



## Remote Estimation of Correlated Sources under Energy Harvesting Constraints

Downloaded from: <https://research.chalmers.se>, 2025-12-05 03:12 UTC

Citation for the original published paper (version of record):

Ozcelikkale, A., McKelvey, T., Viberg, M. (2018). Remote Estimation of Correlated Sources under Energy Harvesting Constraints. IEEE Transactions on Wireless Communications, 17(8): 5300-5313. <http://dx.doi.org/10.1109/TWC.2018.2841390>

N.B. When citing this work, cite the original published paper.

© 2018 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, or reuse of any copyrighted component of this work in other works.

# Remote Estimation of Correlated Sources under Energy Harvesting Constraints

Ayça Özçelikkale, Tomas McKelvey, Mats Viberg

**Abstract**—Remote estimation with an energy harvesting sensor with a limited data and energy buffer is considered. The sensor node observes an unknown temporally correlated field and communicates its observations to a remote fusion center using the energy it harvested. The fusion center employs linear minimum mean-square error (LMMSE) estimation to reconstruct the unknown field. We provide performance guarantees for the estimation error under a block transmission scheme, where at each transmission block, data and energy buffers are completely emptied. Our bounds provide insights into how statistical properties of the energy harvesting process and buffer sizes may affect the estimation error. In particular, these bounds suggest insensitivity of the performance to buffer sizes for signals with low degree of freedom and suggest performance improvements with increasing buffer sizes for signals with relatively higher degree of freedom. Depending only on the mean, variance and finite support of the energy arrival process, these results provide insights for the energy and data buffer sizes for deployment in future energy harvesting wireless sensing systems.

**Index Terms**—energy harvesting, wireless sensor networks, distortion minimization, correlated signals

## I. INTRODUCTION

With the ever increasing number of connected devices, where over 16 billion devices are expected to be connected by 2022, powering of these devices and enabling energy autonomous networked systems is a central concern [1]. Here, energy harvesting provides a promising approach. In energy harvesting (EH) systems, devices are equipped with capabilities to collect energy from renewable sources, such as solar power. EH capabilities not only enable efficient usage of energy sources but also offer enhanced mobility and prolonged network life-times [2–5].

Feasibility of energy harvesting approaches have been investigated and favourable results are obtained for various scenarios, including harvesting from solar energy, mechanical energy sources and radio-frequency (RF) energy [3–5]. Devices with various energy harvesting modalities have recently become commercially available including solar [6], thermoelectric [7] and vibrations [8]. Deployments that utilize energy harvesting solutions have already started to appear for a wide range of applications, including smart buildings and industrial sites [6–8]. Relevant standardisation and commercial solution efforts

that target low power consumption for sensor nodes for internet of things solutions, such as EnOcean wireless standard (ISO/IEC 14543-3-1X), LoRa (<https://loro-alliance.org>) and SigFox (<https://www.sigfox.com>) have recently emerged.

In parallel to these promising developments, there has been a significant effort to understand information transfer capabilities of communication systems with EH capabilities. In the case of energy harvesting from RF sources, the main challenge lies in designing the optimal strategies at the transmitters [9–12]. In the case of systems energy harvesting from natural sources, such as solar power, the key issue is the intermittent nature of the energy supply. The main challenge in these systems is to provide reliable and efficient operation even when the energy supply is unreliable. In this work, we focus on this intermittent nature of EH sources and its effect on the performance of remote estimation systems.

## A. Prior Work

An important distinction in the energy harvesting literature is the one between the offline optimization scheme and the online optimization scheme [2]. In the offline (or deterministic) scheme, profile of the harvested energy is assumed to be known non-causally. In contrast, in the online (or stochastic) scheme, only statistical knowledge about the future energy arrivals is assumed to be known. The offline optimization scheme is relatively well-studied, especially in terms of formulations that adopt communication rate as the performance metric. Analytical results exist for various scenarios, such as point-to-point channels [13], [14], broadcast channels [15] and multiple-access channels [16].

In contrast, online scheme is considered to be less tractable analytically. A typical numerical method here is dynamic programming approach, which utilizes a search over a quantized state space. Unfortunately, this approach not only has high computational complexity, which limits its applicability in low-complexity EH sensors, but it also falls short of providing systematic insight into the effect of system parameters [2]. On the other hand, results that directly provide analytical insight for the online scheme are available only for a limited number of scenarios [17–21]. Structural results for capacity and rate optimization under intermittent energy arrivals are provided in [17–20]. Under a binary decision scheme, where at each time instant the sensor makes a decision to transmit or not, threshold-based policies are proven to be optimal for remote estimation of Markov sources [21]. A learning theoretical approach, where optimal transmission strategies are learned over time without knowledge of statistical parameters

A. Özçelikkale acknowledges the funding support from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 654123 and Swedish Research Council under grant 2015-04011. A. Özçelikkale is with the Division of Signals and Systems, Uppsala University, Sweden. T. McKelvey is with the Dept. of Electrical Engineering, Chalmers University of Technology, Sweden. M. Viberg is with the Dept. of Architecture and Civil Engineering, Chalmers University of Technology, Sweden. E-mails: [ayca.ozcelikkale@angstrom.uu.se](mailto:ayca.ozcelikkale@angstrom.uu.se), [{tomas.mckelvey,mats.viberg}@chalmers.se](mailto:{tomas.mckelvey,mats.viberg}@chalmers.se).

of energy and data arrival processes, is investigated [22]. Here, we contribute to the analysis of online scheme by providing performance guarantees for signal recovery under a block transmission strategy. We further discuss our approach in Section I-B.

Establishing a close connection with estimation of unknown physical fields and in particular degree of sparsity, hence varying degrees of correlation of unknown signals, is an important aspect of performance evaluation for sensing systems. Here sparsity, or equivalently degree of freedom of a signal family, refers to the effective low dimensionality of the unknown signal [23]. In addition to providing a reasonable model for physical fields, sparsity can be utilized to compensate for the unreliable nature of the energy sources in an EH system. Yet, for EH systems, structural results that directly exploit sparsity or correlation characteristics are available only for a limited number of scenarios, such as estimation of a single parameter [24–26], Markov sources [21], [27], [28], circularly wide-sense stationary signals [29], [30], two correlated Gaussian variables [31], and i.i.d. Gaussian sources, as a result of the findings of, for instance, [17–20], [32], [33]. As we will further discuss in Section I-B, here we contribute to this aspect by providing performance guarantees for estimation error which depends on the sparsity of the unknown signal and statistical properties of the energy arrival process.

## B. Contributions

In this work, we consider an EH sensor which observes an unknown correlated field and communicates its observations to a remote fusion center using the energy it harvested. The fusion center employs linear minimum mean-square error (LMMSE) estimation to reconstruct the unknown signal. We consider this problem under a limited data and energy buffer constraint using a block transmission scheme where, at each transmission frame, the data buffer and the battery are completely emptied. Motivated by the high complexity and the high energy cost of source and channel coding operations, we consider an amplify-and-forward strategy as in [25], [26], [32]. We focus on the scheme where the energy used at each transmission is modelled as a random variable, i.e. the online scheme. A preliminary version of this setup is considered in [34], where energy arrival process is restricted to be a Bernoulli process and signal model is restricted to circularly wide-sense stationary signals.

An important contribution of our work stems from our focus on the correlated signal model. Due to this correlated signal model, calculation of the mean-square error requires evaluation of a matrix inverse. Hence the performance criterion, in general, cannot be written as a summation of utilities over time in contrast to the case of formulations based on throughput [14], [15]. Using random matrix theory and compressive sensing tools, we provide performance bounds for this correlated signal set-up under a block transmission scheme. Our results provide insights into how statistical properties of the energy harvesting process and buffer sizes affect the estimation error. Consistent with compressive sensing (CS) results, our bounds suggest insensitivity of the performance to the buffer size for signals

with low degree of freedom, and possible performance gains due to increasing buffer sizes for signals with relatively higher degree of freedom. These performance guarantees, which depend on the sparsity of the signal to be observed and the first and second order statistical properties of the energy arrival process, provide insights into buffer and battery size choices.

An important special case we consider is the case of circularly wide-sense stationary (c.w.s.s.) signals, which are a finite-dimensional analog of wide-sense stationary signals [35], [36]. In addition to the above block transmission scheme, we also consider the strategy of transmission of equidistant samples for the low-pass c.w.s.s. signals. The equidistant sample transmission scheme is motivated by the sampling theorems for c.w.s.s. signals [23], [36]. Our performance guarantees suggest that for low-pass c.w.s.s. signals similar performance can be obtained by both strategies of block transmission (i.e. spreading the energy as much as possible on all samples in the buffer) and sending only equidistant samples with all the energy in the battery at each transmission frame. Our results here complement the results of Ref. [30], where the focus is on the off-line scheme and no high probability results are presented. Together with the off-line results of [30], the performance bounds presented here support the possible flexibility in energy management for sensing of low-pass c.w.s.s. signals under energy harvesting constraints.

The rest of the paper is organized as follows. In Section II, system model is described. Our performance guarantees are presented in Section III. We consider the case of c.w.s.s. signals in Section IV. Discussion of connections to compressive sensing is provided in Section V. In Section VI, numerical illustrations are provided. The paper is concluded in Section VII.

*Notation:* We denote a column vector  $\mathbf{a} \in \mathbb{C}^{N \times 1}$  by  $\mathbf{a} = [a_1; \dots; a_N] \in \mathbb{C}^{N \times 1}$  where semi-colon ; is used to separate the rows. Complex conjugate transpose of a matrix  $A$  is denoted by  $A^\dagger$ . Spectral norm of a matrix  $A$  is denoted by  $\|A\|$ . The  $l^{th}$  row,  $k^{th}$  column element of a matrix  $A$  is denoted by  $[A]_{lk}$ . Positive semi-definite (p.s.d.) partial ordering for Hermitian matrices is denoted by  $\succeq$ .  $I_N$  denotes the identity matrix with  $I_N \in \mathbb{R}^{N \times N}$ . The  $l_2$  norm of a vector  $\mathbf{a}$  is denoted by  $\|\mathbf{a}\|$ . We denote the diagonal matrix whose diagonal elements are the elements of the vector  $\mathbf{a}$  by  $\text{diag}(\mathbf{a})$ . Statistical expectation is denoted by  $\mathbb{E}[\cdot]$ . We denote expectation with respect to (w.r.t.) signals involved with  $\mathbb{E}_S[\cdot]$  and expectation w.r.t. energy arrivals with  $\mathbb{E}_E[\cdot]$  for the sake of clarity when needed.

## II. SYSTEM MODEL

### A. Signal Model

The aim of the remote estimation system is to estimate the unknown complex proper zero mean field  $\mathbf{x} = [x_1; \dots; x_N] \in \mathbb{C}^{N \times 1}$ . Here  $\mathbf{x} \in \mathbb{C}^{N \times 1}$  denotes a field that is defined over time and  $x_t$  denotes the field value at time  $t$ , where  $t = 1, \dots, N$ . The covariance matrix  $K_{\mathbf{x}} = \mathbb{E}[\mathbf{x}\mathbf{x}^\dagger]$  models the possible correlation of the field values in time. Let  $s$  be the number of non-zero eigenvalues of  $K_{\mathbf{x}}$ , i.e. rank of  $K_{\mathbf{x}}$ . Let  $K_{\mathbf{x}} = U_s \Lambda_{x,s} U_s^\dagger$  be the (reduced) eigenvalue decomposition

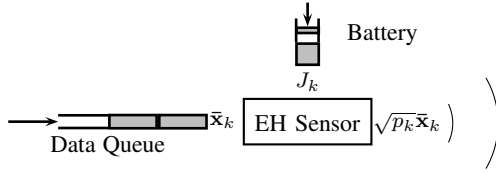


Fig. 1: Energy Harvesting Sensor

(EVD) of  $K_{\mathbf{x}}$ , where  $\Lambda_{x,s} \in \mathbb{C}^{s \times s}$  is the diagonal matrix of non-zero eigenvalues and  $U_s \in \mathbb{C}^{N \times s}$  is the sub-matrix of unitary  $U \in \mathbb{C}^{N \times N}$  corresponding to non-zero eigenvalues. Let  $P_x = \text{tr}[K_{\mathbf{x}}] = \text{tr}[\Lambda_{x,s}]$ . We consider  $\Lambda_{x,s}$ 's of the form  $\Lambda_{x,s} = \frac{P_x}{s} I_s$ . Here  $s$  gives the number of degrees of freedom (d.o.f.), i.e. sparsity level of the signal family.

