Performance Bounds for Remote Estimation with an Energy Harvesting Sensor

Performance Bounds for Remote Estimation with an Energy Harvesting Sensor

Ayça Özçelikkale, Tomas McKelvey, Mats Viberg

***Abstract*—Remote estimation with an energy harvesting sensor with a limited data buffer is considered. The sensor node observes an unknown correlated circularly wide-sense stationary (c.w.s.s.) Gaussian field and communicates its observations to a remote fusion center using the energy it harvested. The fusion center employs minimum mean-square error (MMSE) estimation to reconstruct the unknown field. The distortion minimization problem under the online scheme, where the sensor has only access to the statistical information for the future energy packets is considered. We provide performance bounds on the achievable distortion under a slotted block transmission scheme, where at each transmission time slot, the data and the energy buffer is completely emptied. Our bounds provide insight to the trade-off between the buffer sizes and the achievable distortion. These trade-offs illustrate the insensitivity of the performance to the buffer size for signals with low degree of freedom and suggest performance improvements with increasing buffer size for signals with relatively higher degree of freedom.**

I. Introduction

Energy harvesting (EH) offers a promising perspective for efficient usage of energy sources in sensor networks. Sensors with EH capabilities use alternative available energy sources, such as solar power or mechanical vibrations, instead of completely relying on a fixed battery or the power from the grid. In addition to utilizing energy resources that may otherwise be wasted, EH capabilities offer prolonged battery life-times, eliminates the need for dedicated power cables and offers significant mobility for the nodes in the network [1].

The intermittent nature of the energy available is one of the main challenges in the design of EH systems. Accordingly, performance of EH communication systems have been studied under a broad range of scenarios, but mostly under the performance criterion of rate maximization; see, for instance, [1–4]. Although these works successfully address issues related to reliable communications, they can only provide limited insights about the sensing problem at hand, i.e. the recovery of the unknown field measured by the sensors. In that respect, estimation with EH sensors have been studied in a relatively small number of works, such as parameter estimation [5] and estimation of Markov sources [6], [7]. Nevertheless the interplay between the unreliable characteristics of the EH sources and the estimation performance has only been partially understood.

An important characteristic in estimation problems is the varying degrees of freedom, i.e. degree of sparsity, hence varying degrees of correlation of the unknown signal. In addition to providing a reasonable model for physical fields, the sparsity of the signal can be utilized to compensate for the unreliable nature of the energy sources in an EH system. For EH systems, structural results that directly exploit the sparsity or the correlation characteristics are available only for a limited number of scenarios, such as estimation of a single parameter [5], Markov sources [6], [7], two correlated Gaussian variables [8] and i.i.d. Gaussian sources as a result of the findings of, for instance, [2], [3]. Here, we address this issue by focusing on the trade-offs between the estimation performance, the level of sparsity of the signal and the buffer size.

We consider an EH sensor which observes an unknown correlated c.w.s.s. Gaussian field and communicates its observations to a remote fusion center using the energy it harvested. The fusion center employs MMSE estimation to reconstruct the unknown signal. We consider the problem under a limited data and energy buffer constraint and a slotted transmission scheme where at each transmission time slot the data buffer and the battery is completely emptied. We consider an amplify-and-forward strategy similar to [5]. We focus on the online scheme, where the sensor has only access to the statistical information for the future energy packets. Using random matrix theory and compressive sensing tools, we provide performance bounds that reveal the trade-off between the buffer size and the achievable distortion. As one may infer from compressive sensing results, the performance bounds are observed to be relatively insensitive to the buffer size for signals with low degrees of freedom, i.e. for sparse signals. Our results also quantify the possible performance gain due to increasing buffer sizes for signals with relatively higher degrees of freedom.

The rest of the paper is organized as follows. In Section II, the system model is described. Our performance bounds are presented in Section III. In Section IV, illustration of the bounds is provided. The paper is concluded in Section V.

Notation: The complex conjugate transpose of a matrix A is denoted by $A^\dagger$. The spectral norm of a matrix A is denoted by $||A||$. Positive semi-definite (p.s.d.) ordering for Hermitian matrices is denoted by $\succeq$. I_n denotes the identity matrix with $I_n \in \mathbb{C}^{n\times n}$. The l_2 norm of a vector a is denoted by $\|a\|$.

