Distortion Minimization in Multi-Sensor Estimation With Energy Harvesting

Mojtaba Nourian, *Member, IEEE*, Subhrakanti Dey, *Senior Member, IEEE*, and Anders Ahlén, *Senior Member, IEEE*

Abstract—This paper presents a design methodology for optimal energy allocation to estimate a random source using multiple wireless sensors equipped with energy harvesting technology. In this framework, multiple sensors observe a random process and then transmit an amplified uncoded analog version of the observed signal through Markovian fading wireless channels to a remote station. The sensors have access to an energy harvesting source, which is an everlasting but unreliable random energy source compared to conventional batteries with fixed energy storage. The remote station or so-called fusion centre estimates the realization of the random process by using a best linear unbiased estimator. The objective is to design optimal energy allocation policies at the sensor transmitters for minimizing total distortion over a finite-time horizon or a long term average distortion over an infinite-time horizon subject to energy harvesting constraints. This problem is formulated as a Markov decision process (MDP) based stochastic control problem and the optimal energy allocation policies are obtained by the use of dynamic programming techniques. Using the concept of submodularity, the structure of the optimal energy allocation policies is studied, which leads to an optimal threshold policy for binary energy allocation levels. Motivated by the excessive communication burden for the optimal control solutions where each sensor needs to know the channel gains and harvested energies of all other sensors, suboptimal decentralized strategies are developed where only statistical information about all other sensors' channel gains and harvested energies is required. Numerical simulation results are presented illustrating the performance of the optimal and suboptimal algorithms.

Index Terms—Wireless sensor networks, distributed estimation, best linear unbiased estimator (BLUE), energy/power control, energy harvesting, Markov decision processes, dynamic programming (DP), threshold policy.

I. Introduction

ADVANCES in wireless communications and high-speed low-power electronics technologies have enabled various practical applications of inexpensive, compact and versatile wireless sensor networks (WSNs) in diverse areas such as environmental data gathering [1], industrial process monitoring and controlling [2], mobile robots and autonomous vehicles [3], and monitoring of smart electricity grids [4], among many others. In typical single-hop WSNs, each sensor in the network measures the quantity of interest, creates a local signal and then conveys it to the remote estimator or so-called the fusion centre (FC) where the received signals are processed and a final estimate of the observed quantity computed. One of the important challenges in the area of WSNs is to design efficient strategies for each specific sensor to improve the estimation performance under limited energy resources. This paper addresses the problem of optimal energy allocation to efficiently estimate a random source using distributed wireless sensors equipped with energy harvesting technology.

Designing optimal transmit power/energy allocation strategies for WSNs has been an active area of research in recent years. Two types of sensor transmission strategies are usually considered—the uncoded analog forwarding strategy (motivated by the seminal work of [5] showing the optimality of uncoded analog transmission over a coherent multiple access channel), or using a separate source and channel coding strategy where the sensors compress their information first before transmission over the fading channels (see [6] for a nice survey). In our work, we restrict the discussion to the case of uncoded analog transmission only. Among many papers in this area we first mention here [7] which considers a WSN with an orthogonal multiple-access scheme from sensors to the FC where a best linear unbiased estimator (BLUE) is used (see [8] for the theory of BLUE). For this model, optimal power allocation policies are derived in a way that the total distortion is minimized subject to a sum power constraint at the sensors [7]. The work in [9] considers the same multiple-access scheme as [7] but, instead of the total distortion, total transmission power is minimized under distortion constraints when the statistical knowledge of the source signal is also utilized and more general correlated sources are considered. Sensing noise uncertainty in BLUE using multiple sensors has been considered in [10], and optimal power allocation for linear estimation over coherent multiple access channels has been considered in [11].

So far, most studies on WSNs focused on sensors equipped with conventional batteries with fixed energy storages. But, these batteries are in general hard to replace. In many applications, e.g., biomedical sensors implanted within the human body, the battery lifetime can be prolonged by integrating energy harvesting techniques that can harvest e.g., solar, magnetic, piezoelectric or vibrational energy. In the energy harvesting paradigm the sensors may recharge their batteries by collecting energy from the environment. However, as most renewable energy sources

are unreliable, the amount of energy harvested is random which makes energy harvesting constraints more complicated than conventional fixed energy constraints.

Recently, a considerable amount of research has been devoted to the concept of energy harvesting in the context of information transmission, with a view to minimizing delay or maximizing throughput. In [12], throughput optimal and mean delay optimal energy allocation policies in a single sensor node are studied, while the optimal energy allocation policies that maximize the mutual information of a wireless link were derived in [13] under either causal or non-causal side information available at the transmitters. In [14], the authors investigated an optimal packet scheduling problem for a single-user energy harvesting wireless communication system, where data packets and energy packets arrive at the transmitter in a random manner. They develop optimal off-line scheduling policies for minimizing the delivery time for all packets to the destination in a deterministic setting where the energy harvesting times and the amounts of energy harvested are all known before transmission starts. While no finite battery capacity is assumed in [14], optimal off-line transmission policies with batteries with limited storage capacities are investigated in [15], where a short-term (finite horizon) throughput maximization and the related problem of minimization of the transmission completion time for a given amount of data are studied. These results are further generalized in [16] where also fading channels and optimal online policies are considered. In particular, the authors in [16] considered optimal power allocation problems over static and fading channels that (i) maximize the throughput by a deadline and (ii) minimize transmission completion time of the communication session. Both off-line and online policies are considered along with finite and infinite battery capacities. In the context of estimation, a remote estimation problem with an energy harvesting sensor was considered in [17] where communication scheduling strategies for the sensor and estimation strategies for the estimator are jointly minimized, while [18] focuses on designing novel optimal energy allocation policies at an energy harvesting sensor for efficient remote state estimation of linear stochastic dynamical systems. In a related paper [19], the authors analyze the problem of distortion minimization for an energy harvesting sensor node communicating over a digital fading channel. Here the focus is on minimizing distortion over a finite number of time-slots with total energy, causality and delay constraints. The readers are also referred to [20] in which efficient state estimation algorithms with energy harvesting sensors and fairness control were developed.

In this paper we design optimal energy allocation policies (or equivalently power allocation policies when each transmission slot is normalized to be of unity length) to minimize the total distortion or estimation error of a random Gaussian source measured by multiple sensors. The sensors have access to energy harvesting sources which are everlasting but unreliable energy sources compared to conventional batteries with fixed energy storage. More specifically, we study a WSN with multiple sensors, with Markovian energy harvesting profiles, employing analog transmissions where the noisy sensor observations of a remote Gaussian source are amplified and forwarded to the FC over fading Markovian wireless channels. The FC estimates

the realization of the random process by using a BLUE. We assume that the sensors are measuring the source at a uniform sampling rate and their data queues are always full. Therefore, in contrast to the energy harvesting and fading channel models in [16], the harvested energy processes and channel fading gains here are modelled as discrete-time processes over a slotted time axis. The fading channel gains remain constant over each time slot (same as the sensor sampling period) and can change from one slot to another in a temporally correlated (Markov) manner. We assume that the harvested energy during each slot is stored in the battery (if the battery does not overflow due to finite capacity limitations) and is available for use in the next time slot. The harvested energy from one time slot to another is also allowed to be temporally correlated following a Markov process. Also, in contrast to [19], we consider *multi-sensor estimation* where the sensors employ analog uncoded transmission using an amplify and forward method to the FC.

Our objective is to design optimal energy allocation policies at the transmitters to minimize total distortion over a finite-time horizon or a long term average distortion over an infinite-time horizon subject to energy harvesting constraints. This problem is formulated as an MDP based stochastic control problem and the optimal energy allocation policies are obtained offline by the use of dynamic programming techniques. The optimal transmission energy levels corresponding to a range of discretized values of the sensors' channels, harvested energies and battery levels are stored in a lookup table for use in real-time. Using the concept of submodularity, the structure of the optimal allocation energy policies is studied, simplifying the numerical search for these optimal policies. These policies are based on the assumption that only causal channel state and energy harvesting information is available at transmitters. In the first instance we assume that each sensor knows the values of causal channel gains and harvested energies of all sensors. In practice, the information about channel gains at the sensor transmitter can be achieved via periodic pilot transmissions from the FC and reciprocity between the sensors-to-FC and FC-to-sensors channels (e.g., in a time-division-duplex (TDD) scheme). Each sensor can obtain the information about the channel gains for all other sensors via individual sensors reporting it to the FC and the FC broadcasting this information to all sensors. The information about harvested energies can also be achieved similarly by each sensor reporting them to the FC, followed by the FC broadcasting this information to all sensors via a feedback channel. However, this imposes an excessive communication burden on the sensor network. To circumvent this problem, suboptimal decentralized strategies are developed for the more realistic case where individual sensors only have local information about their channel gains and harvested energies and only statistical information about all other sensors' channel gains and harvested energies. In either case, based on the available information (global or local), the sensor transmitter chooses the corresponding transmission energy level from the stored lookup table as mentioned earlier. Numerical simulation results are presented illustrating the performance of the energy allocation algorithms.

There are a few recent works that are directly relevant to our current work. In an earlier conference version [21] of

this work, we studied, with limited proof details, the optimal energy allocation problem for distributed estimation over a finite-time horizon where fading channel gains and harvested energies vary in an independent and identically distributed (i.i.d.) manner. In parallel, two recent works have also independently considered optimum power allocation for distributed estimation with energy harvesting at the sensors [22], [23]. In [22], the authors study this problem with deterministic energy harvesting models where the harvested energy information is known non-causally. In [23], the authors consider both causal and non-causal information and extend their previous work [13] on throughput maximization of a wireless communication system with an energy harvesting transmitter to a finite horizon estimation problem and derive similar results on the optimum power allocation policy.

The novelty of the current work is that we address the optimal energy allocation problem at the sensor transmitters in a more general setting compared to [21]–[23] where both finite horizon and infinite-time horizon frameworks with causal information are considered and the fading channel gains and harvested energies are modelled as general Markovian processes. Being fully aware of the fact that the optimal solution requires causal channel gains and harvested energy information of all sensors to be known at each sensor transmitter and the excessive communication overhead this generates, we also develop a suboptimal but decentralized framework with a significantly reduced communication overhead where each sensor only uses its local information and the statistics of all other sensors' fading channels and energy harvesting processes.

In summary, the main contributions of this paper are as follows:

i) Unlike the WSN model in [7], we consider energy harvesting sensors that are not constrained by fixed initial battery storages, but rather the randomness of the harvested energy pattern. Energy harvesting is a promising solution to the important problem of energy management in wireless sensor networks. The integration of energy harvesting technology on-board a sensor node is now a practical reality [24]. We thus consider finite and infinite horizon multi-sensor estimation problems with general Markovian energy harvesting processes and Markovian fading channels. The optimal energy allocation problems are formulated as MDP based stochastic control problems and dynamic programming (DP) based techniques are used to obtain the optimal control solutions.

ii) It is well known that the optimal solution obtained by DP techniques, in particular, as a stationary control policy minimizing the infinite horizon control cost is computationally prohibitive. Thus motivated, we provide structural results on the optimal energy allocation policy, which lead to threshold policies which are optimal and yet very simple to implement in some practical cases, e.g., when the sensor is equipped with binary transmission energy levels. Note that most sensors are usually equipped with a finite number of transmission energy/power levels and for flexibility of practical implementation, sensors can be programmed to have only two levels.

iii) Motivated by the communication burden for the optimal control solution where the FC needs to gather and then communicate the information about channel gains and harvested energies to all sensors, we provide a suboptimal solution where only statistical information about all other sensors' channel gains and harvested energies is required. Numerical results are presented to illustrate the performance gaps between the optimal and suboptimal solutions.

The following notation will be used throughout the paper. The set of sensors is denoted by $\mathcal{M} = \{1, \ldots, M\}$ where M is the number of sensors. Discrete time is denoted by $k \in \mathbb{Z}_+ := \{0, 1, 2, \ldots\}$ and the value of sequence X at time slot k is given by X_k or $X(k)$. We use the subscript m as the label for sensor $m \in \mathcal{M}$. The notation $X_m(k)$ or $X_{m,k}$ is used to denote the value of X for sensor m at time k. Boldface letters represent vectors. $\mathbf{A}^T$ denotes the transpose of a vector $\mathbf{A}$. We denote $\mathbf{X}_{(:,k)} = [X_{1,k}, \ldots, X_{M,k}]^T$, $\mathbf{X}_{m,(\cdot)} = [X_{m,0}, X_{m,1}, \ldots]^T$ and $\mathbf{X}_{m,[k]} = [X_{m,0}, \ldots, X_{m,k}]^T$ for $k \in \mathbb{Z}_+$ and $m \in \{1, \ldots, M\}$. We also denote $\mathbf{X}_{(-m,k)} = [X_{1,0}, \ldots, X_{m-1,k}, X_{m+1,k}, \ldots, X_{M,k}]^T$ for $m \in \{1, \ldots, M\}$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. $\mathbb{E}$ denotes the expectation operator. $\frac{\partial \mathcal{L}}{\partial x^*}$ represents the partial differentiation of $\mathcal{L}$ with respect to x evaluated at a specific value $x = x^*$. Finally, $\succeq$ and $\preceq$ represent component-wise inequalities.

The organization of the paper is as follows. The system model is given in Section II. The finite-time horizon optimal energy allocation problem subject to energy harvesting constraints and its solution are given in Section III for both causal and non-causal information patterns. The infinite-time horizon optimal energy allocation problem subject to energy harvesting constraints is formulated, and solved using dynamic programming techniques in Section IV. Section V presents some important structural results of the optimal policies, that can give rise to simple threshold-based energy allocation policies in some special cases. In Section VI, a suboptimal decentralized solution to the case of sensors with only statistical information of other sensors is developed, and Section VII presents the numerical simulation results for the infinite horizon optimal energy allocation problem. Concluding remarks are stated in Section VIII.

II. System Model

A diagram of the system architecture consisting of a WSN with M sensors using an orthogonal multiple-access scheme is shown in Fig. 1. The description of each part of the system is given in detail below.

