Revisiting Effectiveness of Energy Conserving Opportunistic Transmission Schemes in Energy Harvesting Wireless Sensor Networks

Sayan Sen Gupta, *Student Member, IEEE*, and Neelesh B. Mehta, *Fellow, IEEE*

***Abstract*— Opportunistic transmission schemes improve the lifetime of conventional wireless sensor networks (WSNs) by reducing the number of transmissions. However, this comes at the expense of performance since fewer measurements are available. We show that in energy harvesting (EH) WSNs, in which the sensor nodes harvest energy from the environment, this trade-off is fundamentally different. For a general model in which the nodes experience independent and non-identical fading and the EH process at a node is stationary and ergodic, we present lower bounds on the mean squared error (MSE) for two important classes of channel-based opportunistic transmission schemes, namely, censoring and ordered transmissions. For the latter, we present two novel variants that arise depending on whether the energy in the battery of an EH sensor node is accounted for before ordering or not. For censoring, the lower bound leads to an insightful and explicit characterization of the optimum censoring threshold for each node. For ordered transmissions, it helps determine the optimal number of nodes that should be selected to transmit. We find that the ordered transmission schemes can outperform the censoring scheme. We also propose a hybrid scheme that combines the best features of the above schemes for EH WSNs.**

***Index Terms*— Energy harvesting, wireless sensor networks, estimation, censoring, ordered transmissions.**

I. Introduction

WIRELESS sensor networks (WSNs) are finding increasing use in applications such as environmental monitoring, military surveillance, and healthcare due to their low cost and ease of deployment. However, the nodes in a conventional WSN are constrained by their limited on-board battery energy capacity. Hence, devising energy-conserving strategies to increase the lifetime of WSNs is an active and important area of research.

Several lifetime improvement schemes such as clustering [2], data correlation [3], beamforming [4], and opportunistic transmissions [5], [6] have been explored in the literature. Censoring [5] and ordered transmissions [6] are two well known opportunistic transmission schemes that adopt different approaches to improve the energy efficiency. Censoring is a node-specific scheme in which each node decides to transmit or not on the basis of a local metric that it computes. For example, the likelihood ratio of a node's observation is used as the metric in the detection problem considered in [5]. In estimation problems, a node's measurement is the metric in [7], and the node's channel power gain is the metric in [8]. Instead, in ordered transmissions, which is a group-based scheme, a node's decision to transmit depends on the other nodes in the network. In it, the nodes transmit one after another in the decreasing order of their metrics. This can be realized in practice in a distributed manner without any node knowing any other node's metric using the timer-based selection scheme [9]. Ordered transmissions has been studied for detection in [6] and estimation in [10], with the metric being the likelihood ratio of a node's observation.

Energy harvesting (EH) is a green, alternate solution that eliminates the problem of limited lifetime in WSNs. In it, the nodes are equipped with rechargeable batteries and an EH circuitry that enables them to harvest energy from renewable sources, such as solar, vibration, and wind, to replenish their energy buffers. However, since the energy available is random, the nodes can occasionally be unavailable due to lack of energy, which affects performance.

A. Literature Survey on EH WSNs

The literature on EH WSNs has studied problems related to parameter estimation and data transmission. We summarize the most pertinent ones below.

- *Estimation:* The literature on estimation in EH WSNs can be broadly classified into the following two categories:
  - *Single Sensor Estimation:* In [11], for a time-slotted system, mean squared error (MSE) is minimized by optimally adapting the transmit power of the EH sensor node in each slot. This is done for deterministic energy arrivals, in which the amount of energy harvested in a slot is known beforehand, and stochastic energy arrivals, in which the harvested energy is random and is only known up to the current slot. For stochastic energy arrivals, [12] proposes an optimal communication scheduling and estimation

Manuscript received June 28, 2018; revised October 31, 2018; accepted December 5, 2018. Date of publication December 24, 2018; date of current version April 16, 2019. This work was supported in part by the DST-Swarnajayanti Fellowship Award under grant DST/SJF/ETA-01/2014-15. This paper was presented in part at the IEEE Global Communications Conference (Globecom), Abu Dhabi, December 2018 [1]. The associate editor coordinating the review of this paper and approving it for publication was A. S. Cacciapuoti. *(Corresponding author: Sayan Sen Gupta.)*

The authors are with the Department of Electrical Communication Engineering, Indian Institute of Science, Bengaluru 560012, India (e-mail: sayan.riju88@gmail.com; neeleshbmehta@gmail.com).

Digital Object Identifier 10.1109/TCOMM.2018.2889331

strategy that minimizes the expected communication and distortion costs using dynamic programming. In [13], the effect of time correlation between signals is modeled, and the optimal transmit power that minimizes the MSE is determined.

  - *Multi-sensor Estimation:* In [14], the optimal power allocation for finite and infinite horizon estimation problems assuming Markovian energy harvesting processes and Markovian fading channels are solved using dynamic programming. In [15], scheduling of nodes and transmit power control are employed to minimize the MSE averaged over a finite time horizon assuming deterministic energy arrivals. In [16], for stochastic energy arrivals, the optimal transmit power of the nodes when they share their observations with each other and also when they send their accumulated observations to the fusion node (FN) is determined. In [17], for deterministic and stochastic energy arrivals, the goal is to determine the transmit power of each node and the energy that they share with each other in each slot.

- *Data Transmission:* A game-theoretic framework is proposed to optimize the sleep and wake-up probabilities of a single solar-powered EH node in [18]. Learning algorithms are used to determine the optimal transmission policy in [19], and are used to determine the optimal transmit power in [20] and [21] to maximize the number of bits transmitted over a point-to-point communication link. A learning-theoretic framework that optimizes the transmit power to maximize the data packet arrival rate under average delay and delay outage constraints is studied in [22]. A battery energy based censoring scheme is considered in [23], in which a discounted Markov decision process framework is used to maximize the number of messages a node transmits. A single EH node system is considered in [24]. It uses a hidden Markov model for the solar energy, which is obtained from real deployments, and focuses on adapting the transmit power level and the modulation scheme to maximize the number of bits sent.

There is a fundamental difference between the problems of data transmission and parameter estimation. In the former, the goal is to maximize the number of bits transmitted, while in the latter, the goal is to accurately estimate a random physical parameter from the observations received from the EH nodes, which is measured using performance metrics such as the MSE. We focus on the estimation aspect of EH WSN design. The focus of the above papers on this aspect has primarily been on determining the transmit power of the sensor nodes. However, the efficacy of classical energy-conserving opportunistic transmission schemes like censoring and ordered transmissions in EH WSNs has not been thoroughly investigated.

B. Focus and Contributions

We make the following contributions:

- We derive a tractable and insightful lower bound on the MSE for the censoring scheme, and then provide an explicit characterization of the optimal censoring thresholds for each node that jointly minimize the bound. We show that the optimum censoring threshold for an EH node is the point where the node transitions from being energy constrained to being energy unconstrained. Here, an EH node is said to be energy unconstrained if it has sufficient energy to transmit its observation to the FN in a slot with probability 1. Else, it is said to be energy constrained.
- For the channel-based ordering scheme, we present two novel variations that are unique to EH WSNs:
  - *Order then Active (OTA)*: In this scheme, the nodes are ordered based on their channel power gains irrespective of whether they are active or inactive. The K best nodes are selected for transmission. Among the selected nodes, only the active ones transmit. Here, an EH node is active in a slot if it has sufficient energy to transmit its observation to the FN, else it is inactive.
  - *Active then Order (ATO)*: In this scheme, only the active nodes are ordered. Among them, the K best nodes transmit their observations to the FN.

  For both variants, we discuss how these can be physically realized with low complexity. We then derive tractable lower bounds on the MSE for them and use them to numerically determine the optimal subset size that minimizes the bound. A counter-intuitive observation that our study leads to is that OTA outperforms ATO, especially when the nodes are energy constrained.
- We then propose a new two-stage hybrid scheme that combines observation-based censoring with channel-based ordering. It outperforms both OTA and censoring, especially when the average energy harvested per slot is small.

The above results are general in the following two respects. First, they apply to the general class of stationary and ergodic EH processes, which includes the widely used independent and identically distributed (i.i.d.) [17], Bernoulli [25], and Markovian EH models [11], [14], [26]. Second, they apply to the general, practically relevant case where the nodes experience statistically non-identical fading. This happens, for example, when the nodes are at different distances from the FN. While we focus on a WSN with multiple EH nodes, [19], [20], [22], [24] focus on a WSN with one EH node. Furthermore, the focus in [19]–[24] is on data transport, while ours is on estimation, which leads to a fundamentally different model, analysis, and results.

In conventional WSNs, a monotonic trade-off exists between lifetime and performance for these schemes. For example, increasing the censoring threshold reduces the number of times a node transmits its observations to the FN. Doing so increases lifetime, but it also increases the MSE. However, as we show, this no longer holds in EH WSNs for both node-specific and group-based schemes, since an EH node can replenish itself even after its battery energy runs low.

C. Organization and Notations

This paper is organized as follows. Section II presents the system model. Section III analyzes the MSE of the

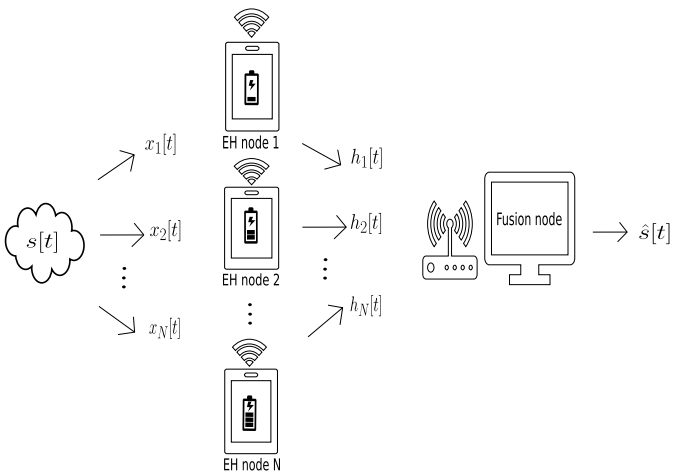

Fig. 1. An EH WSN consisting of N EH nodes that observe an unknown parameter $s[t]$ and subsequently transmit their readings to an FN, which generates an estimate $\hat{s}[t]$.

transmission schemes. Simulation results are presented in Section IV, and are followed by our conclusions in Section V.

Notations: The probability of an event A is denoted by $\Pr[A]$. The conditional probability of A given event B is denoted by $\Pr[A|B]$. The joint probability of events A and B is denoted by $\Pr[A, B]$. The probability density function (PDF) of a random variable (RV) X is denoted by $f_X(\cdot)$. We denote the expectation with respect to X by $\mathbb{E}_X[\cdot]$. Similarly, the expectation conditioned on an event A is denoted by $\mathbb{E}_X[\cdot|A]$. The notation $X \sim \text{CN}(\sigma^2)$ means that X is a zero-mean circularly symmetric complex Gaussian RV with variance σ^2. Matrices and vectors are denoted using boldface characters. $\mathbf{A}^{\text{T}}$ denotes the transpose of a vector $\mathbf{A}$. For a set $\mathcal{B}$, its cardinality is denoted by $|\mathcal{B}|$, its complement by $\mathcal{B}^c$, and its power set by $\mathcal{P}[\mathcal{B}]$. For a complex number c, c^* and $|c|$ denote its complex conjugate and absolute value, respectively. The indicator function $1_{\{a\}}$ equals 1 if a is true and is 0 otherwise.