II. System Model

A. Signal Model

The aim of the remote estimation system is to estimate the unknown complex proper zero mean c.w.s.s. Gaussian field $\mathbf{x} = [x_1, \dots, x_N] \in \mathbb{C}^{N\times 1}$, $\mathbf{x} \sim \mathcal{CN}(0, K_{\mathbf{x}})$ with $K_{\mathbf{x}} = \mathbb{E}[\mathbf{x}\mathbf{x}^\dagger]$. Here, the covariance matrix $K_{\mathbf{x}}$ is circulant [9], [10] and models the possible correlation of the field values in time.

A. Özçelikkale, T. McKelvey and M. Viberg are with Dep. of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden e-mails: {ayca.ozcelikkale, tomas.mckelvey, mats.viberg}@chalmers.se.
A. Özçelikkale acknowledges the support of EU Marie Skłodowska-Curie Fellowship.

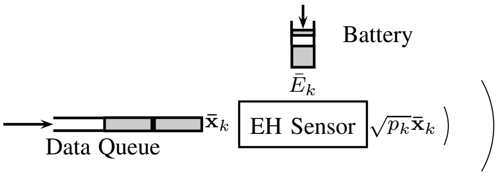


Fig. 1: Energy Harvesting Sensor

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) singular value decomposition 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 submatrix of U corresponding to non-zero eigenvalues with $U \in \mathbb{C}^{N \times N}$ the DFT matrix [10]. Let $P_x = \mathrm{tr}[K_{\mathbf{x}}] = \mathrm{tr}[\Lambda_{x,S}]$. We consider $\Lambda_{x,s}$ 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. the sparsity level of the signal family. 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 [11], [12].

B. Sensing and Communications to the Fusion Center

At time epoch t, the sensor observes the field value at time t, i.e. x_t. The observations are held in a buffer of finite size Q_d before transmission. The buffer contents at the end of transmission time slot k, i.e. at the end of time epoch kQ_d, is given by $\bar{\mathbf{x}}_{\mathbf{k}} = [x_{(k-1)Q_d+1}, x_{(k-1)Q_d+2} \cdots x_{(k-1)Q_d+Q_d}]$. For convenience, $N_T = N/Q_d$ is assumed to be an integer where N_T gives the number of transmission blocks. As illustrated in Fig. 1 and Fig. 2, we consider a block transmission scheme, where at the end of transmission time slot k, the sensor transmits the data in its buffer to a fusion center using an amplify-and-forward strategy

$$\bar{\mathbf{y}}_{\mathbf{k}} = \sqrt{p_k}\bar{\mathbf{x}}_{\mathbf{k}} + \bar{\mathbf{w}}_{\mathbf{k}}, \qquad k = 1, \dots, N_T, \tag{1}$$

where $\mathbf{y}_k$ and $\bar{\mathbf{w}}_k$ denote the received signal at the fusion center and the effective channel noise for transmission time slot k, respectively. Here, $\sqrt{p_k}$ denote the amplification factor adopted by the sensor at transmission time slot k. The channel noise $\bar{\mathbf{w}}_{\mathbf{k}}$ is modelled as complex proper zero mean Gaussian $\bar{\mathbf{w}}_k \in \mathbb{C}^{Q_d \times 1} \sim \mathcal{CN}(0, K_{\bar{\mathbf{w}}_k})$, $K_{\bar{\mathbf{w}}_k} = \mathbb{E}[\bar{\mathbf{w}}_k \bar{\mathbf{w}}_k^{\dagger}]$. Let $\mathbf{w} = [\bar{\mathbf{w}}_{\mathbf{1}}, \dots, \bar{\mathbf{w}}_{\mathbf{N_T}}] \in \mathbb{C}^{\mathbf{N} \times \mathbf{1}}$. The noise for different samples and transmission time slots are assumed to be uncorrelated with $K_{\bar{\mathbf{w}}} = \mathbb{E}[\bar{\mathbf{w}}\bar{\mathbf{w}}^{\dagger}] = \sigma_{\bar{w}}^2 I_N$.

Similar to [5] we consider an amplify-and-forward strategy, motivated by [13] and the high complexity and the high energy cost of source and channel coding operations. The above type of block transmission scheme allows us to study the effect of finite buffer size on estimation and facilitates connections with the off-line uniform allocation strategies optimal for i.i.d. sources [2]. It is also supported by the fact that for devices with low power budgets, it is more energy efficient to send relatively larger amount of data at each transmission [14].