A. Source Model and Sensor Measurements

We assume that time is divided into slots of length equal to the sampling time T_s, which is assumed to be unity without any loss in generality. Therefore the sensors obtain discrete time measurements of a source signal $\{\theta(k)\}$ at time instants $k \geq 0$, where the k-th time slot denotes the period between the time instants k and $k+1$. $\{\theta(k)\}$ is modelled as an i.i.d. Gaussian
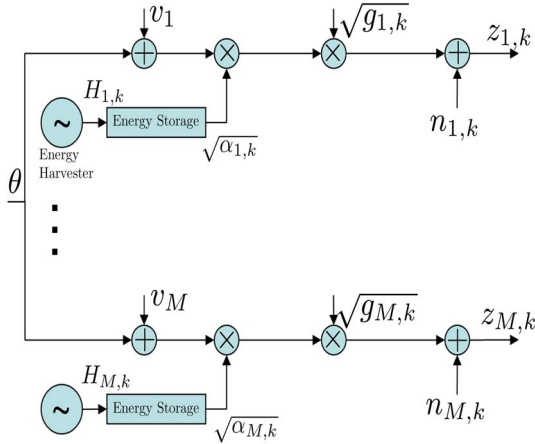
Fig. 1. System architecture.

process with zero mean and variance σ_θ^2. The measurements of the mth sensor, $m \in \mathcal{M} = \{1, \ldots, M\}$, are obtained in the form

$$x_m(k) = \theta(k) + v_m(k), \qquad k \geq 0$$

where $x_m(k)$ is the observation of sensor m at time $k \geq 0$, and $\{\mathbf{v}_{m,(\cdot)}, m \in \mathcal{M}\}$ is a sequence of i.i.d. Gaussian processes, independent of $\{\theta(k)\}_{k \geq 0}$, with zero mean and variance σ_m^2. The sensors transmit their measurements via orthogonal fading channels to an FC where the process $\{\theta(k)\}_{k \geq 0}$ is estimated.

B. Transmitters: Multiple Sensors

We assume that the transmitters adopt an analog amplify and forward uncoded strategy [5]. Thus, the transmitted signal from the mth sensor at each time $k \in \mathbb{Z}_+$ is an amplified version of the measured signal $x_m(k)$ with a power amplifying factor $\alpha_{m,k}$. The energy consumption of sensor m at time k is given by $E_{m,k} = \alpha_{m,k}(\sigma_\theta^2 + \sigma_m^2)$, where $\sigma_\theta^2 + \sigma_m^2$ is the average power of $x_m(k)$ per symbol. Let $g_{m,k}$ be the channel power gain between the mth sensor and the FC in slot k, then the signals $z_m(k) = \sqrt{\alpha_{m,k} g_{m,k}} x_m(k) + n_{m,k}$, $m \in \mathcal{M}$, $k \geq 0$ are received in the FC, where $\{\mathbf{n}_{m,(\cdot)}, m \in \mathcal{M}\}$ denotes a sequence of i.i.d. additive white Gaussian noises (AWGNs) with variance ξ_m^2. The set of fading channel gains $\{\mathbf{g}_{m,(\cdot)} : m \in \mathcal{M}\}$ for the sensors is assumed to be a sequence of independent first-order stationary and homogeneous Markov block-fading processes (see [25]) where the channels remain constant over each fading block (here assumed to be the same as the sampling interval for simplicity), and varying from block to block in a temporally correlated manner. Note that the stationary first-order Markovian modelling includes the case of i.i.d. processes as a special case.

We assume that channel state information is available at the transmitters such that at each time $k \in \mathbb{Z}_+$, the state of all sensors' channel gains $\mathbf{g}_{(:),k}$ is broadcasted to all sensors. In practice, this can be achieved by channel reciprocity between the sensor-to-FC and FC-to-sensor channels (such as in typical time-division-duplex (TDD) based transmissions). In this scenario, the sensor can estimate the channel gain based on pilot signals transmitted from the remote receiver at the beginning of

each fading block. Another possibility (if channel reciprocity does not hold) is to estimate the channel at the receiver based on pilot transmissions from the sensor and send it back to the sensor by channel state feedback. However, transmitting pilot signals consumes energy which should then be taken into account. To conform with our problem formulation, we therefore assume that channel reciprocity holds. Motivated by the communication burden of this formulation, the more practical scenario is studied in Section VI where each sensor only has statistical information about all other sensors' channel gains.

C. Energy Harvestor and Battery Dynamics

Let the unpredictable energy harvesting process of the sensors be denoted by $\{\mathbf{H}_{m,(\cdot)} : m \in \mathcal{M}\}$ which is modelled as a set of independent stationary first-order homogeneous Markov process. This modelling for the harvested energy processes is justified by empirical measurements in the case of solar energy [26]. We also assume that harvested energy processes and channel power gains across the M sensors are mutually independent.

Assume that sensor m consumes energy $E_{m,k}$ at time k from its battery with a maximum storage capacity $\hat{B}_m$, and collects an amount of harvesting energy $H_{m,k}$ which is then stored in the battery for use in the next time slot. The dynamics of the mth sensor's stored battery energy $\mathbf{B}_{m,(\cdot)}$, $m \in \mathcal{M}$, is assumed to follow a first-order Markov model

$$B_{m,k+1} = \min\{B_{m,k} + H_{m,k} - E_{m,k}, \hat{B}_m\}, \ k \geq 0 \qquad (1)$$

with given $0 \leq B_{m,0} \leq \hat{B}_m$, where $\hat{B}_m$ is the maximum stored energy in the battery of sensor m. In this paper, we assume that the sensor only consumes energy for transmission of its measurement to the FC. Note that a constant processing cost for obtaining a measurement at each time-slot is not explicitly considered as this can be easily incorporated by reducing the finite battery capacity by the required amount. We assume that each sensor knows the energy harvesting information of all sensors. This information can be achieved by each sensor reporting to the FC and then FC broadcasting this information to all sensors via feedback. Motivated by the communication burden of this formulation, the more practical scenario is studied in Section VI where each sensor only has statistical information about all other sensors' harvested energies.

Remark 2.1: It should be noted that the assumption of independent channels and energy harvesting processes across the sensors is made only for analytical and computational tractability. In practical settings, it is entirely possible to have energy harvesting processes and fading channels that are correlated across sensors. The solution techniques presented in this paper can handle such correlated models, albeit at the expense of additional analytical complications and computational complexity.

D. Distortion Measure at Fusion Center

At each time $k \in \mathbb{Z}_+$ an estimate $\hat{\theta}(k)$ of the unknown source $\theta(k)$ is obtained at the FC from received signals $\mathbf{z}_{(:,k)}$. We presume that BLUE is utilized at the FC due to its universality and

simplicity (see [7] and references therein). We denote $s_{m,k} = \frac{g_{m,k}}{\xi_m^2 (\sigma_\theta^2 + \sigma_m^2)}$ and $E'_{m,k} = E_{m,k} \sigma_\theta^2$. Then,

$$d_m(E_{m,k}, s_{m,k}) = \frac{E'_{m,k} s_{m,k}}{1 + \gamma_m^{-1} E'_{m,k} s_{m,k}}$$

where $\gamma_m = \frac{\sigma_\theta^2}{\sigma_m^2}$ is the signal-to-noise ratio for sensor m. Then, the achievable distortion at time k (which is the minimum mean square error (MMSE) computed as the variance of the estimate $\hat{\theta}(k)$) is given by

$$D\left(\mathbf{E}_{(:,k)}, \mathbf{s}_{(:,k)}\right) = \frac{\sigma_\theta^2}{\sum_{m=1}^M d_m(E_{m,k}, s_{m,k})} \qquad (2)$$

when $\mathbf{E}_{(:,k)} \succ \mathbf{0}$, and $D(\mathbf{0}, \mathbf{s}_{(:,k)}) = \sigma_\theta^2$ [7]. Note that when $\mathbf{E}_{(:,k)} = \mathbf{0}$ (corresponding to the case when no sensor is transmitting), the best estimate of $\theta(t)$ is simply $\hat{\theta}(k) := \mathbb{E}[\theta(t)] = 0$ which leads to the maximum distortion σ_θ^2.

We establish some notation. Let $\mathbf{g} := (g_1, \ldots, g_M)^T$ be the vector of channel gains defined akin to $(g_{1,k}, \ldots, g_{M,k})$, then $\mathbf{s} := (s_1, \ldots, s_M)^T$ is defined as $s_m = \frac{g_m}{\xi_m^2 (\sigma_\theta^2 + \sigma_m^2)}$ for $m \in \mathcal{M}$.

Theorem 2.1: For given $\mathbf{s} := (s_1, \ldots, s_M)^T$, the distortion function $D(\mathbf{E}, \mathbf{s})$ is a non-increasing convex function of $\mathbf{E} = (E_1, \ldots, E_M)^T \succ \mathbf{0}$.

Proof: We have

$$\frac{\partial D(\mathbf{E}, \mathbf{s})}{\partial E_m} = -\left(D(\mathbf{E}, \mathbf{s})\right)^2 \times \frac{s_m}{\left(1 + \gamma_m^{-1} E_m^{prime} s_m\right)^2} \leq 0$$

$$\frac{\partial D(\mathbf{E}, \mathbf{s})}{\partial E_m \partial E_n} = 2\left(D(\mathbf{E}, \mathbf{s})\right)^3 \times \frac{s_m}{\left(1 + \gamma_m^{-1} E_m^{prime} s_m\right)^2}$$
$$\times \frac{s_n}{\left(1 + \gamma_n^{-1} E'_n s_n\right)^2}, \quad \text{for } n \neq m$$

$$\frac{\partial^2 D(\mathbf{E}, \mathbf{s})}{\partial E_m^2} = 2\left(D(\mathbf{E}, \mathbf{s})\right)^3 \times \left(\frac{s_m^2}{\left(1 + \gamma_m^{-1} E'_m s_m\right)^4}\right.$$
$$\left. + \frac{\gamma_m^{-1} s_m^2}{\left(1 + \gamma_m^{-1} E'_m s_m\right)^3} \times \sum_{m=1}^M d_m(E_m, s_m)\right).$$

We denote $\mathbf{d} := (d_1(E_1, s_1), \ldots, d_M(E_M, s_M))^T$,

$$\mathbf{p} := \left(\frac{s_1}{\left(1 + \gamma_1^{-1} E'_1 s_1\right)^2}, \ldots, \frac{s_M}{\left(1 + \gamma_M^{-1} E'_M s_M\right)^2}\right)^T$$

$$\mathbf{q} := \left(\frac{\gamma_1^{-1} s_1^2}{\left(1 + \gamma_1^{-1} E'_1 s_1\right)^3}, \ldots, \frac{\gamma_M^{-1} s_M^2}{\left(1 + \gamma_M^{-1} E'_M s_M\right)^3}\right)^T$$

then the Hessian of the distortion function is

$$\nabla^2 D(\mathbf{E}, \mathbf{s}) = \frac{2\sigma_\theta^2}{(\mathbf{1}^T \mathbf{d})^3} \left((\mathbf{1}^T \mathbf{d}) \mathrm{diag}(\mathbf{q}) + \mathbf{p}\mathbf{p}^T\right)$$

which is positive semidefinite due to the fact that for all $\mathbf{x} := (x_1, \ldots, x_M)^T$ we have $\mathbf{x}^T \nabla^2 D(\mathbf{E}, \mathbf{s}) \mathbf{x} \geq 0$. Therefore, the distortion function is a non-increasing convex function of $\mathbf{E} \succeq \mathbf{0}$. Note that the discontinuity of the distortion function at $\mathbf{E} = \mathbf{0}$ is not a problem as convex functions can be discontinuous at boundary points. $\qquad \square$

III. FINITE-TIME HORIZON OPTIMAL TRANSMISSION ENERGY ALLOCATION PROBLEM AND ITS SOLUTION

The finite-time horizon optimal transmission energy allocation problem aims to minimize an expected sum distortion over a finite time-window, where the distortion function is given by (2), subject to energy harvesting constraints (1).

$$\min_{\{\mathbf{E}_{(:,k)}: \, 1 \leq k \leq T\}} \sum_{k=1}^T \mathbb{E}\left[D\left(\mathbf{E}_{(:,k)}, \mathbf{s}_{(:,k)}\right)\right]$$
$$\text{s.t. } 0 \leq E_{m,k} \leq B_{m,k} \, (a.s.) \, \forall \, m \in \mathcal{M}, \, 1 \leq k \leq T \qquad (3)$$

where $B_{m,k}$ is the mth sensor's stored battery energy available at time k which satisfies the dynamics (1). The reader is referred to [21] for a similar formulation but with i.i.d. channel gains and harvesting energies.

In the first instance, we consider the realistic scenario of causal information case where the unpredictable future wireless fading channel gains and energy harvesting information are not *a priori* known to the sensors. More precisely, the information available at any sensor at time $k \geq 1$ is given by $I_k = \{(g_{m,[k]}, H_{m,[k]}, B_{m,[k]}) : m \in \mathcal{M}\}$. The causal information I_k is used at each sensor to decide the amount of energy allocation $\mathbf{E}_{(:,k)}$. A policy $\{\mathbf{E}_{m,(\cdot)} : m \in \mathcal{M}\}$ is feasible if the energy harvesting constraint $0 \leq E_{m,k} \leq B_{m,k}$ are almost surely (a.s.) satisfied for all $m \in \mathcal{M}$ and $k \geq 0$. The admissible control set is then given by $\mathcal{U} := \{\mathbf{E}_{m,(\cdot)} : E_{m,k}$ is adapted to sigma-field $\sigma(I_k)$ and $0 \leq E_{m,k} \leq B_{m,k}$ $(a.s.)$ $\forall m \in \mathcal{M}\}$. Intuitively, the admissible controls are the controls that do not use any information about future random processes, which is essentially a causality requirement in stochastic systems.

A. Finite-Time Horizon Optimal Energy Allocation Policies

For the causal information case $\{I_k : 0 \leq k \leq T - 1\}$ where the unpredictable future wireless fading channel gains and energy harvesting information are not *a priori* known to the sensors, the solution to the stochastic control problem (3) is given in the following theorem.