We note that this model covers signal families with a wide range of correlation structures. In particular, signals with rank one correlation matrices where the signal components have a correlation coefficient of one, and white signals with  $K_{\mathbf{x}} = \Lambda_{x,n} = I_n$  are covered with this model. By varying  $s$  and  $U_s$ , this model can be used to represent signals with different correlation structures in between. This type of models have been used to represent signal families that have a low degree of freedom in various signal applications, for instance as a sparse signal model in compressive sensing literature [23], [37].

### B. Sensing and Communications to the Fusion Center

We consider an energy harvesting sensor as shown in Fig. 1. We focus on a slotted discrete-time setting where at each time slot  $t$ , the sensor observes the field value at time slot  $t$ , i.e.  $x_t$ . The observations are held in a buffer of finite size  $Q_d$  before transmission. We consider a block transmission scheme, where time slots  $t$  satisfying  $(k-1)Q_d + 1 \leq t \leq kQ_d$  belongs to transmission frame  $k$  as shown in Fig. 2. Hence, the buffer contents at the end of transmission frame  $k$ , i.e. at the end of time slot  $kQ_d$ , is given by  $\bar{\mathbf{x}}_k = [x_{(k-1)Q_d+1}; x_{(k-1)Q_d+2}; \dots; x_{(k-1)Q_d+Q_d}] \in \mathbb{C}^{Q_d \times 1}$ . For convenience,  $N_T = N/Q_d$  is assumed to be an integer, where  $N_T$  gives the number of transmission frames. At the end of transmission frame  $k$ , the sensor transmits the data in its buffer to a fusion center using an amplify-and-forward block transmission strategy as follows

$$\bar{\mathbf{y}}_k = \sqrt{p_k} \bar{\mathbf{x}}_k + \bar{\mathbf{w}}_k, \quad k = 1, \dots, N_T, \quad (1)$$

where  $p_k$ ,  $\bar{\mathbf{w}}_k$  and  $\bar{\mathbf{y}}_k$  denote the amplification factor, channel noise and the received signal at the fusion center for transmission frame  $k$ , respectively. The channel noise  $\mathbf{w} = [\bar{\mathbf{w}}_1; \dots; \bar{\mathbf{w}}_{N_T}] \in \mathbb{C}^{N \times 1}$  is modeled as complex proper zero mean with  $K_{\mathbf{w}} = \mathbb{E}[\mathbf{w}\mathbf{w}^\dagger] = \sigma_w^2 I_N$ .

The above type of block transmission scheme allows us to spread the energy over multiple signal samples and facilitates connections with uniform power allocation strategies which are optimal for white sources in the offline scheme [17], [30], and the power allocation strategies which match the average arrival rate of the EH process and optimal for white sources in the online scheme [19]. It is also supported by the fact that for devices with low power budgets, it is more energy efficient

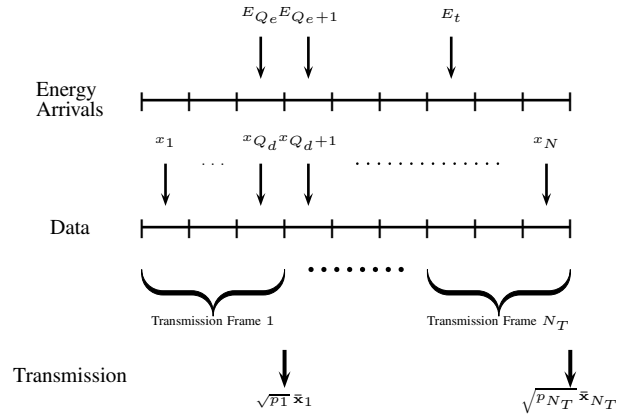


Fig. 2: Time Schedule for the Energy Harvesting Sensor

to send relatively larger amount of data at each transmission [38].

As a second-order characterization of the dynamical range of  $x_t$ , we assume that  $\mathbb{P}(x_t \notin [-\alpha_r \sigma_{x_t}, \alpha_r \sigma_{x_t}]) \leq \epsilon_r \approx 0$ , where  $\sigma_{x_t}^2$  is the variance of  $x_t$  and  $\alpha_r > 0$  is a given constant. For instance, when  $x_t$ 's are modeled as uniform random variables over  $[-a_t, +a_t]$ , one has an exact equality, i.e.  $\epsilon_r = 0$  with  $\alpha_r = \sqrt{12}$ . For the case where  $x_t$ 's are modelled as Gaussian random variables, one has  $\epsilon_r \approx 0$  for a large enough choice of  $\alpha_r$ , where  $\alpha_r = 3$  is a common practical choice used as the effective width of a Gaussian random variable. Hence, we assume that  $x_t$ 's are delimited according to their effective dynamical range with a suitably chosen  $\alpha_r$  and ignore the possible saturation effect. We assume that energy cost of sending a sensor measurement scales with the variance of the random variable but not with its realization. We note that this is consistent with the fact that modern sensors typically provide measurements using analog-to-digital converters and the outputs of such sensors are represented by the same number of bits regardless of the realization of the physical quantity measured. Hence, the energy used by the sensor for communications at frame  $k$  can be written as follows

$$J_k = \beta \sum_{t=1}^{Q_d} p_k \sigma_{x_{(k-1)Q_d+t}}^2, \quad (2)$$

where  $\beta$  is a proportionality constant that includes the time duration per symbol and  $\alpha_r$ . For convenience,  $\beta$  is normalized as  $\beta = 1$  in the rest of the paper. We note that even when sensor outputs are represented by a fixed number of bits, one may need to scale the channel input according to the dynamic range of the random variable to have an effective signal-to-noise ratio on the channel that is consistent with the variable's dynamic range. Hence, we assume that the energy cost of a transmission scales with  $\sigma_{x_t}^2$ . Note that for the scenarios with  $\sigma_{x_t}^2$  constant, such as c.w.s.s. signals, the energy cost only changes with  $p_k$ . A more closer investigation of transmission of sensor outputs where there is no channel coding but fixed-bit sensor readings are transmitted requires investigation of different modulations, quantization schemes, headers; which is beyond the scope of

this work. Here, we use (1) together with (2) as a limited but nevertheless analytically tractable model for transmission of sensor outputs over noisy channels.

We note that our formulation also covers the following related scenario where the energy cost of transmissions are assumed to directly scale with the energy of the realizations: Suppose that during the time slot  $t$ ,  $N_R$  statistically independent realizations of  $x_t$  arrives. Hence, the data buffer holds  $Q_d \times N_R$  values. For instance, with  $Q_d = 2$  for transmission frame  $k$ , the data buffer holds  $2N_R$  values:  $N_R$  realizations of  $x_{(k-1)Q_d+1}$  and  $N_R$  realizations of  $x_{(k-1)Q_d+2}$ . At the end of frame  $k$ , sensor sends  $Q_d \times N_R$  values using the scaling  $\sqrt{p_k}$ . If  $N_R$  is large enough, average energy of the realizations will be proportional to  $\sum_{t=1}^{Q_d} p_k \sigma_{x_{(k-1)Q_d+t}}^2$  regardless of the value of  $Q_d$ , since we then have  $\lim_{N_R \rightarrow \infty} \frac{1}{N_R} \sum_{i=1}^{N_R} |x_{(Q_d-1)k+t}^i|^2 \rightarrow \mathbb{E}[|x_{(Q_d-1)k+t}^i|^2] = \sigma_{x_{(Q_d-1)k+t}}^2$ , where  $x_{(Q_d-1)k+t}^i$  is the  $i$ th realization of  $x_{(Q_d-1)k+t}$  and  $\mathbb{E}[x_{(Q_d-1)k+t}] = 0$ . Hence, our formulation also covers this case with a possibly different constant  $\beta$  in (2).

### C. Energy Constraints at the Sensor:

We consider a battery-aided operation where energy is stored at a battery and used in regular time intervals. Let the initial energy stored at the battery be 0, i.e. the battery is empty. At time slot  $t$ , an energy packet of  $0 \leq E_t < \infty$  arrives to the sensor, where the harvested energy process is an i.i.d. discrete-time stochastic process with mean  $\mu_E$  and variance  $\varrho_E$  and  $E_t \leq E_u < \infty$ , where  $0 < E_u < \infty$  denotes the maximum value of the energy packets. At the end of frame  $k$ , total energy that has arrived to the battery during frame  $k$  is given by  $\bar{E}_k$  as follows

$$\bar{E}_k = \sum_{t=1}^{Q_e} E_{(k-1)Q_e+t}. \quad (3)$$

We assume that the time frames for the data buffer and the battery is synchronized and  $Q_e = Q_d = Q$ . We assume that battery capacity  $C$  satisfies  $C \geq E_u Q$ ,  $\forall Q$  so that a total energy of  $E_u Q$  can be stored in the battery.

In general, the sensor has to operate under energy neutrality conditions:  $\sum_{l=1}^k J_l \leq \sum_{l=1}^k \bar{E}_l$ ,  $k = 1, \dots, N_T$ . These conditions ensure that the energy used at each transmission frame does not exceed available energy. Here, we focus on the case where at each transmission all the energy at the battery is used, i.e.

$$J_k = \bar{E}_k, \quad k = 1, \dots, N_T. \quad (4)$$

Here the left-hand side of (4) depends on the power amplification factor  $p_k$  at transmission frame  $k$  through (2). The right-hand side gives the available energy, i.e. realization of the total energy stored at the battery at the end of transmission frame  $k$ . Performance of linear transmission strategies under such power constraints where available energy is modeled as a deterministic variable have been considered before, see for instance [39–41] for formulations with total energy constraints and [25], [26], [32] for energy harvesting formulations. In this

work, we provide performance guarantees under a stochastic energy arrival model with block transmission.

### D. Estimation at the Fusion Center:

After  $N_T$  transmission frames, i.e. obtaining  $\mathbf{y} = [\bar{\mathbf{y}}_1; \dots; \bar{\mathbf{y}}_{N_T}] \in \mathbb{C}^{N \times 1}$ , the fusion center forms an estimate of  $\mathbf{x}$  using Linear Minimum-Mean Square Error (LMMSE) Estimation. Let us consider a fixed  $E_t$ ,  $t = 1, \dots, N$  realization, where  $p_k$ 's are determined through (4). Hence the LMMSE estimate conditioned on the energy arrivals  $E_t$  can be found as [42, Ch2]

$$\hat{\mathbf{x}} = K_{\mathbf{xy}} K_{\mathbf{y}}^{-1} \mathbf{y}. \quad (5)$$

This is the standard linear mean-square error estimator where the fusion center uses the second-order statistics of the source and noise to form a linear estimate of the unknown variable [42, Ch2]. Let  $\mathbb{E}_S[\cdot]$  denote the statistical expectation with respect to noise and signal statistics, including  $\mathbf{x}, \mathbf{w}$ , but not with respect to energy realizations (and hence not with respect to  $p_k$ 's which are also a function of the energy realizations). Then the mean-square error,  $\varepsilon = \mathbb{E}_S[\|\mathbf{x} - \hat{\mathbf{x}}\|^2]$ , can be expressed as follows [42, Ch2]

$$\varepsilon = \text{tr} \left[ \left( \frac{s}{P_x} I_s + \frac{1}{\sigma_w^2} U_s^\dagger G U_s \right)^{-1} \right], \quad (6)$$

where  $G = \text{diag}(\mathbf{g}) = \text{diag}([p_1 \mathbf{1}_{Q_d}; \dots; p_{N_T} \mathbf{1}_{Q_d}]) \in \mathbb{R}^{N \times N}$ ,  $\mathbf{g} = [g_1; \dots; g_t; \dots; g_N] \in \mathbb{R}^{N \times 1}$  and  $\mathbf{1}_{Q_d} = [1; \dots; 1] \in \mathbb{R}^{Q_d}$  is the vector of ones. Hence  $\mathbf{y} = G^{1/2} \mathbf{x} + \mathbf{w}$ . We note that the possible additional distortion due to the dynamic range limiter for  $x_t$ 's with unbounded support is omitted here. A study of this aspect in the context of estimation under energy harvesting constraints can be found in [24]. For the above standard LMMSE estimation,  $p_k$ 's are assumed to be known at the fusion center as in [25], [26]. Determination of  $p_k$ 's can be seen as a part of the channel estimation process in the communication link between the sensor and the fusion center. We note that here we focus on the reconstruction of the unknown field and the energy cost of this channel estimation operation is not accounted for in our work.