The average energy used by the sensor for communications at transmission slot k can be written as follows:

$$J_k = \tau \mathbb{E}[||\sqrt{p_k}\bar{\mathbf{x}}_{\mathbf{k}}||^2] = \tau p_k \sum_{t=1}^{Q_d} \sigma^2_{x_{kQ_d+t}} = \tau p_k Q_d \frac{P_x}{N}, \tag{2}$$

where we have used $\sigma^2_{x_{kQ_d+t}} = \sigma_x^2 = \mathrm{tr}[K_{\mathbf{x}}]/N = P_x/N$. For convenience, we use time duration as $\tau = 1$ in the rest of the paper.

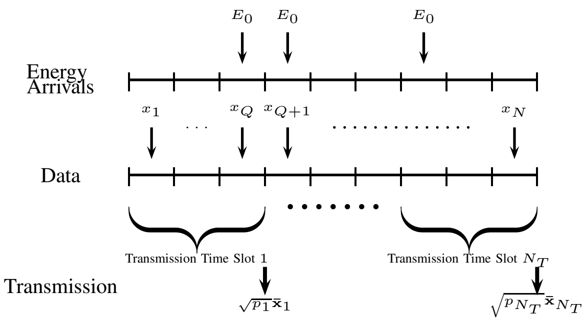


Fig. 2: Time Schedule for the Energy Harvesting Sensor

C. Energy Constraints at the Sensor:

The harvested energy process E_t is an i.i.d. discrete-time stochastic process as follows: $E_t = \delta_t E_0$, where δ_t is a Bernoulli random variable with success probability of p. We consider a battery-aided operation where the energy is stored at a battery and used in regular time intervals. Let the initial energy stored at the battery be 0. At the end of time slot k, the total energy that have arrived to the battery at this time slot is given by $\bar{E}_k = \sum_{t=1}^{Q_e} E_{(k-1)Q_e+t}$. We assume that the transmission time slots for the data buffer and the battery is synchronized and $Q_e = Q_d = Q$. We assume that the battery capacity is large enough so that $E_0 Q$ can be stored at the battery.

In general, the sensor has to operate under the 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 any transmission time slot does not exceed the available energy. Here, we focus on the case where at each transmission all the energy at the battery is used, i.e.

$$J_k = \bar{E}_k, \qquad k = 1, \dots, N_T. \tag{3}$$

D. Estimation at the Fusion Center:

After N_T transmission time slots, 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 $\mathbf{x}$. The MMSE estimate can be found as [15, Ch2]

$$\hat{\mathbf{x}} = \mathbb{E}[\mathbf{x}|\mathbf{y}] = K_{\mathbf{xy}} K_{\mathbf{y}}^{-1} \mathbf{y} \tag{4}$$

The resulting MMSE, $\mathbb{E}[||\mathbf{x} - \mathbb{E}[\mathbf{x}|\mathbf{y}]||^{\mathbf{2}}]$, can be expressed as

$$\varepsilon = \varepsilon(G) = \mathrm{tr}\left[(\frac{s}{P_x} I_s + \frac{1}{\sigma_w^2} U_s^{\dagger} G U_s)^{-1} \right], \tag{5}$$

where $G = \mathrm{diag}([p_1 \mathbf{1}_{Q_d}, \dots, p_{N_T} \mathbf{1}_{Q_d}]) \in \mathbb{C}^{N \times N}$ and $\mathbf{1}_{Q_d} = [1, \dots, 1] \in \mathbb{C}^{Q_d}$ is the vector of ones. We note that here the fusion center knows the source and noise statistics, including the covariance matrices.

We note that in order to perform the MMSE estimation, the fusion center needs to know or estimate the values of p_k. In general, estimation of p_k's can be considered as a part of the channel estimation process. In the above setting the possible energy cost of this operation is not incorporated to the problem set-up. Alternatively, the fusion center can have independent access to the E_t realizations and calculate the strategy p_k, for instance, when it is in the spatial vicinity of the sensor and it can measure solar energy fluctuations E_t itself.