Theorem 3.1: For a given initial condition $I_0 = \{g_{m,0}, H_{m,0}, B_{m,0} : m \in \mathcal{M}\}$ the value of the finite-time horizon minimization problem (3) with causal information is given by $V_0(g_{m,0}, H_{m,0}, B_{m,0})$ which can be computed recursively from the backward Bellman dynamic programming equation

$$V_k(\mathbf{g}, \mathbf{H}, \mathbf{B}) = \min_{\mathbf{0} \preceq \mathbf{E} \preceq \mathbf{B}} \{D(\mathbf{E}, \mathbf{s})$$
$$+ \mathbb{E}\left[V_{k+1}\left(\tilde{\mathbf{g}}, \tilde{\mathbf{H}}, \tilde{\mathbf{B}}\right) \big| \mathbf{g}, \mathbf{H}, \mathbf{E}\right]\}, \quad 0 \leq k \leq T - 1 \quad (4)$$

where $\tilde{B}_m = \min\{B_m + H_m - E_m, \hat{B}_m\}$ for $m \in \mathcal{M}$, and the terminal condition is given by

$$V_T(\mathbf{g}, \mathbf{H}, \mathbf{B}) := \min_{\mathbf{0} \preceq \mathbf{E} \preceq \mathbf{B}} D(\mathbf{E}, \mathbf{s}) = D(\mathbf{B}, \mathbf{s}) \qquad (5)$$

where all available energy is used for transmission in the final time T.

Proof: The proof follows from the optimality equations for finite-time horizon stochastic control problems (see [27]). □

The solution to the stochastic control problem (3) is then given by

$$\mathbf{E}_k^o(\mathbf{g}, \mathbf{H}, \mathbf{B}) = \arg \min_{\mathbf{0} \preceq \mathbf{E} \preceq \mathbf{B}} \{D(\mathbf{E}, \mathbf{s})$$

$$+ \mathbb{E}\left[V_{k+1}(\tilde{\mathbf{g}}, \tilde{\mathbf{H}}, \tilde{\mathbf{B}}) \big| \mathbf{g}, \mathbf{H}, \mathbf{E}\right]\}, \quad 0 \le k \le T-1 \quad (6)$$

where $\tilde{B}_m = \min\{B_m + H_m - E_m, \hat{B}_m\}$ for $m \in \mathcal{M}$, and $V(\cdot, \cdot, \cdot)$ is the solution to the Bellman equation (4).

Remark 3.1: Note that such dynamic programming based optimal energy allocation policies rely on building a look-up table for discretized values for the state (channel, harvested energy and battery level) and action (energy levels) spaces off-line. In real time, based on the exact values of the current state, the optimal energy allocation is obtained via (6). Similar dynamic programming based optimal policies have also been considered under the title of *online transmission policies* in Section VI of [16]. Note however that the energy harvesting and fading channel models of [16] assume temporally independent models, whereas in this work we allow them to evolve according to a temporally correlated Markov process. This generalization requires the application of a more general stochastic control methodology along with the associated Bellman dynamic programming equation given by (4). It is generally known that due to the exponential complexity of the dynamic programming algorithms, such algorithms can be computationally complex. Various suboptimal strategies can be designed by approximating the value function using approximate dynamic programming methods (ADP). Design of suboptimal energy allocation policies based on such ADP techniques is beyond the scope of the current paper and will be treated elsewhere.

The causal information pattern is clearly relevant to the most practical scenario. However, it is also instructive to consider the non-causal information scenario where each sensor has *a priori* information about the fading channel gains $\{\mathbf{g}_{m,(\cdot)} : m \in \mathcal{M}\}$ and harvested energies $\{\mathbf{H}_{m,(\cdot)} : m \in \mathcal{M}\}$ for all time periods including the future ones (see a detailed analysis of this setting in [21] for finite horizon problems and see also [22], [23] for similar results). This may be feasible in the situation of a known/deterministic environment where the wireless fading channel gains and the harvested energies are predictable [13]. More importantly, the performance of the non-causal information case can serve as a benchmark (a lower bound) for the causal case. Indeed, we present some performance comparison between the performances in the causal and the non-causal case in the *Numerical Examples* section. Note also that the finite horizon energy allocation problem for the non-causal case can be formulated as a convex optimization problem similar to the off-line throughput maximization problem with an energy harvesting transmitter as presented in [13], [16]. Below, we present the solution to two such cases with non-causal information. First, note that in general the solution to the dynamic programming equation (6) can only be obtained numerically and there is no closed form solution. To illustrate the properties of the optimal solution, below we provide a closed form solution for a simple example of a horizon $T = 2$

energy allocation problem with non-causal information for a single sensor. Subsequently, we provide the finite-horizon optimal energy allocation results for multiple sensors with non-causal information and infinite battery storage capacity at each sensor. The solution to the non-causal case with finite battery storage capacity can be obtained with the additional constraint that $0 \le E_{m,k} \le B_{m,k}$ (a.s.) $\forall m \in \mathcal{M}$, where $B_{m,k+1} = \min\{B_{m,k} + H_{m,k} - E_{m,k}, \hat{B}_m\}$, $k \ge 0$, following similar convex optimization techniques as used in Section IV of [16] where throughput maximization over fading channel is considered.

B. A Closed Form Solution for the Horizon 2 Single Sensor Problem With Non-Causal Information

To get some insight about the optimal energy allocation strategy (6), we derive the optimal solution in the simple case of a single sensor when $T = 2$. We drop the subscript m in this case and set $\sigma_\theta^2 = 1$, $\xi^2 = 1$ for simplicity. We assume non-causal information case where (B_1, H_1, g_1, g_2) is known to the sensor. Denote $E^* := \frac{\sqrt{g_2}}{\sqrt{g_1} + \sqrt{g_2}}(B_1 + H_1)$, $a := \sqrt{\frac{g_2}{g_1}} H_1$, and $b := \left(\frac{\sqrt{g_1} + \sqrt{g_2}}{\sqrt{g_1}}\right) \hat{B} - H_1$.

Lemma 3.1: Consider the stochastic control problem (3) with $T = 2$ where the information (B_1, H_1, g_1, g_2) is given. Then, the optimal transmission energy for slot 1 is of the form

$$E_1^o = \begin{cases} B_1 & H_1 > \hat{B} \text{ or } B_1 < a \\ E^* & H_1 \le \hat{B} \text{ and } a \le B_1 \le b \\ [B_1 + H_1 - \hat{B}]^+ & H_1 \le \hat{B} \text{ and } B_1 \ge b \end{cases}$$

where $[x]^+ := \max(x, 0)$, and the optimal transmission energy for slot 2 is given by $E_2^o = \min\{B_1 + H_1 - E_1^o, \hat{B}\}$.

Proof: From (5), we get the optimal solution for slot $T = 2$ as $E_2^o = B_2$ which is a function of E_1 by (1). We define

$$W(E) := D\left(E, \frac{g_1}{1 + \sigma^2}\right) + D\left(\min\{B_1 + H_1 - E, \hat{B}\}, \frac{g_2}{1 + \sigma^2}\right).$$

First, suppose $H_1 > \hat{B}$. Then, $\hat{B} = \min\{B_1 + H_1 - E_1, \hat{B}\}$. Thus, the minimum of $W(E_1) = D(E_1, s_1) + D(\hat{B}, s_2)$ where $0 \le E_1 \le B_1$ is achieved by $E_1^o = B_1$. Second, suppose $H_1 \le \hat{B}$. In case $0 \le E_1 \le B_1 + H_1 - \hat{B}$ we have $\min\{B_1 + H_1 - E_1, \hat{B}\} = \hat{B}$ so the optimal E to minimize $W(E)$ such that $0 \le E_1 \le B_1 + H_1 - \hat{B}$ is given by the largest value $B_1 + H_1 - \hat{B}$. On the other hand, in case $[B_1 + H_1 - \hat{B}]^+ \le E_1 \le B_1$ we have $B_1 + H_1 - E_1 = \min\{B_1 + H_1 - E_1, \hat{B}\}$ which yields $W(E_1) = D(E_1, s_1) + D(B_1 + H_1 - E_1, s_2)$. In this case the optimal solution to the unconstrained optimization problem $\min_{E_1} W(E_1)$ is given by $E^* = \frac{\sqrt{g_2}}{\sqrt{g_1} + \sqrt{g_2}}(B_1 + H_1)$. But, by using the convexity of $W(E_1)$ we get

$$E_1^o = \begin{cases} B_1 & \text{if } E^* > B_1 \\ E^* & \text{if } [B_1 + H_1 - \hat{B}]^+ \le E^* \le B_1 \\ [B_1 + H_1 - \hat{B}]^+ & \text{if } E^* < [B_1 + H_1 - \hat{B}]^+ \end{cases}$$

where $H_1 \le \hat{B}$ (see the discussion below Theorem 5.1). By writing the above conditions in terms of B_1, we obtain E_1^o as stated in lemma. This together with (1) implies that $E_2^o = B_2 = \min\{B_1 + H_1 - E_1^o, \hat{B}\}$. □

The above solution implies that: (i) the sensor uses all stored energy if the energy to be harvested is large, $H_1 > \hat{B}$, or the stored energy is small $B_1 < a$, (ii) the sensor saves as much stored energy as possible for slot 2 if the energy to be harvested is small, $H_1 \leq \hat{B}$, and the stored energy is large $B_1 \geq b$, and (iii) in other cases, the sensor tends to spend more energy in the first slot if it knows that the channel gain $\sqrt{g_2}$ is high thus the amount of remaining energy $\frac{\sqrt{g_1}}{\sqrt{g_1}+\sqrt{g_2}}(B_1 + H_1)$ in slot 2 is sufficient to obtain the satisfactory level of total distortion. Note also that this solution for the horizon 2 case can also be used as a basis for constructing simple suboptimal strategies in the case of a finite horizon causal scenario.

C. Sensors With Infinite Energy Storage Capacity

In this section, we provide a brief summary of the the optimal energy allocation results for the finite-time horizon optimization problem (3) where $\hat{B}_m \to \infty$ in (1), and the sensors have non-causal information. Clearly, this is an idealistic and impractical scenario. However, the performance achieved in this setting can be used as a benchmark for the smallest achievable distortion.

As $\hat{B}_m \to \infty$, the dynamics of the mth sensor's stored battery energy $\mathbf{B}_{m,(\cdot)}$, $m \in \mathcal{M}$, in (1) reduces to

$$B_{m,k} = B_{m,0} + \sum_{l=0}^{k-1} H_{m,l} - \sum_{l=0}^{k-1} E_{m,l}. \qquad (7)$$

A non-negative energy allocation is feasible if and only if $B_{m,k} \geq 0$ for $k \in \mathbb{Z}_+$ and $m \in \mathcal{M}$. The finite-time horizon distortion minimization problem (3) with non-causal information can then be formulated as

$$\min_{\{\mathbf{E}_{(:,k)}:\, 1 \leq k \leq T\}} \sum_{k=1}^{T} D\left(\mathbf{E}_{(:,k)}, \mathbf{s}_{(:,k)}\right)$$

$$\text{s.t.} \sum_{l=0}^{k} E_{m,l} - B_{m,0} - \sum_{l=0}^{k-1} H_{m,l} \leq 0 \; (a.s.) \; m \in \mathcal{M} \qquad (8)$$

which is a convex optimization problem (see Theorem 2.1). Note that the expectation operator for the distortion function can be dropped due to the availability of non-causal information about all harvested energy and channel gains. The explicit solution to this problem was provided in [21] and the key steps are summarized below.

Let $\lambda_{m,k} \geq 0$ denote the Lagrange multiplier associated with the constraints in (8), then the Lagrangian associated to (8) is given by

$$\mathcal{L}\left(\{\mathbf{E}_{m,[k]}, \lambda_{m,[k]} : m \in \mathcal{M}\}\right) = \sum_{k=1}^{T} \mathcal{L}_k\left(\mathbf{E}_{(:,k)}, \lambda_{(:,k)}\right)$$

$$= \sum_{k=1}^{T} \Bigg[D\left(\mathbf{E}_{(:,k)}, \mathbf{s}_{(:,k)}\right)$$

$$+ \sum_{m=1}^{M} \lambda_{m,k}\left(\sum_{l=0}^{k} E_{m,l} - B_{m,0} - \sum_{l=0}^{k-1} H_{m,l}\right)\Bigg] \qquad (9)$$

where $E_{m,k} \geq 0$ is the energy allocation of the mth sensor at time k. We note that $\{\mathbf{E}_{m,[k]}^o : m \in \mathcal{M}\}$ and $\{\lambda_{m,[k]}^o : m \in \mathcal{M}\}$ are primal and dual optimal solutions to (9) if and only if they satisfy the following Karush-Kuhn-Tucker (KKT) optimality conditions for $1 \leq k \leq T$ and $m \in \mathcal{M}$: $\sum_{l=0}^{k} E_{m,l} - B_{m,0} - \sum_{l=0}^{k-1} H_{m,l} \leq 0$, $E_{m,k} \geq 0$, $\lambda_{m,k}$, along with $\lambda_{m,k}(\sum_{l=0}^{k} E_{m,l} - B_{m,0} - \sum_{l=0}^{k-1} H_{m,l}) = 0$, and $\frac{\partial \mathcal{L}}{\partial E_{m,k}^o} = 0$ if $E_{m,k}^o > 0$, $\frac{\partial \mathcal{L}}{\partial E_{m,k}^o} \geq 0$ if $E_{m,k}^o = 0$.

Denote $U_{m,k}^* = B_{m,0} + \sum_{l=0}^{k-1} H_{m,l} - \sum_{l=0}^{k-1} E_{m,l}^*$ which represents the largest amount of energy that the m-th sensor can use in slot k. Then, the optimal energy allocation strategy is given in Theorem 3.2 below.

Theorem 3.2: Suppose that the energy capacity of each sensor's battery is infinite. Then, the optimal energy allocation in slot k, $E_{m,k}^o$, is given by

$$E_{m,k}^o = \begin{cases} 0 & \text{if } \left[\sum_{m=1}^{M} d_{m,k}\left(\Omega_{m,k}^*, s_{m,k}\right)\right]^{-1} \geq 1 \\ \Omega_{m,k}^* & \text{if } \left[\sum_{m=1}^{M} d_{m,k}\left(\Omega_{m,k}^*, s_{m,k}\right)\right]^{-1} < 1 \end{cases}$$

where $\Omega_{m,k}^*$ is given by

$$\Omega_{m,k}^* = \begin{cases} 0 & \text{if } \Omega_{m,k} \leq 0 \\ \Omega_{m,k} & \text{if } 0 < \Omega_{m,k} < U_{m,k}^* \\ U_{m,k}^* & \text{if } \Omega_{m,k} \geq U_{m,k}^* \end{cases} \qquad (10)$$

where $\Omega_{m,k} = \frac{\gamma_m}{\sigma_\theta^2 \sqrt{s_{m,k}}}\left(D_k\sqrt{\nu_{m,k}} - \frac{1}{\sqrt{s_{m,k}}}\right)$ and $\nu_{m,k} = \left(\sum_{l=k}^{T} \lambda_{m,l}^o\right)^{-1}$, and D_k stands for a shorthand notation of $D(\mathbf{E}_{(:,k)}^o, \mathbf{s}_{(:,k)})$.

