in \mathcal{X}_m} D_n^{(\mathbf{o}_i)} \right) \right] - \sum_{n \in \mathcal{X}_m} R_n^{(\mathbf{o}_i)} \right\} + \sum_{n \in \mathcal{N}} v_n \left[\frac{R_n^{(\mathbf{o}_i)}}{b} + \mu_{*,n}(\mathbf{P}^{(\mathbf{s}_j)}, \mathbf{s}_j) - \mu_{n,*}(\mathbf{P}^{(\mathbf{s}_j)}, \mathbf{s}_j) \right] + \sum_{n \in \mathcal{N}} \chi_n \left[P_n^{(\mathbf{s}_j)} + P_n^c(R_n^{(\mathbf{o}_i)}) - \tilde{H}_n^{(\mathbf{h}_k)} \right] \quad (19)$$

IV. PROPOSED POLICY

In this section, we propose an algorithm designed following the Lyapunov optimization framework, as developed in [27] and [34], to solve the optimization problem (15). In particular, we aim at finding an online policy π for problem (15) with close-to-optimal performance by using Lyapunov optimization with weight perturbation. The technique of weight perturbation, as proposed in [7], is used to ensure that the energy queues are kept close to a target value. This is done to avoid battery underflow in a way that is reminiscent of the battery management strategies put forth in [14], and it is further discussed in the following.

The proposed policy operates by approximately minimizing, at each time-slot, the one-slot conditional Lyapunov drift plus penalty [34] of the energy and data queues [(12) and (13), respectively] of the network. The optimization is done in an online fashion based on the knowledge of the current channel state $\mathbf{S}(t)$, observation state $\mathbf{O}(t)$, data queue sizes $\mathbf{U}(t)$, and energy queue sizes $\mathbf{E}(t)$. Note that no knowledge of the statistics of the states is required, as it is standard with Lyapunov optimization techniques [27], [34]. Using this approach, we obtained the following online optimization algorithm.

Algorithm: Fix a weight $\boldsymbol{\theta} = [\theta_1, \dots, \theta_N] \in \mathbb{R}_+^N$ and a constant $V > 0$. The parameter θ_n effectively limits the battery size of node n and is hereby considered as a design parameter. On the one hand, this parameter has the advantage of allowing to use Lyapunov optimization techniques, as first demonstrated by [7]; on the other, it enables the evaluation of bounds on the capacity of the batteries that allow to achieve a performance close to the optimal attainable with infinite batteries (see Theorem 5.1). At each time-slot t , based on the values of the queues $\mathbf{E}(t)$ and $\mathbf{U}(t)$, channel states $\mathbf{S}(t)$ and observation states $\mathbf{O}(t)$, perform the following.

- *Energy Harvesting:* For each node $n \in \mathcal{N}$, choose $\tilde{H}_n(t)$ that minimizes $(E_n(t) - \theta_n)\tilde{H}_n(t)$ under the constraint $0 \leq \tilde{H}_n(t) \leq H_n(t)$. That is, if the residual energy that can be stored in the energy buffer is $(\theta_n - E_n(t)) > 0$, perform energy harvesting and store the harvested energy in the energy buffer, i.e., set $\tilde{H}_n(t) = \min\{\theta_n - E_n(t), H_n(t)\}$; otherwise, perform no harvesting, i.e., set $\tilde{H}_n(t) = 0$.
- *Rate-Distortion Optimization:* Choose the source acquisition/compression rate vector $\mathbf{R}(t) = \mathbf{r} = [r_1, \dots, r_N]$ and the distortion levels $\mathbf{D}(t) = \mathbf{d} = [d_1, \dots, d_N]$ to be an optimal solution of the following optimization problem:

$$\underset{\mathbf{r}, \mathbf{d}}{\text{minimize}} \sum_{n \in \mathcal{N}} [U_n(t)r_n - (E_n(t) - \theta_n)P_n^c(r_n) + V\mathfrak{f}_n(d_n)] \quad (23)$$

subject to the rate-distortion region constraint (8), and to the constraints $0 \leq r_n \leq R_{\max}$ and $D_{\min} \leq d_n \leq D_{\max}$, for all $n \in \mathcal{N}$.

- *Power Allocation:* Define the weight of a link (n, m) as

$$W_{n,m}(t) = \max\{U_n(t) - U_m(t) - \delta, 0\} \quad (24)$$

where $\delta = l_{\max}\mu_{\max} + R_{\max}$, and choose $\mathbf{P}(t) = \mathbf{p}$ with entries $p_{n,m}$ for $(n, m) \in \mathcal{L}$ to be an optimal solution of the following optimization problem:

$$\underset{\mathbf{p}}{\text{maximize}} \sum_{n \in \mathcal{N}} \left[\sum_{m \in \mathcal{N} \setminus n} C_{n,m}(\mathbf{p}, \mathbf{S}(t))W_{n,m}(t) + (E_n(t) - \theta_n)p_n \right] \quad (25)$$

where $p_n = \sum_{m \in \mathcal{N} \setminus n} p_{n,m}$, subject to constraints $0 \leq p_n \leq P_{\max}$, for each $n \in \mathcal{N}$;

- *Queues update:* Update $\mathbf{E}(t)$ and $\mathbf{U}(t)$ according to (12) and (13), respectively.

Remark 4.1: In the algorithm proposed above, the energy availability constraint (11) is not explicitly imposed. However, as discussed in Section V, with a proper choice of the weight vector $\boldsymbol{\theta}$, the battery levels are guaranteed to be such that condition (11) is never violated. In other words, the effect of the weight vector $\boldsymbol{\theta}$ is to ensure that, whenever the algorithm requires to draw energy from the batteries for transmission or acquisition/compression, there is energy available at the corresponding nodes to satisfy the request.

Remark 4.2: Solving the optimization problems (23) and (25) requires slot synchronization that, in practical implementations, can be difficult to achieve. For this reason, the performance obtained in this paper represents a lower bound, in terms of both average distortion and queue sizes, to any existing implementation that waives our idealistic assumptions.

A. Discussion About Distributed Optimization and Complexity

Distributed Optimization: While the Energy Harvesting step can be performed independently by all nodes, the Rate-Distortion Optimization problem (23) and the Power Allocation problem (25) require centralized optimization. Decentralized implementations of the Power Allocation problem (25) are discussed in many papers—see, e.g., [36]—while it is possible to solve the Rate-Distortion Optimization problem (23) in a distributed fashion via dual decomposition [37], [38] as discussed in [25] and [39]. It is noted that the dual decomposition solution, due to the rate-distortion constraints (8), would require the centralized update of the dual variables, which might be challenging in practical implementation.