III. Performance Bounds

We will now investigate the effect of different buffer sizes on the system performance. We first provide our main result, and then discuss our results using a lower bound on the average performance and the performance of a related off-line scheme as benchmarks. Let us define

$$f_{bt}(\mu, \varrho, t) \doteq 2s \exp\left(-\frac{\varrho}{\mu^2} h\left(\frac{\mu t}{\varrho}\right)\right) \tag{6}$$

$$f_{bn}(\mu, \varrho, t) \doteq 2s \exp\left(-\frac{t^2/2}{\mu t/3 + \varrho}\right) \tag{7}$$

with $h(a) \doteq (1+a)\log(1+a) - a,\ a \geq 0$.

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

Theorem 3.1: *The performance of the EH system satisfies the following bounds*

I.

$$\mathbb{P}(\varepsilon < \varepsilon_I) \geq 1 - f_{bt}\left(\mu_I, \varrho_I, t\right) \geq 1 - f_{bn}\left(\mu_I, \varrho_I, t\right) \tag{8}$$

for $t \in (0, 1]$, *where* $\eta = s/N$ *and*

$$\varepsilon_I = \frac{1}{1 + \frac{1}{\sigma_w^2}\frac{N}{s} p E_0 (1-t)} P_x \tag{9}$$

$$\mu_I = \max\{1/p - 1, 1\} \min\{Q\eta, 1\} \tag{10}$$

$$\varrho_I = \frac{1}{Q}(1/p - 1)\min\{Q\eta, 1\} \tag{11}$$

II.

$$\mathbb{P}(\varepsilon < \varepsilon_{II}) \geq 1 - \min_{\gamma\in[0,1]} f_{bt}\left(\mu_{II}, \varrho_{II}, t\right) \tag{12}$$

$$\geq 1 - \min_{\gamma\in[0,1]} f_{bn}\left(\mu_{II}, \varrho_{II}, t\right) \tag{13}$$

for $t \in (0, 1]$, *where* $\eta = s/N$ *and*

$$\varepsilon_{II} = \frac{1}{1 + \frac{1}{\sigma_w^2}\frac{N}{s} \bar{p} E_0 \gamma (1-t)} P_x \tag{14}$$

$$\mu_{II} = \max\{1/\bar{p} - 1, 1\} \min\{Q\eta, 1\} \tag{15}$$

$$\varrho_{II} = (1/\bar{p} - 1)\min\{Q\eta, 1\} \tag{16}$$

$$\bar{p} = \sum_{l=\lceil \gamma Q \rceil}^{Q} \binom{n}{l} p^l (1-p)^{Q-l}. \tag{17}$$

Proof: The proof is presented in Section VI.

For $Q = 1$, the energy E_t that arrives to the sensor at time t is immediately used to send the sample x_t. Hence for $Q = 1$, the measurement set-up is the same as the classical compressive sensing setting, and these bounds can be seen as a consequence of eigenvalue bounds provided in the compressive sensing literature [16, Ch.12].

For $Q > 1$, Thm. 3.1 provides a set of novel eigenvalue bounds for the formulation introduced in Section II. As the data 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 will be typically lower compared to the case where the energy is used to send a fewer number of samples, for instance directly sending the sample x_t if positive energy E_t arrives ($Q = 1$). Hence, the bounds 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.

For comparison purposes, we now consider a lower bound on the average error performance over different realizations of the process E_t:

$$\mathbb{E}[\varepsilon] \geq \mathrm{tr}\left[\left(\frac{s}{P_x} I_s + \frac{1}{\sigma_w^2} U_s^\dagger \mathbb{E}[G] U_s\right)^{-1}\right] \tag{18}$$

$$= \frac{1}{1 + \frac{1}{\sigma_w^2}\frac{N}{s} p E_0} P_x \tag{19}$$

where in (18) we have used (5), the Jensen's inequality and the fact that $\mathrm{tr}[X^{-1}]$ is convex for $X \succ 0$. In (19), we have used $\mathbb{E}[G] = \mathbb{E}[p_k] I_N$, $p_k = \frac{1}{\sigma_x^2 Q}\sum_{t=1}^{Q} \delta_{(k-1)Q+t} E_0$, $\mathbb{E}[p_k] = pE_0 N/P_x$ by (2) and (3), and $U_s^\dagger U_s = I_s$.