Proof: From KKT necessary conditions, we have

$$\frac{\partial \mathcal{L}}{\partial E_{m,k}^o} \begin{cases} \geq 0 & \text{if } E_{m,k}^o = 0 \\ = 0 & \text{if } 0 < E_{m,k}^o < U_{m,k}^* \\ \leq 0 & \text{if } E_{m,k}^o = U_{m,k}^*. \end{cases}$$

In case $\mathbf{E}_{(:,k)}^o = \mathbf{0}$, then we can write $\mathcal{L}_k(\mathbf{E}_{(:,k)} = \mathbf{0}) = \sigma_\theta^2 \leq \mathcal{L}_k(\mathbf{E}_{(:,k)} = \Omega_{(:,k)}) = \sigma_\theta^2 [\sum_{m=1}^{M} d_{m,k}(\Omega_{m,k}^*, s_{m,k})]^{-1}$, which implies that $[\sum_{m=1}^{M} d_{m,k}(\Omega_{m,k}^*, s_{m,k})]^{-1} \geq 1$. Now suppose $\mathbf{E}_{(:,k)}^o \neq \mathbf{0}$, then we have

$$\frac{\partial \mathcal{L}}{\partial E_{m,k}^o} = -\left(D\left(\mathbf{E}_{(:,k)}^o, \mathbf{s}_{(:,k)}\right)\right)^2 \times \frac{s_{m,k}}{\left(1 + \gamma_m^{-1}\sigma_\theta^2 E_{m,k}^o s_{m,k}\right)^2} + \frac{1}{\nu_{m,k}}$$

Thus, if $0 < E_{m,k}^o < U_{m,k}^*$ then $E_{m,k}^o = \frac{\gamma_m}{\sigma_\theta^2 \sqrt{s_{m,k}}}\left(D_k\sqrt{\nu_{m,k}} - \frac{1}{\sqrt{s_{m,k}}}\right)$. Note that in this case $\nu_{m,k+1} = \nu_{m,k}$ since the energy causality constraint at slot k is satisfied with inequality and thus $\lambda_{m,k} = 0$ based on the KKT optimality conditions. □

We can also characterize the following obvious property $0 \leq \nu_{m,0} \leq \ldots \leq \nu_{m,T}$, which is helpful for the computation of the optimal solution (10). This follows because $\nu_{m,k} \leq \nu_{m,k+1}$ since $\lambda_{m,k}^o \geq 0$ for all k. Note also that in the last slot T, all the
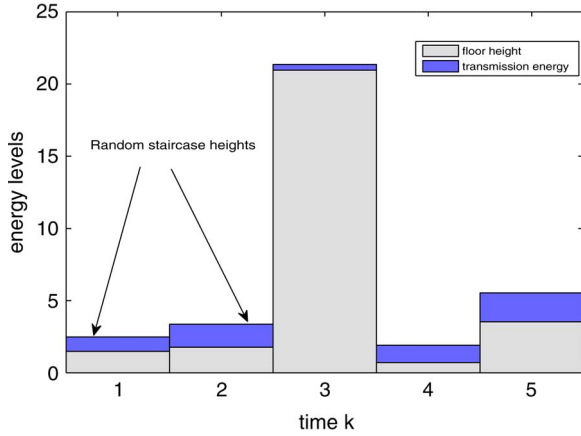
Fig. 2. Illustration of the *random staircase water-filling energy allocation* in a 2-sensor estimation over 5 slots with non-causal information and infinite battery capacity.

remaining energy $B_{m,T}$ must be used up from optimality, thus $E^o_{m,T} = B_{m,T}$.

Remark 3.2: The optimal energy allocation obtained above using a standard convex optimization based Lagrangian formulation follows similar techniques used in [16] and [13] for throughput maximization over fading channels with energy harvesting. The difference in the expressions of the optimal energy allocation (as compared to [13], [16] arise due to the fact that we are considering multiple sensors and the objective function is based on the distortion measure.

Water-filling interpretation: Adapting the term *staircase water-filling* from [13], [23] slightly, this solution can be interpreted as a *random staircase water-filling* algorithm with a flat bottom of height $\frac{\gamma_m}{\sigma^2_\theta s_{m,k}}$ for sensor m at time k. We note that the optimal water level changes over the blocks, and in fact the water level of sensor m at block k is given by $\frac{D(\mathbf{E}^o_{(:,k)}, \mathbf{s}_{(:,k)})}{\sigma^2_m} \sqrt{\frac{v_{m,k}}{s_{m,k}}}$ (since $\frac{\gamma_m}{\sigma^2_\theta} = 1/\sigma^2_m$), where $v_{m,k} = \left(\sum_{l=k}^{T} \lambda^o_{m,l} \right)^{-1}$, and $\mathbf{E}^o_{(:,k)}, \lambda^o_{m,k}$ are the optimal energy allocation and Lagrange multipliers associated with the constraints in (8), respectively. The heights of the steps for this solution are no longer monotonically increasing as in [13], [23], and they depend on the total distortion and the channel gains (hence the word "random"). This solution is also related to the *directional water-filling* power allocation solution presented in [16]. Clearly, due to the energy causality constraints present here, the energy allocation solution presented in (10) also preserves the directional nature of the energy flow. However, unlike [16], the water-levels in each slot are randomly time-varying. Fig. 2 illustrates this solution for a 5-slot optimal energy allocation problem for one of the sensors in a two-sensor distributed estimation scenario where the harvested energies and channel fading gains of the two sensors are i.i.d. with exponential distributions. The blue coloured components show the transmission energy used in each slot. Note that due to the availability of the non-causal information, the specific statistical nature of the energy harvesting process and fading channel gains (as to whether they are i.i.d. or Markov) do not make a difference to the optimal solution.

Remark 3.3: It is easily seen from (7) that the average optimal energy consumption for this scenario given by $\lim_{T \to \infty} \frac{1}{T} \sum_{l=0}^{T-1} E_{m,l}$ is upper bounded by $\bar{H}_m$, the mean of the stationary ergodic energy harvesting process for sensor m (assumed to be finite). For all other scenarios such as with causal information, the associated average energy consumption will be lower due to the fact that the corresponding achieved expected distortion will be higher than the non-causal case.

IV. INFINITE-TIME HORIZON OPTIMAL TRANSMISSION ENERGY ALLOCATION PROBLEM AND ITS SOLUTION

In this section we formulate an infinite-time horizon optimal transmission energy allocation problem subject to energy harvesting constraints (1) to minimize the distortion function (2). In this formulation, the optimal energy allocation policies are computed at the transmitters. It is assumed that the each sensor has perfect causal information about channel gains, harvested energies and maximum battery storages of all sensors. This assumption is feasible by the mechanism that each sensor reports its channel gain and harvested energy to the FC and then FC sends this information to all sensors via feedback (see Sections II-B and II-C). As will be evident later, the optimal energy allocation problem can be solved offline and the optimal solution stored in a lookup table on-board all the sensors, for given (discretized) values of sensors' harvested energies, battery levels and channel gains. Based on the available information, in real time, a sensor can simply select the corresponding transmission energy level from this lookup table.

A. Infinite-Time Horizon Stochastic Control Problem

We now aim to find the optimal energy allocation strategy in $\mathcal{U}$ that minimizes the expected total distortion measure over infinite-time horizon. This optimization problem is formulated as a Markov decision process based stochastic control problem:

$$\min_{\{\mathbf{E}_{(:,k)}: \, k \geq 0\}} \limsup_{T \to \infty} \frac{1}{T} \sum_{k=1}^{T} \mathbb{E}\left[D\left(\mathbf{E}_{(:,k)}, \mathbf{s}_{(:,k)} \right) \right]$$

$$\text{s.t. } 0 \leq E_{m,k} \leq B_{m,k} \quad (a.s.) \qquad \forall m \in \mathcal{M}, \, k \geq 0 \quad (11)$$

where $B_{m,k}$ is the mth sensor's stored battery energy available at time k which satisfies the dynamics (1). Note that the expectation in (11) is computed over all random variables $\{\mathbf{g}_{m,(\cdot):m \in \mathcal{M}}\}$, $\{\mathbf{H}_{m,(\cdot):m \in \mathcal{M}}\}$.

It is evident from (11) that for any sensor $m \in \mathcal{M}$ the transmission energy used in slot k, $E_{m,k}$, affects its amount of stored energy $B_{m,k+1}$ available at time $k+1$ which in turn affects its transmission energy $E_{m,k+1}$ since $0 \leq E_{m,k+1} \leq B_{m,k+1} = \min\{B_{m,k} + H_{m,k} - E_{m,k}, \hat{B}_m\}$ by (1). Moreover, $E_{m,k}$ also affects energy allocation strategies of all other sensors, $E_{n,k}$ where $n \neq m$, due to the coupling in distortion function $D(\cdot, \cdot)$ defined in (2). Note that (11) is a multi-sensor stochastic control problem with centralized information. Similar to the finite-horizon scenario, we assume the availability of causal information at each sensor.

B. Stationary Optimal Energy Allocation Policies

The stochastic control problem (11) with centralized information $\{I_k\}_{k \geq 0}$ can be regarded as a Markov Decision Process (MDP) formulation [27], [28]. The stationary energy allocation policy is computed offline from the Bellman dynamic programming equations given below.

Some notation is presented first. Given the vectors of fading channel gains $\mathbf{g} := (g_1, \ldots, g_M)^T$, harvested energies $\mathbf{H} := (H_1, \ldots, H_M)^T$, battery storages $\mathbf{B} := (B_1, \ldots, B_M)^T$ and sensors' energy consumptions $\mathbf{E} := (E_1, \ldots, E_M)^T$ at time $k \geq 0$, we denote the corresponding fading channel gains, harvested energies and battery storages at time $k+1$ by $\tilde{\mathbf{g}} := (\tilde{g}_1, \ldots, \tilde{g}_M)^T$, $\tilde{\mathbf{H}} := (\tilde{H}_1, \ldots, \tilde{H}_M)^T$ and $\tilde{\mathbf{B}} := (\tilde{B}_1, \ldots, \tilde{B}_M)^T$, respectively. We recall that for any sensor $m \in \mathcal{M}$ both fading channel gains $\mathbf{g}_{m,(\cdot)}$ and harvested energies $\mathbf{H}_{m,(\cdot)}$ are modelled as first-order homogeneous Markov processes, and the dynamics of its stored batteries $\mathbf{B}_{m,(\cdot)}$ is given by (1) as $\tilde{B}_m = \min\{B_m + H_m - E_m, \hat{B}_m\}$.

The solution to the stochastic control problem (11) is now presented in the following theorem.

Theorem 4.1: Independent of the initial condition $\{g_{m,0}, H_{m,0}, B_{m,0} : m \in \mathcal{M}\}$, the value of the infinite-time horizon stochastic control problem (11) is given by ρ which is the solution of the average-cost optimality (Bellman) equation

$$\rho + V(\mathbf{g}, \mathbf{H}, \mathbf{B})$$
$$= \min_{\mathbf{0} \preceq \mathbf{E} \preceq \mathbf{B}} \left\{ D(\mathbf{E}, \mathbf{s}) + \mathbb{E}\left[V(\tilde{\mathbf{g}}, \tilde{\mathbf{H}}, \tilde{\mathbf{B}}) \big| \mathbf{g}, \mathbf{H}, \mathbf{E} \right] \right\} \quad (12)$$

where V is called the relative value function.

Proof: See the Appendix. □

The stationary optimal solution to the stochastic control problem (11) is then given by

$$\mathbf{E}^o(\mathbf{g}, \mathbf{H}, \mathbf{B})$$
$$= \arg \min_{\mathbf{0} \preceq \mathbf{E} \preceq \mathbf{B}} \left\{ D(\mathbf{E}, \mathbf{s}) + \mathbb{E}\left[V(\tilde{\mathbf{g}}, \tilde{\mathbf{H}}, \tilde{\mathbf{B}}) \big| \mathbf{g}, \mathbf{H}, \mathbf{E} \right] \right\} \quad (13)$$

where $V(\cdot, \cdot, \cdot)$ is the solution to the average cost Bellman equation (12).

Remark 4.1: Equation (12) together with the control policy $\mathbf{E}^o$ defined in (13) is known as the average cost optimality equations. If a control policy $\mathbf{E}^o$, a measurable function V, and a constant ρ exist which solve (12), (13), then the control $\mathbf{E}^o$ is optimal, and ρ is the optimal cost in the sense that

$$\limsup_{T \to \infty} \frac{1}{T} \sum_{k=0}^{T-1} \mathbb{E}\left[D\left(\mathbf{E}^o_{(:,k)}, \mathbf{s}_{(:,k)} \right) \right] = \rho$$

and for any other control policy $\{\mathbf{E}_{(:,k)} : k \geq 0\}$ such that $0 \leq E_{m,k} \leq B_{m,k}$ (a.s.) $\forall m \in \mathcal{M}$ and $k \geq 0$, we have

$$\limsup_{T \to \infty} \frac{1}{T} \sum_{k=0}^{T-1} \mathbb{E}\left[D\left(\mathbf{E}_{(:,k)}, \mathbf{s}_{(:,k)} \right) \right] \geq \rho.$$

The reader is referred to [29] for a proof of the average cost optimality equations and related results. □

We now simplify the terms in (12). First, we have

$$D(\mathbf{E}, \mathbf{s}) = \sigma_\theta^2 \times \left(\sum_{m=1}^{M} \frac{\sigma_\theta^2 E_m s_m}{1 + \gamma_m^{-1} \sigma_\theta^2 E_m s_m} \right)^{-1}$$

where $s_m = \frac{g_m}{\xi_m^2(\sigma_\theta^2 + \sigma_m^2)}$ and $0 \leq E_m \leq B_m$.