Complexity: It can be noticed that the Rate-Distortion optimization (23) is computationally intensive since it involves an exponential number of constraints to be satisfied [i.e., (8)]. A general approach to keep this number of constraints as small as possible would be to divide large networks into clusters of fixed maximum size and apply the distributed source coding technique within the cluster. Moreover, in the Rate-Distortion optimization (23), for fixed d_n , the polyhedron described by the constraints in (8) is a contra-polymatroid [40], and thus from [40, Lemma 3.3], a greedy algorithm solves (23) optimally (see, e.g., [25] and [26]).

V. PERFORMANCE ANALYSIS

In this section, we provide analytical insights into the performance of the policy proposed in Section IV. To this end, we define the parameters $\beta_n = \min\{\alpha_n, 1\}$ [recall (10)] and $\gamma_n = \sup_{D_{\min} \leq d_n \leq D_{\max}} \left[\frac{f_n(d_n) - f_n(D_{\max})}{\log(d_n/D_{\max})} \right]$, which is finite under the given assumptions.

Theorem 5.1: Under the proposed algorithm with $\theta = [\theta_1, \dots, \theta_N]$, where $\theta_n = \frac{\gamma_n}{\beta_n}V + \alpha_n R_{\max} + P_{\max}$, we have the following.

- 1) The data queue and the energy queue of all nodes are bounded as

$$0 \leq E_n(t) \leq \theta_n \quad (26)$$

$$0 \leq U_n(t) \leq \gamma_n V + R_{\max} \quad (27)$$

respectively, for all nodes $n \in \mathcal{N}$ and all times t .

- 2) When a node $n \in \mathcal{N}$ allocates a nonzero power to any of its outgoing links (i.e., $P_n(t) > 0$), and/or when it chooses a nonzero source acquisition rate (i.e., $R_n(t) > 0$), thus expending energy for source acquisition/compression, we have that

$$E_n(t) \geq \alpha_n R_{\max} + P_{\max}. \quad (28)$$

This condition guarantees that the energy availability constraint (11) is satisfied for all nodes $n \in \mathcal{N}$ and all times t (see Remarks 4.1 and 5.2).

- 3) The overall cost F_0^π (15) achieved by the proposed scheme satisfies the bound

$$F_0^\pi = \sum_{n \in \mathcal{N}} F_n^\pi \leq F_0^* + \frac{B}{V} \quad (29)$$

where F_0^* is the optimal cost of problem (15) and the finite constant B is $B = N(\mu_{\max}(\mu_{\max} + R_{\max}) + R_{\max}^2/2) + N/2(H_{\max}^2 + \alpha_{\max}^2 R_{\max}^2 + P_{\max}^2 + 2\alpha_{\max} R_{\max} P_{\max}) + N(\delta l_{\max} \mu_{\max} + H_{\max}^2/4)$.

Proof: See Appendix A. ■

Remark 5.2: The fact that (28) implies that the proposed algorithm satisfies the energy availability constraint (11) at each time-slot follows since each node $n \in \mathcal{N}$ cannot consume an energy larger than $\alpha_n R_{\max} + P_{\max}$ in a time-slot. In fact, $\alpha_n R_{\max}$ is the maximum energy spent for compressing the acquired data, and P_{\max} is the maximum transmission energy consumption.

Remark 5.3: Following [16], under the modified stability requirement $\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^{T-1} U_n(t) < \infty$, for all $n \in \mathcal{N}$, the proposed algorithm can be proven to guarantee near-optimal performance with probability one.

Remark 5.4: To elaborate on the technical challenges solved in the proof of Theorem 5.1, we observe that a key step is the determination of a suitable vector of thresholds θ_n that allows the desired performance described by (26)–(29) to be attained. We recall that the parameters θ_n are used by each node $n \in \mathcal{N}$ to manage its energy queue. This problem was tackled in [15] for the case where the sources have to be delivered losslessly and are independent, so that distortion constraints are not an issue and distributed source coding is not applicable. The scenario

considered here is instead more involved due to the fact that the rate-distortion constraints (28), arising from the use of distributed source coding, couple the rate-distortion optimization and the power allocation problem at all nodes. We have tackled this problem by relaxing the rate-distortion constraints (8) and finding a way to bound the optimal Lagrangian multiplier of the corresponding dual problem. This is discussed in Lemma A.1 in Appendix A [see conditions (41) and (42)].

VI. EXTENSION WITH SIDE INFORMATION AT THE SINK

We now consider an extended version of the problem studied thus far, in which the sink node d , rather than being the final destination for the sources measured at the sensors, acts as a cluster head and communicates to a network collector node c (see Fig. 1), on a communication link modeled as for any other pair of nodes (see Section II-B). The key novel aspect of this extended model is that node d can measure a source correlated with that of the sensors and use such side information to improve the system performance. Specifically, thanks to the side information available at node d , the rate requirements for communication from the sensors to d can be reduced. However, node d , which is powered by energy harvesting as are all the sensors, also needs to communicate with node c . Therefore, a new tradeoff arises between the energy allocated by d to acquire side information and that used by d to communicate with c . With the following extension, we investigate the impact of allowing the acquisition of side information at the sink, where the amount of side information can be controlled subject to the sensing and compression costs.

We now discuss how the model discussed in Section II needs to be modified in order to account for the different setting of interest here. First, the destination d acquires a source signal that is correlated with the sensor's measures with a rate $R_d(t)$. This affects the rate-distortion constraints (8) in that the entropy function $g(\mathcal{X}, \mathbf{O}(t))$ should now be conditioned on the side information available at the receiver (see, e.g., [41]). This leads to modified rate-distortion constraints (8) with a function $g(\mathcal{X}, \mathbf{O}(t), R_d(t))$ that also depends on $R_d(t)$. An example of this function will be given in Section VII. The energy used for acquiring the side information is given by $P_d^c(R_d(t)) = \alpha_d R_d(t)$ and the slot duration, similar to all other nodes. Moreover, the data queue at node d evolves as

$$U_d(t+1) \leq \max\{U_d(t) - \mu_{d,c}(t), 0\} + \mu_{*,d}(t) \quad (30)$$

where $\mu_{d,c}(t)$ and $\mu_{*,d}(t)$ represent, respectively, the transmitted and received data at time t , and transmission is to the collector node c . Note that no other node is connected to the network collector c apart from d . The energy queue $E_d(t)$, instead, evolves according to (12). Finally, $\mathbf{P}(t)$ and $\mathbf{S}(t)$ are extended to consider the additional link $(d, c) \in \mathcal{L}$ and the rate achievable on that link is given by $C_{d,c}(\mathbf{P}(t), \mathbf{S}(t))$, which is assumed to have the same properties as for all other links (see Section II). We refer to the power used for transmission by node d as P_d .