II. System Model

Fig. 1 shows a time-slotted EH WSN consisting of a set $\mathcal{N} = \{1, 2, \ldots, N\}$ of N EH nodes and an FN. The models for sensor node readings, energy harvesting storage, channel fading processes, transmission, and reception are as follows.

A. Sensor Readings

The observation $x_i[t]$ at the i^{th} EH node in the t^{th} time slot is

$$x_i[t] = s[t] + v_i[t], \quad \text{for } 1 \leq i \leq N, \tag{1}$$

where $s[t]$ is the parameter to be estimated in the t^{th} time slot. It has a mean of zero and a variance of σ_s^2, and is i.i.d. across t. The observation noise $v_i[t] \sim \text{CN}(\sigma_v^2)$ is independent of $s[t]$ and is i.i.d. across i and t. The observation model in (1) is justified when the physical phenomenon to be sensed remains unchanged over the region where the EH nodes are being deployed. It has also been studied in [14], [15], [17], and [27].

B. EH and Storage Model

The EH process at a node is assumed to be stationary and ergodic with mean $\bar{H} > 0$ per slot, and is i.i.d. across nodes. As mentioned, this general model encompasses several other models studied in the literature [14], [16], [28]. The energy is harvested at the beginning of a slot and is stored in an energy buffer for use in that slot and subsequent slots. The capacity of the buffer is assumed to be infinite in order to ensure analytical tractability [11], [16].

C. Channel Model

Let $h_i[t]$ be the channel power gain between the i^{th} node and the FN in the t^{th} time slot. It is an exponential RV with mean λ_i, which models Rayleigh fading. We assume a general block fading model in which the channel fades are i.i.d. across time and independent across nodes, and different nodes have different mean channel power gains. This is justified since the nodes are geographically separated and are at different distances from the FN. We also assume that the EH and channel fading processes are mutually independent [14]. Let $\mathbf{h}(t) = (h_1(t), \ldots, h_N(t))$ denote the vector of channel power gains in the t^{th} time slot.

D. Transmission Model

The nodes amplify and forward their observations to the FN over a set of orthogonal channels.[1] Each node transmits its observation with a fixed power P. This model enables an energy-efficient design of power amplifiers, which is an important issue in sensor nodes [30]. When a node transmits its observation to the FN, the energy it consumes is given by PT_{tx}, where T_{tx} is the transmission duration. Let

$$\rho = \frac{\bar{H}}{PT_{\text{tx}}}. \tag{2}$$

As mentioned, a node is active in a slot if it has sufficient energy to transmit, else it is inactive. A node i knows $h_i[t]$ in slot t perfectly [14], [15], [27].

The opportunistic transmission schemes are as follows:

- *Censoring*: In this scheme, the i^{th} node first checks if its channel power gain exceeds a censoring threshold θ_i. If so, it transmits provided it is active. Let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_N)$ denote the censoring vector. Since an EH node only compares its channel power gain with a threshold, the scheme is simple to implement.
- *Ordered Transmissions*: Two variants arise for this scheme in EH WSNs:
  - *Order then Active (OTA)*: It can be implemented in a distributed manner with low complexity as follows. Each node maintains a timer that is a monotone non-increasing function of its channel power gain. When its timer expires, the node transmits in its orthogonal channel for a duration T_{tx} with power P

[1] We do not consider the alternate model in which all nodes transmit simultaneously over the same channel, such that their signals arrive simultaneously and combine coherently at the FN [29]. This model requires tight synchronization between sensor nodes and the FN, and is difficult to realize in practice in WSNs.

its corresponding observation if it is active, and a low energy pilot signal if it is inactive.[2] Once the FN receives K data or pilot transmissions, it sends out a broadcast message asking the remaining nodes to halt their timers. Thus, only the K best nodes can transmit.

– *Active then Order (ATO)*: In contrast with OTA, in a slot, only the active nodes start their timers. Once a node's timer expires, it transmits its observations to the FN for a duration T_{tx} with power P in its orthogonal channel. Upon receiving K data transmissions, the FN broadcasts a message informing the remaining nodes to stop their timers. If the number of active nodes is less than K, the FN ends up waiting until the end of the slot and the number of data transmissions it receives is less than K. Thus, ATO can also be implemented in a distributed manner with low complexity.

The probability that two timers expire very close to each other is assumed to be negligible [6].

E. Reception Model

The signal received by the FN from the i^{th} node in the t^{th} time slot is given by

$$y_i[t] = \sqrt{\alpha h_i[t]} x_i[t]\, e^{j\phi_i[t]} + w_i[t]\,, \quad \text{for } 1 \le i \le N, \quad (3)$$

where $\alpha = P/\left(\sigma_s^2 + \sigma_v^2\right)$, $\phi_i[t]$ is the phase of the i^{th} channel, and $w_i[t] \sim \text{CN}(\sigma_w^2)$ is the noise at the receiver for the i^{th} channel in the t^{th} time slot. The noise at the receiver is i.i.d. across nodes and slots. Since the signal, noise, channel fading, and EH processes are stationary, we drop the time parameter t henceforth. Let $\mathcal{T}$ denote the set of nodes that transmit in a slot.

F. Estimation at FN

Based on the received signal vector, the FN computes a linear minimum mean squared error (LMMSE) estimate of s based on its knowledge of h_i, for $i \in \mathcal{T}$, and the first and second order statistics of the signal and noise. We denote the instantaneous MSE, which averages over the signal and noise but not the channel fading, as $\mathsf{MSE}(\mathcal{T})$. It equals [31, Ch. IV]

$$\mathsf{MSE}(\mathcal{T}) = \mathbb{E}_{s,\hat{s}}\left[|s-\hat{s}|^2 | \mathcal{T}\right]. \quad (4)$$

III. MSE Analysis

For the i^{th} node, let A_i be the event that the node is active, S_i be the event that it is selected, and T_i be the event that it transmits. Let $\zeta_i = \Pr[\text{A}_i]$. Node i is said to be *energy unconstrained* if $\zeta_i = 1$. Else, it is *energy constrained*.

[2]The energy consumed by the pilot signal is much lower than PT_{tx}. The sensor node can maintain a small energy reserve for supporting pilot signal transmissions even if it is inactive for several time slots.

A. Censoring Scheme

When the nodes in $\mathcal{T}$ transmit their observations to the FN, the MSE is given by [31, Ch. IV]

$$\mathsf{MSE}(\mathcal{T}) = \sigma_s^2 \left(\frac{\sigma_s^2}{\sigma_v^2}\sum_{i\in\mathcal{T}} f(h_i) + 1\right)^{-1}, \quad (5)$$

where

$$f(h_i) = \frac{h_i}{h_i + \frac{\sigma_w^2}{P}\left(\frac{\sigma_s^2}{\sigma_v^2}+1\right)}.$$

In (5), there are two sources of randomness: $\mathcal{T}$, which is affected by the randomness in the EH processes, and the channel power gains. The average MSE, denoted by $\overline{\mathsf{MSE}}$, is

$$\overline{\mathsf{MSE}} = \sum_{\mathcal{T}\subseteq\mathcal{N}} \mathbb{E}_{\mathbf{h}}\left[\mathsf{MSE}(\mathcal{T}) \,\middle|\, \mathcal{T}\right] \Pr[\mathcal{T}]\,. \quad (6)$$

Evaluating (6) is analytically intractable. To gain insights, we derive a lower bound on $\overline{\mathsf{MSE}}$. We first state the following lemma.

Lemma 1: Node i is energy constrained if

$$\theta_i < -\lambda_i \log \rho, \quad (7)$$

and is energy unconstrained otherwise. Furthermore, the probability $\Pr[\text{T}_i]$ that node i transmits is given by

$$\Pr[\text{T}_i] = \begin{cases} \rho, & \text{if } \theta_i < -\lambda_i \log\rho, \\ e^{-\frac{\theta_i}{\lambda_i}}, & \text{otherwise.} \end{cases} \quad (8)$$

Proof: The proof is given in Appendix A. ∎

We see that as θ_i increases, i.e., node i is censored more, it becomes energy unconstrained. This is because its energy usage decreases. Lemma 1 leads to the following lower bound.

Result 1: Let $\mathcal{C}$ be the set of energy constrained nodes and $\mathcal{U}$ be the set of energy unconstrained nodes, as determined from Lemma 1. Then, $\overline{\mathsf{MSE}}$ is lower bounded by

$$\overline{\mathsf{MSE}} \ge \Upsilon, \quad (9)$$

where

$$\Upsilon = \sigma_s^2 \left(\frac{\sigma_s^2}{\sigma_v^2}\sum_{i\in\mathcal{C}} m(\theta_i,\lambda_i)\,\rho + \frac{\sigma_s^2}{\sigma_v^2}\sum_{j\in\mathcal{U}} m(\theta_j,\lambda_j)\, e^{-\frac{\theta_j}{\lambda_j}} + 1\right)^{-1}, \quad (10)$$

and

$$m(\theta,\lambda) = \frac{1}{\lambda}\int_{\theta}^{\infty} f(x) e^{-\frac{(x-\theta)}{\lambda}}\, dx. \quad (11)$$

Proof: The proof is given in Appendix B. ∎

Using Gauss-Laguerre quadrature [32, Ch. 25], $m(\theta,\lambda)$ is computed using a finite series as follows: $m(\theta,\lambda) \approx \sum_{i=1}^{n} z_i g(a_i)$, where $g(y) = f(\theta+\lambda y)$, z_i and a_i are the weights and abscissas, respectively, and n is the number

of terms. The lower bound also leads to the following elegant characterization of the optimal censoring threshold vector.

Result 2: The optimal censoring vector $\boldsymbol{\theta}^*_{\text{LB}}$ that minimizes the lower bound Υ is

$$\boldsymbol{\theta}^*_{\text{LB}} = (-\lambda_1 \log\rho, -\lambda_2 \log\rho, \ldots, -\lambda_N \log\rho). \quad (12)$$

Proof: The proof is given in Appendix C. ∎

Remark 1: From Result 2 and Lemma 1 we see that at $\boldsymbol{\theta}^*_{\text{LB}}$, each node becomes energy unconstrained. Note that the optimal censoring thresholds for different nodes are different.

1) Insights From I.I.D. Case: To gain more insights, we now consider the special case where the channel fading processes of the nodes are statistically identical. Hence, $\lambda_1 = \cdots = \lambda_N = \lambda$. By symmetry, we also have $\theta_1 = \cdots = \theta_N = \theta$. From Results 1 and 2, we get the following.

Corollary 1: The $\overline{\mathsf{MSE}}$ lower bound Υ simplifies to

$$\Upsilon = \begin{cases} \sigma_s^2 \left(\dfrac{N\rho\sigma_s^2}{\sigma_v^2} m(\theta,\lambda) + 1 \right)^{-1}, & \text{if } \theta < -\lambda\log\rho, \\ \sigma_s^2 \left(\dfrac{Ne^{-\frac{\theta}{\lambda}}\sigma_s^2}{\sigma_v^2} m(\theta,\lambda) + 1 \right)^{-1}, & \text{otherwise.} \end{cases} \quad (13)$$

The censoring threshold vector that minimizes Υ is given by $\boldsymbol{\theta}^*_{\text{LB}} = (-\lambda\log\rho, \ldots, -\lambda\log\rho)$.