Here  $G$  is a random vector due to random energy arrivals; hence our setting is different from the offline scheme where the performance is evaluated under known energy values. Furthermore, calculation of the mean-square error in (6), in general, requires evaluation of a matrix inverse as opposed to a direct sum of utility functions over time, such as in the case of throughput based formulations. Our block transmission scheme provides a possibly sub-optimal but low-complexity strategy for this correlated signal setting.

## III. PERFORMANCE BOUNDS

Let us define

$$f_{bt}(\mu, \varrho, r) \triangleq 2s \exp \left( -\frac{\varrho}{\mu^2} h \left( \frac{\mu r}{\varrho} \right) \right) \quad (7)$$

$$f_{bn}(\mu, \varrho, r) \triangleq 2s \exp \left( -\frac{r^2/2}{\mu r/3 + \varrho} \right) \quad (8)$$

with  $h(a) \triangleq (1+a) \ln(1+a) - a$ ,  $a \geq 0$ .

We now present our main result, i.e. guarantees on the error performance that hold with high probability:

**Theorem 3.1:** Let  $u_i \in \mathbb{C}^{s \times 1}$  denote the  $i^{th}$  column of the matrix  $U_s^\dagger$ . Let  $\eta_L = \min_i \|u_i\|^2$ , and  $\eta_U = \max_i \|u_i\|^2$ . Let  $E_u$  be parametrized as  $E_u = r_E \mu_E$ ,  $r_E \geq 1$ . Performance of the EH system satisfies the following bounds

I.

$$\mathbb{P}(\varepsilon < \varepsilon_I) \geq 1 - f_{bt}(\mu_I, \varrho_I, r) \geq 1 - f_{bn}(\mu_I, \varrho_I, r) \quad (9)$$

for  $r \in (0, \frac{1}{\eta_U}]$ , where

$$\varepsilon_I = \frac{1}{1 + \frac{1}{\sigma_w^2} \mu_E (\frac{1}{\eta_U} - r)} P_x \quad (10)$$

$$\mu_I = \frac{1}{\eta_L} \max\{r_E - 1, 1\} \min\{Q\eta_U, 1\} \quad (11)$$

$$\varrho_I = \frac{\varrho_E}{\mu_E^2} \frac{1}{\eta_L^2} \frac{1}{Q} \min\{Q\eta_U, 1\} \quad (12)$$

II.

$$\mathbb{P}(\varepsilon < \varepsilon_{II}) \geq 1 - f_{bt}(\mu_{II}, \varrho_{II}, r) \geq 1 - f_{bn}(\mu_{II}, \varrho_{II}, r) \quad (13)$$

for  $r \in (0, \frac{1}{\eta_U}]$ ,  $\gamma \in [0, Qr_E]$ , where

$$\varepsilon_{II} = \frac{1}{1 + \frac{1}{\sigma_w^2} \frac{1}{Q} \bar{p} \gamma \mu_E (\frac{1}{\eta_U} - r)} P_x \quad (14)$$

$$\mu_{II} = \frac{1}{\eta_L} \max\{\frac{1}{\bar{p}} - 1, 1\} \min\{Q\eta_U, 1\} \quad (15)$$

$$\varrho_{II} = \frac{1}{\eta_L^2} (\frac{1}{\bar{p}} - 1) \min\{Q\eta_U, 1\} \quad (16)$$

$$\bar{p} = \mathbb{P}(\bar{E}_k \geq \gamma \mu_E) \quad (17)$$

**Proof:** The proof is presented in Section VIII.

Both Bound I in (9)-(12) and Bound II in (13)-(17) provide performance guarantees (i.e. upper bounds) for the mean-square error performance. For instance, Bound I states the following: The mean-square error  $\varepsilon$  is guaranteed to be lower than  $\varepsilon_I$  with probability greater than  $1 - f_{bt}(\mu_I, \varrho_I, r)$ , where the parameters  $\varepsilon_I, \mu_I, \varrho_I, r$  are related through (10)-(12). Bound II has a similar form with the parameters  $\varepsilon_{II}, \mu_{II}, \varrho_{II}, r$ . We note that whether Bound I or Bound II is tighter depends on the system parameters. This aspect is illustrated in Fig. 3 and Fig. 4 in Section VI. Bound I and Bound II can be seen as performance guarantees based on average number of samples that can be transmitted. These discussions, together with connections with compressive sensing, are provided in Section V.

**Remark 3.1:** For  $Q = 1$ , energy  $E_t$  that arrives to the sensor at time  $t$  is immediately used to send the sample  $x_t$ . As the buffer size  $Q > 1$  gets larger, the probability of sending the samples in the buffer (with non-zero power) increases since the probability of the battery being charged with nonzero energy also increases while waiting for the data buffer to be full. On the other hand, the power used to send each sample may be lower compared to the case where the energy is used to send a fewer number of samples, for instance compared to

the scenario of directly sending the sample  $x_t$  with energy  $E_t$  if an energy packet of  $E_t > 0$  arrives ( $Q = 1$ ). Hence, the bounds presented here can be interpreted as an exploration of the trade-off between using a small number of samples with high signal-to-noise ratio (SNR), i.e. high power, and a high number of samples with low SNR in the estimation process.

**Remark 3.2:** Energy allocations which are as uniform as possible, or alternatively as balanced as possible are optimal for white sources [17], [19], [30]. In our formulation, the data buffer and energy buffer allows us to mimic these uniform-like allocations, where larger buffer sizes allow energy allocations that are more uniform over the whole time duration of interest, i.e.  $1 \leq t \leq N$ . Hence, varying  $Q$  values allows us to study the effect of different buffer sizes, or equivalently the effect of balanced energy allocations for signals that are not necessarily exactly white. Note that we assume the battery capacity is large enough so that  $C \geq E_u Q$ ,  $\forall Q$ . Hence, the observations here are in comparison to the maximum size of the energy packet that can arrive at each time slot. In particular, a large  $Q$  value means the device has a large enough battery so that  $Q$  of the energy packets can be stored.

We note that the block transmission scheme considered in this work is possibly sub-optimal in the sense that there may be other transmission strategies that use only statistical knowledge of future energy arrivals, but can guarantee smaller error values for a given fixed probability.

#### A. Comparison with the average performance

For comparison purposes, we now present a lower bound on the average error performance over different realizations of the energy arrival process  $E_t$ .

**Lemma 3.1:** The following lower bound on the average error holds with  $\eta_L = \min_i \|u_i\|^2$

$$\mathbb{E}_E[\varepsilon] \geq \frac{1}{1 + \frac{1}{\sigma_w^2} \mu_E \frac{1}{\eta_L}} P_x. \quad (18)$$

The proof is provided in Section IX. We note that this bound does not depend on  $Q$ .

**Remark 3.3:** Comparing (18) and the error expressions in Thm. (3.1) we observe that both expressions provide error expressions in the form  $\frac{1}{1 + \text{SNR}_{\text{eff}}} P_x$  where  $\text{SNR}_{\text{eff}}$  takes the form  $\text{SNR}_{\text{eff}}^D = \frac{1}{\sigma_w^2} \frac{1}{\eta_L} \mu_E$  for (18) and it takes the form  $\text{SNR}_{\text{eff}}^P = \frac{1}{\sigma_w^2} \mu_E (\frac{1}{\eta_U} - r)$ , for instance, for (10). Hence the error expressions in Thm. (3.1) provide different operating points for how close one can operate to (18) and with which probability through the variable  $r$ .

#### B. Comparison with the off-line scheme with a total energy constraint

As a benchmark for our bounds in Thm. 3.1, we now consider an associated off-line scheme [30]. In particular, we consider the case where amplification factors are not modeled as random variables that depend on the energy arrivals but deterministic variables to be optimized. Let us consider the case where each component  $x_t$  is sent as follows:

$$y_t = \sqrt{b_t} x_t + w_t, \quad t = 1, \dots, N. \quad (19)$$

Here we introduced the notation  $b_t \geq 0$  to denote amplification factors to emphasize that these are modeled as deterministic variables as opposed to random variables. In contrast to the setting of Section II and hence the setting of Thm. 3.1, here a block transmission constraint is not imposed onto the set of admissible sensor strategies (hence  $Q = 1$ ). Let us denote the error as follows

$$\bar{\varepsilon}(B) = \text{tr} \left[ \left( \frac{s}{P_x} I_s + \frac{1}{\sigma_w^2} U_s^\dagger B U_s \right)^{-1} \right], \quad (20)$$

where  $b_t \geq 0, \forall t$  and  $B = \text{diag}(\mathbf{b}) = \text{diag}([b_1; \dots; b_N]) \in \mathbb{R}^{N \times N}$ . We consider the following optimization problem

$$\varepsilon_d = \min_B \bar{\varepsilon}(B) \quad \text{s.t.} \quad \sum_{l=1}^N b_l \sigma_{x_l}^2 = E_{\text{tot}}. \quad (21)$$

In this deterministic scheme, the sensor has a total energy of  $E_{\text{tot}}$  and it can freely distribute this energy on the samples in order to minimize the error. We note the following result:

**Lemma 3.2:** [30, Lemma 3.5] *An optimal strategy for (21) is given by uniform  $b_t$  with  $b_t = E_{\text{tot}}/P_x, \forall t$ . The optimum value is given by  $\varepsilon_d = \frac{1}{1 + \frac{1}{\sigma_w^2} \mu_E \frac{N}{s}} P_x$ .*

Consider the case with  $E_{\text{tot}} = \mu_E N$ , which is the total energy that would have been obtained if an energy packet of  $\mu_E$  were harvested at each time slot. We note that  $\mu_E$  is the mean of the energy arrival process. In this sense, Lemma 3.2 may be used as a deterministic benchmark. Hence, the benchmark becomes

$$\varepsilon_d = \frac{1}{1 + \frac{1}{\sigma_w^2} \mu_E \frac{N}{s}} P_x. \quad (22)$$

Similar to the lower bound of (18), (22) is a benchmark for Thm. 3.1 in terms of how close one can operate to this value. We note that, in general,  $\frac{1}{\eta_L} \neq \frac{N}{s}$ , hence (18) and (22) provide different benchmarks.

#### IV. CIRCULARLY WIDE-SENSE STATIONARY SIGNALS

In this section, we specialize to the case of circularly wide-sense stationary signals, which constitute a finite dimensional analog of wide-sense stationary signals [35], [36]. Covariance matrices of c.w.s.s. signals are circulant by definition, i.e. covariance matrix of a c.w.s.s. signal is determined by its first row as  $[K_{\mathbf{x}}]_{tk} = [K_1]_{\text{mod}_N(k-t)}$ , where  $K_1 \in \mathbb{C}^{1 \times N}$  is the first row of  $K_{\mathbf{x}}$  [35], [36]. The unitary matrix  $U$  in the EVD of covariance matrices of c.w.s.s. signals is given by the Discrete Fourier Transform (DFT) matrix [35], [36]. Let  $F^N$  denote the DFT matrix of size  $N \times N$ , i.e.  $[F^N]_{tk} = (1/\sqrt{N}) \exp(-j \frac{2\pi}{N} (t-1)(k-1))$ ,  $1 \leq t, k \leq N$ , where  $j = \sqrt{-1}$ . Hence, the reduced EVD of  $K_{\mathbf{x}}$  is given by  $K_{\mathbf{x}} = F_{\Omega}^N \Lambda_{x,s} F_{\Omega}^{N\dagger}$ , where  $\Lambda_{x,s} = \text{diag}(\lambda_k) = \frac{P_x}{s} I_s \in \mathbb{R}^{s \times s}$  and  $F_{\Omega}^N \in \mathbb{C}^{N \times s}$  is the matrix that consists of  $s$  columns of  $F^N$  corresponding to the non-zero eigenvalues. Due to the circulant covariance matrix structure, the variances of the components of a c.w.s.s. signal satisfy  $\sigma_{x_t}^2 = \sigma_x^2 = P_x/N, \forall t$ .