Let $\mathbb{G}$ and $\mathbb{H}$ be the time-invariant probability transition laws of the Markovian fading channel process $\{\mathbf{g}_{m,(\cdot)} : m \in \mathcal{M}\}$ and the Markovian harvested energy process $\{\mathbf{H}_{m,(\cdot)} : m \in \mathcal{M}\}$, respectively. Since the mutually independent processes $\mathbf{g} = (g_1, \ldots, g_M)^T$ and $\mathbf{H} = (H_1, \ldots, H_M)^T$ are independent of other processes and random variables, we may write

$$\mathbb{E}\left[V(\tilde{\mathbf{g}}, \tilde{\mathbf{H}}, \tilde{\mathbf{B}}) \big| \mathbf{g}, \mathbf{H}, \mathbf{E} \right]$$
$$= \int_{\tilde{\mathbf{g}}, \tilde{\mathbf{H}}} V(\tilde{\mathbf{g}}, \tilde{\mathbf{H}}, \tilde{\mathbf{B}}) \mathbb{G}(\tilde{\mathbf{g}}|\mathbf{g}) \mathbb{H}(\tilde{\mathbf{H}}|\mathbf{H}) d\tilde{\mathbf{g}} d\tilde{\mathbf{H}} \quad (14)$$

where $\tilde{\mathbf{B}} = (\tilde{B}_1, \ldots, \tilde{B}_M)^T$ such that $\tilde{B}_m = \min\{B_m + H_m - E_m, \hat{B}_m\}$ for any sensor $m \in \mathcal{M}$.

Remark 4.2: The expression (14) can be simplified further in the three following cases:

(i) If the processes $\{\mathbf{g}_{m,(\cdot)} : m \in \mathcal{M}\}$ and $\{\mathbf{H}_{m,(\cdot)} : m \in \mathcal{M}\}$ are finite state Markov chains with N_g and N_H states labeled as $\{1, \cdots N_g\}$ and $\{1, \cdots, N_H\}$, respectively, then the right term in (14) becomes

$$\sum_{i=1}^{N_g} \sum_{j=1}^{N_H} V\left(i, j, \tilde{\mathbf{B}}(j)\right) \times (\mathbb{P}(\mathbf{g})\mathbb{G})_i \times (\mathbb{P}(\mathbf{H})\mathbb{H})_j$$

where $\tilde{\mathbf{B}}(j) := \min\{\mathbf{B} - \mathbf{E} + \mathbf{H}(j), \hat{\mathbf{B}}\}$, $\mathbb{P}(\mathbf{g}) := [\mathbb{P}(\mathbf{g}(1)), \ldots, \mathbb{P}(\mathbf{g}(N_g))]$, $\mathbb{P}(\mathbf{H}) := [\mathbb{P}(\mathbf{H}(1)), \ldots, \mathbb{P}(\mathbf{H}(N_H))]$, $\mathbb{G}$ and $\mathbb{H}$ are the probability transition matrices for $\mathbf{g} = (g_1, \ldots, g_M)^T$ and $\mathbf{H} = (H_1, \ldots, H_M)^T$, respectively, and $(\mathbb{P}(\mathbf{g})\mathbb{G})_i$ denotes the i-th component of the vector $\mathbb{P}(\mathbf{g})\mathbb{G}$.

(ii) If Markovian processes $\{\mathbf{g}_{m,(\cdot)} : m \in \mathcal{M}\}$ and $\{\mathbf{H}_{m,(\cdot)} : m \in \mathcal{M}\}$ are independent processes across sensors, then the right hand side term in (14) becomes

$$\int_{\tilde{\mathbf{g}}, \tilde{\mathbf{H}}} V(\tilde{\mathbf{g}}, \tilde{\mathbf{H}}, \tilde{\mathbf{B}}) \times \prod_{m=1}^{M} \left(\mathbb{G}_m(\tilde{g}_m|g_m) \times \mathbb{H}_m(\tilde{H}_m|H_m) \right) d\tilde{\mathbf{g}} d\tilde{\mathbf{H}}$$

where $\mathbb{G}_m$ and $\mathbb{H}_m$ are the probability transition laws of the mth sensor's Markovian processes $\mathbf{g}_{m,(\cdot)}$ and $\mathbf{H}_{m,(\cdot)}$, respectively.

(iii) If processes $\{\mathbf{g}_{m,(\cdot)} : m \in \mathcal{M}\}$ and $\{\mathbf{H}_{m,(\cdot)} : m \in \mathcal{M}\}$ are i.i.d. processes over time and across the sensors, then the right hand side term in (14) becomes

$$\int_{\tilde{\mathbf{g}}, \tilde{\mathbf{H}}} V(\tilde{\mathbf{g}}, \tilde{\mathbf{H}}, \tilde{\mathbf{B}}) \times \prod_{m=1}^{M} \left(\mathbb{P}(\tilde{g}_m) \times \mathbb{P}(\tilde{H}_m) \right) d\tilde{\mathbf{g}} d\tilde{\mathbf{H}}.$$

□

We solve the Bellman equation (12) by the use of *relative value iteration* algorithms (see Chapter 7 in [27]). Note that a discretized version of the Bellman equation (12) is used to

facilitate numerical computation for finding suboptimal solutions to the stochastic control problem (11), where the fading channel, harvested energy and battery level state spaces and the energy allocation space are discretized. As the number of discretization levels increases, it is expected that these discretized (suboptimal) solutions converge to the optimal solutions [30]. As mentioned earlier, the optimal energy allocation solution as a function of these discretized state values of $(\mathbf{g}, \mathbf{H}, \mathbf{B})$ is stored in a lookup table on-board all the sensors, which can then select the appropriate energy level in real-time based on available causal information about $(\mathbf{g}, \mathbf{H}, \mathbf{B})$.

V. SOME STRUCTURAL RESULTS ON THE OPTIMAL ENERGY ALLOCATION POLICIES

In this section the structure of the optimal transmission energy allocation policy (13) is studied for the infinite-time horizon stochastic control problem (11). The idea is to apply the notion of submodularity (see [31]) to the recursive Bellman equation (12). Following the same arguments one can show similar structural results for the finite-time horizon optimal energy allocation strategy (6).

We now present the relative value iteration algorithm to solve the Bellman equation (12). It is used to construct structural results for the optimal transmission policy. First, we consider the Bellman equation for the finite T-horizon stochastic control problem:

$$V_t(\mathbf{g}, \mathbf{H}, \mathbf{B})$$
$$= \min_{\mathbf{0} \preceq \mathbf{E} \preceq \mathbf{B}} \left\{ D(\mathbf{E}, \mathbf{s}) + \mathbb{E}\left[V_{t+1}(\tilde{\mathbf{g}}, \tilde{\mathbf{H}}, \tilde{\mathbf{B}}) \big| \mathbf{g}, \mathbf{H}, \mathbf{E}\right] \right\} \quad (15)$$

with terminal condition $V_{T+1}(\mathbf{g}, \mathbf{H}, \mathbf{B}) = 0$. We now define the function

$$H_t(\cdot, \cdot, \cdot) := V_t(\cdot, \cdot, \cdot) - V_t(\mathbf{g}_f, \mathbf{H}_f, \mathbf{B}_f), \quad 0 \le t \le T \quad (16)$$

where $(\mathbf{g}_f, \mathbf{H}_f, \mathbf{B}_f) \ne (\mathbf{g}_0, \mathbf{H}_0, \mathbf{B}_0)$ is fixed. We then have the following relative value iteration algorithm recursion

$$H_t(\mathbf{g}, \mathbf{H}, \mathbf{B})$$
$$= \min_{\mathbf{0} \preceq \mathbf{E} \preceq \mathbf{B}} \left\{ D(\mathbf{E}, \mathbf{s}) + \mathbb{E}\left[V_{t+1}(\tilde{\mathbf{g}}, \tilde{\mathbf{H}}, \tilde{\mathbf{B}}) \big| \mathbf{g}, \mathbf{H}, \mathbf{E}\right] \right\}$$
$$- \min_{\mathbf{0} \preceq \mathbf{E} \preceq \mathbf{B}} \left\{ \mathbb{E}\left[D(\mathbf{E}, \mathbf{s}) \big| \mathbf{g} = \mathbf{g}_f\right] \right.$$
$$\left. + \mathbb{E}\left[V_{t+1}(\tilde{\mathbf{g}}, \tilde{\mathbf{H}}, \tilde{\mathbf{B}}) \big| \mathbf{g} = \mathbf{g}_f, \mathbf{H} = \mathbf{H}_f, \mathbf{E} = \mathbf{E}_f\right] \right\}$$
$$\quad (17)$$

for $0 \le t \le T-1$. It can be shown that the relative value recursion (17) converges to the optimal solution ρ of the infinite-time horizon average cost Bellman equation (12) as T goes to infinity (see the discussion on page 391 in Chapter 7 of [27]).

Lemma 5.1: Given $\mathbf{g}$ and $\mathbf{H}$, the value function $V_t(\mathbf{g}, \mathbf{H}, \mathbf{B})$ in (15) is convex in $\mathbf{B}$ for $0 \le t \le T$.

Proof: First, note that, for given $\mathbf{g}$ and $\mathbf{H}$, the final time value function

$$V_T(\mathbf{g}, \mathbf{H}, \mathbf{B}) = \min_{\mathbf{0} \preceq \mathbf{E} \preceq \mathbf{B}} D(\mathbf{E}, \mathbf{s}) = D(\mathbf{B}, \mathbf{s})$$

is a convex function in $\mathbf{B}$ by Theorem 2.1. Now assume that $V_{t+1}(\mathbf{g}, \mathbf{H}, \mathbf{B})$, $0 \le t \le T-1$, is convex in $\mathbf{B}$ for given $\mathbf{g}$ and $\mathbf{H}$. Then, for given $\mathbf{H}$ and $\mathbf{E}$, the function

$$V_{t+1}\left(\mathbf{g}, \mathbf{H}, \min\{\mathbf{B} - \mathbf{E} + \mathbf{H}, \hat{\mathbf{B}}\}\right)$$

is convex in $\mathbf{B}$, since it is the minimum of $V_{t+1}(\mathbf{g}, \mathbf{H}, \hat{\mathbf{B}})$ which is a constant independent of $\mathbf{B}$, and by the induction hypothesis the convex function $V_{t+1}(\mathbf{g}, \mathbf{H}, \mathbf{B} - \mathbf{E} + \mathbf{H})$ in $\mathbf{B}$. Since the expectation operator preserves convexity,

$$\mathbb{E}\left[V_{t+1}\left(\tilde{\mathbf{g}}, \tilde{\mathbf{H}}, \min\{\mathbf{B} - \mathbf{E} + \mathbf{H}, \hat{\mathbf{B}}\}\right) \big| \mathbf{g}, \mathbf{H}, \mathbf{E}\right]$$

given in (15) is a convex function in $\mathbf{B}$. But, $V_t(\mathbf{g}, \mathbf{H}, \mathbf{B})$ in (15) is the infimal convolution of two convex functions in $\mathbf{B}$ for given $\mathbf{g}$ and $\mathbf{H}$ and hence is convex in $\mathbf{B}$ (see the proof of Theorem 1 in [13]). □

Definition 5.1: [[32] after [31]] A function $F(x, y) : X \times Y \to S$ is submodular in (x, y) if $F(x_1, y_1) + F(x_2, y_2) \le F(x_1, y_2) + F(x_2, y_1)$ for all $x_1, x_2 \in X$ and $y_1, y_2 \in Y$ such that $x_1 \ge x_2$ and $y_1 \ge y_2$. □

It is important to note that the submodularity is a sufficient condition for optimality of monotone increasing policies. Specifically, if $F(x, y)$ defined above is submodular in (x, y) then $y(x) = \arg\min_y F(x, y)$ is non-decreasing in x. We now present the main Theorem of this section which gives structural results on the optimal energy allocation policies (13).

Theorem 5.1: Given $\mathbf{g}$ and $\mathbf{H}$, the optimal energy allocation policy of any sensor $m \in \mathcal{M}$, $E_m^o(\mathbf{g}, \mathbf{H}, \mathbf{B})$, given in (13) is non-decreasing in B_m regardless of all other sensors' policies, $E_n^o(\mathbf{g}, \mathbf{H}, \mathbf{B})$, $n \ne m$, for all $m \in \mathcal{M}$.

Proof: For given sensor $m \in \mathcal{M}$, we assume $\mathbf{B}_{(-m)} = (B_1, \ldots, B_{m-1}, B_{m+1}, \ldots, B_M)$ and $\mathbf{E}_{(-m)} = (E_1, \ldots, E_{m-1}, E_{m+1}, \ldots, E_M)$ are fixed. Then, for fixed $\mathbf{g}$ and $\mathbf{H}$ the optimal energy policy of the sensor m can be written as

$$E_m^o(\mathbf{g}, \mathbf{H}, \mathbf{B})$$
$$= \arg \min_{\mathbf{0} \preceq \mathbf{E} \preceq \mathbf{B}} \left\{ D(\mathbf{E}, \mathbf{s}) \right.$$
$$\left. + \mathbb{E}\left[V\left(\tilde{\mathbf{g}}, \tilde{\mathbf{H}}, \min\{\mathbf{B} - \mathbf{E} + \mathbf{H}, \hat{\mathbf{B}}\}\right) \big| \mathbf{g}, \mathbf{H}, \mathbf{E}\right] \right\}$$
$$=: L(B_m, E_m) \quad (18)$$

by the fact that B_n and E_n are fixed for all $n \ne m$, and E_m is only constrained by its own battery energy storage B_m. We aim to show that $L(B_m, E_m)$ is submodular in (B_m, E_m) based on Definition 5.1, i.e., for every $E_m' \ge E_m$ and $B_m' \ge B_m$, we have

$$L\left(B_m', E_m'\right) - L\left(B_m, E_m'\right) \le L\left(B_m', E_m\right) - L(B_m, E_m). \quad (19)$$

It is evident that $\mathbb{E}[D(\mathbf{E}, \mathbf{s}) | \mathbf{g}]$ is submodular in (B_m, E_m) since it is independent of B_m. Let $\mathbf{x} = (x_1, \ldots x_M)$ such that $x_i = B_i - E_i$ for $i \in \{1, \ldots, M\}$, then we denote

$$Z(x_m) := \mathbb{E}\left[V\left(\tilde{\mathbf{g}}, \tilde{\mathbf{H}}, \min\{\mathbf{B} - \mathbf{E} + \mathbf{H}, \hat{\mathbf{B}}\}\right) \big| \mathbf{g}, \mathbf{H}, \mathbf{E}\right]$$

where $x_n = B_n - E_n$ is fixed for all $n \ne m$. Since $Z(x_m)$ is convex in x_m (by Lemma 5.1) we have

$$Z(x + \varepsilon) - Z(x) \le Z(y + \varepsilon) - Z(y), \quad x \le y, \ \varepsilon \ge 0$$

(see Proposition 2.2.6 in [33]). Now let $x = B_m - E'_m$, $y = B_m - E_m$ and $\varepsilon = B'_m - B_m$. Then, we have the submodularity condition (19) for $Z(x_m)$ [13]. Therefore, $L(B_m, E_m)$ is submodular in (B_m, E_m). Note that submodularity is a sufficient condition for optimality of monotone increasing policies, i.e., since $L(B_m, E_m)$ is submodular in (B_m, E_m) then $E^o(B_m) = \arg\min_{E_m} L(B_m, E_m)$ is non-decreasing in B_m (see [31]). $\square$

The structural result of Theorem 5.1 implies that, for fixed $\mathbf{g}$, $\mathbf{H}$, $\mathbf{B}_{(-m)} = (B_1, \ldots, B_{m-1}, B_{m+1}, \ldots, B_M)$ and $\mathbf{E}_{(-m)} = (E_1, \ldots, E_{m-1}, E_{m+1}, \ldots, E_M)$, if E^*_m is the unique solution to the convex unconstrained minimization problem

$$E^o_m(\mathbf{g}, \mathbf{H}, \mathbf{B}) = \arg\min_{E_m} \{D(\mathbf{E}, \mathbf{s})$$
$$+ \mathbb{E}\left[V\left(\tilde{\mathbf{g}}, \tilde{\mathbf{H}}, \min\{\mathbf{B} - \mathbf{E} + \mathbf{H}, \hat{\mathbf{B}}\}\right) \middle| \mathbf{g}, \mathbf{H}, \mathbf{E}\right]\}$$

which can be easily solved using numerical techniques such as a bisection search. Then, the solution to the problem (13) where $0 \le E_m \le B_m$ will be of the form

$$E^o_m(\mathbf{g}, \mathbf{H}, \mathbf{B}) = \begin{cases} 0 & \text{if } E^*_m \le 0 \\ E^*_m & \text{if } 0 < E^*_m < B_m \\ B_m & \text{if } E^*_m \ge B_m. \end{cases}$$

In other words, the search for the optimal E^o_m can be carried out in one direction for a given B_m. Note that since E^o_m is a function of E^o_n, $n \ne m$, the computation of the optimal energy policy of sensor m, E^o_m, cannot be carried out independently from those of other sensors E^o_n, $n \ne m$.