In what follows, we modify the algorithm proposed in Section IV in order to address the new setting outlined above. The modified algorithm works as follows.

- *Energy Harvesting*: Follow the same procedure as for the algorithm discussed in Section IV, for all nodes including node d .
- *Rate-Distortion Allocation*: Choose $R_n(t)$ and $D_n(t)$, $n = 1, \dots, N$, and $R_d(t)$ to be the optimal solution of the following optimization problem:

$$\begin{aligned} \underset{(r,d),r_d}{\text{minimize}} \quad & \sum_{n \in \mathcal{N}} [U_n(t)r_n - (E_n(t) - \theta_n)P_n^c(r_n) + Vf_n(d_n)] \\ & + (E_d(t) - \theta_d)P_d^c(r_d) \end{aligned} \quad (31)$$

subject to $\sum_{n \in \mathcal{X}} r_n \geq g(\mathcal{X}, \mathbf{O}(t), r_d) - \log((2\pi e)^{|\mathcal{X}|} \prod_{n \in \mathcal{X}} d_n)$, $\forall \mathcal{X} \subseteq \mathcal{N}$, $0 \leq r_n \leq R_{\max}$ and $D_{\min} \leq d_n \leq D_{\max}$, $n \in \mathcal{N}$ and $0 \leq r_d \leq R_{\max}$;

- *Power Allocation*: Define the weight of a link $(n, m) \in \mathcal{L}$ as⁴ (24) and choose $\mathbf{P}(t) = \mathbf{p}$ with entries $p_{n,m}$ for $(n, m) \in \mathcal{L}$ to be an optimal solution of the following optimization problem:

$$\begin{aligned} \underset{\mathbf{P}}{\text{maximize}} \quad & \sum_{n \in \mathcal{N}} \left[\sum_{m \in \mathcal{N} \setminus n} C_{n,m}(\mathbf{p}, \mathbf{S}(t))W_{n,m}(t) \right. \\ & \left. + (E_n(t) - \theta_n)p_n \right] + C_{d,c}(\mathbf{p}, \mathbf{S}(t))W_{d,c}(t) \\ & + (E_d(t) - \theta_d)p_d \end{aligned} \quad (32)$$

subject to $0 \leq p_n \leq P_{\max}$, for each $n \in \mathcal{N} \cup \{d\}$.

- *Queues Update*: Update $\mathbf{E}(t)$ and $E_d(t)$ according to (12), $\mathbf{U}(t)$ according to (13), and $U_d(t)$ using (30).

The algorithm proposed above is a simple modification of the algorithm proposed in Section IV that accounts for the need to allocate rate and power also for node d . It can be proven that this algorithm has similar optimality properties as the algorithm of Section IV, as summarized in Theorem 5.1. We omit a formal statement of this result here since it is a straightforward extension of Theorem 5.1.

VII. NUMERICAL RESULTS

In this section, we provide further insights into the performance of the system under study, via some numerical results. We consider the network topology of Fig. 1, where the set \mathcal{N} of nodes gathers spatially correlated data and transmits them to the sink node d . We first consider the setup without side information at the sink described in Section II. We assume that nodes $\{1, 2, 3\}$ collect the measurements, while nodes $\{4, 5\}$ are only used as relays (or equivalently measure zero-power sources). The signal samples measured at nodes $\{1, 2, 3\}$ are jointly Gaussian with zero mean and time-independent correlation matrix

$$\mathbf{O}(t) = \begin{bmatrix} 1 & \omega & \omega \\ \omega & 1 & \omega \\ \omega & \omega & 1 \end{bmatrix} \quad (33)$$

where $\omega \in [-1, 1]$ is the spatial correlation coefficient. The channel-state matrix $\mathbf{S}(t)$ has independent entries that for every

⁴We remind that \mathcal{L} is extended to consider the link (d, c) .

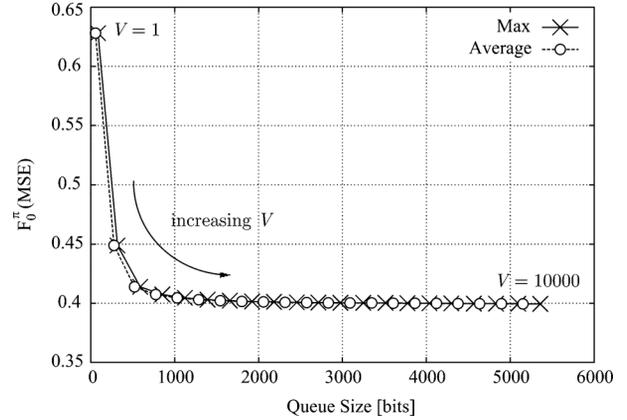


Fig. 4. F_0^π versus maximum and average network queue size for $1 \leq V \leq 10000$. Each pair of values for sum-distortion and queue size is obtained for a different value of V , from 1 to 10 000, with step length 500. ($\omega = 0.5$).

link (n, m) are Rayleigh distributed with the same fading power $\mathbb{E}[S_{n,m}(t)] = 1$, while the energy-harvesting vector $\mathbf{H}(t)$ has independent entries that are uniformly distributed in $[0, H_{\max}]$, with $H_{\max} = 3$. Both channel and energy-harvesting statistics are i.i.d. across time-slots.

For the channel capacity function, we consider the Shannon capacity given by (3), while the entropy function is given by (9) and the cost function is $f_n(D_n(t)) = D_n(t)$ for all $n \in \mathcal{N}$. Moreover, we set the numerical values $\alpha_n = 1$ for all $n \in \mathcal{N}$, $D_{\min} = 0.001$, and $P_{\max} = \alpha_{\max} R_{\max}$, with $R_{\max} = g(\{1, 2, 3\}, \mathbf{O}(t)) - \log((2\pi e D_{\min})^3)$. In what follows, we refer to *network queue size* as the sum of the queue sizes of all nodes in \mathcal{N} .