Hence,  $J_k = \sum_{t=1}^{Q_d} p_k \sigma_{x_{(k-1)Q_d+t}}^2 = p_k Q P_x / N$ , and by (3), (4), we have the following

$$p_k = \frac{N}{P_x} \frac{1}{Q} \sum_{t=1}^Q E_{(k-1)Q+t}. \quad (23)$$

For c.w.s.s. signals, we have  $\eta_L = \min_i \|u_i\|^2 = \frac{s}{N}$ , and  $\eta_U = \max_i \|u_i\|^2 = \frac{s}{N}$  due to the DFT matrix. Hence, (10)-(12) can be expressed as

$$\varepsilon_I = \frac{1}{1 + \frac{1}{\sigma_w^2} \mu_E \frac{N}{s} (1 - \tilde{r})} P_x, \quad (24)$$

$$\mu_I = \max\{r_E - 1, 1\} \min\{Q \frac{s}{N}, 1\}, \quad (25)$$

$$\varrho_I = \frac{\varrho_E}{\mu_E^2} \frac{1}{Q} \min\{Q \frac{s}{N}, 1\}, \quad (26)$$

where  $\tilde{r} \in (0, 1]$ . Here we have scaled  $r, \mu_I, \varrho_I$  while going from Eqn. (10)-(12) to Eqn. (24)-(26), since  $f_{bt}(\cdot)$  and  $f_{bn}(\cdot)$  only depend on the ratios between  $r, \mu_I, \varrho_I$ . Eqn. (14)-(16) can be specialized to the case of c.w.s.s. signals, similarly. We note that these bounds also hold for other signal families for which  $\|u_i\|^2$  is constant for all  $i$ , such as unitary Hadamard matrices.

##### A. Equidistant sampling of low-pass c.w.s.s. signals

We now focus on the case of low-pass c.w.s.s. signals, i.e. c.w.s.s. signals for which eigenvalues of  $K_{\mathbf{x}}$  that correspond to the low frequency components indexed by  $\Omega = \{0, 1, \dots, s-1\}$  are possibly non-zero and the rest of the eigenvalues are zero. Hence, for c.w.s.s. signals only eigenvalues that are possibly non-zero are the ones associated with the frequencies  $\exp(-j \frac{2\pi}{N} r)$ ,  $r = 0, \dots, s-1$ . Such signals can be recovered from their uniformly taken samples with zero mean-square error when the total number of (complex-valued) samples is larger than the number of non-zero eigenvalues [23]. This property, which is consistent with the deterministic sampling theorems and the sampling theorems for wide-sense stationary signals [36], [43], motivates us to study strategies that send equidistant samples under our EH framework. In particular, we are interested in understanding which of the following is a better strategy: i) sending all of the observations of the sensor with as equal energy as possible as suggested by the energy harvesting literature; or ii) sending only the equidistant samples as suggested by the sampling theorems. The block transmission scheme of Section II provides a low-complexity approach for implementing strategies similar to (i). We study the strategy in (ii) below.

In particular, we consider strategies that send one sample out of every  $Q = N/s$  samples as follows: Let  $t_d \in \{0, \dots, Q-1\}$  be the fixed initial delay before sending the first sample and  $N_T = N/Q \in \mathbb{Z}$  be the number of transmissions as before. Recall that  $y_t = \sqrt{g_t} x_t + w_t$ . Hence, under uniform sampling we have  $g_t \geq 0$ , if  $t = Q(k-1) + t_d + 1$ ,  $1 \leq k \leq N_T$ , and  $g_t = 0$  otherwise. Hence, the received signal at transmission frame  $k$  is the single sample  $x_{Q(k-1)+t_d+1}$  as follows

$$y_k = \sqrt{p_k} x_{Q(k-1)+t_d+1} + w_k, \quad (27)$$

where  $\sqrt{p_k}$  denotes the amplification factor,  $w_k \in \mathbb{C}$ ,  $\mathbb{E}[w_k w_k^\dagger] = \sigma_w^2$  denotes the i.i.d. complex proper zero-mean channel noise, as before. Energy used by the sensor for communications at transmission frame  $k$  can be written as follows:

$$J_k = p_k \sigma_{x(k-1)Q_d+t_d+1}^2 = p_k \frac{P_x}{N}, \quad (28)$$

where we have used the fact that for c.w.s.s. signals  $\sigma_{x_t}^2 = \sigma_x^2 = P_x/N, \forall t$ . By (4), we again have  $J_k = \bar{E}_k, \forall k$ . At each transmission frame  $k$ , this scheme uses all the energy in the battery to send only one  $x_t$  value and discards all the other samples in the data buffer. We obtain the following bound for the performance of this system:

**Theorem 4.1:** *Performance of the equidistant sample transmission strategy of (27) for low-pass c.w.s.s. signals satisfies the following bound with  $r \in (0, 1)$*

$$\mathbb{P}(\varepsilon < \varepsilon_I^u) \geq 1 - f_{bt}(\mu_I^u, \varrho_I^u, r^u) \geq 1 - f_{bn}(\mu_I^u, \varrho_I^u, r^u) \quad (29)$$

$$\varepsilon_I^u = \frac{1}{1 + \frac{1}{\sigma_w^2} \mu_E \frac{N}{s} (1-r)} P_x \quad (30)$$

$$\mu_I^u = \max\{r_E - 1, 1\} \quad (31)$$

$$\varrho_I^u = \frac{\varrho_E}{\mu_E^2} \frac{s}{N}. \quad (32)$$

**Proof:** The proof is presented in Section XI.

Comparing (30)-(32) with (24)-(26) for  $Q = \frac{N}{s}$  reveals that for low-pass c.w.s.s. signals, both the strategy of Thm. 3.1, which spreads the energy accumulated in the battery evenly on the samples in the buffer, and the equidistant sample transmission strategy of Thm. 4.1, which uses the energy only on one sample from the buffer, results in the same performance guarantees. This property is consistent with the performance of the associated strategies in the off-line scenario under a total energy constraint as discussed below:

*Comparison with the off-line scheme under equidistant sample transmission strategy:* Let us consider the equidistant sample transmission scheme under a total energy constraint:

$$\varepsilon_{de} = \min_{B_u} \bar{\varepsilon}(B_u) \quad \text{s.t.} \quad \sum_{l=1}^N b_l \sigma_{x_l}^2 = E_{tot}, \quad (33)$$

under the condition  $b_t \geq 0$ , if  $t = Q(k-1) + t_d + 1, 1 \leq k \leq N_T$ , and  $b_t = 0$  otherwise; and  $Q = N/s$ ,  $B_u = \text{diag}([b_1; \dots; b_N]) \in \mathbb{R}^{N \times N}$ .

**Lemma 4.1:** [30, Corollary 3.3] *An optimal strategy for (33) is given by  $b_t = \frac{E_{tot}}{P_x} \frac{N}{s}$ , if  $t = Q(k-1) + t_d + 1, 1 \leq k \leq N_T$  and  $b_t = 0$  otherwise. The optimum value is given by  $\varepsilon_{de} = \frac{1}{1 + \frac{1}{\sigma_w^2} \frac{E_{tot}}{s}} P_x$ .*

Hence, under the off-line scheme with a total energy constraint, performance of the uniform power allocation over all  $x_t$ 's, which is given by Lemma 3.2, and performance of the equidistant sample transmission strategy given by Lemma 4.1 are the same. In this work, we have shown that performance bounds in the online case for block transmission scheme of Thm. 3.1 (specialized to c.w.s.s. signals in (24)-(26)) and the

performance bounds for the equidistant sample transmission scheme of Thm. 4.1 are also the same. These two set of results together suggest flexibility in energy allocation for estimation of low-pass c.w.s.s signals in energy harvesting systems. Nevertheless, we note that Thm. 3.1 and Thm. 4.1 provide upper bounds, i.e. guarantees for signal recovery with a given error with a given probability. Hence, insights and guidelines derived from these results should take this point into consideration.

## V. CONNECTIONS TO COMPRESSIVE SENSING

The scenario of  $Q = 1$  in (1) is closely related to the classical compressive sensing setting. In particular, consider the case where the energy arrival process can be modeled as an i.i.d. Bernoulli random process. A typical compressive sensing set-up is the scenario where the measurement process is modeled as an i.i.d. Bernoulli process where a measurement is made, for instance, when the Bernoulli random variable is 1 and is dropped when the Bernoulli random variable is 0. Hence for  $Q = 1$ , the bounds presented here are closely related to the eigenvalue bounds provided in compressive sensing literature [44, Ch.12]. In particular, consider the scenario of  $Q = 1$  with static  $\sigma_{x_t}^2 = \sigma_x^2$ , (such as in the case of circularly wide-sense stationary signals) and Bernoulli energy arrivals. Then, the bounds in Thm. 3.1 can be seen as a consequence of the eigenvalue bounds in the CS literature, see for instance [44, Ch.12], [45, Thm. 1.2], [23]. For  $Q > 1$  or non-uniform  $\sigma_{x_t}^2$ , Thm. 3.1 provides a set of novel eigenvalue bounds for the formulation introduced in Section II.

To further elaborate on connections to compressive sensing, we now focus on the scenario where there is no noise on the channel, i.e.  $\sigma_w^2 = 0$ . Hence, the system model becomes

$$\mathbf{y} = G^{1/2} \mathbf{x}, \quad (34)$$

where  $G$  is the diagonal matrix of amplification factors as defined in Section II-D. Let us consider the following question: "For which energy arrival rates, sparsity levels and queue sizes, can we recover  $\mathbf{x}$  from the observations  $\mathbf{y}$  with zero mean-square error (with high probability)?"

Recall that,  $K_{\mathbf{x}} = U_s \Lambda_{x,s} U_s^\dagger$ , hence  $\mathbf{x} \in \mathbb{C}^{N \times 1}$  belongs to a signal family of low degree of freedom  $\mathbf{x} = U_s \bar{\mathbf{x}}$  where  $\bar{\mathbf{x}} \in \mathbb{C}^{s \times 1}$  and the covariance matrix  $K_{\bar{\mathbf{x}}} = \Lambda_{x,s}$ . Hence, we have a setting that is similar to typical compressive sensing set-ups. Nevertheless, note that in typical CS scenarios, support of the signal is not known during the signal recovery whereas here we consider a scenario where the support is known. Note that locations of the measurements (i.e. which rows of  $G$  are non-zero) are modelled as random both in our setting and in compressive sensing scenarios. Thm. 3.1 has the following corollary:

**Corollary 5.1:** *Fix  $N$ . Consider the energy arrival process  $E_t$  with  $E_t = \kappa_t E_b$ ,  $\kappa_t \sim \text{Bernoulli}(p)$ ;  $E_b > 0$ . Let  $\sigma_w^2 = 0$ ,  $0 \leq p \leq 1/2$ ,  $Q \leq N/s$ ,  $\bar{p} = 1 - (1-p)^Q$ ,  $\bar{p} \leq 1/2$ ,  $\delta \in (0, 1]$  and  $U$  be the DFT matrix. Suppose that at least one of the following conditions is satisfied*

$$C_{e,I} \times (Q/3 + 1) \times s \times \ln(2s/\delta) \leq N \times p \quad (35)$$

$$C_{e,II} \times Q \times s \times \ln(2s/\delta) \leq N \times \bar{p} \quad (36)$$



where  $C_{e,I} > 2$ , and  $C_{e,II} > 8/3$ , are numerical constants. Then, the mean-square error  $\varepsilon$  is zero with probability at least  $1 - \delta$ .

The proof is provided in Section X. In the above,  $C_{e,I} \approx 2$  and  $C_{e,II} \approx 8/3$ , please see Section X for details. We note that the right hand side of (35)-(36) increases with increasing energy arrival success rate and can be interpreted as the average number of transmitted samples. Hence, (35)-(36) asserts that if  $s$  is small enough and the energy arrival rate is high enough, then the signal can be recovered from its samples with high probability. This is analogous to the compressive sensing results where sufficient number of measurements for recovery of sparse signals are presented. In particular, consider the following sufficient condition from [45, Thm. 1.1]

$$C_c \times s \times \ln(N/\delta) \leq M \quad (37)$$

and  $C'_c \ln^2(N/\delta) \leq M$  where  $C_c$  and  $C'_c$  are fixed numerical constants and  $M$  is the number of measurements whose locations (i.e. which rows of  $G$  are non-zero) are chosen randomly. If (37) holds, then with probability at least  $1 - \delta$ , an arbitrary signal (with random signs) with support of size  $s$  can be recovered from randomly selected  $M$  measurements [45, Thm. 1.1]. Note that  $Q = 1$  here. Comparing (35)-(36) and (37), we observe that both conditions give the (average) number of observations that guarantee signal recoverability.