A. Threshold Policy for Binary Energy Allocation Levels

Note that while solving for the optimal energy allocation level in the Bellman equation requires not only discretization of the state space but also that of the action space. However, the discretization of the action space to a finite number of energy allocation levels is not often an issue as in practice, a sensor transmitter can be programmed to have a finite number of transmission power/energy levels. In fact, for simplicity of implementation, often a sensor can be equipped with only two power/energy levels for transmission. Thus it is perfectly natural to consider the scenario where the energy allocation space is binary.

In case that the transmission energy allocation control of sensor m, E_m, belongs to a two element set $\{u_0, u_1\}$ where $u_0 < u_1$, the monotonicity of Theorem 5.1 yields a threshold structure. This threshold structure implies that, for fixed $\mathbf{g}$, $\mathbf{H}$, $\mathbf{B}_{(-m)}$ and $\mathbf{E}_{(-m)}$, the optimal transmission energy allocation policy is threshold of the form

$$E^o_m(\mathbf{g}, \mathbf{H}, \mathbf{B}) = \begin{cases} u_0 & \text{if } B_m \le B^*_m(\mathbf{g}, \mathbf{H}, \mathbf{B}) \\ u_1 & \text{otherwise} \end{cases} \quad (20)$$

where $B^*_m(\mathbf{g}, \mathbf{H}, \mathbf{B})$ is the battery storage threshold.

The threshold structure of the optimal transmission energy allocation policy in the case of a binary energy allocation set simplifies the implementation of the optimal energy allocation significantly. However, this requires the knowledge of the battery energy threshold B^* above. In general, there is no closed form expression for B^*, but it can be found via iterative search

algorithms. Here we present a gradient algorithm based on Algorithm 1 in [34] to find the threshold.

First, we establish some notation. For fixed $\mathbf{g}$, $\mathbf{H}$, $\mathbf{B}_{(-m)}$ and $\mathbf{E}_{(-m)}$ denote

$$J_k(B^*_m) := \{D(\mathbf{E}, \mathbf{s})$$
$$+ \mathbb{E}\left[V_k\left(\tilde{\mathbf{g}}, \tilde{\mathbf{H}}, \min\{\mathbf{B} - \mathbf{E} + \mathbf{H}, \hat{\mathbf{B}}\}\right) \middle| \mathbf{g}, \mathbf{H}, \mathbf{E}\right]\} \quad (21)$$

where the threshold policy E_m is defined based on (20) as

$$E^o_m = \begin{cases} u_0 & \text{if } B_m \le B^*_m \\ u_1 & \text{otherwise.} \end{cases}$$

The term $J_k(\cdot)$ above is the first term of right hand side expression in (17) without the minimization, where the threshold policy E^o_m, depending on the threshold policy B^*, is used in the relative value iteration.

Set a pre-specified accuracy parameter ε. For $n \in \mathbb{N}$, $0.5 < \kappa \le 1$ and $\omega, \varsigma > 0$ we denote $\omega_n := \frac{\omega}{(n+1)^\kappa}$ and $\varsigma_n := \frac{\varsigma}{(n+1)^\kappa}$. Then for fixed $\mathbf{g}$, $\mathbf{H}$, $\mathbf{B}_{(-m)}$ and $\mathbf{E}_{(-m)}$, the computation of the optimal battery storage threshold B^*_m is carried out by the following algorithm:

Algorithm 1 Gradient algorithm for computing the threshold

1: Initialize the battery storage threshold $B^{(0)}_m$.
2: **repeat**
3: For $n = 0, 1, \ldots$
    1) Compute the gradient:

$$\partial_{B_m} J^{(n)}_k := \frac{J_k\left(B^{(n)}_m + \omega_n\right) - J_k\left(B^{(n)}_m - \omega_n\right)}{2\omega_n}. \quad (22)$$

    2) Update the battery storage threshold of sensor m via

$$B^{(n+1)}_m = B^{(n)}_m - \varsigma_n \partial_{B_m} J^{(n)}_k$$

    which gives

$$E^{(n+1)}_m(\mathbf{g}, \mathbf{H}, \mathbf{B}) = \begin{cases} u_0 & \text{if } B_m \le B^{(n+1)}_m \\ u_1 & \text{otherwise.} \end{cases}$$

    3) Compute $J_k(B^{(n+1)}_m)$ using (21).
4: **until** Convergence: $\frac{|B^{n+1}_m - B^n_m|}{B^n_m} \le \varepsilon$.

The above algorithm is a gradient-estimate based algorithm (see [35]) for estimating the optimal threshold B^* where only measurements of the loss function is available (i.e., no gradient information). We note that (22) evaluates an approximation to the gradient. This algorithm generates a sequence of estimates for the threshold policy B^* which converges to a local minimum with corresponding energy allocation u^*. The reader is referred to [35] for associated convergence analysis of this and other related algorithms (see e.g., Theorem 7.1 in [35]). Note that gradient-estimate based algorithms are sensitive to initial conditions and should be evaluated for several distinct initial conditions to find the best local minimum.

VI. SUBOPTIMAL DECENTRALIZED SOLUTIONS: SENSORS WITH ONLY STATISTICAL INFORMATION ABOUT CHANNEL GAINS AND HARVESTING ENERGIES

The models in Sections III and IV and assume that each sensor has perfect causal information about the fading channels gains and harvested energies of all sensors. In this section, we consider the more realistic scenario where each sensor has only statistical information about all other sensors' channel gains and harvested energies.

We now aim to construct a suboptimal decentralized strategy for a generic sensor $n \in \mathcal{M}$. Note that any other sensor applies similar suboptimal algorithm in a decentralized manner independently of all other sensors.

For the generic sensor $n \in \mathcal{M}$, we define

$$\bar{D}_n \left(\mathbf{E}_{(:,k)}, g_{n,k}, \mathbb{P}\left(\mathbf{g}_{(-n,k)} \right) \right)$$
$$= \sigma_\theta^2 \int_{\mathbf{g}_{(-n,k)}} \left(\sum_{m=1}^{M} \frac{\sigma_\theta^2 E_{m,k} s_{m,k}}{1 + \gamma_m^{-1} \sigma_\theta^2 E_{m,k} s_{m,k}} \right)^{-1}$$
$$\times \mathbb{P}\left(\mathbf{g}_{(-n,k)} \right) d\mathbf{g}_{(-n,k)}$$

where we recall that $\mathbf{g}_{(-n,k)} = [g_{1,0}, \dots, g_{n-1,k}, g_{n+1,k}, \dots, g_{M,k}]^T$, $s_{m,k} = \frac{g_{m,k}}{\xi_m^2 (\sigma_\theta^2 + \sigma_m^2)}$.

Sensor n applies the following decentralized model for the battery storage of any other sensor $m \in \mathcal{M}$, $\bar{B}_{m,(\cdot)}$:

$$\bar{B}_{m,k+1} = \min \left\{ \bar{B}_{m,k} + \bar{H}_{m,k} - E_{m,k}, \hat{B}_m \right\}, \tag{23}$$

where $\bar{B}_{m,0} = B_{m,0}$, and $\bar{H}_{m,k} = \mathbb{E}(H_{m,k})$ is the mean of the harvested energy of sensor m at time k computed by sensor n by the use of the density function $\mathbb{P}(H_{m,k})$. Since sensor n has perfect information about its own harvested energy, we have $\bar{H}_{n,k} = H_{n,k}$ which yields (1).

We present some notation. Let μ be a probability density function on the state of fading channel gains $\mathbf{g}_{(-n)} = (g_1, \dots, g_{n-1}, g_{n+1}, \dots, g_M)^T$, and ν be a probability density function on the state of harvested energies $\mathbf{H}_{(-n)} = (H_1, \dots, H_{n-1}, H_{n+1}, \dots, H_M)^T$ at time $k \geq 0$. Then, we denote the corresponding density functions of the fading channel gains and harvested energies at time $k+1$ by $\tilde{\mu}$ and $\tilde{\nu}$, respectively. Since for any sensor $m \in \mathcal{M}$ both fading channel gains $\mathbf{g}_{m,(\cdot)}$ and harvested energies $\mathbf{H}_{m,(\cdot)}$ are modelled as first-order homogeneous Markov processes, we may write $\tilde{\mu} = \tilde{\mathbb{G}} \mu$, and $\tilde{\nu} = \tilde{\mathbb{H}} \nu$ where $\tilde{\mathbb{G}}$ and $\tilde{\mathbb{H}}$ are the probability transition matrices for $\mathbf{g}_{(-n)}$ and $\mathbf{H}_{(-n)}$, respectively. Given the vector of battery storages $\bar{\mathbf{B}} := (\bar{B}_1, \dots, \bar{B}_M)^T$ and the sensors' energy consumptions $\mathbf{E} := (E_1, \dots, E_M)^T$ at time $k \geq 0$, we denote the corresponding battery storages at time $k+1$ by $\tilde{\mathbf{B}} := (\tilde{B}_1, \dots, \tilde{B}_M)^T$ such that the dynamics of $\tilde{B}_{m,(\cdot)}$ is given by (23) as $\tilde{B}_m = \min\{\bar{B}_m + \mathbb{E}(H_m) - E_m, \hat{B}_m\}$.

We now present the solution to the decentralized version of stochastic control problem (11) in the following theorem.

Theorem 6.1: For any sensor $n \in \mathcal{M}$, independent of the initial condition $\{g_{m,0}, H_{m,0}, B_{m,0} : m \in \mathcal{M}\}$, the value of the infinite-time horizon stochastic control problem (11) with de-

centralized information is given by ρ which is the solution of the average-cost optimality (Bellman) equation

$$\rho + V(g, H, \mu, \nu, \bar{\mathbf{B}}) = \min_{0 \preceq \mathbf{E} \preceq \bar{\mathbf{B}}} \left\{ \mathbb{E}\left[D_n(\mathbf{E}, \mathbf{s}) | g, \mu \right] \right.$$
$$\left. + \mathbb{E}\left[V(\tilde{g}, \tilde{H}, \tilde{\mu}, \tilde{\nu}, \tilde{\mathbf{B}}) \,\middle|\, g, H, \mu, \nu, \mathbf{E} \right] \right\} \tag{24}$$

where V is called the relative value function.

Proof: The proof is similar to the proof of Theorem 4.1 and is omitted. □

The stationary decentralized optimal solution computed by sensor n is then given by

$$\bar{\mathbf{E}}^o(g, H, \mu, \nu, \bar{\mathbf{B}}) = \arg \min_{0 \preceq \mathbf{E} \preceq \bar{\mathbf{B}}} \left\{ \mathbb{E}\left[D_n(\mathbf{E}, \mathbf{s}) | g, \mu \right] \right.$$
$$\left. + \mathbb{E}\left[V(\tilde{g}, \tilde{H}, \tilde{\mu}, \tilde{\nu}, \tilde{\mathbf{B}}) \,\middle|\, g, H, \mu, \nu, \mathbf{E} \right] \right\} \tag{25}$$

such that $\bar{\mathbf{E}}^o = (\bar{E}_1^o, \cdots, \bar{E}_n^o, \cdots, \bar{E}_M^o)$ which is the optimal strategies for all sensors from the perspective of sensor n using perfect information about its own fading channel gain and harvested energy (g, H) together with only statistical information about the other sensors' fading channel gains and harvested energies (μ, ν). Note that sensor n only applies its own control E_n^o.