B. OTA Scheme

As mentioned earlier, in this scheme, the K best nodes are selected and the ones that are active among them transmit. Using order statistics notation, let $h_{i:\mathcal{N}}$ be the i^{th} largest channel power gain among all the nodes $\mathcal{N}$ in the network and $i:\mathcal{N}$ be its corresponding index. Thus, the instantaneous MSE, denoted by MSE, is given by

$$\mathsf{MSE} = \sigma_s^2 \left(\frac{\sigma_s^2}{\sigma_v^2} \sum_{i=1}^{K} f(h_{i:\mathcal{N}}) 1_{\{A_{i:\mathcal{N}}\}} + 1 \right)^{-1}, \quad (14)$$

where, as per our notation, $A_{i:\mathcal{N}}$ is the event that the i^{th} best node in $\mathcal{N}$ is active. The indicator term $1_{\{A_{i:\mathcal{N}}\}}$ arises because the i^{th} best node $i{:}\mathcal{N}$ transmits only if it is active. Averaging over $h_{1:\mathcal{N}}, \ldots, h_{K:\mathcal{N}}$ and $1_{\{\mathrm{A}_{1:\mathcal{N}}\}}, \ldots, 1_{\{\mathrm{A}_{K:\mathcal{N}}\}}$ yields

$$\overline{\mathsf{MSE}} = \mathbb{E}_{\mathbf{h}_{K:\mathcal{N}}, \mathbf{a}_{K:\mathcal{N}}} \left[\sigma_s^2 \left(\frac{\sigma_s^2}{\sigma_v^2} \sum_{i=1}^{K} f(h_{i:\mathcal{N}}) 1_{\{A_{i:\mathcal{N}}\}} + 1 \right)^{-1} \right], \quad (15)$$

where $\mathbf{h}_{K:\mathcal{N}} = (h_{1:\mathcal{N}}, \ldots, h_{K:\mathcal{N}})$ and $\mathbf{a}_{K:\mathcal{N}} = \left(1_{\{\mathrm{A}_{1:\mathcal{N}}\}}, \ldots, 1_{\{\mathrm{A}_{K:\mathcal{N}}\}}\right)$.

Evaluating (15) is analytically intractable. We, therefore, derive a lower bound for it. Let the probability $\Pr[i = (j:\mathcal{N})]$ that the i^{th} node is the j^{th} best in $\mathcal{N}$ be denoted by ν_{ij}. A closed-form expression for ν_{ij} is given by the following lemma.

Lemma 2: ν_{ij} is given by

$$\nu_{ij} = \frac{1}{\lambda_i} \sum_{\mathcal{R} \in \mathcal{G}_{j-1}[\mathcal{N}\setminus\{i\}]} \left[\sum_{\mathcal{Q} \in \mathcal{P}[\mathcal{N}\setminus(\mathcal{R}\cup\{i\})]} (-1)^{|\mathcal{Q}|} \times \left(\frac{1}{\lambda_i} + \sum_{m\in\mathcal{R}} \frac{1}{\lambda_m} + \sum_{p\in\mathcal{Q}} \frac{1}{\lambda_p} \right)^{-1} \right], \quad (16)$$

where $\mathcal{G}_{j-1}[\mathcal{N}\setminus\{i\}]$ denotes the set of all subsets of cardinality $j-1$ of the set $\mathcal{N}\setminus\{i\}$. The probability ζ_i that node i is active is then given by

$$\zeta_i = \min\left\{ 1, \rho \left(\sum_{j=1}^{K} \nu_{ij} \right)^{-1} \right\}. \quad (17)$$

Proof: The proof is given in Appendix D. ∎

Let $\mathcal{S}_{\mathcal{N}}$ be the set of all permutations of the vector $(\lambda_i : i \in \mathcal{N})$ and $\eta \in \mathcal{S}_{\mathcal{N}}$ be a particular permutation that permutes the vector $(\lambda_1, \ldots, \lambda_N)$ to $(\eta_1, \ldots, \eta_N)$. Using Lemma 2, we get the following lower bound on $\overline{\mathsf{MSE}}$ of OTA.

Result 3: $\overline{\mathsf{MSE}}$ of OTA is lower bounded as

$$\overline{\mathsf{MSE}} \geq \Omega_K = \sigma_s^2 \left(1 + \frac{\sigma_s^2}{\sigma_v^2} \sum_{i=1}^{K} f(\mathbb{E}[h_{i:\mathcal{N}}]) \sum_{j=1}^{N} \zeta_j \nu_{ji} \right)^{-1}, \quad (18)$$

where

$$\mathbb{E}[h_{i:\mathcal{N}}] = \sum_{n=i}^{N} \sum_{\eta\in\mathcal{S}_{\mathcal{N}}} \prod_{q=1}^{N} \left[\eta_q^{-1} \left(\sum_{r=1}^{q} \eta_r^{-1} \right)^{-1} \right] \left(\sum_{p=1}^{n} \eta_p^{-1} \right)^{-1}. \quad (19)$$

Proof: The proof is given in Appendix E. ∎

The optimal K that minimizes Ω_K is then found numerically.

1) Insights From I.I.D. Case: As before, to gain insights, we consider the special case where $\lambda_1 = \cdots = \lambda_N = \lambda$. This implies $\zeta_1 = \cdots = \zeta_N = \zeta$. Then, (19) simplifies to

$$\mathbb{E}[h_{i:\mathcal{N}}] = \lambda \sum_{j=i}^{N} j^{-1}. \quad (20)$$

Furthermore, (16) simplifies to $\nu_{ij} = 1/N$. For the i.i.d. case, we get the following elegant characterization of the regions in which the nodes are energy constrained and unconstrained.

Lemma 3: All the nodes are energy constrained if and only if $K > N\rho$, in which case the probability that a node is active is given by $\zeta = N\rho/K$.

Proof: The proof is similar to that for Lemma 1, and is not repeated here. ∎

The simplified lower bound for the i.i.d. case is as follows.

Corollary 2: $\overline{\mathsf{MSE}}$ is lower bounded as

$$\overline{\mathsf{MSE}} \geq \Omega_K = \begin{cases} \sigma_s^2 \left(\dfrac{\sigma_s^2}{\sigma_v^2} \sum_{i=1}^{K} a_i + 1 \right)^{-1}, & \text{if } K \leq N\rho, \\ \sigma_s^2 \left(\dfrac{N\rho\sigma_s^2}{K\sigma_v^2} \sum_{i=1}^{K} a_i + 1 \right)^{-1}, & \text{otherwise,} \end{cases} \quad (21)$$

where $a_i = \left(\lambda \sum_{j=i}^{N} j^{-1}\right) / \left(\lambda \sum_{j=i}^{N} j^{-1} + \frac{\sigma_w^2}{P}\left(\frac{\sigma_s^2}{\sigma_v^2}+1\right)\right)$. From Corollary 2, when $K \leq N\rho$, we see that Ω_K decreases as K increases. However, when $K > N\rho$, Ω_K increases as as K increases. Thus, the optimal value of K is either $\lfloor N\rho \rfloor$ or $\lceil N\rho \rceil$. This can be understood as follows. For small K, fewer nodes transmit, as a result of which the estimation error is large. When K increases, a node is selected more often, which eventually decreases the probability that it is active and, thus, increases the estimation error. The optimal K provides a balance between the above two conflicting trends.

C. ATO Scheme

Recall that in this scheme, the K best active nodes transmit their observations to the FN. Let $\mathcal{N}_a$ be the set of all active nodes. If $|\mathcal{N}_a| \leq K$, then all the active nodes transmit their observations to the FN. Let $\mathbf{h}_{\mathcal{N}_a}$ be the vector of channel power gains of nodes that are active and $h_{i:\mathcal{N}_a}$ be the i^{th} largest element in $\mathbf{h}_{\mathcal{N}_a}$. Let $\mathbf{h}_{K:\mathcal{N}_a} = (h_{1:\mathcal{N}_a}, \ldots, h_{K:\mathcal{N}_a})$. For the ATO scheme, the instantaneous MSE, denoted by $\mathsf{MSE}(\mathcal{N}_a)$, is given by

$$\mathsf{MSE}(\mathcal{N}_a) = \begin{cases} \sigma_s^2 \left(\frac{\sigma_s^2}{\sigma_v^2} \sum_{i \in \mathcal{N}_a} f(h_i) + 1\right)^{-1}, & \text{if } |\mathcal{N}_a| \leq K, \\ \sigma_s^2 \left(\frac{\sigma_s^2}{\sigma_v^2} \sum_{i=1}^{K} f(h_{i:\mathcal{N}_a}) + 1\right)^{-1}, & \text{otherwise.} \end{cases} \tag{22}$$

The above two cases arise because when $|\mathcal{N}_a| \leq K$, ordering the nodes does not affect the MSE. Hence, averaging over the channel power gains and the set of active nodes, we get

$$\begin{aligned} \overline{\mathsf{MSE}} = & \sum_{\substack{\mathcal{N}_a \subseteq \mathcal{N}: \\ 0 \leq |\mathcal{N}_a| \leq K}} \mathbb{E}_{\mathbf{h}_{\mathcal{N}_a}} \left[\sigma_s^2 \left(\frac{\sigma_s^2}{\sigma_v^2} \sum_{i \in \mathcal{N}_a} f(h_i) + 1\right)^{-1} \Big| \mathcal{N}_a \right] \\ & \times \Pr[\mathcal{N}_a] + \sum_{\substack{\mathcal{N}_a \subseteq \mathcal{N}: \\ K+1 \leq |\mathcal{N}_a| \leq N}} \Pr[\mathcal{N}_a] \\ & \times \mathbb{E}_{\mathbf{h}_{K:\mathcal{N}_a}} \left[\sigma_s^2 \left(\frac{\sigma_s^2}{\sigma_v^2} \sum_{i=1}^{K} f(h_{i:\mathcal{N}_a}) + 1\right)^{-1} \Big| \mathcal{N}_a \right]. \end{aligned} \tag{23}$$

In ATO, it is intractable to even derive $\Pr[\mathcal{N}_a]$ because the battery energies of the EH nodes are coupled, which is unlike OTA. This is because the selection of a node, which is a function of its battery energy, depends on how many other nodes are active and, thus, their battery energy levels. To circumvent this challenge, we employ the widely used decoupling approximation [33], as per which the event that node i is active is independent of whether the other nodes are active. As we shall see below, the coupling between the nodes gets captured by the statistical parameters. Therefore,

$$\Pr[\mathcal{N}_a] = \left(\prod_{j \in \mathcal{N}_a} \zeta_j\right) \prod_{l \in \mathcal{N}_a^c} (1 - \zeta_l). \tag{24}$$

Next, we derive a lower bound on the average MSE given in (23). The first step in this comes from the following two lemmas, which leads to a set of equations solving which yields ζ_i. In a slot, let $\mathcal{A}$ be a set of active nodes such that $\mathcal{A} \subseteq \mathcal{N} \setminus \{i\}$. The conditional probability $\Pr[\mathrm{S}_i | \mathcal{A} \cup \{i\}]$ that node i is selected when the set of active nodes is $\mathcal{A} \cup \{i\}$ is given by the following lemma.