We also consider the performance of an associated off-line scheme. We focus on the following optimization problem

$$\varepsilon_d = \min_{\mathrm{tr}[G] \leq P_T} \varepsilon(G) \tag{20}$$

where $G = \mathrm{diag}(g) = \mathrm{diag}([g_1, \dots, g_N]) \in \mathbb{C}^{N\times N}$. Here the sensor has a total energy of $\bar{P}_T = P_T \sigma_x^2$ and it can freely distribute this energy on the samples in order to minimize the error. We note that in contrast to the setting in Section II, here a block transmission constraint is not imposed onto the set of admissible sensor strategies. We set $\bar{P}_T = pE_0 N$, i.e. $P_T = pE_0 N/\sigma_x^2$ for comparison purposes. We obtain the following result:

Lemma 3.1: *An optimal strategy for* (20) *is given by the uniform power allocation* $G = \mathrm{diag}(P_T/N) = \mathrm{diag}(pE_0 N/P_x)$, *and the optimum value is given by*

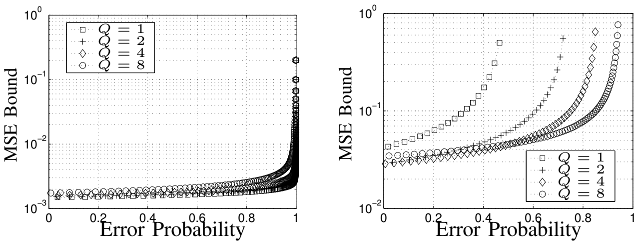
$$\varepsilon_d = \frac{1}{1 + \frac{1}{\sigma_w^2}\frac{N}{s} p E_0} P_x \tag{21}$$

Proof: The proof is presented in Section VII.

A total energy of $\bar{P}_T = pE_0 N$ corresponds to the total energy that will be obtained if an energy packet of pE_0 were harvested at each time step, which coincides with the average energy of our Bernoulli scheme. In this sense, (21) can be interpreted as a deterministic benchmark. Comparing our two benchmark schemes, we observe that both (19) and (21) provide error expressions in the same form $\frac{1}{1+SNR_{eff}} P_x$ where $SNR_{eff} = \frac{1}{\sigma_w^2}\frac{N}{s} pE_0$. Comparing the bounds in Thm. 3.1 with these, we observe that through a variable $t \in [0, 1)$, the bounds in Thm. 3.1 provide different operating points for how close one can operate to these benchmarks (for instance $\frac{1}{1+SNR_{eff}(1-t)} P_x$ in (9)) and with which probability.

IV. Numerical Results

We now illustrate the performance of our bounds. Let $N = 256$, $p = 0.4$, $E_0 = 1$, $P_x = 1$, $\sigma_w^2 = 10^{-4} P_x$. The resulting bounds are presented in Fig. 3a and Fig. 3b, for $s = 4$ and $s = 16$, respectively. Here the y-axis corresponds to the error bound as provided by ε_I/ε_{II} and the x-axis corresponds to the probability on the right-hand side of (8)/(12). While plotting



(a) s = 4 (b) s = 16

Fig. 3: MSE bound versus probability of error

the bounds, for a given probability we present the tightest of Bound I and Bound II, i.e. the bound that provides the smallest error value with a given probability.

In both figures as the desired performance becomes more demanding, i.e. the mean-square error (MSE) value decreases, the probability that this error can be guaranteed becomes smaller. When the degree of freedom of the signal is sufficiently low ($s=4$, Fig. 3a), the performance bound is observed to be relatively insensitive to the buffer size. On the other hand, when the degree of freedom is higher ($s = 16$, Fig. 3b), the bound becomes more sensitive to the buffer size. For $s=16$, with small buffer sizes $Q = 1, 2$, the bound cannot provide any guarantees that hold with high probability; whereas with higher buffer sizes, small values of error can be guaranteed with high probability (for instance with probability higher than 0.8 in Fig. 3b). We note that here it is the Bound II that illustrates the behaviour with $s = 16$. We observe that as s becomes larger, the signal can be said to be more close to an i.i.d. source, with the limiting case of $s = N$ corresponding to an exactly i.i.d source. Hence these results are consistent with the results of [2], which show that for i.i.d. sources the strategies that spread energy on the samples as much as possible is an optimum strategy for the offline scheme. Here, the buffers and the slotted transmission scheme facilitate strategies more close to such a uniform allocation. In that respect, evaluation of exponentially decreasing online energy allocation scheme of [3] in our remote estimation framework is considered future work.