We first examine the effect of constant V , which was shown in Theorem 5.1 to characterize the $(V, 1/V)$ tradeoff between the network queue size and the additive gap with respect to the lower bound of Theorem 3.1. To this end, in Fig. 4, we set $\omega = 0.5$ and plot the average sum-distortion F_0^π as a function of the maximum and average network queue size for different value of V . Confirming the results of Theorem 5.1, we observe that the sum-distortion F_0^π gradually converges to the lower bound set by the optimal value F_0^* for increasing V . A closer inspection of the results also reveals an almost linear increase of the maximum and time average network queue size with respect to V , as suggested by Theorem 5.1 (not shown).

Next, we evaluate the impact of the spatial correlation parameter ω . As discussed, an increasing ω is expected to lead to a reduction in the energy consumption for the same reconstruction accuracy at the sink thanks to the spatial energy tradeoffs enabled by distributed source coding. This is confirmed by the results in Fig. 5, where we plot the sum-distortion F_0^π versus the average and maximum network queue size, where each point is obtained for a different value of the correlation ω in $[0, 1]$. We note that an increasing ω leads to a reduction of both the network queue size and F_0^π . Note that the performance with $\omega = 0$ corresponds to that of a conventional source coding system (i.e., not leveraging distributed source coding) as, in this case, distributed source coding does not offer any advantage and reduces to conventional compression. Thus, comparison between the performance with $\omega = 0$ and $\omega > 0$ reveals the gain of leveraging

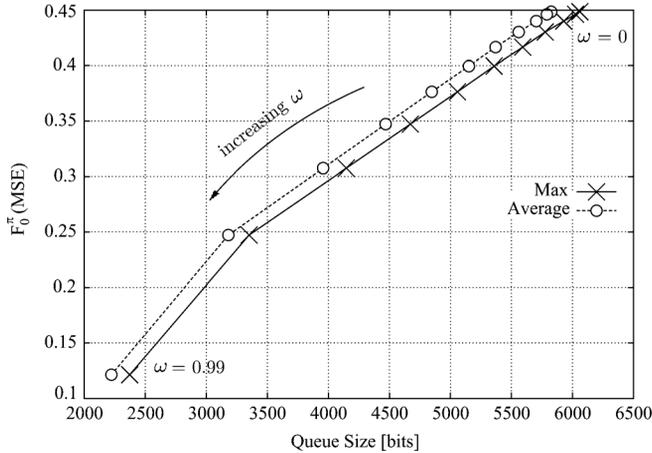


Fig. 5. F_0^π versus maximum and average network queue size for different values of source correlation ω , from 0 to 0.99, with step length 0.1. ($V = 10\,000$).

distributed source coding. Note that this gain is quite substantial, leading in the best case ($\omega \rightarrow 1$) to a decrease of a factor 3.7 in terms of distortion and of a factor 2.6 in terms of queue size at the nodes.

Finally, we evaluate the performance in the scenario of Section VI, where the sink node d acts as a cluster head, that measures a source correlated with that of the remaining sensors and communicates the gathered data to node c (see Fig. 1). To this end, we replace the entropy function $g(\mathcal{X}, \mathbf{O}(t))$ with a function $g(\mathcal{X}, \mathbf{O}(t), R_d(t))$ that takes into account the side information obtained by d with rate $R_d(t)$. We recall that $R_d(t)$ is a decision variable of the new problem; see (31). Following [41], the function $g(\mathcal{X}, \mathbf{O}(t), R_d(t))$ is given by (9) where the correlation matrix (33) should now be conditioned on the side information available at the destination [41]. According to the simple source model described in Appendix C, we assume that this conditional covariance matrix is given by

$$\mathbf{O}(t) = \begin{bmatrix} 1 - \omega\omega_d(t) & \omega(1 - \omega_d(t)) & \omega(1 - \omega_d(t)) \\ \omega(1 - \omega_d(t)) & 1 - \omega\omega_d(t) & \omega(1 - \omega_d(t)) \\ \omega(1 - \omega_d(t)) & \omega(1 - \omega_d(t)) & 1 - \omega\omega_d(t) \end{bmatrix} \quad (34)$$

where $\omega_d(t) = 1 - 2^{-R_d(t)}$. We consider the same simulation parameters as above and we additionally set $\alpha_d = 1$ and, only for node d , $H_{\max} = 12$.

Fig. 6 shows the sum-distortion F_0^π and the average network queue size versus $\omega \in [0, 1)$. As a reference, we compare the performance of the proposed algorithm to that of a scheme that sets $R_d(t) = 0$. This scheme, therefore, does not acquire side information at the sink and instead utilizes all the available energy at the sink for transmission to node c . As expected, having additional side information at the sink is beneficial in terms of overall performance gains (queues size and distortion). From the results of Fig. 6, it can be seen that this leads to improvements of 44% and 32% in terms of queue size length and distortion F_0^π , respectively.

VIII. CONCLUSION

Energy harvesting poses new challenges in terms of energy management of wireless networks. In sensor networks, these

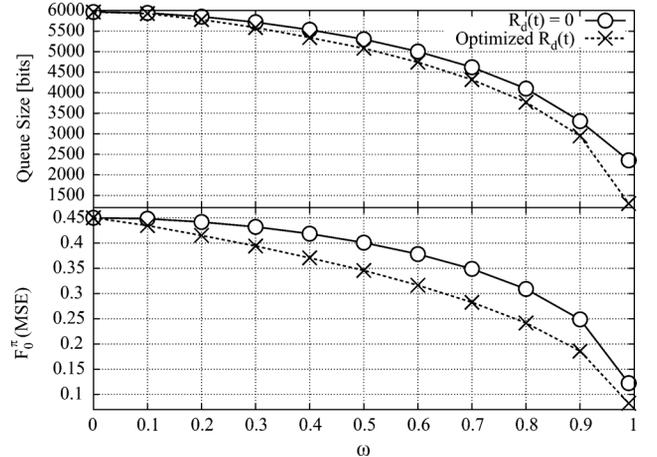


Fig. 6. F_0^π and average network queue size versus source correlation ω . ($V = 10\,000$).