Lemma 4: The conditional probability that node i is the j^{th} best node in the set of active nodes $\mathcal{A} \cup \{i\}$, $\Pr[i = j : (\mathcal{A} \cup \{i\}) | \mathcal{A} \cup \{i\}]$, is as follows:

$$\begin{aligned} & \Pr[i = j : (\mathcal{A} \cup \{i\}) | \mathcal{A} \cup \{i\}] \\ & = \frac{1}{\lambda_i} \sum_{\mathcal{R} \in \mathcal{G}_{j-1}[\mathcal{A}]} \left[\sum_{\mathcal{Q} \in \mathcal{P}[\mathcal{A} \setminus \mathcal{R} \cup \{i\}]} (-1)^{|\mathcal{Q}|} \right. \\ & \qquad \left. \times \left(\frac{1}{\lambda_i} + \sum_{m \in \mathcal{R}} \frac{1}{\lambda_m} + \sum_{p \in \mathcal{Q}} \frac{1}{\lambda_p} \right)^{-1} \right]. \end{aligned} \tag{25}$$

Recall that $\mathcal{G}_{j-1}[\mathcal{A}]$ denotes the set of all subsets of cardinality $j-1$ of the set $\mathcal{A}$. Furthermore, conditioned on the set of active nodes $\mathcal{A} \cup \{i\}$, probability that node i is selected is given by

$$\Pr[\mathrm{S}_i | \mathcal{A} \cup \{i\}] = \sum_{j=1}^{K} \Pr[i = j : (\mathcal{A} \cup \{i\}) | \mathcal{A} \cup \{i\}]. \tag{26}$$

Proof: The proof is similar to Appendix D, except that the probabilities are calculated with respect to the set $\mathcal{A} \cup \{i\}$ instead of $\mathcal{N}$. We skip its details to conserve space. ■

Lemma 5: The probability $\Pr[\mathrm{T}_i]$ that node i transmits is

$$\begin{aligned} \Pr[\mathrm{T}_i] = & \sum_{\substack{\mathcal{A} \subseteq \mathcal{N} \setminus \{i\}: \\ 0 \leq |\mathcal{A}| \leq K-1}} \Pr[\mathcal{A} \cup \{i\}] \\ & + \sum_{\substack{\mathcal{A} \subseteq \mathcal{N} \setminus \{i\}: \\ K \leq |\mathcal{A}| \leq N-1}} \Pr[\mathrm{S}_i | \mathcal{A} \cup \{i\}] \Pr[\mathcal{A} \cup \{i\}], \end{aligned} \tag{27}$$

where $\Pr[\mathrm{S}_i | \mathcal{A} \cup \{i\}]$ is given in Lemma 4 and $\Pr[\mathcal{A} \cup \{i\}] = \left(\prod_{j \in \mathcal{A} \cup \{i\}} \zeta_j\right) \prod_{l \in \mathcal{N} \setminus (\mathcal{A} \cup \{i\})} (1 - \zeta_l)$.

Proof: The proof is given in Appendix F. ■

When node i is energy constrained, along lines similar to Lemma 1, it can be shown that $\Pr[\mathrm{T}_i] = \rho$. Then, Lemma 5 leads to N non-linear equations in $\zeta_1, \ldots, \zeta_N$, which are solved numerically as follows. It is easy to see that nodes with higher mean channel power gains will be selected more often and will, thus, become inactive with a higher probability. Formally, if $\lambda_m < \lambda_n$ then $\zeta_m \geq \zeta_n$, for any $m, n \in \mathcal{N}$ and $m \neq n$. Therefore, if $m \in \mathcal{C}$ then $n \in \mathcal{C}$. We first set $\mathcal{C} = \mathcal{N}$. If this leads to a consistent solution in the numerical computation, then we terminate. Else, we successively exclude the nodes in the increasing order of λ_i, for $i \in \mathcal{N}$, from $\mathcal{C}$, and repeat the procedure until a consistent solution is found. Using Lemmas 4 and 5, we obtain the following lower bound on $\overline{\mathsf{MSE}}$ for ATO.

Result 4: $\overline{\mathsf{MSE}}$ for ATO is lower bounded as

$$\begin{aligned}\overline{\mathsf{MSE}} &\geq \Psi_K \\ &= \sum_{\substack{\mathcal{N}_a\subseteq\mathcal{N}:\\ 0\leq|\mathcal{N}_a|\leq K}} \sigma_s^2\left(\frac{\sigma_s^2}{\sigma_v^2}\sum_{i\in\mathcal{N}_a} m\left(0,\lambda_i\right)+1\right)^{-1}\Pr[\mathcal{N}_a] \\ &\quad+\sum_{\substack{\mathcal{N}_a\subseteq\mathcal{N}:\\ K+1\leq|\mathcal{N}_a|\leq N}}\Pr[\mathcal{N}_a] \\ &\quad\times\sigma_s^2\left(\frac{\sigma_s^2}{\sigma_v^2}\sum_{i=1}^{K} f(\mathbb{E}[h_{i:\mathcal{N}_a}|\mathcal{N}_a])+1\right)^{-1}. \end{aligned} \tag{28}$$

Here, $m\left(0,\lambda_i\right)$ is given by (11), $\Pr[\mathcal{N}_a]$ is given by (24), and

$$\mathbb{E}\left[h_{i:\mathcal{N}_a}|\mathcal{N}_a\right]=\sum_{n=i}^{|\mathcal{N}_a|}\sum_{\eta\in\mathcal{S}_{\mathcal{N}_a}}\frac{\prod_{q=1}^{|\mathcal{N}_a|}\left[\eta_q^{-1}\left(\sum_{r=1}^{q}\eta_r^{-1}\right)^{-1}\right]}{\sum_{p=1}^{n}\eta_p^{-1}}, \tag{29}$$

where $\mathcal{S}_{\mathcal{N}_a}$ is the set of all permutations of the vector $(\lambda_i : i\in\mathcal{N}_a)$.

Proof: The proof is given in Appendix G. ∎

1) Insights From I.I.D. Case: In this case, $\lambda_1=\cdots=\lambda_N=\lambda$. Hence, $\zeta_1=\cdots=\zeta_N=\zeta$. Similar to OTA, the nodes will be energy unconstrained if and only if $K\leq\lfloor N\rho\rfloor$. Utilizing Lemma 5 and $\Pr[\mathrm{T}_i]=\rho$, we get

$$\begin{aligned}\rho=\zeta\left(1-\zeta\right)^{N-1}&\left[\sum_{i=0}^{K-1}\binom{N-1}{i}\left(\frac{\zeta}{1-\zeta}\right)^{i}\right. \\ &\left.+K\sum_{i=K}^{N-1}\frac{1}{i+1}\binom{N-1}{i}\left(\frac{\zeta}{1-\zeta}\right)^{i}\right]. \end{aligned} \tag{30}$$

ζ is obtained by solving (30) numerically, using, for example, fsolve in Matlab.

Remark 2: From (30), it can be seen that $\rho>\zeta K/N$. This yields $\zeta<N\rho/K$. Comparing this with the expression for ζ for OTA that is obtained from Lemma 3, we see that when all the nodes are energy constrained, the probability of them being active is greater in OTA than ATO. This happens because the probability that an active node is selected is greater in ATO, which makes the EH nodes inactive more often.

Substituting $\lambda_1=\cdots=\lambda_N=\lambda$ and $\zeta_1=\cdots=\zeta_N=\zeta$ in Result 4, we get the following.

Corollary 3: $\overline{\mathsf{MSE}}$ for ATO is lower bounded by

$$\overline{\mathsf{MSE}}\geq\Psi_K=\begin{cases}\sigma_s^2\left(\dfrac{\sigma_s^2}{\sigma_v^2}\displaystyle\sum_{i=1}^{K}a_i+1\right)^{-1}, & \text{if } K\leq N\rho,\\ \sigma_s^2\left(\displaystyle\sum_{j=0}^{K}\frac{C_j}{b_j}+\sum_{j=K+1}^{N}\frac{C_j}{d_j}\right), & \text{otherwise,}\end{cases} \tag{31}$$

where $b_j=1+\left(j\sigma_s^2 m\left(0,\lambda\right)/\sigma_v^2\right)$, $C_j=\binom{N}{j}\zeta^j(1-\zeta_j)^{N-j}$, $d_j=\frac{\sigma_s^2}{\sigma_v^2}\sum_{i=1}^{K}\frac{\lambda\sum_{r=i}^{j}r^{-1}}{\lambda\sum_{r=i}^{j}r^{-1}+\frac{\sigma_w^2}{P}\left(\frac{\sigma_s^2}{\sigma_v^2}+1\right)}+1$, and a_i is given in Corollary 2.

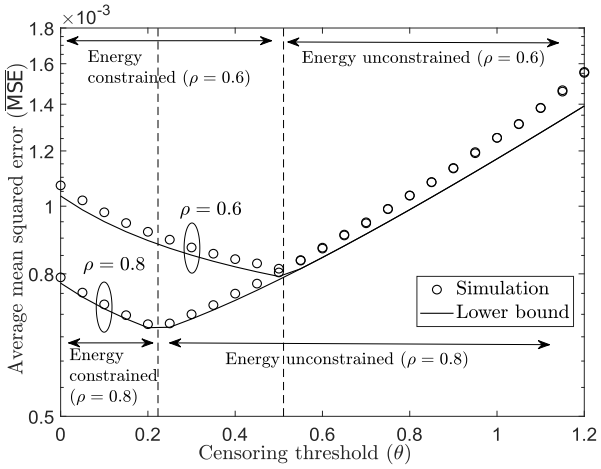

Fig. 2. Censoring, i.i.d. case: $\overline{\mathsf{MSE}}$ as a function of θ for different values of ρ ($N=30$).

D. Comments

We see from the analysis that the parameter ρ in (2) drives the performance of the opportunistic transmission schemes. It is a function of only the first moment $\bar{H}$ of the EH process. Therefore, the above analysis can be applied to a quasi-static scenario in which the time scale at which the statistics of the EH process change is larger than that of small-scale fading. In practical WSNs, static energy consumption also occurs to keep the nodes active. This can be approximately incorporated in our framework by replacing $\rho=\bar{H}/\left(PT_{\text{tx}}\right)$ with $\rho=\left(\bar{H}-E_{\text{st}}\right)/\left(PT_{\text{tx}}\right)$, where E_{st} is the static energy consumed per slot.

IV. Numerical Results

We now present results from Monte Carlo simulations and compare them with our analysis. The simulations are run for a duration of 10^5 time slots and the Bernoulli EH model is simulated [30]. For the purpose of illustration, we set $s\sim\mathrm{CN}(1)$, $\sigma_s^2/\sigma_v^2=20$ dB, $P/\sigma_w^2=23$ dB, and $T_{\text{tx}}=50$ ms.[3] To model independent and non-identically distributed (i.n.i.d.) fading, we set the mean channel power gain of node i as $\lambda_i=\beta^i$, where $\beta\leq 1$. The smaller the β, the more statistically non-identical are the channel power gains of the nodes. We first study the schemes separately and then compare them.

A. Censoring Scheme

1) I.I.D. Case: Fig. 2 plots $\overline{\mathsf{MSE}}$ as a function of the censoring threshold θ for $N=30$ and two values of ρ, which correspond to two values of $\bar{H}$. Also plotted is the lower bound derived in (13). We can see from the plot that the lower bounds track the simulation curves well. Furthermore, the censoring thresholds that minimize the lower bounds are the same as

[3]For example, this corresponds to a transmit power of 10 dBm, carrier frequency of 2.4 GHz, path-loss exponent of 3.8, reference distance of 1 m, and the distance between the transmitter and receiver of 40 m. At the receiver, the noise figure is 10 dB, the bandwidth is 100 kHz, and the temperature is 300 K.