An important step in the derivation of compressive sensing results is the derivation of eigenvalue bounds. In particular, consider the following type of sufficient condition [45, Thm. 1.2], [44, Thm. 12.12],

$$C''_c \times s \times \ln(2s/\delta) \leq M, \quad (38)$$

where  $C''_c > 8/3$  is a numerical constant and  $M$  is, again, the number of measurements. If (38) holds, the matrix  $U_s^\dagger G U_s$  is invertible with probability at least  $1 - \delta$ . Hence, the mean-square error will be zero in (34). Note that (38) is derived under the assumption that support is fixed and known, as in our set-up. As discussed in the beginning of this section, for  $Q = 1$ , (35)-(36) and (38) are the same.

Different from (38), our results in (35)-(36) reveal how the eigenvalue bounds depend on the queue size parameter  $Q$ , which is included in the system formulation due to the energy harvesting aspect. We note that (35), which was derived from Bound I, suggests smaller buffer lengths are preferable (in the sense that for fixed  $s, p, N$  values, larger  $Q$  values will not satisfy (35)). On the other hand,  $\bar{p}$  that appears on the right hand side of (36) also depends on  $Q$ . Note that (36) was derived from Bound II. Whether (36) (and Bound II) favors smaller or larger  $Q$  values depends on the system parameters. Further investigation of this point is provided in Section VI.

#### A. Discussions

In a wide range of sensing applications, there exist unknown physical quantities that we would like to estimate, such as temperature values in a smart building application or flow rates in an industrial application. Typically, sensors make measurements of these parameters and these measurements are collected at a remote central decision center wirelessly.

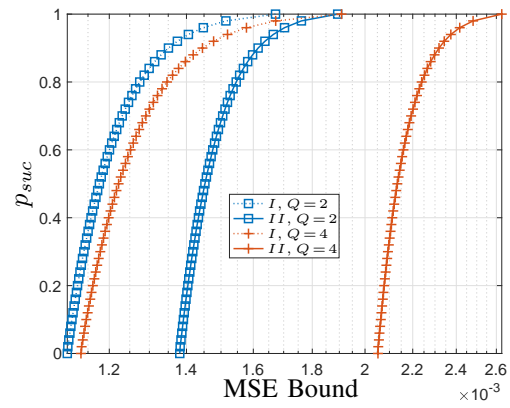


Fig. 3:  $p_{suc}$  versus Bound I and Bound II, Bernoulli energy arrivals,  $s = 4$ ,  $p = 0.5$ .

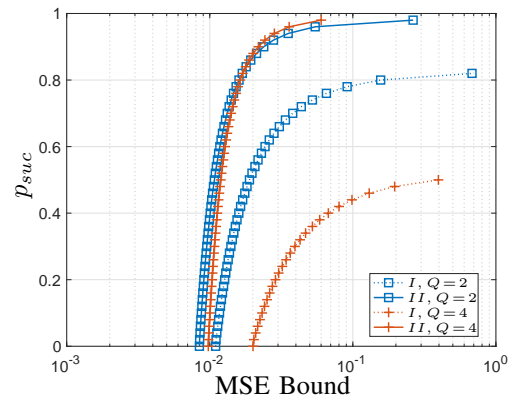


Fig. 4:  $p_{suc}$  versus Bound I and Bound II, Bernoulli energy arrivals,  $s = 8$ ,  $p = 0.3$ .

Practical sensor deployments that utilize energy harvesting sensors for such remote estimation tasks have started to emerge for various applications, including oil and gas industries, consumer electronics, chemical processing, steel manufacturing [6–8]. Nevertheless, fundamental performance limits of these systems from an estimation theoretic framework are not fully investigated. Our work here contributes to this aspect by studying such a remote estimation problem under a LMMSE framework.

Compressive sensing based approaches provide us an attractive set of tools for investigating these remote estimation problems. In particular, concept of sparsity allows us to study a large class of signals including correlated signals. Moreover, the tools developed for studying the effect of random measurements in compressive sensing literature provide promising candidates for studying the unreliable nature of available energy in EH systems, as illustrated in this work. Hence, we believe that compressive sensing based approaches will be instrumental to study fundamental sensing trade-offs for future EH sensing systems.

## VI. NUMERICAL RESULTS

We now illustrate our bounds by presenting the trade-offs between the guaranteed MSE and the probability of obtaining

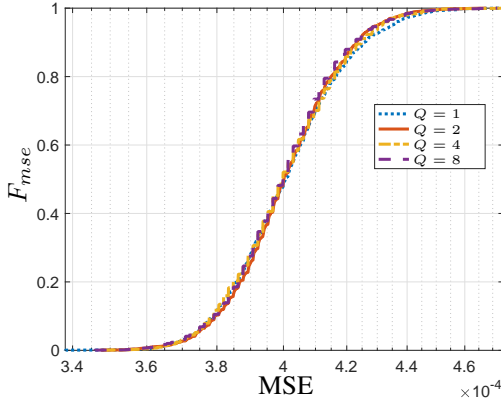


Fig. 5: Empirical CDF of MSE ( $F_{mse}$ ), Bernoulli energy arrivals,  $s = 2$ ,  $p = 0.5$ .

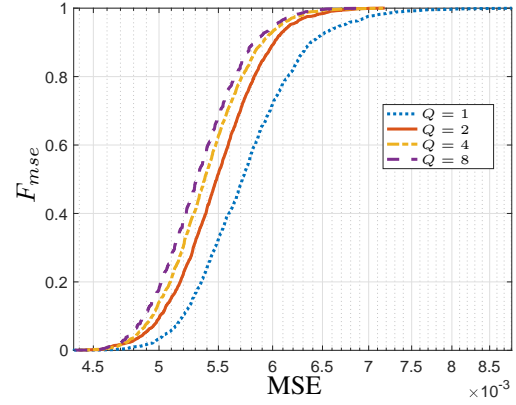


Fig. 6: Empirical CDF of MSE ( $F_{mse}$ ), Bernoulli energy arrivals,  $s = 16$ ,  $p = 0.3$ .

that MSE. The horizontal axis corresponds to the error bound as provided by  $\varepsilon_I$  or  $\varepsilon_{II}$  and the vertical axis corresponds to the probability on the right-hand side of (9)/(13), which is referred as  $p_{suc}$ . We use  $f_{bt}(\cdot)$  for the calculation of  $p_{suc}$ . Hence, the horizontal axis value of each point on the plot shows a particular error value and the vertical axis value shows the probability with which we can guarantee the MSE to be smaller than that particular error value. We normalize the error bounds with the total uncertainty in the signal, i.e. we report  $\varepsilon_I/P_x$  and  $\varepsilon_{II}/P_x$ . Unless otherwise stated, we consider the energy arrival process  $E_t$  i.i.d. with  $E_t = \kappa_t E_b$ ,  $\kappa_t \sim \text{Bernoulli}(p)$ ;  $E_b = 1$  and  $U$  is the DFT matrix. Hence, the benchmark of (18)/(22) is equal to  $\varepsilon_d = \frac{1}{1 + \frac{1}{\sigma_w^2} p \frac{N}{s}} P_x$ .

*Comparison of Bound I and Bound II:* We note that both Bound I and Bound II are upper bounds. Which bound is tighter (i.e. which bound guarantees a given error value with the highest probability) depends on the system parameters. We now illustrate this point. Let  $N = 256$ ,  $P_x = N$ ,  $\sigma_w^2 = 10^{-4} P_x$ . Both Bound I and Bound II are presented in Fig. 3 and Fig. 4, for  $s = 4$ ,  $p = 0.5$  and  $s = 8$ ,  $p = 0.3$ , respectively. In Fig. 3, Bound I is tighter whereas in Fig. 4 Bound II is tighter. This behaviour is consistent with our other numerical investigations where Bound I is observed to be typically tighter for small  $s/N$  ratios and high energy arrival rates and vice-versa for Bound II. For instance, if we decrease  $p$  value for Fig. 4, Bound I can no longer provide a guarantees with non-zero probability. In the rest of this section, while plotting the bounds, for a given probability value  $p_{suc}$  we present the tightest of Bound I and Bound II, i.e. the bound that guarantees a given error value with the highest probability.

*Comparison with Empirical Performance:* For comparison purposes, we first present the empirical cumulative distribution function (CDF) of the mean-square error in (6). In these experiments, we fix the support (i.e. locations of the non-zero eigenvalues) and look at the empirical CDF of the mean-square error with random energy arrivals over  $N_{sim} = 2000$  realizations. Let  $N = 512$ ,  $P_x = N$ ,  $\sigma_w^2 = 10^{-4} P_x$ . Empirical CDF values are presented in Fig. 5 and Fig. 6 for  $s = 2$ ,  $p = 0.5$  and  $s = 8$ ,  $p = 0.3$  respectively. The corresponding bounds are presented in Fig. 7 and Fig. 8. The benchmark of

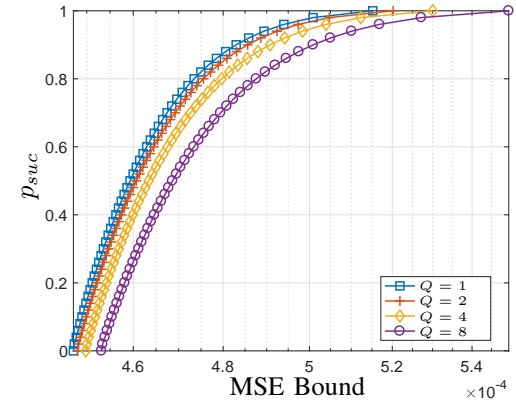


Fig. 7:  $p_{suc}$  versus MSE Bound, Bernoulli energy arrivals,  $s = 2$ ,  $p = 0.5$ .

(18)/(22) is  $\varepsilon_d \approx 4 \times 10^{-4}$  is for Fig. 5 and  $\varepsilon_d \approx 5.3 \times 10^{-3}$  is for Fig. 6. Comparing Fig. 5 and Fig. 7, we observe that both the empirical results and the bounds show that we will operate close to these benchmarks with high probability for this scenario, where  $s/N$  is small and energy arrival rate  $p$  is high. On the other hand, the gap between the bounds and the empirical results, and also the gap between the empirical results and the benchmarks of (18)/(22) are larger for Fig. 6 and Fig. 8 where the ratio  $s/N$  is larger and  $p$  is smaller. These observations are consistent with compressive sensing literature where signal recovery guarantees are provided only for sparse signals (low  $s/N$  values).

*Effect of System Parameters on Performance Guarantees:* In both Fig. 7 and Fig. 8, as the target performance becomes more demanding, i.e. the error value decreases, the probability that this error can be guaranteed becomes smaller. When the degree of freedom of the signal is sufficiently low ( $s=2$ , Fig. 7), the performance bound is observed to be relatively insensitive to the buffer size. On the other hand, when the degree of freedom is higher and energy arrival rate is smaller, ( $s=16$ , Fig. 8) the bound becomes more sensitive to the buffer size. For  $s=16$ , with  $Q=1$ , the bound cannot provide any guarantees that hold with probability higher than 0.9; whereas with higher buffer sizes, relatively small values of error can be guaranteed with

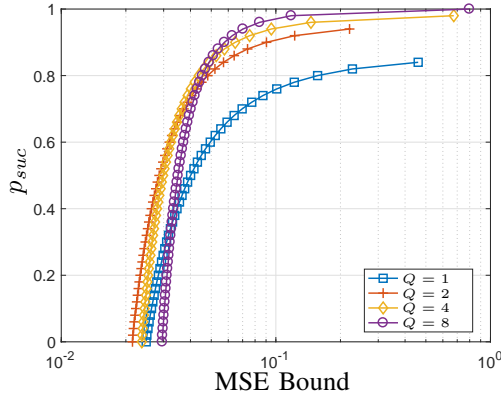


Fig. 8:  $p_{suc}$  versus MSE Bound, Bernoulli energy arrivals,  $s = 16$ ,  $p = 0.3$ .