challenges are compounded by the need for balancing the energy consumed by source coding tasks (i.e., data compression) against that used for transmission. Moreover, the correlation among the data readings collected by different sensors, if leveraged via distributed source coding, makes it possible to exploit spatial energy tradeoffs across the sensors, thus allowing for better performance in terms of memory usage and distortion at the sink. Based on the above, this work has proposed a dynamic online optimization strategy for multihop wireless sensor networks with energy-harvesting capabilities. This strategy jointly optimizes source coding and data transmission activities for time-varying sources and channels by ensuring queue stability at the nodes and energy neutrality. The proposed technique, based on Lyapunov optimization, has been analytically shown to be characterized by a $(V, 1/V)$ tradeoff, revealing a linear relationship for queue and battery size (V) and an inverse proportionality in terms of optimality gap ($1/V$), where V is a tunable parameter of the algorithm. Numerical results have demonstrated the key role of source correlation and distributed source coding in the system performance. In addition, they have shown substantial room for improvement in terms of reduction of queue sizes and corresponding increased data reconstruction accuracy when distributed source coding is applied in conjunction with transmission scheduling techniques. These facts motivate the design of practical algorithms to achieve such goals. An aspect that calls for further research is, for instance, the design of algorithms that operate with partial information about the source correlation structure. As further possible avenues for future work, one can consider the distributed implementation of the proposed strategies, compression techniques that aggregate data at intermediate nodes, and joint source-channel coding transmission techniques exploiting analog transmissions [42], [43].

APPENDIX A PROOF OF THEOREM 5.1

Proof:

1) From the energy-harvesting part of the algorithm, we have that $E_n(t) \leq \theta_n$ since harvesting is performed only when $E_n(t) < \theta_n$ and the maximum amount of harvested energy

in that case is $\theta_n - E_n(t)$. This proves (26). We now prove (27) by induction on t . Inequality (27) holds for $t = 0$ since $U_n(0) = 0$ for all n . Then, assuming that (27) is satisfied for all n at time t , we show that it holds also for time $t + 1$. To this end, we consider separately the different possible cases in which a node n receives or does not receive data from other nodes (i.e., endogenous data) and/or acquires or does not acquire its measurement (i.e., exogenous data). First, if node n receives neither endogenous nor exogenous data, then we have that $U_n(t + 1) \leq U_n(t) \leq \gamma_n V + R_{\max}$, which proves the claim. Second, assume that node $n \in \mathcal{N}$ receives endogenous, but not exogenous, data. It follows from (25) that, for some node $m \in \mathcal{N}$, with $m \neq n$, we must have

$$U_n(t) \leq U_m(t) - \delta \leq \gamma_n V + R_{\max} - \delta. \quad (35)$$

However, since any node can receive at most $l_{\max}\mu_{\max}$ bits per channel use of endogenous data, we have from (35) and the definition of δ that $U_n(t + 1) \leq \gamma_n V \leq \gamma_n V + R_{\max}$, which proves the claim.

We now analyze the case where node n receives exogenous, but not endogenous, data. This implies that $r_n > 0$ is obtained from the solution of problem (23). We define the corresponding Lagrangian function as

$$\begin{aligned} \mathcal{L}(\mathbf{r}, \mathbf{d}, \boldsymbol{\lambda}, \mathbf{v}) &= \sum_{n \in \mathcal{N}} [U_n(t)r_n - (E_n(t) - \theta_n)P_n^c(r_n) + V f_n(d_n)] \\ &+ \sum_m \lambda_m \left[g(\mathcal{X}_m, \mathbf{O}(t)) - \log \left((2\pi e)^{|\mathcal{X}_m|} \prod_{l \in \mathcal{X}_m} d_l \right) \right. \\ &\quad \left. - \sum_{l \in \mathcal{X}_m} r_l \right] + \sum_{n \in \mathcal{N}} v_n (d_n - D_{\max}) \end{aligned} \quad (36)$$

where we have relaxed the constraints (8) and constraints $d_n \leq D_{\max}$. The Lagrange dual function is given by

$$G(\boldsymbol{\lambda}, \mathbf{v}) = \inf_{\mathbf{r}, \mathbf{d}} \mathcal{L}(\mathbf{r}, \mathbf{d}, \boldsymbol{\lambda}, \mathbf{v}) \quad (37)$$

where the infimum is taken with the constraints $0 \leq r_n \leq R_{\max}$ and $d_n \geq 0$, and the dual problem is given by

$$\text{maximize}_{\boldsymbol{\lambda} \geq 0, \mathbf{v} \geq 0} G(\boldsymbol{\lambda}, \mathbf{v}). \quad (38)$$

Lemma A.1: Any dual optimal vector $\boldsymbol{\lambda}^*$ [i.e., a vector $\boldsymbol{\lambda}$ maximizing (38)] satisfies the conditions

$$\sum_{m: n \in \mathcal{X}_m} \lambda_m^* \leq \gamma_n V \quad (39)$$

for all $n \in \mathcal{N}$. Moreover, any primal optimal r_n^* satisfies the condition

$$\begin{aligned} r_n^* &= \operatorname{argmin}_{0 \leq r_n \leq R_{\max}} U_n(t)r_n - (E_n(t) - \theta_n)P_n^c(r_n) \\ &\quad - r_n \sum_{m: n \in \mathcal{X}_m} \lambda_m^*. \end{aligned} \quad (40)$$

The proof of Lemma A.1 can be found in Appendix B.

According to (40), we have that $r_n^* > 0$ is an optimal solution of problem (23) only if the value of the right-hand side of (40)

evaluated at $r_n = 0$ is larger than the value obtained by evaluating it at r_n^* , which can be expressed, using (10), as

$$U_n(t)r_n^* + (\theta_n - E_n(t))\alpha_n r_n^* - r_n^* \sum_{m: n \in \mathcal{X}_m} \lambda_m^* \leq 0. \quad (41)$$

From (26), (39), and (41), we further obtain

$$U_n(t) \leq \sum_{m: n \in \mathcal{X}_m} \lambda_m^* \leq \gamma_n V \quad (42)$$

which implies that a node n receives exogenous data from outside the network only when $U_n(t) \leq \gamma_n V$. Hence, recalling that $R_n(t) \leq R_{\max}$, we obtain the desired result $U_n(t + 1) \leq \gamma_n V + R_{\max}$.

Finally, if a node n receives both endogenous and exogenous data, we have from (35) that $U_n(t) \leq \gamma_n V - l_{\max}\mu_{\max}$. Since a node n can receive at most $l_{\max}\mu_{\max}$ bits per channel use of endogenous data and R_{\max} bits per channel use of exogenous data, we have the desired inequality $U_n(t + 1) \leq \gamma_n V + R_{\max}$, which completes the proof of part 1).