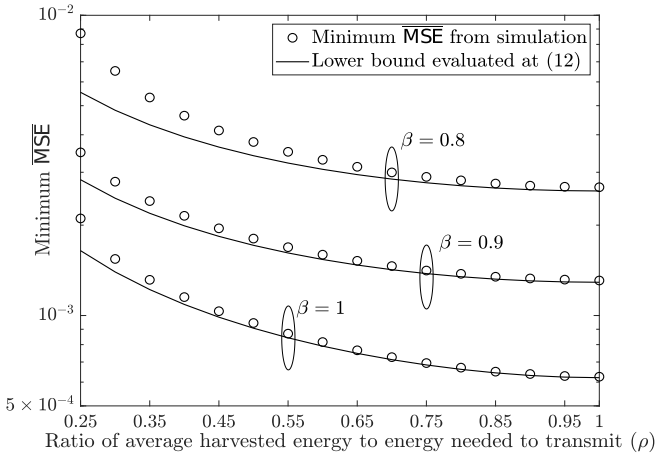

Fig. 3. Censoring, i.n.i.d. case: Minimum $\overline{\mathsf{MSE}}$ as a function of ρ for different values of β ($N = 30$).

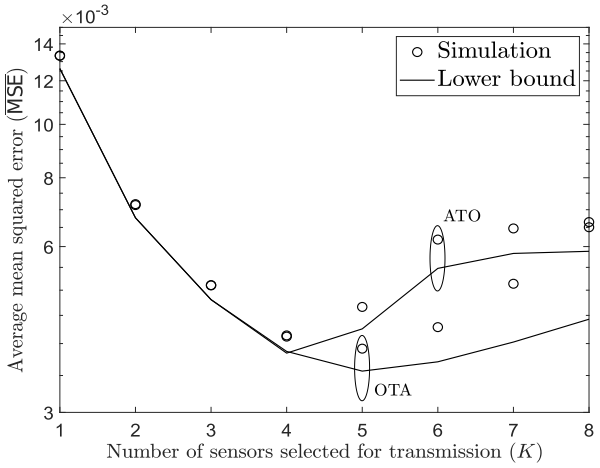

Fig. 4. Ordered transmissions, i.n.i.d. case: $\overline{\mathsf{MSE}}$ as a function of K for ATO and OTA ($N = 8$, $\beta = 0.9$, and $\rho = 0.6$).

the ones that minimize $\overline{\mathsf{MSE}}$, which illustrates the utility of optimizing the tractable lower bound. Vertical lines in the figure demarcate the regions in which all the nodes are energy constrained or unconstrained. We see that when θ increases, $\overline{\mathsf{MSE}}$ decreases, which is unlike conventional WSNs. This is because as the nodes are censored more, they conserve energy and become available for transmission more often. The optimal censoring threshold lies at the boundary between energy constrained and unconstrained regions. Beyond this point, $\overline{\mathsf{MSE}}$ increases because censoring prevents the nodes from transmitting. As ρ decreases, the minimum $\overline{\mathsf{MSE}}$ increases due to less energy being harvested. Notably, as ρ decreases, the optimal censoring threshold increases, as can also be inferred from Lemma 1.

2) I.N.I.D. Case: Fig. 3 plots the minimum $\overline{\mathsf{MSE}}$ as a function of ρ for different values of β for $N = 30$. The minimum $\overline{\mathsf{MSE}}$ is obtained by numerically finding the optimal censoring threshold vector for each β and ρ, which turns out to be the same as that in (12). Also shown is the lower bound, in which (10) is evaluated at the optimal censoring threshold vector given in (12). Similar to Fig. 2, we see that the lower bound tracks the simulation curve well. Furthermore, the gap between the two decreases as ρ increases. As β decreases, the channels become weaker, due to which the minimum $\overline{\mathsf{MSE}}$ increases. Similar to the i.i.d. case, as ρ increases, the minimum $\overline{\mathsf{MSE}}$ decreases.

B. Ordered Transmissions Scheme

1) I.N.I.D. Case: Fig. 4 plots $\overline{\mathsf{MSE}}$ as a function of K for ATO and OTA for $N = 8$. Although the lower bounds become loose for large K, we see that they both track the simulation curves well and accurately identify the optimal K. Notably, OTA yields a lower minimum $\overline{\mathsf{MSE}}$ than ATO. This is counter-intuitive at first sight because ATO uses extra information about the nodes' battery energy levels before ordering them. We investigate this further in Fig. 6.

2) I.I.D. Case: Fig. 5 plots $\overline{\mathsf{MSE}}$ as a function of K for OTA and ATO for $N = 30$. As in Fig. 4, the lower bound

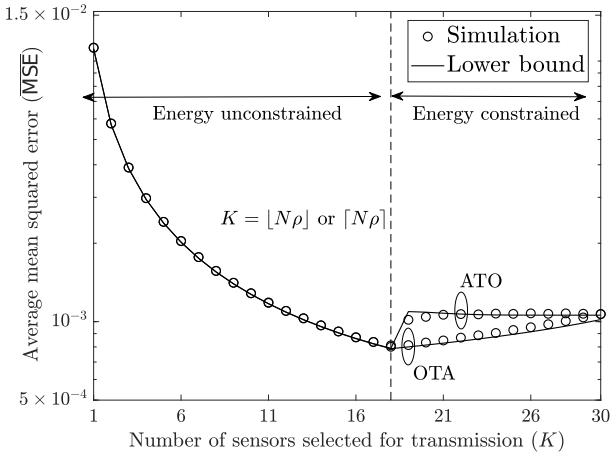

Fig. 5. Ordered transmissions, i.i.d. case: $\overline{\mathsf{MSE}}$ as a function of K for OTA and ATO ($N = 30$ and $\rho = 0.6$).

tracks the simulation curves well and accurately identifies the optimal point for both. Now, all the nodes are either energy constrained or energy unconstrained. For small K, increasing it decreases $\overline{\mathsf{MSE}}$ because more readings reach the fusion node. However, once all the nodes become energy constrained, increasing K increases $\overline{\mathsf{MSE}}$ because it increases the probability that they are inactive. When the nodes are energy constrained, i.e., when $K \geq 19$, OTA outperforms ATO. Otherwise, the performance of ATO and OTA is the same.

Fig. 6 plots the probability ζ from simulations that a node is active as a function of K for OTA and ATO. When the nodes are energy unconstrained or when $K = N$, ζ is the same for both schemes and so is $\overline{\mathsf{MSE}}$. However, as stated in Remark 2, once the nodes become energy constrained, i.e., when $K > N\rho$, the probability of a node being active is greater for OTA than for ATO. Furthermore, when the nodes are energy constrained, ζ for OTA is more sensitive to K than ATO. This explains why $\overline{\mathsf{MSE}}$ of OTA is more sensitive to K than ATO when $K > N\rho$.

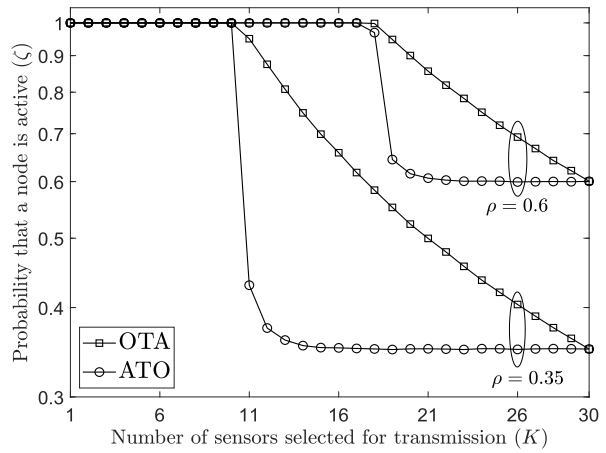

Fig. 6. Ordered transmissions, i.i.d. case: ζ as a function of K for OTA and ATO for different values of ρ ($N = 30$).

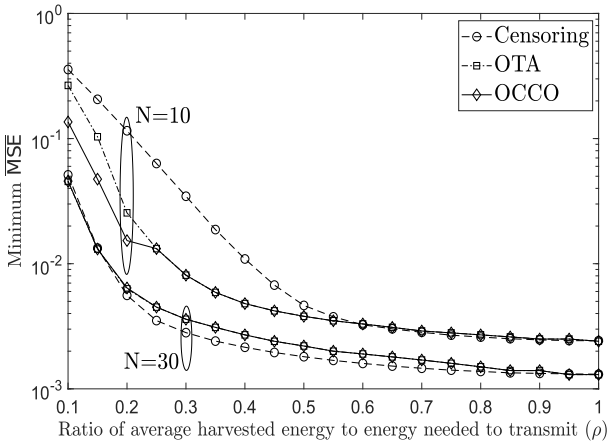

Fig. 7. Performance benchmarking: Minimum $\overline{\text{MSE}}$ as a function of ρ for censoring, OTA, and OCCO for different values of N ($\beta = 0.9$).

C. Hybrid Scheme and Performance Benchmarking

Having studied the node-specific and group-based schemes separately, we now compare all of them. We also propose a new two-stage hybrid scheme that combines observation-based censoring with channel-based ordering (OCCO). First, the nodes are censored based on their observations. Node i will not transmit if $|x_i| < \delta$, where δ is the observation-based censoring threshold. The second stage is similar to OTA, where the nodes are ordered in the decreasing order of their channel power gains, and among the K best, the active ones transmit. From the observations, the FN computes the LMMSE estimate, which depends on δ and the first and second order statistics of the signal and noise. Similar to OTA, OCCO can also be realized in a distributed manner using timer-based selection.

Fig. 7 plots the minimum $\overline{\text{MSE}}$ for censoring, OTA, and OCCO as a function of ρ. For OCCO, the minimum $\overline{\text{MSE}}$ is obtained by jointly optimizing δ and K numerically for each value of ρ. As N and ρ increase, the performance of all the schemes improves. When N is small and ρ is also small, i.e., when the energy harvested is scarce, OCCO outperforms OTA, which, in turn, outperforms censoring. The reasons for this are as follows. When energy harvested is scarce, the probability that no node transmits is the dominant factor that determines the minimum $\overline{\text{MSE}}$. From Lemma 1, this turns out to be $(1-\rho)^N$ for censoring. For small ρ and N, it is significantly larger than the probability that no node transmits for OTA. Hence, OTA yields a lower minimum $\overline{\text{MSE}}$ than censoring, although the average number of transmissions for both of them are approximately equal. Unlike OTA, OCCO ensures that the magnitudes of the observations that reach the FN contribute to a better estimate of the signal, which lowers the minimum $\overline{\text{MSE}}$. For larger $N = 30$ and $\rho > 0.20$, censoring outperforms OCCO and OTA, but this difference is negligible.

V. Conclusions

We analyzed the performance of both node-specific censoring and group-based ordered opportunistic transmission schemes for an EH WSN. We saw that the performance of the OTA and ATO ordered transmission schemes, which differed in their consideration of the battery energy state at the time of ordering and selection, was different. We derived insightful lower bounds on $\overline{\text{MSE}}$ for the above schemes for the general class of stationary and ergodic EH processes, which led to a characterization of the optimal operating point for censoring and an easy computation of the optimal number of nodes to be selected for OTA and ATO. When the EH nodes were energy starved, OTA outperformed ATO and censoring, which implied that group-based techniques that took the channel states of all the nodes into account in a distributed manner were preferable. We also saw that combining observation-based censoring with channel-based ordering lowered the average MSE further in this region, but had a negligible impact when the nodes were not energy starved or large in number.

Several interesting avenues for research exist. One avenue is characterizing the performance of the opportunistic transmission schemes for more general correlation models, such as the full-rank and rank-one models of [29]. Another avenue is incorporating the effect of imperfect channel state information.