V. Conclusions

We have considered remote estimation of a correlated c.w.s.s. Gaussian field with an EH sensor with a limited data and energy buffer. We have provided performance bounds on the achievable distortion under a slotted block transmission scheme. Our bounds illustrate the insensitivity of the performance to the buffer size for signals with low degree of freedom and the possible performance gain due to increasing buffer sizes for signals with relatively higher degree of freedom.

VI. Proof of Thm.3.1

We first prove the first family of bounds indexed by I in (8)-(11). We first note that

$$\varepsilon = \sum_{i=1}^{s} \frac{1}{\lambda_i(\frac{s}{P_x} I_s + \frac{1}{\sigma_w^2} U_s^\dagger G U_s)}, \tag{22}$$

$$\leq \frac{1}{1 + \frac{1}{\sigma_w^2}\frac{P_x}{s}\lambda_{min}(U_s^\dagger G U_s)} P_x. \tag{23}$$

Let

$$\tilde{p}_k = \frac{p_k}{\mathbb{E}[p_k]} - 1, \tag{24}$$

$$\bar{Z}_k = \sum_{t=1}^{Q} u_{(k-1)Q+t} u_{(k-1)Q+t}^\dagger, \tag{25}$$

$$Z_k \doteq \tilde{p}_k \bar{Z}_k, \qquad k = 1, \dots, N_T. \tag{26}$$

where $u_i \in \mathbb{C}^s$ is the i^{th} column of the matrix $U_s^\dagger$. Hence

$$\sum_{k=1}^{N_T} Z_k = \frac{1}{\mathbb{E}[p_k]} \sum_{k=1}^{N_T} p_k \bar{Z}_k - \sum_{k=1}^{N_T} \bar{Z}_k, \tag{27}$$

$$= \frac{1}{\mathbb{E}[p_k]} U_s^\dagger G U_s - U_s^\dagger U_s, \tag{28}$$

$$= \frac{1}{\mathbb{E}[p_k]} U_s^\dagger G U_s - I_s. \tag{29}$$

We will now use the Matrix Bernstein Inequality on Z_k to find lower bounds for the eigenvalues of the first term in (29). We will then use these in (23) to bound the estimation error.

Lemma 6.1: *[Matrix Bernstein Inequality [16, Ch.8]] Let $V_1, \dots, V_M \in \mathbb{C}^s$ be independent zero-mean Hermitian random matrices. Assume that $\|V_l\| \leq \mu_V$, $\forall l \in \{1, \dots, M\}$ almost surely. Let $\varrho_V \doteq \|\sum_{l=1}^{M} \mathbb{E}[V_l^2]\|$. Then, for $t > 0$*

$$\mathbb{P}(\|\sum_{l=1}^{M} V_l\| \geq t) \leq f_{bt}(\mu_V, \varrho_V, t) \leq f_{bn}(\mu_V, \varrho_V, t) \tag{30}$$

with $f_{bt}(.)$ and $f_{bn}(.)$ as defined in (6)-(7).

We note that Z_k in (26) are statistically independent random matrices with $\mathbb{E}[Z_k] = 0$. We bound the spectral norm of Z_k as follows

$$\|Z_k\| \leq \max_k \left| \frac{p_k}{\mathbb{E}[p_k]} - 1 \right| \|\bar{Z}_k\|. \tag{31}$$

Here we have the following bound for $\|\bar{Z}_k\|$

$$\|\bar{Z}_k\| = \|\sum_{t=1}^{Q} u_{(k-1)Q+t} u_{(k-1)Q+t}^\dagger\|, \tag{32}$$

$$\leq Q \max_{k,t} \|u_{(k-1)Q+t} u_{(k-1)Q+t}^\dagger\|, \tag{33}$$

$$= Q \max_{k,t} \|u_{(k-1)Q+t}\|^2, \tag{34}$$

$$= Q\frac{s}{N}, \tag{35}$$

where (35) follows since U is the DFT matrix. We also have

$$\|\bar{Z}_k\| \leq \|\sum_{k=1}^{N_T} \bar{Z}_k\| = \|I_s\| = 1, \tag{36}$$

where (36) follows from the fact that for $A \succeq 0$ and $B \succeq 0$, $\lambda_{max}(A) \leq \lambda_{max}(A+B)$. Combining (35) and (36), we have