2) To prove the claim, we need to show that if

$$E_n(t) < \alpha_n R_{\max} + P_{\max} \quad (43)$$

then the following two conditions must be satisfied.

- The Rate-Distortion problem (23) is minimized by choosing $R_n(t) = r_n^* = 0$ (which implies $P_n^c(t) = 0$) for all $n \in \mathcal{N}$.
- The Power Allocation problem (25) selects a power matrix $\mathbf{P}(t)$ such that $P_n(t) = 0$ for all $n \in \mathcal{N}$.

From Lemma A.1, and in particular from (40), condition a) is verified if

$$U_n(t)r_n - (E_n - \theta_n)P_n^c(r_n) - r_n \sum_{m: n \in \mathcal{X}_m} \lambda_m^* > 0, \quad \text{for all } r_n > 0 \quad (44)$$

where we recall that $\boldsymbol{\lambda}^*$ is any optimal dual vector of problem (38). This is proved by the following inequalities:

$$\begin{aligned} &U_n(t)r_n - (E_n - \theta_n)\alpha_n r_n - r_n \sum_{m: n \in \mathcal{X}_m} \lambda_m^* \\ &> U_n(t)r_n + \frac{\gamma_n}{\beta_n} V \alpha_n r_n - r_n \sum_{m: n \in \mathcal{X}_m} \lambda_m^* \\ &\geq U_n(t)r_n + \frac{\gamma_n}{\beta_n} V \alpha_n r_n - r_n \gamma_n V \\ &= U_n(t)r_n + \gamma_n V \frac{(\alpha_n - \beta_n)r_n}{\beta_n} \geq 0 \end{aligned}$$

where the first inequality follows from (43) and the assumption of Theorem 5.1 that $\theta_n = \frac{\gamma_n}{\beta_n} V + \alpha_n R_{\max} + P_{\max}$; the second from (39); and the last inequality follows from $U_n(t) \geq 0$, $r_n > 0$ and from the definition of β_n . This proves (44) and thus that condition a) is satisfied if (43) holds.

To prove b) we first note that the bound (27) implies that the weight (24) satisfies the inequality

$$\begin{aligned} W_{n,m}(t) &= \max\{U_n(t) - U_m(t) - \delta, 0\} \\ &\leq \gamma_n V - l_{\max}\mu_{\max} \end{aligned} \quad (45)$$

for all $(n, m) \in \mathcal{L}$ and for all time t .

We now show by contradiction that condition b) holds when (43) is satisfied. To this end, assume that the power allocation vector \mathbf{P}^* that maximizes (25) at time t is such that some entry $P_{n,m}^*$ is positive. Starting from \mathbf{P}^* , we now obtain a new power allocation vector \mathbf{P} , in which we set $P_{n,m} = 0$. Clearly, the power matrix \mathbf{P} is also feasible. We demonstrate that the objective function of (25) when evaluated at \mathbf{P}^* is smaller than at \mathbf{P} , thus leading to a contradiction. Denoting as $G(\mathbf{P})$ the objective function of (25), this is shown by the following inequalities:

$$\begin{aligned} G(\mathbf{P}^*) - G(\mathbf{P}) &= \sum_{n \in \mathcal{N}} \sum_{l \in \mathcal{N} \setminus n} [C_{n,l}(\mathbf{P}^*, \mathbf{S}(t)) - C_{n,l}(\mathbf{P}, \mathbf{S}(t))] W_{n,l}(t) \\ &\quad + (E_n(t) - \theta_n) P_{n,m}^* \\ &\leq C_{n,m}(\mathbf{P}^*, \mathbf{S}(t)) W_{n,m}(t) + (E_n(t) - \theta_n) P_{n,m}^* \\ &\leq C_{n,m}(\mathbf{P}^*, \mathbf{S}(t)) (\gamma_n V - l_{\max} \mu_{\max}) + (E_n(t) - \theta_n) P_{n,m}^* \\ &\leq (\gamma_n V - l_{\max} \mu_{\max}) \xi P_{n,m}^* + (E_n(t) - \theta_n) P_{n,m}^* \\ &< (\gamma_n V - l_{\max} \mu_{\max}) \xi P_{n,m}^* - \frac{\gamma_n}{\beta_n} V P_{n,m}^* < 0 \end{aligned}$$

where the first inequality derives from $\mu_{n,l}(\mathbf{P}^*, \mathbf{S}(t)) - \mu_{n,l}(\mathbf{P}, \mathbf{S}(t)) \leq 0$ for all $l \neq m$ (Property 2), the second from (45), the third from Property 1 and the fourth from (43). This shows that \mathbf{P}^* is not optimal for (25), thus leading to a contradiction, which completes the proof of 2).

3) The proof of 3) is a relatively simple application of the general theory of [27] and [34]. The complete derivation can be found in our technical report [35].

APPENDIX B PROOF OF LEMMA A.1

Proof: Let $\boldsymbol{\lambda}^*$ and \mathbf{v}^* be an optimal solution of the dual problem (38), and $\mathbf{r}^* = [r_1^*, \dots, r_N^*]$ and $\mathbf{d}^* = [d_1^*, \dots, d_N^*]$ be an optimal solution of the (primal) problem (23). Existence of $(\mathbf{r}^*, \mathbf{d}^*)$ and $(\boldsymbol{\lambda}^*, \mathbf{v}^*)$ is guaranteed by Weierstrass theorem [44, Proposition 2.1.1] and by Slater's condition [44, Proposition 3.5.4, part a)]. By [44, Proposition 6.1.1], the following conditions must be satisfied by \mathbf{d}^* and $(\boldsymbol{\lambda}^*, \mathbf{v}^*)$: primal feasibility, namely $d_n^* \leq D_{\max}$, and the complementary slackness conditions $v_n^* (d_n^* - D_{\max}) = 0$ for all $n \in \mathcal{N}$, and $(\mathbf{r}^*, \mathbf{d}^*) = \operatorname{argmin} \mathcal{L}(\mathbf{r}, \mathbf{d}, \boldsymbol{\lambda}^*, \mathbf{v}^*)$ where the minimization is taken under the constraints $d_n \geq D_{\min}$ and $0 \leq r_n^* \leq R_{\max}$ for all $n \in \mathcal{N}$. From (36), the given conditions imply that