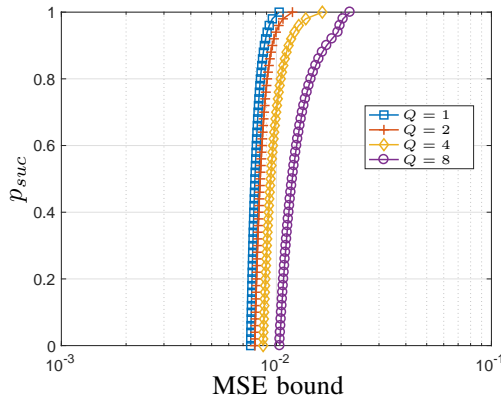


Fig. 9:  $p_{suc}$  versus MSE Bound, uniformly distributed energy packets,  $s = 16$ .

high probability (for instance with probability higher than 0.9). We observe that as  $s$  becomes larger, signal can be said to be more close to a white source, with the limiting case of  $s = N$  corresponding to an exactly white source. Hence, these results are consistent with the results of [17], [19], [30] discussed in Remark 3.2.

We now consider the scenario with  $E_t$  i.i.d. with uniformly distributed energy packets, i.e.  $E_t \sim \text{Uniform}[0, E_u]$ ,  $E_u = 0.6$  in Fig. 9. Here  $E_u$  is chosen so that the uniform arrival case here has the same average energy with the Bernoulli arrival scenario of Fig. 8. Comparing Fig. 8 and Fig. 9, we observe that in the uniform energy arrival case of Fig. 9, long buffer lengths do not offer performance gains as they provide in the Bernoulli energy arrival case of Fig. 8. This is consistent with the fact that in the case of uniformly distributed arrivals the variance of energy packets is smaller and the need to spread the energy over samples by the use of a buffer is expected to be less prominent compared to Bernoulli energy arrival case.

## VII. CONCLUSIONS

We have considered remote estimation of an unknown field with an EH sensor with a limited data and energy buffer. In contrast to much of the existing work, we have focused on a correlated signal model. We have provided structural results

in terms of performance guarantees on the achievable distortion under random energy arrivals with a block transmission scheme. Our performance guarantees provide insights into the trade-offs between the size of buffers, statistical properties of the energy arrival process, degree of freedom of the signal and the achievable distortion. These results also have the advantage that their calculation requires only knowledge of the mean, variance and finite support about the energy arrival process, whose exact probability distribution can be difficult to reliably estimate in practice. Generalizations of our approach into settings that allow energy saving between transmission blocks and applications to fading environments are considered as important future research directions.

## VIII. PROOF OF THM. 3.1

The proof relies on the Matrix Bernstein Inequality, a fundamental random matrix theory tool used in compressive sensing [44, Ch.8]. We first prove the first family of bounds indexed by I in (9)-(12). We first note that

$$\varepsilon = \sum_{i=1}^s \frac{1}{\lambda_i \left( \frac{s}{P_x} I_s + \frac{1}{\sigma_w^2} U_s^\dagger G U_s \right)}, \quad (39)$$

$$\leq \frac{1}{1 + \frac{1}{\sigma_w^2} \frac{P_x}{s} \lambda_{\min}(U_s^\dagger G U_s)} P_x. \quad (40)$$

In the remaining of the section, we let  $k = 1, \dots, N_T$ ,  $t = 1, \dots, Q$  and use the indexing  $z_{k,t} = z_{(k-1)Q+t}$  for any variable  $z_i$ ,  $i = 1, \dots, N$ . Let

$$S_k \triangleq \sum_{t=1}^Q \sigma_{x_{k,t}}^2 \quad (41)$$

where  $\sigma_{x_{k,t}}^2 = \sigma_{x_{(k-1)Q+t}}^2$ . By (4) and (41), we have

$$p_k = \frac{1}{S_k} \bar{E}_k = \frac{1}{S_k} \sum_{l=1}^Q E_{k,l}, \quad (42)$$

where  $E_{k,l} = E_{(k-1)Q+l}$ . Let  $u_i \in \mathbb{C}^{s \times 1}$  denote the  $i^{\text{th}}$  column of the matrix  $U_s^\dagger$ . Let  $Y_{k,t} \triangleq u_{k,t} u_{k,t}^\dagger \in \mathbb{C}^{s \times s}$ , with  $u_{k,t} = u_{(k-1)Q+t}$ . Let us consider

$$\bar{p}_k \triangleq p_k - \mathbb{E}[p_k], \quad (43)$$

$$W_k \triangleq \sum_{t=1}^Q Y_{k,t} \quad (44)$$

$$Z_k \triangleq \bar{p}_k W_k \quad (45)$$

Hence

$$\sum_{k=1}^{N_T} Z_k = \sum_{k=1}^{N_T} p_k \sum_{t=1}^Q Y_{k,t} - \sum_{k=1}^{N_T} \mathbb{E}[p_k] \sum_{t=1}^Q Y_{k,t}, \quad (46)$$

$$= U_s^\dagger G U_s - U_s^\dagger \bar{G} U_s, \quad (47)$$

where  $G = \text{diag}([p_1 \mathbf{1}_{Q_d}, \dots, p_{N_T} \mathbf{1}_{Q_d}]) \in \mathbb{R}^{N \times N}$ ,  $\bar{G} = \text{diag}([\mathbb{E}[p_1] \mathbf{1}_Q, \dots, \mathbb{E}[p_{N_T}] \mathbf{1}_Q]) \in \mathbb{R}^{N \times N}$  and  $\mathbf{1}_Q = [1, \dots, 1] \in \mathbb{R}^Q$  is the vector of ones. We will now use the Matrix Bernstein Inequality on  $Z_k$  to find lower bounds for the eigenvalues of the first term in (47). We will then use these in (40) to bound the estimation error.

**Lemma 8.1:** [Matrix Bernstein Inequality [44, Ch.8]] Let  $V_1, \dots, V_M \in \mathbb{C}^{s \times s}$  be independent zero-mean Hermitian random matrices. Assume that  $\|V_l\| \leq \mu_V, \forall l \in \{1, \dots, M\}$  almost surely. Let  $\varrho_V \triangleq \|\sum_{l=1}^M \mathbb{E}[V_l^2]\|$ . Then, for  $t > 0$

$$\mathbb{P}(\|\sum_{l=1}^M V_l\| \geq t) \leq f_{bt}(\mu_V, \varrho_V, t) \leq f_{bn}(\mu_V, \varrho_V, t) \quad (48)$$

with  $f_{bt}(\cdot)$  and  $f_{bn}(\cdot)$  as defined in (7)-(8).

We note that  $Z_k$  in (45) are statistically independent Hermitian random matrices with  $\mathbb{E}[Z_k] = 0$ . We bound the spectral norm of  $Z_k$  as follows

$$\|Z_k\| = \|\bar{p}_k W_k\| \leq \left( \max_k |\bar{p}_k| \right) \|W_k\|. \quad (49)$$

We obtain the following bound for  $\|W_k\|$

$$\|W_k\| = \left\| \sum_{t=1}^Q u_{k,t} u_{k,t}^\dagger \right\|, \quad (50)$$

$$\leq Q \max_{k,t} \|u_{k,t} u_{k,t}^\dagger\|, \quad (51)$$

$$= Q \max_{k,t} \|u_{k,t}\|^2, \quad (52)$$

$$= Q\eta_U, \quad (53)$$

where

$$\eta_U \triangleq \max_{k,t} \|u_{k,t}\|^2. \quad (54)$$

We also have the following

$$\|W_k\| \leq \left\| \sum_{k=1}^{N_T} W_k \right\| = \|I_s\| = 1, \quad (55)$$

where (55) follows from the fact that for  $A \succeq 0$  and  $B \succeq 0$ ,  $\lambda_{\max}(A) \leq \lambda_{\max}(A + B)$ . By (53) and (55), we have the following

$$\|W_k\| \leq \min\{Q\eta_U, 1\}. \quad (56)$$

We now consider the term with  $\bar{p}_k = p_k - \mathbb{E}[p_k]$  in (49)

$$\max_k |p_k - \mathbb{E}[p_k]| \leq \max_k \max\{p_k - \mathbb{E}[p_k], \mathbb{E}[p_k]\} \quad (57)$$

$$\leq \max_k \max\left\{ \frac{QE_u - Q\mu_E}{S_k}, \frac{Q\mu_E}{S_k} \right\} \quad (58)$$

$$\leq Q \max\{E_u - \mu_E, \mu_E\} \frac{1}{\min S_k} \quad (59)$$

$$\leq \mu_E \max\{r_E - 1, 1\} \frac{1}{\eta_L} \frac{s}{P_x}, \quad (60)$$

where we have used  $\mathbb{E}[p_k] = Q\mathbb{E}[E_k] = Q\mu_E$ ,  $p_k \leq QE_u$  and  $E_u = r_E \mu_E$ . Here (60) follows from

$$S_k = \sum_{t=1}^Q \sigma_{x_{k,t}}^2 \geq Q \min_{k,t} \sigma_{x_{k,t}}^2 = Q\eta_L \frac{P_x}{s}, \quad (61)$$

where  $\sigma_{x_{k,t}}^2 = \frac{P_x}{s} \|u_{k,t}\|^2$  and

$$\eta_L \triangleq \min_{k,t} \|u_{k,t}\|^2. \quad (62)$$

Hence by (49), (56) and (60)

$$\|Z_k\| \leq \mu_E \frac{s}{P_x} \frac{1}{\eta_L} \max\{r_E - 1, 1\} \min\{Q\eta_U, 1\} \triangleq \bar{\mu}_I, \quad \forall k. \quad (63)$$

We now consider the variance term, i.e.,

$$\left\| \sum_{k=1}^{N_T} \mathbb{E}[Z_k^2] \right\| = \left\| \sum_{k=1}^{N_T} \mathbb{E}[\bar{p}_k^2] W_k^2 \right\|, \quad (64)$$

$$\leq \max_k \mathbb{E}[\bar{p}_k^2] \left\| \sum_{k=1}^{N_T} W_k^2 \right\|, \quad (65)$$

where we have used  $\mathbb{E}[\bar{p}_k^2] W_k^2 \preceq (\max_k \mathbb{E}[\bar{p}_k^2]) W_k^2$  and  $\sum_{k=1}^{N_T} \mathbb{E}[\bar{p}_k^2] W_k^2 \preceq (\max_k \mathbb{E}[\bar{p}_k^2]) \sum_{k=1}^{N_T} W_k^2$ . Here (65) follows from the fact that for Hermitian  $A, B$  with  $A \succeq B$ , we have  $\lambda_k(A) \geq \lambda_k(B)$ , where  $\lambda_k(\cdot)$  denote the ordered eigenvalues [46, Cor. 7.7.4].

The spectral norm term in (65) can be bounded as

$$\left\| \sum_{k=1}^{N_T} W_k^2 \right\| \leq \max_k \|W_k\| \left\| \sum_{k=1}^{N_T} W_k \right\|, \quad (66)$$

$$\leq \min\{Q\eta_U, 1\} \quad (67)$$

where (66) follows from the fact that  $W_k \succeq 0$ , see for instance [47, Sec. 2], and (67) follows from (56) and (55).

We now consider  $\mathbb{E}[\bar{p}_k^2]$  in (64). We have the following

$$\mathbb{E}[\bar{p}_k^2] = \frac{1}{S_k^2} \sum_{l=1}^Q \mathbb{E}[(E_{k,l} - \mathbb{E}[E_{k,l}])^2] \quad (68)$$

$$= \frac{Q}{S_k^2} \mathbb{E}[(E_{k,l} - \mathbb{E}[E_{k,l}])^2] \quad (69)$$

$$= \frac{Q\varrho_E}{S_k^2} \quad (70)$$

$$\leq \frac{Q\varrho_E}{Q\eta_L^2} \left( \frac{s}{P_x} \right)^2, \quad (71)$$

where  $\varrho_E$  is the variance of the energy arrival process as defined before, (68) follows from the fact that  $\bar{p}_k$  is a sum of statistically independent zero mean variables and (71) follows from  $S_k^2 \geq Q^2 (\min_{k,t} \sigma_{x_{k,t}}^2)^2 = Q^2 \eta_L^2 \left( \frac{P_x}{s} \right)^2$ .