We first simplify the terms in (24). First, we have

$$\mathbb{E}\left[D_n(\mathbf{E}, \mathbf{s}) | g, \mu \right] = \bar{D}_n(\mathbf{E}, g, \mu)$$
$$= \sigma_\theta^2 \times \int_{\mathbf{g}_{(-n)}} \left(\sum_{m=1}^{M} \frac{\sigma_\theta^2 E_m s_m}{1 + \gamma_m^{-1} \sigma_\theta^2 E_m s_m} \right)^{-1} \times \mu(\mathbf{g}_{(-n)}) d\mathbf{g}_{(-n)}$$

where $s_m = \frac{g_m}{\xi_m^2 (\sigma_\theta^2 + \sigma_m^2)}$ and $0 \leq E_m \leq \bar{B}_m$. Second, we may write

$$\mathbb{E}\left[V(\tilde{g}, \tilde{H}, \tilde{\mu}, \tilde{\nu}, \tilde{\mathbf{B}}) \,\middle|\, g, H, \mu, \nu, \mathbf{E} \right]$$
$$= \int_{\tilde{\mathbf{g}}, \tilde{\mathbf{H}}} V(\tilde{g}, \tilde{H}, \tilde{\mu}, \tilde{\nu}, \tilde{\mathbf{B}}) \tilde{\mu}(\tilde{\mathbf{g}}_{-n}) \tilde{\nu}(\tilde{\mathbf{H}}_{-n}) \mathbb{P}(\tilde{g}|g) \mathbb{P}(\tilde{H}|H) d\tilde{\mathbf{g}} d\tilde{\mathbf{H}}$$

where $\tilde{\mu} = \tilde{\mathbb{G}}\mu$, $\tilde{\nu} = \tilde{\mathbb{H}}\nu$, and $\tilde{\mathbf{B}} = (\tilde{B}_1, \dots, \tilde{B}_M)^T$ such that $\tilde{B}_m = \min\{\bar{B}_m + \mathbb{E}(H_m) - E_m, \hat{B}_m\}$ for $m \in \mathcal{M}$.

We note that this decentralized solution is obviously suboptimal but reduces the communication overhead and is hence more practical.

VII. NUMERICAL RESULTS

In this section, we present some numerical results on the performance of the optimal energy allocation strategies for both causal and non-causal side information. We focus on the infinite-horizon case primarily because finite horizon results were presented in [21]. Note that for the infinite-horizon case, the results for the non-causal scenario was obtained from the finite-horizon optimal energy allocation with non-causal information, with a suitably long time horizon used to obtain a long-term average for comparison purposes. For obtaining these numerical results, we let the variance of the source signal σ_θ^2 be equal to 1 mW. First, we consider the case of a single sensor with the following parameters: $\xi^2 = 0.01$ mW, $\sigma^2 = 0.01$ mW and $\gamma = 100$ defined in Section II. Both channel power gains
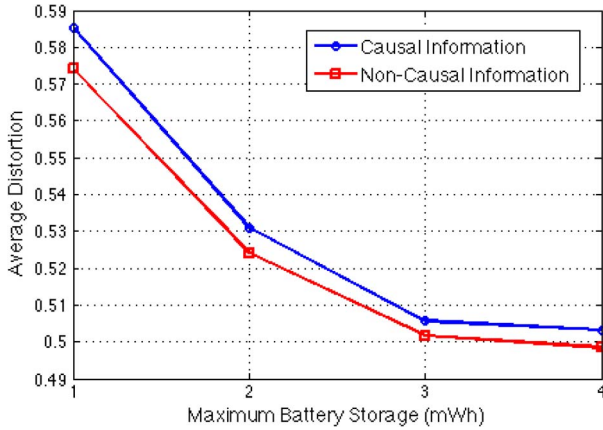
Fig. 3. One sensor case with i.i.d channel gains and harvesting energies: Infinite-time horizon average distortion versus the maximum battery storage (mWh).
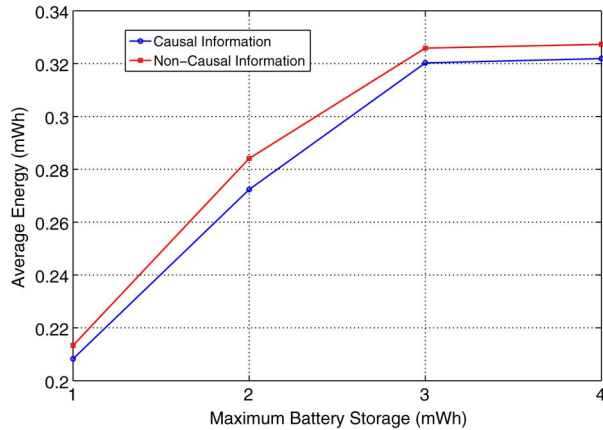

Fig. 4. One sensor case with i.i.d channel gains and harvesting energies: Infinite-time horizon average transmission energy (mWh) versus the maximum battery storage (mWh).

$\{g_k\}$ and harvested energies $\{H_k\}$ are assumed to be i.i.d. exponentially distributed with probability density function (p.d.f) of the forms $\mathbb{P}(g_k) = \frac{1}{\bar{g}} \exp(-g_k/\bar{g})$ and $\mathbb{P}(H_k) = \frac{1}{\bar{H}} \exp(-H_k/\bar{H})$, respectively, with $\bar{g}$ and $\bar{H}$ being their means. We now fix the mean of the fading channel gains to $\bar{g} = 0.1$ (or -10 dBm) and the mean of harvested energy to $\bar{H} = 4$ milliwatt hour (mWh). Then, we plot in Fig. 3 the average distortion versus the maximum battery storage energy for the infinite-time horizon formulation (11) for both cases of causal and non-causal information. We note that the performance gets better as the maximum battery storage energy increases in both cases since the sensor has more flexibility in choosing transmission energy levels. Fig. 3 also shows that in this setting the performance for the non-causal information case is generally better than the performance of system with causal information. However, the performance gap decreases as the battery capacity increases (e.g., from 17.54% at a battery capacity of 1 mWh to 8.8% at a battery capacity of 4 mWh). For this case we also plot in Fig. 4 the corresponding average transmission energy (mWh) versus the maximum battery storage energy. Note that as the maximum battery storage increases, the average energy consumption also increases. However, since the average energy consumption is upper bounded by the mean of the harvested energy process
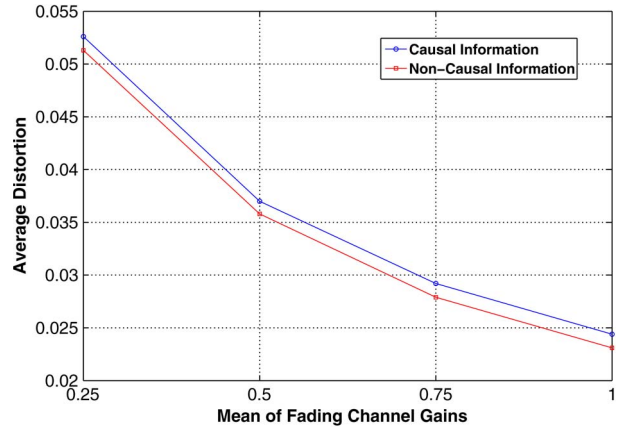

Fig. 5. One sensor case with i.i.d channel gains and harvesting energies: Infinite-time horizon average transmission energy (mWh) versus the mean of fading channel gains.

(see Remark 3.3), it eventually converges to the average energy consumption of the infinite maximum energy storage capacity case (see Section III-C). The non-causal information pattern allows the sensor to spend more energy on average to achieve a smaller distortion.

We now fix the mean of harvested energy to $\bar{H} = 1$ (mW) and the maximum battery storage energy to 2 (mWh). For the infinite-time horizon formulation (11) the average distortion versus the mean of the fading channel gains is plotted in Fig. 5 for both cases of causal and non-causal information. As shown in Fig. 5 the performance gets better as the mean of the fading channel gains increases in both cases. For example, in the causal information case, the average distortion decreases by almost 33% as the average channel gain increases by 2 dB (from 0.5 to 1).

Second, we take multiple sensors with following parameters $\xi_m^2 = 0.01$ mW and $\sigma_1^2 = \sigma_3^2 = 0.01$ mW, whereas $\sigma_2^2 = \sigma_4^2 = 0.0125$ mW. This results in $\gamma_1 = \gamma_3 = 100$ and $\gamma_2 = \gamma_4 = 80$, as defined in Section II. The fading channel gains and harvested energies are assumed to be 2-level discrete Markov chains with values $\{0, 0.05\}$ and $\{0, 0.25\}$, respectively, and the transition matrix $\begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix}$. We assume uniform maximum storage energy capacity $\hat{B}_m = \hat{B}$ for all m. A comparative performance evaluation of multi-sensors for the infinite-time horizon formulation (11) with causal fading channel gains and energy harvesting information is illustrated in Fig. 6 which shows the performance gets better as the maximum battery storage increases. Fig. 6 also illustrates the advantage of multiple sensors as the distortion performance can be improved by increasing the number of sensors.

Similar simulation results are shown in Fig. 7 in the case of 3-level discrete Markov chains for fading channel gains and harvested energies with values $\{0, 0.05, 0.1\}$ and $\{0, 0.25, 0.5\}$, respectively. The common transition matrix for these Markov chains is given by $\mathbb{T} = \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.5 & 0.3 & 0.2 \end{bmatrix}$. We can see that in the optimal energy allocation scheme, the allocated energy values will depend not only on the current channel gain and
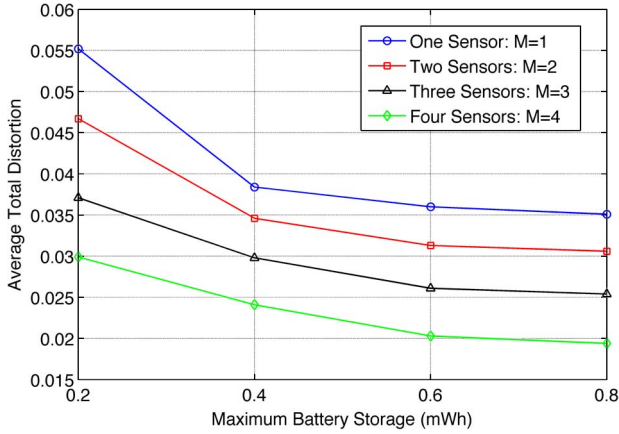
Fig. 6. Multi-sensor case with 2-level Markovian channel gains and harvesting energies: Infinite-time horizon average distortion versus the maximum battery storage (mWh).
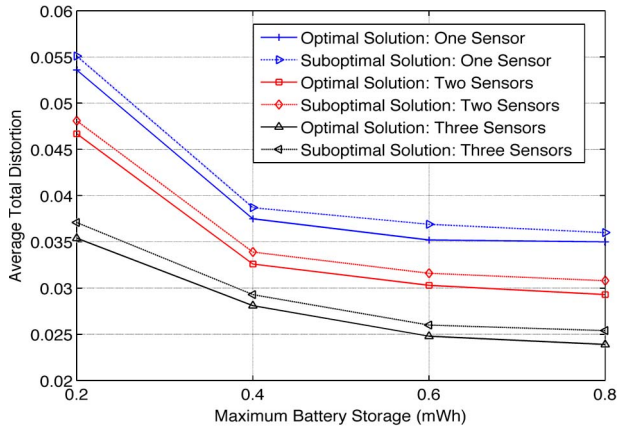

Fig. 7. Optimal and suboptimal solutions to multi-sensor case with 3-level Markovian channel gains and harvesting energies: Infinite-time horizon average distortion versus the maximum battery storage (mWh).

current distortion but also on the energy harvesting through the battery storage. The allocated energy tends to be higher when the distortion is larger, provided the corresponding channel gain and the battery storage are not too small.

In Fig. 7 we further plot the suboptimal solutions where each sensor only knows its own channel and harvested energy together with the statistics of all other sensors' channels and harvested energies. Similar to Fig. 6 the performance gets better as the maximum battery storage increases. Fig. 7 also indicates that, as expected, the performance for the optimal solution with full exact information of other sensors' channels and harvested energies is better than the performance of the suboptimal solution with only statistical information about channel gains and harvested energies of other sensors (see Section VI). Note that the decentralized suboptimal solution reduces the communication overhead and is hence more practical.

Threshold Policy: We now consider the case that the transmission energy allocation control belongs to a two element set $\{0, 0.5\}$ in the infinite-time horizon formulation (11). As explained at the end of Section V the optimal transmission energy allocation policy is threshold of the form

$$E^o(g, H, B) = \begin{cases} 0 & \text{if } B \leq B^*(g, H, B) \\ 0.5 & \text{otherwise} \end{cases}$$
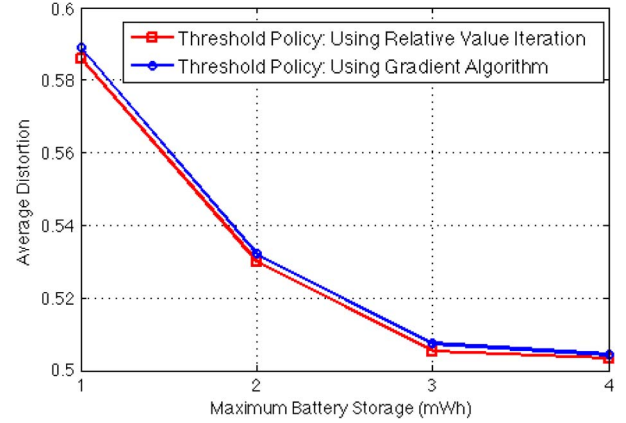

Fig. 8. Performance of threshold policy using (i) relative value iteration algorithm, and (ii) gradient algorithm.

where $B^*(\cdot, \cdot, \cdot)$ is the corresponding battery storage threshold. We assume one sensor with i.i.d. exponentially distributed channel gains and harvested energies where $\bar{g} = 4$ (dB) and $\bar{H} = 4$ (mWh). The maximum battery storage is $\hat{B} = 1$ (mWh). Fig. 8 shows the simulation results where the relative value iteration algorithm and the threshold policy based algorithm are used. In the gradient algorithm, the parameters of Section V-A are set to $\omega = 0.1$, $\varsigma = 0.5$ and $\kappa = 1$. It can be seen that there is a small gap between the simulation results obtained via the two methods. This can be attributed to the fact that the optimal threshold is not exactly calculated by the gradient algorithm which only converges to a local minimum.

VIII. CONCLUSION

We considered the problem of minimizing the distortion incurred in estimating a random source via multiple sensors equipped with energy harvesting technology. This problem is formulated as a Markov decision process and then optimal energy allocation policies are obtained by the use of dynamic programming techniques. Using the concept of submodularity, the structure of the optimal allocation energy policies is studied. Suboptimal solutions are also discussed for the more realistic case that each sensor only knows its own channel and harvested energy together with the statistics of all other sensors' channels and harvested energies. Numerical results were presented to illustrate the distortion performances corresponding to these scenarios. Future generalizations of this work will consider, amongst others, battery storage imperfections in charging and discharging, non-orthogonal multi-access schemes and mutual interference amongst the sensor transmissions.