$$\begin{aligned} V f_n(D_{\max}) - \log(D_{\max}) \sum_{m:n \in \mathcal{X}_m} \lambda_m^* \\ - \left(V f_n(d_n^*) - \log(d_n^*) \sum_{m:n \in \mathcal{X}_m} \lambda_m^* \right) \geq 0, \quad (46) \end{aligned}$$

must be satisfied. This is because the Lagrangian $\mathcal{L}(\mathbf{r}, \mathbf{d}, \boldsymbol{\lambda}^*, \mathbf{v}^*)$ when evaluated at $d_n = d_n^*$ should be no larger than that evaluated in $d_n = D_{\max}$. We thus have the inequalities

$$\begin{aligned} \sum_{m:n \in \mathcal{X}_m} \lambda_m^* &\leq \frac{f_n(d_n^*) - f_n(D_{\max})}{\log(d_n^*/D_{\max})} V \\ &\leq \sup_{D_{\min} \leq d_n \leq D_{\max}} \left[\frac{f_n(d_n) - f_n(D_{\max})}{\log(d_n/D_{\max})} \right] V \\ &= \gamma_n V \end{aligned}$$

where the second inequality follows since $D_{\min} \leq d_n^* \leq D_{\max}$, and the third from the definition of γ_n .

APPENDIX C SOURCE MODEL

Here, we present a simple source model for which we determine numerical results in Section VII. Let the source signals measured at sensors in \mathcal{N} be spatially correlated with parameter ω . Since the measurements are Gaussian, we can write for the i th sensor $X_i = \sqrt{\omega}A + \sqrt{1-\omega}B_i$, with A and B_i independent Gaussian random variables with zero mean and unitary variance. Moreover, we assume that the sink is able to measure A with an accuracy that depends on the rate R_d used for acquisition. From standard rate-distortion theory, we have the relationship $R_d = I(A; Y)$, where Y is the side information available at the sink. By choosing the optimal test channel $Y = \sqrt{\omega_d}A + \sqrt{1-\omega_d}C$, where ω_d denotes the correlation between the measurement Y at the sink and A (see, e.g., [5]), we obtain the conditional covariance matrix (34) and the associated entropy function $g(\mathcal{X}, \mathbf{O}(t), R_d(t))$.

REFERENCES

- [1] W. Dargie and C. Poellabauer, *Fundamentals of Wireless Sensor Networks: Theory and Practice*. Hoboken, NJ, USA: Wiley, 2010.
- [2] J. Paradiso and T. Starner, "Energy scavenging for mobile and wireless electronics," *IEEE Pervasive Comput.*, vol. 4, no. 1, pp. 18–27, Jan.–Mar. 2005.
- [3] S. Chalasani and J. Conrad, "A survey of energy harvesting sources for embedded systems," in *Proc. IEEE Southeastcon*, Huntsville, AL, USA, Apr. 2008, pp. 442–447.
- [4] K. C. Barr and K. Asanovic, "Energy-aware lossless data compression," *Trans. Comput. Syst.*, vol. 24, pp. 250–291, Aug. 2006.
- [5] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2012.
- [6] D. Zordan, G. Quer, M. Zorzi, and M. Rossi, "Modeling and generation of space-time correlated signals for sensor network fields," in *Proc. IEEE GLOBECOM*, Houston, TX, USA, Dec. 2011, pp. 1–6.
- [7] L. Huang and M. J. Neely, "Utility optimal scheduling in energy harvesting networks," in *Proc. 12th ACM MobiHoc*, Paris, France, May 2011, Art. no. 21.
- [8] P. Castiglione, O. Simeone, E. Erkip, and T. Zemen, "Energy-neutral source-channel coding for cyber-physical systems," in *Proc. IEEE CAMSAP*, San Juan, Puerto Rico, Dec. 2011.
- [9] 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.
- [10] C. K. Ho and R. Zhang, "Optimal energy allocation for wireless communications with energy harvesting constraints," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4808–4818, Sep. 2012.
- [11] B. Devillers and D. Gündüz, "A general framework for the optimization of energy harvesting communication systems with battery imperfections," *J. Commun. Netw.*, vol. 14, no. 2, pp. 130–139, Apr. 2012.
- [12] S. Chen, P. Sinha, N. Shroff, and C. Joo, "Finite-horizon energy allocation and routing scheme in rechargeable sensor networks," in *Proc. IEEE INFOCOM*, Shanghai, China, Apr. 2011, pp. 2273–2281.
- [13] V. Sharma, U. Mukherji, V. Joseph, and S. Gupta, "Optimal energy management policies for energy harvesting sensor nodes," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1326–1336, Apr. 2010.
- [14] R. Srivastava and C. E. Koksal, "Basic tradeoffs for energy management in rechargeable sensor networks," ArXiv e-prints, Sep. 2010.
- [15] L. Huang and M. J. Neely, "Utility optimal scheduling in processing networks," *Perform. Eval.*, vol. 68, no. 11, pp. 1002–1021, Nov. 2011.
- [16] Z. Mao, C. E. Koksal, and N. B. Shroff, "Near optimal power and rate control of multi-hop sensor networks with energy replenishment: Basic limitations with finite energy and data storage," *IEEE Trans. Autom. Control*, vol. 57, no. 4, pp. 815–829, Apr. 2012.