Hence the variance term in (64) can be bounded as follows

$$\left\| \sum_{k=1}^{N_T} \mathbb{E}[Z_k^2] \right\| \leq \frac{Q\varrho_E}{Q\eta_L^2} \left( \frac{s}{P_x} \right)^2 \min\{Q\eta_U, 1\} \triangleq \bar{\varrho}_I. \quad (72)$$

Using (63), (72) and the Matrix Bernstein Inequality reveals that for  $\bar{r} > 0$ ,  $\|\sum_{k=1}^{N_T} Z_k\| < \bar{r}$  holds with probability greater than  $p_{bt} = 1 - f_{bt}(\bar{\mu}_I, \bar{\varrho}_I, \bar{r})$ . We note that for Hermitian  $A, B$ ,  $\|A - B\| < \bar{r}$  implies  $\lambda_{\min}(A) > \lambda_{\min}(B) - \bar{r}$ . Therefore, using (47), with probability greater than  $p_{bt}$

$$\lambda_{\min}(U_s^\dagger G U_s) > \lambda_{\min}(U_s^\dagger \bar{G} U_s) - \bar{r} \quad (73)$$

$$\geq \min_k \mathbb{E}[p_k] - \bar{r} \quad (74)$$

$$= \frac{\mu_E}{\eta_U} \frac{s}{P_x} - \bar{r} \quad (75)$$

where (74) follows from the fact that  $U_s^\dagger \bar{G} U_s = \sum_{t=1}^Q \mathbb{E}[p_{k,t}] Y_{k,t}$  with  $Y_{k,t} \succeq 0$  and

$$\sum_{t=1}^Q \mathbb{E}[p_{k,t}] Y_{k,t} \succeq \sum_{t=1}^Q \min_{t,k} \mathbb{E}[p_{k,t}] Y_{k,t} = \min_{t,k} \mathbb{E}[p_{k,t}] I_s \quad (76)$$

so that  $U_s^\dagger \bar{G} U_s \succeq \min_{t,k} \mathbb{E}[p_{k,t}] I_s$ . Hence  $\lambda_{\min}(U_s^\dagger \bar{G} U_s) \geq \min_{t,k} \mathbb{E}[p_{k,t}]$  due to the fact that for Hermitian  $A, B$  with  $A \succeq B$ , we have  $\lambda_l(A) \geq \lambda_l(B)$ , where  $\lambda_l(\cdot)$  denote the ordered eigenvalues [46, Cor. 7.7.4]. Here (75) follows from  $\mathbb{E}[p_k] = \frac{Q\mu_E}{S_k} \geq \frac{\mu_E}{\eta_U} \frac{s}{P_x}$  where  $S_k$  is bounded as follows

$$S_k = \sum_{t=1}^Q \sigma_{x_{k,t}}^2 \leq Q \max_{k,t} \sigma_{x_{k,t}}^2 = Q \eta_U \frac{P_x}{s}. \quad (77)$$

Let us introduce  $r, \mu_I, \varrho_I$ , such that  $\bar{r} = \mu_E \frac{s}{P_x} r, \bar{\mu}_I = \mu_E \frac{s}{P_x} \mu_I, \bar{\varrho}_I = (\mu_E \frac{s}{P_x})^2 \varrho_I$ . Hence, (75) is expressed as

$$\lambda_{\min}(U_s^\dagger \bar{G} U_s) > \mu_E \frac{s}{P_x} \left( \frac{1}{\eta_U} - r \right). \quad (78)$$

We note that  $\bar{\mu}_I, \bar{\varrho}_I, \bar{r}$  can be scaled as above without a change in the value of  $f_{bt}(\cdot)$  and  $f_{bn}(\cdot)$ , i.e.  $f_{bt}(\bar{\mu}_I, \bar{\varrho}_I, \bar{r}) = f_{bt}(\mu_I, \varrho_I, r)$  and  $f_{bn}(\bar{\mu}_I, \bar{\varrho}_I, \bar{r}) = f_{bn}(\mu_I, \varrho_I, r)$ . Using  $r, \mu_I, \varrho_I$ , (78) and (40) leads to the bounds in (9)-(12).

We now consider the second set of bounds given in (13)-(17). We first consider the event  $\bar{E}_k \geq \gamma \mu_E$  and define a new Bernoulli random variable  $\bar{\delta}_k = \mathbb{1}_{\bar{E}_k \geq \gamma \mu_E}$ , where  $\gamma \in [0, Q r_E]$  and  $\mathbb{1}$  is the indicator function. We define the probability  $\bar{p}$  as follows

$$\bar{p} \triangleq \mathbb{P}(\bar{E}_k \geq \gamma \mu_E) \quad (79)$$

Hence  $\mathbb{P}(\bar{\delta}_k = 1) = \bar{p}$  and  $\mathbb{P}(\bar{\delta}_k = 0) = 1 - \bar{p}$ . Let us define

$$p_k^L \triangleq \frac{\gamma \mu_E}{S_k} \bar{\delta}_k. \quad (80)$$

We note that  $p_k^L$  provides a lower bound for  $p_k$ ,  $\forall k$ . Hence, we have  $p_k W_k \succeq p_k^L W_k$ ,  $\forall k$ , and we have

$$\sum_{k=1}^{N_T} p_k W_k \succeq \sum_{k=1}^{N_T} p_k^L W_k. \quad (81)$$

Hence the minimum eigenvalue of  $\sum_{k=1}^{N_T} p_k^L W_k$  provides a lower bound for the minimum eigenvalue of  $\sum_{k=1}^{N_T} p_k W_k$  [46, Cor. 7.7.4]. Now re-iterating the steps for the proof of bounds in (9)-(12) reveals a set of bounds similar to (9)-(12), but that also depend on  $\gamma$ . Here the variables related to  $p_k$  are replaced with variables related to  $p_k^L$ . In particular,  $\bar{p}_k$  is replaced by  $\bar{p}_k^L = p_k^L - \mathbb{E}[p_k^L] = p_k^L - \frac{\bar{p} \gamma \mu_E}{S_k}$  and (60) becomes  $\bar{p} \gamma \mu_E \max\{\frac{1}{\bar{p}} - 1, 1\} \frac{1}{Q \eta_L} \frac{s}{P_x}$ . Similarly, (71) becomes  $(\bar{p} \gamma \mu_E)^2 (\frac{1}{\bar{p}} - 1) \frac{1}{Q^2 \eta_L^2} (\frac{s}{P_x})^2$  and (75) becomes  $\frac{\bar{p} \gamma \mu_E}{Q \eta_U} \frac{s}{P_x} - \bar{r}$ . Using these values, and normalizing  $\bar{\mu}_I, \bar{\varrho}_I, \bar{r}$  appropriately as before, we arrive at the bounds in (13)-(17).

## IX. PROOF OF LEMMA 3.1

Let the notation be the same with Section VIII. We note that  $\mathbb{E}[p_k] = \frac{Q\mu_E}{S_k} \leq \frac{\mu_E}{\eta_L} \frac{s}{P_x}$  where  $S_k$  is bounded by (61). Hence,

$$\mathbb{E}[G] = \text{diag}(\mathbb{E}[p_1] \mathbf{1}_Q; \dots; \mathbb{E}[p_{N_T}] \mathbf{1}_Q) \preceq \text{diag}\left(\frac{\mu_E}{\eta_L \frac{P_x}{s}}\right) \quad (82)$$

Hence, the error can be bounded as follows

$$\mathbb{E}_E[\varepsilon] \geq \text{tr} \left[ \left( \frac{s}{P_x} I_s + \frac{1}{\sigma_w^2} U_s^\dagger \mathbb{E}[G] U_s \right)^{-1} \right] \quad (83)$$

where (83) follows from (6), Jensen's inequality and  $\text{tr}[X^{-1}]$  is convex for  $X \succ 0$ . The result follows from (82) and properties of p.s.d. ordering [46, 7.7.2], [46, 7.7.4].

## X. PROOF OF COROLLARY 5.1

We first focus on (9) to derive (35). We define the event  $S_{\sigma_w^2}$  as  $S_{\sigma_w^2} = \{\varepsilon < \frac{1}{1 + \frac{1}{\sigma_w^2} \mu_E (\frac{1}{\eta_U} - r)} P_x = \frac{\sigma_w^2}{\sigma_w^2 + \mu_E (\frac{1}{\eta_U} - r)} P_x\}$ . Suppose that  $\frac{1}{\eta_U} - r$  is non-zero. Then,

$$\lim_{\sigma_w^2 \rightarrow 0} \mathbb{P}(S_{\sigma_w^2}) = \lim_{\sigma_w^2 \rightarrow 0} \mathbb{E}[\mathbb{1}_{S_{\sigma_w^2}}] = \mathbb{E}[\lim_{\sigma_w^2 \rightarrow 0} \mathbb{1}_{S_{\sigma_w^2}}] = \mathbb{P}(\varepsilon = 0)$$

where  $\mathbb{1}$  is the indicator function and we changed the order of the expectation and the limit due to Dominated Convergence Theorem. Hence, we can use (9) to find the probability that  $\varepsilon = 0$  is zero. To have  $\varepsilon = 0$ , it is sufficient to have the expression  $\frac{1}{\eta_U} - r$  bounded away from zero.

Let us consider a given failure probability  $\delta$ , so that  $1 - \delta \leq 1 - f_{bn}(\mu_I, \varrho_I, r)$ . Hence, we have  $\frac{r^2}{\mu_I r/3 + \varrho_I} \geq 2 \ln(\frac{2s}{\delta})$ . We now re-parametrize  $r \in (0, \frac{1}{\eta_U}]$ , as  $r = \tilde{r} \frac{1}{\eta_U}$ , where  $0 < \tilde{r} \leq 1$ . Using (11), (12),  $\eta_L = \eta_U = \frac{s}{N}$ , we obtain

$$2(1/\tilde{r}^2)(\max\{1/p - 1, 1\} \tilde{r} Q/3 + 1/p - 1) s \ln(2s/\delta) \leq N \quad (84)$$

A sufficient condition for (84) is (35), hence we conclude that (35) is a sufficient condition for  $\varepsilon = 0$ . Similarly, (13) leads to the condition

$$2(1/\tilde{r}^2)((\tilde{r}/3) \max(a, 1) + a) Q s \ln(2s/\delta) \leq N \quad (85)$$

where  $a = (1/\bar{p} - 1)$ . Choosing  $\gamma$  of (17) as  $1/p$  (at least one energy packet arrives during  $Q$  time slots), we set  $\bar{p} = 1 - (1 - p)^Q$ . Using  $\bar{p} \leq 1/2$ , hence  $a \geq 1$ , a sufficient condition for (85) is (36). The result follows.

## XI. PROOF OF THM. 4.1

We bound the estimation error using (40). We recall that  $U_s = F_\Omega^N \in \mathbb{C}^{N \times s}$  consists of the first  $s$  columns of the size  $N$  DFT matrix  $F^N \in \mathbb{C}^{N \times N}$ . The proof relies on the fact that equidistantly row sampled  $F_\Omega^N$  can be associated with the DFT matrix of size  $s$ ,  $F^s \in \mathbb{C}^{s \times s}$ . Let  $f_N = \exp(-j \frac{2\pi}{N})$ . The entries of the row-sampled  $F_\Omega^N$  every  $Q = N/s$  rows can be expressed in terms of the entries of  $F^s$  as follows

$$[F_\Omega^N]_{(N/s)l+t_d+1, k+1} = (1/\sqrt{N}) f_N^{((N/s)l+t_d)k} \quad (86)$$

$$= (1/\sqrt{N}) f_s^{lk} f_N^{t_d k} \quad (87)$$

$$= \sqrt{s/N} [F^s]_{l+1, k+1} f_N^{t_d k}, \quad (88)$$

where  $0 \leq k \leq s-1, 0 \leq l \leq s-1$ . Now we adopt arguments similar to Section VIII. By (4) and (28), we have  $p_k = \bar{E}_k \frac{N}{P_x}$ , where  $\mathbb{E}[p_k] = Q \mu_E \frac{N}{P_x}$ . Let  $\bar{p}_k \triangleq p_k - \mathbb{E}[p_k]$ .