APPENDIX

Proof of Theorem 4.1: The proof follows from the average-cost optimality equations for stochastic control problems (see e.g., [27]). We first show that the following inequality holds:

$$\rho + V(\mathbf{g}, \mathbf{H}, \mathbf{B}) \geq \min_{\mathbf{0} \preceq \mathbf{E} \preceq \mathbf{B}} \left\{ D(\mathbf{E}, \mathbf{s}) + \mathbb{E}\left[V(\tilde{\mathbf{g}}, \tilde{\mathbf{H}}, \tilde{\mathbf{B}}) \middle| \mathbf{g}, \mathbf{H}, \mathbf{E} \right] \right\}$$
$$\text{(A.1)}$$

by verifying conditions (W) and (B) of [36] that guarantee the existence of solutions to (A.1) for MDPs with general state space. Denote the state space $\mathcal{S}$ and action space $\mathcal{A}$, i.e., $(\mathbf{g}_{(:,k)}, \mathbf{H}_{(:,k)}, \mathbf{B}_{(:,k)}) \in \mathcal{S}$ and $\mathbf{E}_{(:,k)} \in \mathcal{A}$. Condition (W) of [36] refers to some compactness conditions on the state and action space and some continuity conditions on the transition probability functions for specific state and action pairs, and the cost function. In our notation these conditions can be summarised as:

0) The state space $\mathcal{S}$ is locally compact.
1) Let $U(\cdot)$ be the mapping that assigns to each $(\mathbf{g}_{(:,k)}, \mathbf{H}_{(:,k)}, \mathbf{B}_{(:,k)})$ the nonempty set of available actions. Then $U((\mathbf{g}_{(:,k)}, \mathbf{H}_{(:,k)}, \mathbf{B}_{(:,k)})$ lies in a compact subset of $\mathcal{A}$ and $U(\cdot)$ is upper semicontinuous.
2) The transition probabilities are weakly continuous.
3) The one-step cost $D(\cdot, \cdot)$ is lower semicontinuous.

By $\mathbf{E}_{(:,k)} \preceq \mathbf{B}_{(:,k)}$ where $B_m \leq B_{max}$ for all $m \in \mathcal{M}$, conditions 0) and 1) of (W) can be easily verified. Condition 2) follows from (1), while condition 3) follows from the definition of the distortion (2) which is bounded.

A popular technique for proving average optimality is by defining the following discounted cost and then letting the discount factor $\rho \to 1$:

$$v_\rho(\mathbf{g}_0, \mathbf{H}_0, \mathbf{B}_0) = \inf_{\{\mathbf{E}_{(:,k)} : k \geq 0\}} \mathbb{E}\left[\sum_{k=0}^{\infty} \rho^k \, \mathbb{E}\left[D(\mathbf{E}, \mathbf{s}) | \mathbf{g}\right] \middle| \mathbf{g}_0, \mathbf{H}_0, \mathbf{B}_0\right]$$

where $0 < \rho < 1$. Define also

$$m_\rho = \inf_{(\mathbf{g}_0, \mathbf{H}_0, \mathbf{B}_0)} v_\rho(\mathbf{g}_0, \mathbf{H}_0, \mathbf{B}_0)$$

$$w_\rho(\mathbf{g}_0, \mathbf{H}_0, \mathbf{B}_0) = v_\rho(\mathbf{g}_0, \mathbf{H}_0, \mathbf{B}_0) - m_\rho$$

Then Condition (B) of [36] in our notation implies that

$$\sup_{\rho < 1} w_\rho(\mathbf{g}_0, \mathbf{H}_0, \mathbf{B}_0) < \infty, \qquad \forall (\mathbf{g}_0, \mathbf{H}_0, \mathbf{B}_0).$$

This can be verified easily based on the fact that distortion function (2) is bounded by σ_θ^2. Hence condition (B) of [36] is satisfied and a solution to (A.1) exists.

To show equality in (A.1), we will require a further equicontinuity property on $v_\rho(\cdot, \cdot, \cdot)$ to be satisfied for all $\rho \in (0,1)$. This can be verified by a similar argument as in the proof of Proposition 3.2 of [37] (see Appendix A of [37]). Finally, one needs to verify certain further conditions to prove that there exists a bounded solution to the average cost optimality (12). Here we simply summarize the relevant theorem from [38] that leads us to this proof, namely, Theorem 5.5.4. The assumptions required for this proof rely on non-negativity, lower semi-continuity and inf-compactness of the distortion function and continuity of the transition probability functions (See Assumption 4.2.1 in [38]), which are guaranteed in our case. Also, one needs Assumption 5.5.1 of [38] to be satisfied, which requires some boundedness and equicontinuity conditions on the solutions to the discounted-cost optimality equation associated with $v_\rho(\mathbf{g}_0, \mathbf{H}_0, \mathbf{B}_0)$ defined above. For our case, this assumption can also be verified. Due to space restrictions, the lengthy details of these assumptions and the corresponding proof of Theorem 5.5.4 in [38] cannot be provided here fully. The readers are referred to Section 5.5 of [38] for further details.

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102–114, Aug. 2002.

[2] V. C. Gungor and G. P. Hancke, "Industrial wireless sensor networks: Challenges, design principles, and technical approaches," *IEEE Trans. Ind. Electron.*, vol. 56, no. 10, pp. 4258–4265, Oct. 2009.

[3] C.-Y. Chong and S. P. Kumar, "Sensor networks: Evolution, opportunities and challenges," *Proc. IEEE*, vol. 91, no. 8, pp. 1247–1256, Aug. 2003.

[4] V. C. Gungor, B. Lu, and G. P. Hancke, "Opportunities and challenges of wireless sensor networks in smart grid," *IEEE Trans. Ind. Electron.*, vol. 57, no. 10, pp. 3557–3564, Oct. 2010.

[5] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: Lossy source-channel communication revisited," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1147–1158, May 2003.

[6] J.-J. Xiao, A. Ribeiro, Z.-Q. Luo, and G. Giannakis, "Distributed compression-estimation using wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 23, no. 4, pp. 27–41, Jul. 2006.

[7] S. Cui, J. Xiao, A. J. Goldsmith, Z. Luo, and H. V. Poor, "Estimation diversity and energy effiiency in distributed sensing," *IEEE Trans. Signal Process.*, vol. 55, no. 9, pp. 4683–4695, Sep. 2007.

[8] J. M. Mendel, *Lessons in Estimation Theory For Signal Processing, Communications and Control*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1995.

[9] I. Bahceci and A. K. Khandani, "Linear estimation of correlated data in wireless sensor networks with optimum power allocation and analog modulation," *IEEE Trans. Commun.*, vol. 56, no. 7, pp. 1146–1156, Jul. 2008.

[10] J.-Y. Wu and T.-Y. Wang, "Power allocation for robust distributed best-linear-unbiased estimation against sensing noise variance uncertainty," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2853–2869, Jun. 2013.

[11] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. Goldsmith, "Linear coherent decentralized estimation," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 757–770, Feb. 2008.

[12] 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.

[13] 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.

[14] 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.

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

[16] 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. Area Commun.*, vol. 29, no. 8, pp. 1732–1743, Sep. 2011.

[17] A. Nayyar, T. Basar, D. Teneketzis, and V. V. Veeravalli, "Optimal strategies for communication and remote estimation with an energy harvesting sensor," *IEEE Trans. Autom. Control*, vol. 58, no. 9, pp. 2246–2260, Sep. 2013.

[18] M. Nourian, A. S. Leong, and S. Dey, "Optimal energy allocation for Kalman filtering over packet dropping links with imperfect acknowledgments and energy harvesting constraints," *IEEE Trans. Autom. Control*, vol. 59, no. 8, pp. 2128–2143, Aug. 2014.

[19] D. G. O. Orhan and E. Erkip, "Delay-constrained distortion minimization for energy harvesting transmission over a fading channel," in *IEEE. Int. Symp. Inf. Theory*, Istanbul, Turkey, Jul. 2013, pp. 1794–1798.

[20] L. Vijayandran, M. B-Pearce, K. Kansanen, and T. Ekman, "Efficient state estimation with energy harvesting and fairness control using stochastic optimization," in *Proc. IEEE GLOBECOM*, Houston, TX, USA, Dec. 2011, pp. 1–6.

[21] A. Limmanee, S. Dey, and A. Ahlén, "Distortion minimization via multiple sensors with energy harvesting," in *Proc. 14th IEEE SPAWC*, Darmstadt, Germany, Jun. 2013, pp. 225–229.

[22] C. Huang, Y. Zhou, T. Jiang, P. Zhang, and S. Cui, "Power allocation for joint estimation with energy harvesting constraints," in *Proc. IEEE ICASSP*, Vancouver, BC, Canada, May 2013, pp. 4804–4808.

[23] Y. Zhao, B. Chen, and R. Zhang, "Optimal power allocation for an energy harvesting estimation system," in *Proc. IEEE ICASSP*, Vancouver, BC, Canada, May 2013, pp. 4549–4553.

[24] S. Roundy, P. K. Wright, and J. M. Rabaey, *Energy Scavenging for Wireless Sensor Networks*. New York, NY, USA: Springer-Verlag, 2004.

[25] D. E. Quevedo, A. Ahlén, and K. H. Johannson, "State estimation over sensor networks with correlated wireless fading channels," *IEEE Trans. Autom. Control*, vol. 58, no. 3, pp. 581–593, Mar. 2013.

[26] C. K. Ho, P. D. Khoa, and P. C. Ming, "Markovian models for harvested energy in wireless communications," in *IEEE ICCS*, Singapore, Nov. 2010, pp. 311–315.

[27] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 2nd ed., vol. 1. Belmont, MA, USA: Athena Scientific, 1995.

[28] E. Altman, *Constrained Markov Decision Processes*, vol. 7. Boca Raton, FL, USA: CRC, 1999.

[29] A. Arapostathis, V. S. Borkar, E. Fernández-Gaucherand, M. K. Ghosh, and S. I. Marcus, "Discrete-time controlled Markov processes with average cost criterion: A survey," *SIAM J. Control Optim.*, vol. 31, no. 2, pp. 282–344, Mar. 1993.

[30] H. Yu and D. P. Bertsekas, "Discretized approximations for POMDP with average cost," in *Proc. 20th Conf. Uncertainty Artif. Intell.*, Banff, AB, Canada, 2004, pp. 619–627.

[31] D. M. Topkis, *Supermodularity and Complementarity*. Princeton, NJ, USA: Princeton Univ. Press, 2001.

[32] M. H. Ngo and V. Krishnamurthy, "Optimality of threshold policies for transmission scheduling in correlated fading channels," *IEEE Trans. Commun.*, vol. 57, no. 8, pp. 2474–2483, Aug. 2009.

[33] D. Simchi-Levi, X. Chen, and J. Bramel, *The Logic of Logistics: Theory, Algorithms, and Applications for Logistics and Supply Chain Management*. New York, NY, USA: Springer-Verlag, 2004.

[34] V. Krishnamurthy, R. Bitmead, M. Gevers, and E. Miehling, "Sequential detection with mutual information stopping cost," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 700–714, Feb. 2012.

[35] J. C. Spall, *Introduction to Stochastic Search and Optimization: Estimation, Simulation, and Control*. New York, NY, USA: Wiley, 2005.

[36] M. Schäl, "Average optimality in dynamic programming with general state space," *Math. Oper. Res.*, vol. 18, no. 1, pp. 163–172, Feb. 1993.

[37] M. Huang and S. Dey, "Dynamic quantizer design for hidden Markov state estimation via multiple sensors with fusion center feedback," *IEEE Trans. Signal Process.*, vol. 54, no. 8, pp. 2887–2896, Aug. 2006.

[38] O. Hernández-Lerma and J. B. Lasserre, *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. New York, NY, USA: Springer-Verlag, 1996.

Mojtaba Nourian (S'06–M'12) received dual B.Sc. degrees in applied mathematics and in electrical engineering and the M.Sc. degree in applied mathematics from Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran, in 2005 and 2007, respectively. He obtained the Ph.D. degree in electrical engineering from McGill University, Montreal, Canada, in September 2012, under the supervision of Peter E. Caines.

He was a Research Fellow at the Department of Electrical and Electronic Engineering, University of Melbourne, Australia during 2012–2013. His current research interests include mean field stochastic control, signal processing for sensor networks and networked system estimation and control.

Subhrakanti Dey (M'96–SM'06) was born in India, in 1968. He received the B.Tech. and M.Tech. degrees from the Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology, Kharagpur, in 1991 and 1993, respectively, and the Ph.D. degree from the Department of Systems Engineering, Research School of Information Sciences and Engineering, Australian National University, Canberra, in 1996.

He is currently a Professor with the Department of Engineering Sciences, Uppsala University, Sweden. Prior to this, he was a Professor with the Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, Australia, from 2000 until early 2013. From September 1995 to September 1997, and September 1998 to February 2000, he was a Postdoctoral Research Fellow with the Department of Systems Engineering, Australian National University. From September 1997 to September 1998, he was a Postdoctoral Research Associate with the Institute for Systems Research, University of Maryland, College Park, MD, USA. His current research interests include networked control systems, wireless communications and networks, signal processing for sensor networks, and stochastic and adaptive estimation and control.

Professor Dey currently serves on the Editorial Board of IEEE TRANSACTIONS ON SIGNAL PROCESSING and Elsevier *Systems and Control Letters*. He was also an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING during 2007–2010 and the IEEE TRANSACTIONS ON AUTOMATIC CONTROL during 2004–2007.

Anders Ahlén (S'80–M'84–SM'90) was born in Kalmar, Sweden, and received the Ph.D. degree in automatic control from Uppsala University. He is a Full Professor and holds the Chair in Signal Processing at Uppsala University where he is also the Head of the Signals and Systems Division of The Department of Engineering Sciences. He was with the Systems and Control Group, Uppsala University, from 1984 to 1992 as an Assistant and Associate Professor in automatic control. During 1991 he was a Visiting Researcher at the Department of Electrical and Computer Engineering, The University of Newcastle, Australia. He was a Visiting Professor at the same university in 2008. In 1992, he was appointed Associate Professor of Signal Processing at Uppsala University. During 2001–2004, he was the CEO of Dirac Research AB, a company offering state-of-the-art audio signal processing solutions. He is currently the Chairman of the board of the same company. He was a member of the Center of Excellence in Wireless Sensor Networks, WISENET, from 2007–2013. His research interests, which include signal processing, communications and control, are currently focused on signal processing for wireless communications, wireless systems beyond 3G, wireless sensor networks, wireless control, and audio signal processing. From 1998 to 2004, he was the Editor of Signal and Modulation Design for the IEEE TRANSACTIONS ON COMMUNICATIONS.