$$\|\bar{Z}_k\| \leq \min\{Qs/N, 1\}. \tag{37}$$

We now consider the term with p_k in (31). We note that by (2) and (3), $p_k = \bar{E}_k \frac{1}{Q}\frac{N}{P_x}$, hence $p_k = \frac{1}{Q}\sum_{t=1}^{Q} \delta_{(k-1)Q+t} \bar{E}_0$ where $\bar{E}_0 = E_0 \frac{N}{P_x}$. Here $\mathbb{E}[p_k] = p\bar{E}_0$. Hence

$$\max_k \left|\frac{p_k}{\mathbb{E}[p_k]} - 1\right| \leq \max\{\frac{\max_k p_k}{\mathbb{E}[p_k]} - 1, 1\} = \max\{\frac{1}{p} - 1, 1\}. \tag{38}$$

Hence by (31), (37) and (38)

$$\|Z_k\| \leq \max\{1/p-1, 1\} \min\{Qs/N, 1\} \doteq \mu_I, \quad \forall k. \quad (39)$$

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

$$\|\sum_{k=1}^{N_T} \mathbb{E}[Z_k^2]\| = \mathbb{E}[\tilde{p}_k^2] \, \|\sum_{k=1}^{N_T} (\bar{Z}_k)^2\|. \quad (40)$$

Here the spectral norm term can be bounded as

$$\|\sum_{k=1}^{N_T} (\bar{Z}_k)^2\| \leq \max_k \|\bar{Z}_k\| \|\sum_{k=1}^{N_T} \bar{Z}_k\|, \quad (41)$$

$$\leq \max_k \|\bar{Z}_k\|, \quad (42)$$

$$\leq \min\{Qs/N, 1\}, \quad (43)$$

where (41) follows from the fact that $Z_k \succeq 0$, see for instance [17, Sec. 2], (42) follows from $\|\sum_{k=1}^{N_T} \bar{Z}_k\| = \|I_s\| = 1$ and (43) follows from (37).

We now consider the term with $\tilde{p}_k^2$ in (40)

$$\mathbb{E}[\tilde{p}_k^2] = \mathbb{E}[(\frac{p_k}{\mathbb{E}[p_k]} - 1)^2] = \mathbb{E}[\frac{p_k^2}{\mathbb{E}[p_k]^2} - 2\frac{p_k}{\mathbb{E}[p_k]} + 1], \quad (44)$$

$$= \frac{\mathbb{E}[p_k^2]}{\mathbb{E}[p_k]^2} - 1. \quad (45)$$

We have

$$\mathbb{E}[p_k^2] = \mathbb{E}[(\frac{1}{Q} \sum_{t=1}^{Q} \delta_{(k-1)Q+t} \bar{E}_0)^2], \quad (46)$$

$$= (\frac{\bar{E}_0}{Q})^2 (\sum_{t=1}^{Q} \mathbb{E}[\delta_t^2] + \sum_{1 \leq t, l \leq Q, \; t \neq l} \mathbb{E}[\delta_t \delta_l]), \quad (47)$$

$$= (\frac{\bar{E}_0}{Q})^2 (Qp + (Q^2 - Q)p^2). \quad (48)$$

Using (45) and (48), we obtain $\mathbb{E}[\tilde{p}_k^2] = \frac{1}{Q}(\frac{1}{p} - 1)$. Hence the variance term in (40) can be bounded as

$$\|\sum_{k=1}^{N_T} \mathbb{E}[Z_k^2]\| \leq \frac{1}{Q}(\frac{1}{p} - 1) \min\{Qs/N, 1\} \doteq \varrho_I. \quad (49)$$

Using (39), (49) and the Matrix Bernstein Inequality reveals that $\|\sum_{k=1}^{N_T} Z_k\| < t$ holds with probability greater than $1 - f_{bt}(\mu_I, \varrho_I, t)$. We note that for Hermitian A, $\|A - I\| < t$ implies $\lambda_{min}(A) > (1-t)$. Therefore, using (29),

$$\lambda_{min}(U_s^\dagger G U_s) > (1-t)\mathbb{E}[p_k] = (1-t)p\frac{N}{P_x} E_0 \quad (50)$$

with probability greater than $1 - f_{bt}(\mu_I, \varrho_I, t)$. Using this eigenvalue bound in (23) leads to the bounds in (8)-(11).