In contrast to Section VIII, due to the equidistant sample transmission setting, here we define  $W_k$  as follows

$$W_k \triangleq Y_{k, t_d+1}, \quad (89)$$

where  $Y_{k,t} \triangleq u_{k,t} u_{k,t}^\dagger$  as in Section VIII. Hence, with  $Z_k \triangleq (p_k - \mathbb{E}[p_k])W_k$ , we have

$$\begin{aligned} \sum_{k=1}^{N_T} Z_k &= \sum_{k=1}^{N_T} p_k Y_{k,t_d+1} - \sum_{k=1}^{N_T} \mathbb{E}[p_k] Y_{k,t_d+1}, \\ &= U_s^\dagger G U_s - U_s^\dagger \bar{G} U_s, \end{aligned} \quad (90)$$

where  $G = \text{diag}(g_t) \in \mathbb{R}^{N \times N}$ ,  $\bar{G} = \mathbb{E}[G] \in \mathbb{R}^{N \times N}$  with  $g_t = p_k$ , if  $t = Q(k-1) + t_d + 1$ ,  $1 \leq k \leq N_T$ , and  $g_t = 0$  otherwise; as dictated by the equidistant sampling strategy.

We note that  $\eta_U$  and  $\eta_L$  as defined in (54) and (62) is given by  $\eta_U = \eta_L = \frac{s}{N}$  due to the fact that here  $U_s$  is given by  $U_s = F_\Omega^N$ . Hence, we have  $\|W_k\| = \|u_{k,t_d+1} u_{k,t_d+1}^\dagger\| \leq \eta_U = s/N$ . Due to (88), we also have  $\|W_k\| \leq \|\sum_{k=1}^{N_T} W_k\| = \|\frac{s}{N} F^s F^{s\dagger}\| = \|\frac{s}{N} I_s\| = \frac{s}{N}$ . Hence, we bound  $\|W_k\|$  as  $\|W_k\| \leq \frac{s}{N}$ .

Similar to (60),  $|\bar{p}_k|$  can be bounded as  $|\bar{p}_k| \leq \max_k |p_k - \mathbb{E}[p_k]| \leq Q\mu_E \max\{r_E - 1, 1\} \frac{N}{P_x}$ . Hence we have

$$\|Z_k\| \leq Q\mu_E \frac{s}{P_x} \max\{r_E - 1, 1\} \triangleq \bar{\mu}_I^u, \quad \forall k. \quad (91)$$

For  $\bar{\varrho}_I^u$ , we note that  $\|\sum_{k=1}^{N_T} W_k^2\| \leq (s/N)^2$  and  $\mathbb{E}[\bar{p}_k^2] \leq Q\varrho_E (\frac{N}{P_x})^2$ . Hence using (65), we have

$$\|\sum_{k=1}^{N_T} \mathbb{E}[Z_k^2]\| \leq \varrho_E Q (\frac{s}{P_x})^2 \triangleq \bar{\varrho}_I^u. \quad (92)$$

Using (91), (92) and the Matrix Bernstein Inequality shows that  $\|\sum_{k=1}^{N_T} Z_k\| < \bar{r}^u$  holds with probability greater than  $p_{bt} = 1 - f_{bt}(\bar{\mu}_I^u, \bar{\varrho}_I^u, \bar{r}^u)$ . Therefore, we have the following

$$\lambda_{\min}(U_s^\dagger G U_s) > \lambda_{\min}(U_s^\dagger \bar{G} U_s) - \bar{r}^u \quad (93)$$

$$= \frac{s}{N} \lambda_{\min}(F^{s\dagger} \text{diag}(\mathbb{E}[p_k]) F^s) - \bar{r}^u \quad (94)$$

$$= \mu_E Q \frac{s}{P_x} - \bar{r}^u \quad (95)$$

where (94) follows from (88); and (95) follows from the fact that  $\mathbb{E}[p_k] = \mu_E Q N / P_x$  and  $F^{s\dagger} F^s = I_s$ . Set  $c = \mu_E Q \frac{s}{P_x}$  where  $Q = N/s$ . We conclude the proof by rescaling  $\bar{\mu}_I^u$ ,  $\bar{\varrho}_I^u$ ,  $\bar{r}^u$  with  $c$ ,  $c^2$  and  $c$ , respectively.

## REFERENCES

- [1] Ericsson, "Ericsson mobility report: On the pulse of networked society," 2016.
- [2] D. Gündüz, K. Stamatiou, N. Michelusi, and M. Zorzi, "Designing intelligent energy harvesting communication systems," *IEEE Commun. Mag.*, vol. 52, no. 1, pp. 210–216, 2014.
- [3] J. M. Gilbert and F. Balouchi, "Comparison of energy harvesting systems for wireless sensor networks," *Int. Journal of Autom. and Comput.*, vol. 5, pp. 334–347, Oct. 2008.
- [4] R. Vullers, R. Schaijk, H. Vissers, J. Penders, and C. Hoof, "Energy Harvesting for Autonomous Wireless Sensor Networks," *IEEE Solid-State Circuits Mag.*, vol. 2, no. 2, pp. 29–38, 2010.
- [5] M. Gorlatova, A. Wallwater, and G. Zussman, "Networking low-power energy harvesting devices: Measurements and algorithms," *IEEE Trans. Mobile Comput.*, vol. 12, no. 9, pp. 1853–1865, 2013.
- [6] <https://www.enocean.com>.
- [7] <http://perpetuapower.com>.
- [8] <https://perpetuum.com>.
- [9] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, pp. 1989–2001, May 2013.
- [10] K. Huang and E. Larsson, "Simultaneous information and power transfer for broadband wireless systems," *IEEE Trans. Signal Process.*, pp. 5972–5986, Dec. 2013.
- [11] J. Park and B. Clerckx, "Joint wireless information and energy transfer in a  $K$ -user MIMO interference channel," *IEEE Trans. Wireless Commun.*, vol. 13, pp. 5781–5796, Oct. 2014.
- [12] A. Özçelikkale and T. M. Duman, "Linear precoder design for simultaneous information and energy transfer over two-user MIMO interference channels," *IEEE Trans. Wireless Commun.*, vol. 14, pp. 5836–5847, Oct 2015.
- [13] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, "Transmission with Energy Harvesting Nodes in Fading Wireless Channels: Optimal Policies," *IEEE J. Sel. Areas Commun.*, vol. 29, pp. 1732–1743, Sept. 2011.
- [14] K. Tutuncuoglu and A. Yener, "Optimum transmission policies for battery limited energy harvesting nodes," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 1180–1189, March 2012.
- [15] M. A. Antepi, E. Uysal-Biyikoglu, and H. Erkal, "Optimal packet scheduling on an energy harvesting broadcast link," *IEEE J. Sel. Areas Commun.*, vol. 29, pp. 1721–1731, Sept. 2011.
- [16] J. Yang and S. Ulukus, "Optimal packet scheduling in a multiple access channel with energy harvesting transmitters," *Journal of Communications and Networks*, vol. 14, no. 2, pp. 140–150, 2012.
- [17] O. Ozel and S. Ulukus, "Achieving AWGN capacity under stochastic energy harvesting," *IEEE Trans. Inf. Theory*, vol. 58, pp. 6471–6483, Oct 2012.
- [18] Y. Dong, F. Farnia, and A. Özgür, "Near optimal energy control and approximate capacity of energy harvesting communication," *IEEE J. Sel. Areas Commun.*, vol. 33, pp. 540–557, March 2015.
- [19] R. Srivastava and C. E. Koksal, "Basic performance limits and tradeoffs in energy-harvesting sensor nodes with finite data and energy storage," *IEEE/ACM Trans. on Networking*, vol. 21, pp. 1049–1062, Aug. 2013.
- [20] A. Kazerouni and A. Özgür, "Optimal online strategies for an energy harvesting system with Bernoulli energy recharges," in *2015 Inter. Symp. on Modeling and Opt. in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, pp. 235–242, 2015.
- [21] A. Nayyar, T. Başar, D. Teneketzis, and V. Veeravalli, "Optimal strategies for communication and remote estimation with an energy harvesting sensor," *IEEE Trans. Autom. Control*, vol. 58, pp. 2246–2260, Sept 2013.
- [22] P. Blasco, D. Gündüz, and M. Dohler, "A Learning Theoretic Approach to Energy Harvesting Communication System Optimization," *IEEE Trans. Commun.*, vol. 12, pp. 1872–1882, Apr. 2013.
- [23] A. Özçelikkale, S. Yüksel, and H. M. Ozaktas, "Unitary precoding and basis dependency of MMSE performance for Gaussian erasure channels," *IEEE Trans. Inf. Theory*, vol. 60, pp. 7186–7203, Nov 2014.
- [24] Y.-W. P. Hong, "Distributed estimation with analog forwarding in energy-harvesting wireless sensor networks," in *IEEE Int. Conf. Comm. Sys. (ICCS)*, pp. 142–146, 2014.
- [25] M. Nourian, S. Dey, and A. Ahlén, "Distortion minimization in multi-sensor estimation with energy harvesting," *IEEE J. Sel. Areas Commun.*, vol. 33, pp. 524–539, Mar. 2015.
- [26] 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, pp. 2848–2863, June 2015.
- [27] M. Nourian, A. 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, 2014.
- [28] M. Calvo-Fullana, J. Matamoros, and C. Antón-Haro, "Reconstruction of correlated sources with energy harvesting constraints in delay-constrained and delay-tolerant communication scenarios," *IEEE Trans. Wireless Commun.*, vol. 16, pp. 1974–1986, March 2017.
- [29] A. Özçelikkale, T. McKelvey, and M. Viberg, "Transmission strategies for remote estimation under energy harvesting constraints," in *Proc. European Signal Process. Conf. (EUSIPCO)*, pp. 572 – 576, 2016.
- [30] A. Özçelikkale, T. McKelvey, and M. Viberg, "Transmission strategies for remote estimation with an energy harvesting sensor," *IEEE Trans. Wireless Commun.*, vol. 16, pp. 4390–4403, July 2017.
- [31] R. Gangula, D. Gündüz, and D. Gesbert, "Distributed compression and transmission with energy harvesting sensors," in *2015 IEEE International Symposium on Information Theory (ISIT)*, pp. 1139–1143, 2015.
- [32] Y. Zhao, B. Chen, and R. Zhang, "Optimal power management for remote estimation with an energy harvesting sensor," *IEEE Trans. Wireless Commun.*, vol. 14, pp. 6471–6480, Nov. 2015.



- [33] O. Orhan, D. Gündüz, and E. Erkip, "Source-channel coding under energy, delay, and buffer constraints," *IEEE Trans. Wireless Commun.*, vol. 14, pp. 3836–3849, July 2015.
- [34] A. Özçelikkale, T. McKelvey, and M. Viberg, "Performance bounds for remote estimation with an energy harvesting sensor," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, pp. 460–464, 2016.
- [35] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, 1993.
- [36] R. M. Gray, *Toeplitz and Circulant Matrices: a Review*. Now Publishers Inc., 2006.
- [37] A. Tulino, G. Caire, S. Verdú, and S. Shamai, "Support recovery with sparsely sampled free random matrices," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4243–4271, 2013.
- [38] T. Tirronen, A. Larmo, J. Sachs, B. Lindoff, and N. Wiberg, "Machine-to-machine communication with long-term evolution with reduced device energy consumption," *Trans. on Emerging Telecomm. Tech.*, vol. 24, pp. 413–426, June 2013.
- [39] I. Bahceci and A. Khandani, "Linear estimation of correlated data in wireless sensor networks with optimum power allocation and analog modulation," *IEEE Trans. Commun.*, vol. 56, pp. 1146–1156, July 2008.
- [40] K.-H. Lee and D. Petersen, "Optimal linear coding for vector channels," *IEEE Trans. Commun.*, vol. 24, no. 12, pp. 1283–1290, 1976.
- [41] O. Tan, D. Gündüz, and J. Gómez-Vilardebó, "Linear transmission of composite Gaussian measurements over a fading channel under delay constraints," *IEEE Trans. Wireless Commun.*, vol. 15, pp. 4335–4347, June 2016.
- [42] B. D. O. Anderson and J. B. Moore, *Optimal filtering*. Prentice-Hall, 1979.
- [43] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, 3rd ed., 1991.
- [44] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. Springer, 2013.
- [45] E. J. Candes and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Problems*, vol. 23, pp. 969–985, June 2007.
- [46] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1990.
- [47] J. A. Tropp, "The expected norm of a sum of independent random matrices: An elementary approach," in *High Dimensional Probability VII*, pp. 173–202, Springer, 2016.