Appendix

A. Proof of Lemma 1

Let $\bar{E}_i$ be the average energy consumed by node i in a slot. Clearly, $\bar{E}_i = PT_{\text{tx}}\Pr[\text{T}_i]$. A node transmits if it is active and it is not censored. Hence,

$$\Pr[\text{T}_i] = \Pr[\text{A}_i \cap \text{S}_i]\,. \tag{32}$$

At the beginning of a slot, a node's battery energy level is a function of its channel power gains in the previous slots. As the channel fading process is independent across time and the EH and fading processes are independent of each other, $\Pr[\text{A}_i \cap \text{S}_i] = \Pr[\text{A}_i]\Pr[\text{S}_i]$. Since h_i is an exponential RV with mean λ_i, the probability that node i is not censored is given by $\Pr[\text{S}_i] = \Pr[h_i > \theta_i] = e^{-\theta_i/\lambda_i}$. Hence, (32) becomes

$$\Pr[\text{T}_i] = \zeta_i e^{-\frac{\theta_i}{\lambda_i}}. \tag{33}$$

From the law of conservation of energy, we know that

$$\bar{E}_i \leq \bar{H}. \tag{34}$$

Substituting (33) in (34) yields $PT_{\text{tx}}\zeta_i e^{-\theta_i/\lambda_i} \leq \bar{H}$. Consider the following three cases:

1) When $\theta_i > -\lambda_i \log \rho$: Substituting $\rho = \bar{H}/(PT_{\text{tx}})$ and rearranging the terms in the inequality $\theta_i > -\lambda_i \log \rho$, we get $\bar{E}_i < \bar{H}\zeta_i$. We now prove that $\zeta_i = 1$. Assume that $\zeta_i < 1$. Then, $\bar{E}_i < \bar{H}$. Thus, node i accumulates an average energy of $\bar{H} - \bar{E}_i > 0$ in its battery in every slot. Hence, the energy stored in its battery will become infinite almost surely and the node is energy unconstrained. This contradicts our assumption that $\zeta_i < 1$. Hence, $\zeta_i = 1$.

2) When $\theta_i = -\lambda_i \log \rho$: Rearranging terms and substituting $\rho = \bar{H}/(PT_{\text{tx}})$ yields $\bar{E}_i = \bar{H}\zeta_i$. As above, we can prove that $\zeta_i = 1$.

3) When $0 \leq \theta_i < -\lambda_i \log \rho$: In this case, $\bar{E}_i > \bar{H}\zeta_i$. It is easy to see that $\zeta_i = 1$ is not possible since this violates the inequality in (34). Thus, $\zeta_i < 1$. Also, from the law of conservation of energy, we get $\bar{E}_i = \bar{H}$.

Evaluating $Pr[T_i]$: When node i is energy unconstrained, by definition $\zeta_i = 1$. Therefore, from (33), we get $\Pr[\mathrm{T}_i] = e^{-\theta_i/\lambda_i}$. On the other hand, when node i is energy constrained, $\bar{E}_i = \bar{H}$. Using $\bar{E}_i = PT_{\text{tx}}\Pr[\mathrm{T}_i]$ and rearranging the terms gives $\zeta_i = \rho e^{\theta_i/\lambda_i}$. Substituting this in (33) yields (8).

B. Proof of Result 1

For a positive-valued RV X and an event A, from Jensen's inequality, we know that

$$\mathbb{E}_X\left[\frac{1}{X}\Big|A\right] \geq \frac{1}{\mathbb{E}_X[X|A]}. \tag{35}$$

Hence,

$$\mathbb{E}_{\mathbf{h}}\left[\sigma_s^2\left(\frac{\sigma_s^2}{\sigma_v^2}\sum_{i\in\mathcal{T}} f(h_i)+1\right)^{-1}\Big|\mathcal{T}\right] \geq \sigma_s^2\left(\frac{\sigma_s^2}{\sigma_v^2}\sum_{i\in\mathcal{T}}\mathbb{E}_{h_i}\left[f(h_i)\Big|h_i\geq\theta_i\right]+1\right)^{-1}. \tag{36}$$

Since h_i is an exponential RV with mean λ_i, we get

$$\mathbb{E}_{h_i}\left[f(h_i)\Big|h_i\geq\theta_i\right] = \frac{1}{\lambda_i}\int_{\theta_i}^{\infty} f(x)e^{-\frac{(x-\theta_i)}{\lambda_i}} \triangleq m(\theta_i,\lambda_i). \tag{37}$$

Thus, from (6), (36), and (37), we get

$$\overline{\mathsf{MSE}} \geq \mathbb{E}_{\mathcal{T}}\left[\sigma_s^2\left(\frac{\sigma_s^2}{\sigma_v^2}\sum_{i\in\mathcal{T}} m(\theta_i,\lambda_i)+1\right)^{-1}\right]. \tag{38}$$

Considering all possibilities that track whether each node transmits or not, which can be written in terms of indicator functions $1_{\{\mathrm{T}_1\}},\ldots,1_{\{\mathrm{T}_N\}}$, (38) can be recast as

$$\overline{\mathsf{MSE}} \geq \mathbb{E}_{1_{\{\mathrm{T}_1\}},\ldots,1_{\{\mathrm{T}_N\}}}\left[\sigma_s^2\left(\frac{\sigma_s^2}{\sigma_v^2}\sum_{i=1}^{N} m(\theta_i,\lambda_i)\,1_{\{\mathrm{T}_i\}}+1\right)^{-1}\right]. \tag{39}$$

Using (35) and the equality $\mathbb{E}\left[1_{\{\mathrm{T}_i\}}\right] = \Pr[\mathrm{T}_i]$, we get

$$\overline{\mathsf{MSE}} \geq \sigma_s^2\left(\frac{\sigma_s^2}{\sigma_v^2}\sum_{i=1}^{N} m(\theta_i,\lambda_i)\Pr[\mathrm{T}_i]+1\right)^{-1}. \tag{40}$$

Substituting the expression for $\Pr[\mathrm{T}_i]$ from Lemma 1 in (40) yields the lower bound in (10).

C. Proof of Result 2

Minimizing Υ is the same as maximizing its denominator, which is given by

$$\frac{\sigma_s^2}{\sigma_v^2}\sum_{i\in\mathcal{C}} m(\theta_i,\lambda_i)\,\rho + \frac{\sigma_s^2}{\sigma_v^2}\sum_{j\in\mathcal{U}} m(\theta_j,\lambda_j)\,e^{-\frac{\theta_j}{\lambda_j}} + 1. \tag{41}$$

It is easy to see that

$$\max_{\theta_1,\ldots,\theta_N}\left\{\rho\sum_{i\in\mathcal{C}} m(\theta_i,\lambda_i) + \sum_{j\in\mathcal{U}} m(\theta_j,\lambda_j)\,e^{-\frac{\theta_j}{\lambda_j}}\right\} \leq \rho\sum_{i\in\mathcal{C}}\max_{\theta_i}\{m(\theta_i,\lambda_i)\} + \sum_{j\in\mathcal{U}}\max_{\theta_j}\left\{m(\theta_j,\lambda_j)\,e^{-\frac{\theta_j}{\lambda_j}}\right\}. \tag{42}$$

As shown below, this upper bound is achievable when

$$\theta_i = -\lambda_i\log\rho, \quad \text{for } 1\leq i\leq N. \tag{43}$$

Lemma 6: $m(\theta_i,\lambda_i)$ is a monotonically increasing function of θ_i, for $i\in\mathcal{C}$, and $m(\theta_j,\lambda_j)\,e^{-\theta_j/\lambda_j}$ is a monotonically decreasing function of θ_j, for $j\in\mathcal{U}$.

Proof: Differentiating $m(\theta_i,\lambda_i)$ with respect to θ_i, for $i\in\mathcal{C}$, and subsequently performing some algebraic manipulations, we get

$$\frac{\partial m(\theta_i,\lambda_i)}{\partial\theta_i} = \frac{1}{\lambda_i^2}\int_{\theta_i}^{\infty} f(x)e^{-\frac{(x-\theta_i)}{\lambda_i}}dx - \frac{f(\theta_i)}{\lambda_i}. \tag{44}$$

We can recast (44) as follows:

$$\frac{\partial m(\theta_i,\lambda_i)}{\partial\theta_i} = \frac{1}{\lambda_i^2}\int_{\theta_i}^{\infty}(f(x)-f(\theta_i))\,e^{-\frac{(x-\theta_i)}{\lambda_i}}dx \geq 0. \tag{45}$$

Hence, $m(\theta_i,\lambda_i)$ is increasing in θ_i for $i\in\mathcal{C}$.

To show that $m(\theta_j,\lambda_j)\,e^{-\theta_j/\lambda_j}$ is decreasing in θ_j, for $j\in\mathcal{U}$, consider $\frac{\partial}{\partial\theta_j}\left(m(\theta_j,\lambda_j)\,e^{-\theta_j/\lambda_j}\right)$:

$$\begin{aligned}\frac{\partial}{\partial\theta_j}\left(m(\theta_j,\lambda_j)\,e^{-\frac{\theta_j}{\lambda_j}}\right) &= \frac{\partial}{\partial\theta_j}\left(\frac{e^{-\frac{\theta_j}{\lambda_j}}}{\lambda_j}\int_{\theta_j}^{\infty} f(x)e^{-\frac{1}{\lambda_j}(x-\theta_j)}dx\right),\\ &= -\frac{e^{-\frac{\theta_j}{\lambda_j}}}{\lambda_j}f(\theta_j) \leq 0.\end{aligned} \tag{46}$$

The inequality follows because $\theta_j \geq 0$. Hence, the result follows. ■

D. Proof of Lemma 2

For the i^{th} node to be the j^{th} best, the channel power gains of exactly $j-1$ nodes must exceed h_i and those of the remaining $N-j$ nodes should be less than h_i. Let these $j-1$ nodes constitute the set $\mathcal{R}$. To obtain ν_{ij}, we sum over all possible $\mathcal{R}$. Thus,

$$\nu_{ij} = \sum_{\mathcal{R}\in\mathcal{G}_{j-1}[\mathcal{N}\setminus\{i\}]} \int_0^\infty \left[\prod_{m\in\mathcal{R}} \Pr[h_m > y | h_i = y]\right] \times \left[\prod_{l\in\mathcal{N}\setminus(\mathcal{R}\cup\{i\})} \Pr[h_l < y | h_i = y]\right] p_{h_i}(y)\, dy, \quad (47)$$

where $\mathcal{G}_{j-1}[\mathcal{N}\setminus\{i\}]$ is the collection of all subsets of $\mathcal{N}\setminus\{i\}$ of size $j-1$. Since h_i is an exponential RV with mean λ_i,

$$\nu_{ij} = \sum_{\mathcal{R}\in\mathcal{G}_{j-1}[\mathcal{N}\setminus\{i\}]} \frac{1}{\lambda_i} \int_0^\infty \left[\prod_{m\in\mathcal{R}} e^{-\frac{y}{\lambda_m}}\right] \times \left[\prod_{l\in\mathcal{N}\setminus(\mathcal{R}\cup\{i\})} \left(1 - e^{-\frac{y}{\lambda_l}}\right)\right] e^{-\frac{y}{\lambda_i}} dy. \quad (48)$$

Expanding the product terms in (48) gives

$$\nu_{ij} = \sum_{\mathcal{R}\in\mathcal{G}_{j-1}[\mathcal{N}\setminus\{i\}]} \frac{1}{\lambda_i} \int_0^\infty e^{-\left(\frac{y}{\lambda_i} + \sum_{m\in\mathcal{R}} \frac{y}{\lambda_m}\right)} \times \left[\sum_{\mathcal{Q}\in\mathcal{P}[\mathcal{N}\setminus(\mathcal{R}\cup\{i\})]} (-1)^{|\mathcal{Q}|} \left(e^{-\sum_{p\in\mathcal{Q}} \frac{y}{\lambda_p}}\right)\right] dy, \quad (49)$$

where $\mathcal{P}[\mathcal{N}\setminus(\mathcal{R}\cup\{i\})]$ is the power set of $\mathcal{N}\setminus(\mathcal{R}\cup\{i\})$. Integrating (49) and rearranging the terms yields (16).