We now consider the second set of bounds given in (14)-(17). We first consider the probability that $p_k \geq \gamma \bar{E}_0$

$$\bar{p} \doteq \mathbb{P}(p_k \geq \gamma \bar{E}_0) = \mathbb{P}(\frac{1}{Q} \sum_{t=1}^{Q} \delta_{(k-1)Q+t} \geq \gamma) \quad (51)$$

for $\gamma \in [0, 1]$. We define a new Bernoulli random variable $\bar{\delta}_k = \mathbb{1}_{p_k \geq \gamma \bar{E}_0}$, where $\mathbb{1}$ is the indicator function. Hence $\mathbb{P}(\bar{\delta}_k = 1) = \bar{p}$ and $\mathbb{P}(\bar{\delta}_k = 0) = 1 - \bar{p}$. Let $\bar{p}_k = \gamma \bar{E}_0 \bar{\delta}_k$. We note that $\bar{p}_k$ provides a lower bound for $p_k, \forall k$. Hence the minimum eigenvalue of $\sum_{k=1}^{N_T} \bar{p}_k \bar{Z}_k$ provides a lower bound for the minimum eigenvalue of $\sum_{k=1}^{N_T} p_k \bar{Z}_k$. Now re-iterating the steps for the proof of bounds in (8)-(11) reveals a set of bounds similar to (8)-(11), but that also depends on γ. By expressing the value of $\bar{p}$ in (51) terms of success probabilities of a binomial variable, we arrive at the bounds in (12)-(17).

VII. Proof of Lemma 3.1

Since the error is a decreasing function of $\mathrm{tr}[G]$, we have $\mathrm{tr}[G] = P_T$ for an optimum G. Let us consider a point G, $\mathrm{tr}[G] = \mathrm{tr}[U_s^\dagger G U_s] = P_T$. Let $G_a = (P_T/N) I_N$. We note that $\mathrm{tr}[G_a] = \mathrm{tr}[U_s^\dagger G_a U_s] = P_T$ and G_a is also feasible. Since $\varepsilon(G)$ is a symmetric and convex function of $\lambda_i(U_s^\dagger G U_s)$, $\varepsilon(G) \geq \varepsilon(G_a)$ [18]. Hence G_a provides a lower bound for the performance of all such G and it is an optimal strategy.

References

[1] D. Gündüz, K. Stamatiou, N. Michelusi, and M. Zorzi, "Designing intelligent energy harvesting communication systems," *IEEE Communications Magazine*, vol. 52, no. 1, pp. 210–216, 2014.

[2] O. Ozel and S. Ulukus, "Achieving AWGN capacity under stochastic energy harvesting," *IEEE Trans. Inf. Theory*, vol. 58, pp. 6471–6483, Oct 2012.

[3] 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.

[4] 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.

[5] M. Nourian, S. Dey, and A. Ahlen, "Distortion Minimization in Multi-Sensor Estimation With Energy Harvesting," *IEEE J. Sel. Areas Commun.*, vol. 33, pp. 524–539, Mar. 2015.

[6] 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.

[7] 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, pp. 2128–2143, Aug 2014.

[8] 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.

[9] 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.

[10] R. M. Gray, *Toeplitz and Circulant Matrices: a Review*. Now Publishers Inc., 2006.

[11] A. Tulino, G. Caire, S. Verdu, and S. Shamai, "Support recovery with sparsely sampled free random matrices," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4243–4271, 2013.

[12] A. Özçelikkale, S. Yüksel, and H. Ozaktas, "Unitary precoding and basis dependency of MMSE performance for Gaussian erasure channels," *IEEE Trans. Inf. Theory*, vol. 60, pp. 7186–7203, Nov 2014.

[13] M. Gastpar, "Uncoded transmission is exactly optimal for a simple Gaussian sensor network," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5247–5251, 2008.

[14] 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.

[15] B. D. O. Anderson and J. B. Moore, *Optimal filtering*. Prentice-Hall, 1979.

[16] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. Springer, 2013.

[17] J. A. Tropp, "The expected norm of a sum of independent random matrices: An elementary approach," *arXiv preprint*, 2015.

[18] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*. Academic Press, 1979.