- [17] 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.
- [18] L. Lin, N. B. Shroff, and R. Srikant, "Asymptotically optimal energy-aware routing for multihop wireless networks with renewable energy sources," *IEEE/ACM Trans. Netw.*, vol. 15, no. 5, pp. 1021–1034, Oct. 2007.
- [19] L. Lin, N. B. Shroff, and R. Srikant, "Energy-aware routing in sensor networks: A large system approach," *Ad Hoc Netw.*, vol. 5, pp. 818–831, Aug. 2007.
- [20] V. Joseph, V. Sharma, and U. Mukherji, "Joint power control, scheduling and routing for multihop energy harvesting sensor networks," in *Proc. ACM MSWiM*, Tenerife, Spain, Oct. 2009, pp. 128–136.
- [21] C. Luna, Y. Eisenberg, R. Berry, T. Pappas, and A. Katsaggelos, "Joint source coding and data rate adaptation for energy efficient wireless video streaming," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 10, pp. 1710–1720, Dec. 2003.
- [22] E. Akyol and M. v. d. Schaar, "Compression-aware energy optimization for video decoding systems with passive power," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 9, pp. 1300–1306, Sep. 2008.
- [23] M. J. Neely and A. Sharma, "Dynamic data compression with distortion constraints for wireless transmission over a fading channel," ArXiv e-prints, Jul. 2008.
- [24] A. B. Sharma, L. Golubchik, R. Govindan, and M. J. Neely, "Dynamic data compression in multi-hop wireless networks," in *Proc. 11th Int. Joint Conf. Meas. Modeling Comput. Syst.*, Seattle, WA, USA, Jun. 2009, pp. 145–156.
- [25] R. Cristescu and B. Beferull-Lozano, "Lossy network correlated data gathering with high-resolution coding," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2817–2824, Jun. 2006.
- [26] T. Cui, T. Ho, and L. Chen, "On distributed distortion optimization for correlated sources," in *Proc. IEEE ISIT*, Nice, France, Jun. 2007, pp. 2731–2735.
- [27] L. Georgiadis, M. J. Neely, and L. Tassiulas, "Resource allocation and cross-layer control in wireless networks," *Found. Trends Netw.*, vol. 1, no. 1, pp. 1–144, Apr. 2006.
- [28] I. Shomorony, A. S. Avestimehr, H. Asnani, and T. Weissman, "Worst-case source for distributed compression with quadratic distortion," in *Proc. IEEE Inf. Theory Workshop*, Lausanne, Switzerland, Sep. 2012, pp. 187–191.
- [29] R. Zamir and T. Berger, "Multiterminal source coding with high resolution," *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp. 106–117, Jan. 1999.
- [30] A. Gupta, S. Verdú, and T. Weissman, "Rate-distortion in near-linear time," in *Proc. IEEE ISIT*, Toronto, ON, Canada, Jul. 2008, pp. 847–851.
- [31] S. Jalali, A. Montanari, and T. Weissman, "Lossy compression of discrete sources via the Viterbi algorithm," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2475–2489, Apr. 2012.
- [32] R. Venkataramanan and S. Tatikonda, "Sparse regression codes for multi-terminal source and channel coding," in *Proc. Allerton*, Monticello, IL, USA, Oct. 2012, pp. 1966–1974.
- [33] S. Cui, A. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1089–1098, Aug. 2004.
- [34] M. J. Neely, *Stochastic network optimization with application to communication and queueing systems*. San Francisco, CA, USA: Morgan & Claypool, 2010.
- [35] D. Tapparello, M. Simeone, and M. Rossi, "Dynamic compression-transmission for energy-harvesting multihop networks with correlated sources," ArXiv e-prints, Mar. 2012.
- [36] Y. Li, B. Vucetic, Z. Zhou, and M. Dohler, "Distributed adaptive power allocation for wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 948–958, Mar. 2007.
- [37] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA, USA: Athena Scientific, 2005.
- [38] D. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1439–1451, Aug. 2006.
- [39] R. Cristescu, B. Beferull-Lozano, and M. Vetterli, "Networked Slepian-Wolf: theory, algorithms, and scaling laws," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4057–4073, Dec. 2005.
- [40] D. N. C. Tse and S. V. Hanly, "Multi-access fading channels—part i: polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2796–2815, Nov. 1998.
- [41] M. Gastpar, "The Wyner-Ziv problem with multiple sources," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2762–2768, Nov. 2004.
- [42] R. Rajesh and V. Sharma, "Amplify and forward for correlated data gathering over hierarchical sensor networks," in *Proc. IEEE WCNC*, Budapest, Hungary, Apr. 2009, pp. 1–6.
- [43] Y. Huang and Y. Hua, "Energy planning for progressive estimation in multihop sensor networks," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 4052–4065, Oct. 2009.
- [44] D. P. Bertsekas, *Convex Analysis and Optimization*. Belmont, MA, USA: Athena Scientific, 2003.



Cristiano Tapparello (M'13) received the M.Sc. degree (with honors) in computer engineering and Ph.D. degree in information engineering from the University of Padova, Padova, Italy, in 2008 and 2012, respectively.

In 2011, he visited the Center for Wireless Communication and Signal Processing Research (CWCSRP), New Jersey Institute of Technology (NJIT), Newark, NJ, USA, where he performed research on the design of networking protocols for energy-harvesting wireless networks. Since 2012, he has been a Postdoctoral Research Associate with the SIGNET group, Department of Information Engineering (DEI), University of Padova. His current research interests include stochastic modeling and optimization of wireless systems, energy-scavenging solutions for wireless sensor networks, and the design and implementation of practical algorithms for constrained embedded systems.



Osvaldo Simeone (M'06) received the M.Sc. (with honors) and Ph.D. degrees in information engineering from Politecnico di Milano, Milan, Italy, in 2001 and 2005, respectively.

He is currently with the Center for Wireless Communications and Signal Processing Research (CWCSRP), New Jersey Institute of Technology (NJIT), Newark, NJ, USA, where he is an Associate Professor. His research interests include wireless communications, information theory, data compression, and machine learning.

Dr. Simeone serves as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS. He is a co-recipient of Best Paper awards of the IEEE SPAWC 2007 and IEEE WRECOM 2007.



Michele Rossi (S'02–M'04–SM'13) received the Laurea degree in electrical engineering and Ph.D. degree in information engineering from the University of Ferrara, Ferrara, Italy, in 2000 and 2004, respectively.

From 2000 to 2005, he was a Research Fellow with the Department of Engineering, University of Ferrara. During 2003, he was on leave with the Center for Wireless Communications (CWC), University of California, San Diego (UCSD), La Jolla, CA, USA, where he performed research on wireless sensor networks.

In 2005, he joined the Department of Information Engineering, University of Padova, Padova, Italy, where he is now an Assistant Professor. He is currently involved in the EU-funded IoT-A and SWAP projects, both on wireless sensor networking. His research interests are centered around stochastic modeling and optimization of telecommunication systems, with particular emphasis on the dissemination of data in wireless ad hoc and sensor networks, the design of integrated MAC/routing schemes, spatio-temporal compression techniques, and cooperative routing policies. His current research activity also includes the design of energy-scavenging solutions for wireless sensor networks and embedded wireless systems in general.

Dr. Rossi currently serves on the Editorial Board of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He is a co-recipient of four Best Paper awards from the IEEE and several local recognitions for the high quality of his research.