Evaluating Pr[S_i]: Node i will be selected to transmit if it is the j^{th} best node, for some $1 \le j \le K$. Therefore,

$$\Pr[\mathrm{S}_i] = \Pr\left[(i = (1{:}\mathcal{N})) \cup \cdots \cup (i = (K{:}\mathcal{N}))\right]. \quad (50)$$

The events $i = (1{:}\mathcal{N}), \ldots, i = (K{:}\mathcal{N})$ are mutually exclusive. Thus, applying the law of total probability to (50) gives

$$\Pr[\mathrm{S}_i] = \sum_{j=1}^{K} \Pr[i = (j{:}\mathcal{N})] = \sum_{j=1}^{K} \nu_{ij}. \quad (51)$$

Evaluating Pr[A_i]: As we saw before, S_i and A_i are independent. Hence, $\bar{E}_i$ can be written as

$$\bar{E}_i = PT_{\text{tx}} \Pr[\mathrm{A}_i] \Pr[\mathrm{S}_i] = PT_{\text{tx}} \zeta_i \Pr[\mathrm{S}_i]. \quad (52)$$

Thus, (34) can be recast as

$$PT_{\text{tx}} \zeta_i \Pr[\mathrm{S}_i] \le \bar{H}. \quad (53)$$

When node i is energy unconstrained, $\zeta_i = 1$. When it is energy constrained, $\zeta_i < 1$ and $\bar{E}_i = \bar{H}$. Substituting this in (52), using (51), and finally solving for ζ_i gives (17).

E. Proof of Result 3

From (15), we know that

$$\overline{\mathsf{MSE}} = \mathbb{E}_{\mathbf{h}_{K:\mathcal{N}}, \mathbf{a}_{K:\mathcal{N}}} \left[\sigma_s^2 \left(\frac{\sigma_s^2}{\sigma_v^2} \sum_{i=1}^{K} f(h_{i:\mathcal{N}}) 1_{\{A_{i:\mathcal{N}}\}} + 1\right)^{-1}\right].$$

Using (35),

$$\overline{\mathsf{MSE}} \ge \sigma_s^2 \left(\frac{\sigma_s^2}{\sigma_v^2} \sum_{i=1}^{K} \mathbb{E}_{h_{i:\mathcal{N}}, 1_{\{A_{i:\mathcal{N}}\}}} \left[f(h_{i:\mathcal{N}}) 1_{\{A_{i:\mathcal{N}}\}}\right] + 1\right)^{-1}. \quad (54)$$

The RV $h_{i:\mathcal{N}}$ is determined by the channel power gains of all the nodes in the current slot. On the other hand, the RV $1_{\{A_{i:\mathcal{N}}\}}$ is a function of the channel power gains in the previous slots and the EH process upto the current slot. Since the channel fading process is independent across slots and is also independent of the EH process, the RVs $h_{1:\mathcal{N}}$ and $1_{\{\mathrm{A}_{i:\mathcal{N}}\}}$ are independent. Hence,

$$\mathbb{E}_{h_{i:\mathcal{N}}, 1_{\{A_{i:\mathcal{N}}\}}} \left[f(h_{i:\mathcal{N}}) 1_{\{A_{i:\mathcal{N}}\}}\right] = \mathbb{E}[f(h_{i:\mathcal{N}})]\, \mathbb{E}\left[1_{\{A_{i:\mathcal{N}}\}}\right]. \quad (55)$$

Since $f(h_{i:\mathcal{N}})$ is a concave function of $h_{i:\mathcal{N}}$, applying Jensen's inequality in (55) gives

$$\mathbb{E}_{h_{i:\mathcal{N}}}[f(h_{i:\mathcal{N}})]\, \mathbb{E}\left[1_{\{A_{i:\mathcal{N}}\}}\right] \le f(\mathbb{E}[h_{i:\mathcal{N}}]) \Pr[A_{i:\mathcal{N}}]. \quad (56)$$

From the results on order statistics of i.n.i.d. exponential RVs in [34], it can be shown that

$$\mathbb{E}[h_{i:\mathcal{N}}] = \sum_{n=i}^{N} \sum_{\eta\in\mathcal{S}_{\mathcal{N}}} \prod_{q=1}^{N} \left[\eta_q^{-1} \left(\sum_{r=1}^{q} \eta_r^{-1}\right)^{-1}\right] \left(\sum_{p=1}^{n} \eta_p^{-1}\right)^{-1}. \quad (57)$$

Using (55) and (56) in (54) yields

$$\overline{\mathsf{MSE}} \ge \sigma_s^2 \left(\frac{\sigma_s^2}{\sigma_v^2} \sum_{i=1}^{K} f(\mathbb{E}[h_{i:\mathcal{N}}]) \Pr[A_{i:\mathcal{N}}] + 1\right)^{-1}. \quad (58)$$

Evaluating Pr[$A_{i:\mathcal{N}}$]: The probability of the event $A_{i:\mathcal{N}}$ that the i^{th} best node in $\mathcal{N}$ is active can be written as

$$\Pr[A_{i:\mathcal{N}}] = \Pr\left[((1 = (i{:}\mathcal{N})) \cap \mathrm{A}_1) \cup \cdots \cup ((N = (i{:}\mathcal{N})) \cap \mathrm{A}_N)\right]. \quad (59)$$

The events $(1 = (i{:}\mathcal{N})) \cap \mathrm{A}_1, \ldots, (N = (i{:}\mathcal{N})) \cap \mathrm{A}_N$ are mutually exclusive. Hence,

$$\Pr[A_{i:\mathcal{N}}] = \sum_{j=1}^{N} \Pr[(j = (i{:}\mathcal{N})) \cap \mathrm{A}_j]. \quad (60)$$

For any node j, the event $j = (i{:}\mathcal{N})$ is a function of the channel power gains of all the nodes in the current slot. On the other hand, A_j is determined by the channel power gains in the previous slots and the EH process upto the current slot. Hence, by the same reasoning as above about the independence between $h_{i:\mathcal{N}}$ and $1_{\{A_{i:\mathcal{N}}\}}$, it can be inferred that for any

node j, the events $j = (i{:}\mathcal{N})$ and A_j are independent. Thus, (60) becomes

$$\Pr[A_{i:\mathcal{N}}] = \sum_{j=1}^{N} \Pr[j = (i:\mathcal{N})] \Pr[A_j]\,. \tag{61}$$

Substituting the expressions for $\Pr[j = (i{:}\mathcal{N})]$ and $\Pr[\mathrm{A}_j]$ from (16) and (17), respectively, in (61), and then substituting the expression for $\Pr[A_{i:\mathcal{N}}]$ in (58) yields (18).

F. Proof of Lemma 5

For the ATO scheme, the probability $\Pr[\mathrm{T}_i]$ that node i transmits can be written as

$$\Pr[\mathrm{T}_i] = \Pr[\mathrm{S}_i|\mathrm{A}_i] \Pr[\mathrm{A}_i] = \zeta_i \Pr[\mathrm{S}_i|\mathrm{A}_i]\,. \tag{62}$$

Recall that $\mathcal{A}$ is a set of active nodes such that $\mathcal{A} \subseteq \mathcal{N} \setminus \{i\}$. Hence, writing in terms of the number of nodes m other than i that are active, we get

$$\Pr[\mathrm{S}_i|\mathrm{A}_i] = \sum_{m=0}^{N-1} \Pr[\mathrm{S}_i, |\mathcal{A}| = m|\mathrm{A}_i]. \tag{63}$$

We can rewrite (63) using the relation $\Pr[\mathrm{S}_i, |\mathcal{A}| = m|\mathrm{A}_i] = \Pr[\mathrm{S}_i|\mathrm{A}_i, |\mathcal{A}| = m] \Pr[|\mathcal{A}| = m|\mathrm{A}_i]$. Thus,

$$\Pr[\mathrm{S}_i|\mathrm{A}_i] = \sum_{m=0}^{N-1} \Pr[\mathrm{S}_i|\mathrm{A}_i, |\mathcal{A}| = m] \Pr[|\mathcal{A}| = m|\mathrm{A}_i]. \tag{64}$$

We break the summation in (64) into two parts. The first one is for the cases in which $|\mathcal{A}| \leq K-1$, and the second one is for $|\mathcal{A}| \geq K$. Thus,

$$\begin{aligned}\Pr[\mathrm{S}_i|\mathrm{A}_i] = &\sum_{m=0}^{K-1} \Pr[\mathrm{S}_i|\mathrm{A}_i, |\mathcal{A}| = m] \Pr[|\mathcal{A}| = m|\mathrm{A}_i] \\ &+ \sum_{m=K}^{N-1} \Pr[\mathrm{S}_i|\mathrm{A}_i, |\mathcal{A}| = m] \Pr[|\mathcal{A}| = m|\mathrm{A}_i].\end{aligned} \tag{65}$$

As per the decoupling approximation, $\Pr[\mathcal{A}|\mathrm{A}_i] = \Pr[\mathcal{A}]$ since $i \notin \mathcal{A}$. Furthermore, when $m \leq K-1$, we have $\Pr[\mathrm{S}_i|\mathrm{A}_i, |\mathcal{A}| = m] = 1$. Thus, (65) can be written as

$$\begin{aligned}\Pr[\mathrm{S}_i|\mathrm{A}_i] = &\sum_{m=0}^{K-1} \Pr[|\mathcal{A}| = m] \\ &+ \sum_{m=K}^{N-1} \Pr[\mathrm{S}_i|\mathrm{A}_i, |\mathcal{A}| = m] \Pr[|\mathcal{A}| = m]\,.\end{aligned} \tag{66}$$

For $m \leq K-1$,

$$\Pr[|\mathcal{A}| = m] = \sum_{\substack{\mathcal{A} \subseteq \mathcal{N}\setminus\{i\}: \\ |\mathcal{A}| = m}} \left(\prod_{j\in\mathcal{A}} \zeta_j\right) \prod_{l \in \mathcal{N}\setminus(\mathcal{A}\cup\{i\})} (1-\zeta_l)\,. \tag{67}$$

Similarly, for $m \geq K$, we can show that

$$\begin{aligned}&\Pr[\mathrm{S}_i|\mathrm{A}_i, |\mathcal{A}| = m] \Pr[|\mathcal{A}| = m] \\ &= \sum_{\substack{\mathcal{A} \subseteq \mathcal{N}\setminus\{i\}: \\ |\mathcal{A}| = m}} \Pr[\mathrm{S}_i|\mathcal{A}\cup\{i\}] \left(\prod_{j\in\mathcal{A}} \zeta_j\right) \prod_{l \in \mathcal{N}\setminus(\mathcal{A}\cup\{i\})} (1-\zeta_l)\,.\end{aligned} \tag{68}$$

Therefore, (66) simplifies to

$$\begin{aligned}\Pr[\mathrm{S}_i|\mathrm{A}_i] = &\sum_{\substack{\mathcal{A} \subseteq \mathcal{N}\setminus\{i\}: \\ 0 \leq |\mathcal{A}| \leq K-1}} \left(\prod_{j\in\mathcal{A}} \zeta_j\right) \prod_{l \in \mathcal{N}\setminus(\mathcal{A}\cup\{i\})} (1-\zeta_l) \\ &+ \sum_{\substack{\mathcal{A} \subseteq \mathcal{N}\setminus\{i\}: \\ K \leq |\mathcal{A}| \leq N-1}} \Pr[\mathrm{S}_i|\mathcal{A}\cup\{i\}] \left(\prod_{j\in\mathcal{A}} \zeta_j\right) \\ &\times \prod_{l \in \mathcal{N}\setminus(\mathcal{A}\cup\{i\})} (1-\zeta_l)\,.\end{aligned} \tag{69}$$

Substituting (69) in (62) yields the desired result in (27).

G. Proof of Result 4

Using the inequality in (35), the first term in (23) yields

$$\begin{aligned}\mathbb{E}_{\mathbf{h}_{\mathcal{N}_a}}&\left[\sigma_s^2 \left(\frac{\sigma_s^2}{\sigma_v^2} \sum_{i\in\mathcal{N}_a} f(h_i) + 1\right)^{-1} \Big| \mathcal{N}_a\right] \\ &\geq \sigma_s^2 \left(\frac{\sigma_s^2}{\sigma_v^2} \sum_{i\in\mathcal{N}_a} \mathbb{E}_{h_i}[f(h_i)] + 1\right)^{-1}, \qquad (70) \\ &= \sigma_s^2 \left(\frac{\sigma_s^2}{\sigma_v^2} \sum_{i\in\mathcal{N}_a} m\,(0,\lambda_i) + 1\right)^{-1}. \qquad (71)\end{aligned}$$

Note that in (70), the conditioning on $\mathcal{N}_a$ is dropped because, as argued before, h_i is independent of $\mathcal{N}_a$. Applying Jensen's inequality to the second term in (23), we get

$$\begin{aligned}\mathbb{E}_{\mathbf{h}_{K:\mathcal{N}_a}}&\left[\sigma_s^2 \left(\frac{\sigma_s^2}{\sigma_v^2} \sum_{i=1}^{K} f(h_{i:\mathcal{N}_a}) + 1\right)^{-1} \Big| \mathcal{N}_a\right] \\ &\geq \sigma_s^2 \left(\frac{\sigma_s^2}{\sigma_v^2} \sum_{i=1}^{K} \mathbb{E}_{h_{i:\mathcal{N}_a}}\left[f(h_{i:\mathcal{N}_a})\Big|\mathcal{N}_a\right] + 1\right)^{-1}. \qquad (72)\end{aligned}$$

Since $f(h_{i:\mathcal{N}_a})$ is a concave function of $h_{i:\mathcal{N}_a}$, applying Jensen's inequality to (72) gives

$$\begin{aligned}&\sigma_s^2 \left(\frac{\sigma_s^2}{\sigma_v^2} \sum_{i=1}^{K} \mathbb{E}_{h_{i:\mathcal{N}_a}}\left[f(h_{i:\mathcal{N}_a})\Big|\mathcal{N}_a\right] + 1\right)^{-1} \\ &\qquad \geq \sigma_s^2 \left(\frac{\sigma_s^2}{\sigma_v^2} \sum_{i=1}^{K} f(\mathbb{E}[h_{i:\mathcal{N}_a}|\mathcal{N}_a]) + 1\right)^{-1}. \qquad (73)\end{aligned}$$

The expression for $\mathbb{E}[h_{i:\mathcal{N}_a}|\mathcal{N}_a]$ is given in (29).

Furthermore, from the decoupling approximation, $\Pr[\mathcal{N}_a] = \left(\prod_{j\in\mathcal{N}_a} \zeta_j\right) \prod_{l\in\mathcal{N}_a^c} (1-\zeta_l)$. Substituting this expression and using (71), (72), and (73) in (23) yields (28).

REFERENCES

[1] S. Sen Gupta and N. B. Mehta, "Revisiting censoring in energy harvesting wireless sensor networks," in *Proc. IEEE Globecom*, Dec. 2018, pp. 1–6.

[2] O. Younis, M. Krunz, and S. Ramasubramanian, "Node clustering in wireless sensor networks: Recent developments and deployment challenges," *IEEE Netw.*, vol. 20, no. 3, pp. 20–25, May 2006.

[3] S. He, J. Chen, X. Li, X. Shen, and Y. Sun, "Leveraging prediction to improve the coverage of wireless sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 4, pp. 701–712, Apr. 2012.

[4] Z. Han and H. V. Poor, "Lifetime improvement of wireless sensor networks by collaborative beamforming and cooperative transmission," in *Proc. ICC*, Jun. 2007, pp. 3954–3958.

[5] S. Appadwedula, V. V. Veeravalli, and D. L. Jones, "Decentralized detection with censoring sensors," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1362–1373, Apr. 2008.

[6] R. S. Blum and B. M. Sadler, "Energy efficient signal detection in sensor networks using ordered transmissions," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3229–3235, Jul. 2008.

[7] E. J. Msechu and G. B. Giannakis, "Sensor-centric data reduction for estimation with WSNs via censoring and quantization," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 400–414, Jan. 2012.

[8] J. Matamoros and C. Antón-Haro, "Opportunistic power allocation and sensor selection schemes for wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 534–539, Feb. 2010.

[9] V. Shah, N. B. Mehta, and R. Yim, "Optimal timer based selection schemes," *IEEE Trans. Commun.*, vol. 58, no. 6, pp. 1814–1823, Jun. 2010.

[10] R. S. Blum, "Ordering for estimation and optimization in energy efficient sensor networks," *IEEE Trans. Signal Process.*, vol. 59, no. 6, pp. 2847–2856, Jun. 2011.

[11] Y. Zhao, B. Chen, and R. Zhang, "Optimal power management for remote estimation with an energy harvesting sensor," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6471–6480, Nov. 2015.

[12] 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.

[13] A. Özçelikkale, T. McKelvey, and M. Viberg, "Transmission strategies for remote estimation with an energy harvesting sensor," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4390–4403, Jul. 2017.

[14] M. Nourian, S. Dey, and A. Ahlén, "Distortion minimization in multi-sensor estimation with energy harvesting," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 3, pp. 524–539, Mar. 2015.

[15] P. Du, Q. Yang, Z. Shen, and K. S. Kwak, "Distortion minimization in wireless sensor networks with energy harvesting," *IEEE Commun. Lett.*, vol. 21, no. 6, pp. 1393–1396, Jun. 2017.

[16] S. Zhang, S. Liu, V. Sharma, and P. K. Varshney, "Optimal sensor collaboration for parameter tracking using energy harvesting sensors," *IEEE Trans. Signal Process.*, vol. 66, no. 12, pp. 3339–3353, Jun. 2018.

[17] S. Knorn, S. Dey, A. Ahlén, and D. E. Quevedo, "Distortion minimization in multi-sensor estimation using energy harvesting and energy sharing," *IEEE Trans. Signal Process.*, vol. 63, no. 11, pp. 2848–2863, Jun. 2015.

[18] D. Niyato, E. Hossain, M. M. Rashid, and V. K. Bhargava, "Wireless sensor networks with energy harvesting technologies: A game-theoretic approach to optimal energy management," *IEEE Wireless Commun.*, vol. 14, no. 4, pp. 90–96, Aug. 2007.

[19] P. Blasco, D. Gündüz, and M. Dohler, "A learning theoretic approach to energy harvesting communication system optimization," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1872–1882, Apr. 2013.

[20] A. Ortiz, H. Al-Shatri, X. Li, T. Weber, and A. Klein, "Reinforcement learning for energy harvesting point-to-point communications," in *Proc. ICC*, May 2016, pp. 1–6.

[21] A. Masadeh, Z. Wang, and A. E. Kamal, "Reinforcement learning exploration algorithms for energy harvesting communications systems," in *Proc. ICC*, May 2018, pp. 1–6.

[22] I. Ahmed, K. T. Phan, and T. Le-Ngoc, "Optimal stochastic power control for energy harvesting systems with delay constraints," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3512–3527, Dec. 2016.

[23] J. Fernandez-Bes, J. Cid-Sueiro, and A. G. Marques, "An MDP model for censoring in harvesting sensors: Optimal and approximated solutions," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 8, pp. 1717–1729, Aug. 2015.

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

[25] A. Kazerouni and A. Özgür, "Optimal online strategies for an energy harvesting system with Bernoulli energy recharges," in *Proc. WiOpt*, May 2015, pp. 235–242.

[26] A. Seyedi and B. Sikdar, "Energy efficient transmission strategies for body sensor networks with energy harvesting," *IEEE Trans. Commun.*, vol. 58, no. 7, pp. 2116–2126, Jul. 2010.

[27] H. Zhou, T. Jiang, C. Gong, and Y. Zhou, "Optimal estimation in wireless sensor networks with energy harvesting," *IEEE Trans. Veh. Technol.*, vol. 65, no. 11, pp. 9386–9396, Nov. 2016.

[28] K. Ramachandran and B. Sikdar, "A population based approach to model the lifetime and energy distribution in battery constrained wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 4, pp. 576–586, May 2010.

[29] M. F. A. Ahmed, T. Y. Al-Naffouri, M.-S. Alouini, and G. Turkiyyah, "The effect of correlated observations on the performance of distributed estimation," *IEEE Trans. Signal Process.*, vol. 61, no. 24, pp. 6264–6275, Dec. 2013.

[30] S. Rao and N. B. Mehta, "Energy harvesting WSNs for accurately estimating the maximum sensor reading: Trade-offs and optimal design," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4562–4573, Aug. 2015.

[31] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York, NY, USA: Springer-Verlag, 1994.

[32] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, 9th ed. New York, NY, USA: Dover, 1972.

[33] B. Medepally and N. B. Mehta, "Voluntary energy harvesting relays and selection in cooperative wireless networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3543–3553, Nov. 2010.

[34] M. Z. Win and J. H. Winters, "Analysis of hybrid selection/maximal-ratio combining of diversity branches with unequal SNR in Rayleigh fading," in *Proc. VTC*, May 1999, pp. 215–220.

Sayan Sen Gupta (S'18) received the B.Tech. degree in electronics and communications from the West Bengal University of Technology, India, in 2011, and the M.Eng. degree in electronics and telecommunications from Jadavpur University, India, in 2013. He is currently pursuing the Ph.D. degree with the Department of Electrical Communication Engineering, Indian Institute of Science, Bengaluru. From 2013 to 2015, he was with Tata Consultancy Services Ltd. His research interests include design and analysis of energy harvesting wireless sensor networks.

Neelesh B. Mehta (S'98–M'01–SM'06–F'19) received the B.Tech. degree in electronics and communications engineering from IIT Madras in 1996, and the M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology, Pasadena, CA, USA, in 1997 and 2001, respectively. He is currently a Professor with the Department of Electrical Communication Engineering, Indian Institute of Science, Bengaluru. He is a fellow of the Indian National Academy of Engineering and the National Academy of Sciences India. He is a recipient of the Shanti Swarup Bhatnagar Award 2017 and the Swarnjayanti Fellowship. He has served on the Board of Governors of IEEE ComSoc from 2012 to 2015. He has served on the Executive Editorial Committee of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2014 to 2017, and served as its chair from 2017 to 2018. He is currently an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS.