

# Transmission Strategies for Remote Estimation under Energy Harvesting Constraints

Ayça Özçelikkale, Tomas McKelvey, Mats Viberg

**Abstract**—We consider the remote estimation of a time-correlated field using an energy harvesting (EH) sensor. The sensor observes the unknown field and communicates its observations to a remote fusion center using an amplify-forward strategy. We consider the design of optimal transmission strategies in order to minimize the mean-square error (MSE) at the fusion center. Contrary to traditional approaches, the degree of correlation between the field values constitutes an important aspect of our formulation. We provide the optimal power allocation strategies for a number of illustrative scenarios, including the circularly wide-sense stationary (c.w.s.s.) signals with static correlation coefficient and the sampled low-pass c.w.s.s. signals. Based on these results, we propose low-complexity policies for the general case. Numerical evaluations illustrate the performance of the optimal and the low-complexity policies.

## I. INTRODUCTION

Energy harvesting solutions offer a promising framework for future wireless sensor networks. Instead of completely relying on a fixed battery, sensors with EH capabilities can collect energy from the environment, such as solar power or mechanical vibrations. In addition to enabling energy autonomous sensing systems, EH capabilities also offer prolonged network lifetimes and significant mobility for the nodes in the network [1].

One of the key issues in the design of EH systems is the intermittent nature of the energy supply. Here, due to unreliable nature of the energy supply, the EH device has to find the optimal trade-off between using all the available energy for the current operations and saving all of it for the future. In that respect, reliable communications with EH nodes have been studied under a broad range of scenarios under rate maximization criterion [1–3]. Here we adopt an alternative approach and focus on the estimation aspect of problem, i.e. recovery of the unknown field measured by the sensors.

Currently the treatment of the estimation aspect, in particular the effect of the possible statistical correlation between the unknown field values, for the EH sensing systems is quite limited. Optimal strategies for the case of i.i.d. Gaussian sources follow from the findings of [3]. The parameter estimation problems considered in [4], [5] provide insights about the limiting case where the unknown value is fully spatially correlated across sensors. Investigations in [6–8] provide guidelines for Markov sources. The scenario with two correlated Gaussian variables is studied in [9].

A. Özçelikkale, T. McKelvey and M. Viberg are with the 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.

Here we address this issue by focusing on the estimation of a time-correlated field using an EH sensor. The EH sensor observes the unknown field and communicates its observations to the remote fusion center using an amplify-forward strategy under energy harvesting constraints. We consider the problem of design of transmission strategies in order to minimize the MSE at the fusion center. We adopt the off-line optimization scheme, i.e. the scheme where the sensor has acausal access to arrival information for the energy packets. We provide the optimal power allocation strategies for a number of illustrative scenarios. In particular, we show that most majorized power allocation strategies, i.e. power allocations as uniform as possible, are optimal regardless of the degree of correlation in the case of c.w.s.s. signals with static correlation coefficient and sampled low-pass c.w.s.s. signals. Motivated by these results, we propose low-complexity policies for the general case. Numerical evaluations illustrate the trade-offs offered by the effective degrees of freedom of the signal and the performance of the proposed policies.

The rest of the paper is organized as follows. We present the problem formulation in Section II. In Section III, the optimal strategies for a number of illustrative scenarios are provided. In Section IV, low-complexity strategies for the general case are proposed. Numerical evaluations are provided in Section V. The paper is concluded in Section VI.

*Notation:* The complex conjugate transpose of a matrix  $A$  is denoted by  $A^\dagger$ . The  $i^{\text{th}}$  row,  $k^{\text{th}}$  column element of a matrix  $A$  is denoted by  $[A]_{tk}$ . The positive semi-definite (p.s.d.) ordering is denoted by  $\succeq$ .  $I_n \in \mathbb{C}^{n \times n}$  denotes the identity matrix.

## II. SYSTEM MODEL AND PROBLEM STATEMENT

### A. Signal Model

The aim of the remote estimation system is to estimate the unknown complex proper zero-mean 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  $K_{\mathbf{x}} = \mathbb{E}[\mathbf{x}\mathbf{x}^\dagger]$ ,  $P_x \triangleq \text{tr}[K_{\mathbf{x}}]$ . Let  $s$  be the number of non-zero eigenvalues of  $K_{\mathbf{x}}$ , i.e. rank of  $K_{\mathbf{x}}$ . Let  $\Omega$  denote the index set of non-zero eigenvalues. Here  $K_{\mathbf{x}} = U_\Omega \Lambda_{x,s} U_\Omega^\dagger$  is the reduced singular value decomposition (s.v.d.) of  $K_{\mathbf{x}}$  where  $\Lambda_{x,s} \in \mathbb{C}^{s \times s}$  is the diagonal matrix of non-zero eigenvalues and  $U_\Omega \in \mathbb{C}^{n \times s}$  is the sub-matrix formed by the columns of  $U$  corresponding to the non-zero eigenvalues.

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

At time slot  $t$ , the sensor measures  $x_t$ , the field value at time  $t$  and communicates it to the fusion center as follows:

$$y_t = \sqrt{a_t} x_t + w_t, \quad t = 1, \dots, n \quad (1)$$

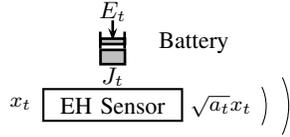


Fig. 1: Energy Harvesting Sensor

where  $\sqrt{a_t}$ ,  $y_t$  and  $w_t$  denote the amplification factor adopted by the sensor, the received signal at the fusion center, and the channel noise respectively. Here  $\mathbf{w} = [w_1, \dots, w_n] \in \mathbb{C}^{n \times 1}$  is complex proper zero-mean Gaussian with  $\mathbf{w} \in \mathbb{C}^{n \times 1} \sim \mathcal{CN}(0, K_{\mathbf{w}})$ ,  $K_{\mathbf{w}} = \sigma_w^2 I_n$ .

### C. Energy Constraints at the Sensor

The average energy used by the sensor during transmission of  $x_t$  can be written as follows:

$$J_t = \tau \mathbb{E}[|\sqrt{a_t}x_t|^2] = \tau a_t \sigma_{x_t}^2, \quad (2)$$

where the transmit duration  $\tau = 1$  in the rest of the paper. At each time slot  $t$ , an energy packet of  $E_t$  arrives to the battery. The sensor operates under the following energy neutrality conditions

$$\sum_{l=1}^t J_l \leq \sum_{l=1}^t E_l, \quad t = 1, \dots, n. \quad (3)$$

where the initial energy at the battery is zero. These conditions ensure that the energy used at any time does not exceed the available energy. Here we consider a device with a large enough battery capacity so that no energy packet  $E_t$  has to be dropped.

### D. Estimation at the Fusion Center

After receiving  $\mathbf{y} = [y_1, \dots, y_n] \in \mathbb{C}^{n \times 1}$ , the fusion center forms the minimum MSE (MMSE) estimate of  $\mathbf{x}$ , i.e.  $\hat{\mathbf{x}} = \mathbb{E}[\mathbf{x}|\mathbf{y}]$ . The resulting MMSE can be expressed as [10, Ch2]

$$\begin{aligned} \varepsilon(A) &= \text{tr}[K_{\mathbf{x}} - K_{\mathbf{x}\mathbf{y}}K_{\mathbf{y}}^{-1}K_{\mathbf{y}\mathbf{x}}^\dagger] \\ &= \text{tr} \left[ (\Lambda_{x,s}^{-1} + \gamma U_{\Omega}^\dagger \text{diag}(a_t) U_{\Omega})^{-1} \right], \end{aligned} \quad (4) \quad (5)$$

where  $K_{\mathbf{x}\mathbf{y}} = \mathbb{E}[\mathbf{x}\mathbf{y}^\dagger] = K_{\mathbf{x}}A^\dagger$ ,  $K_{\mathbf{y}} = AK_{\mathbf{x}}A^\dagger + K_{\mathbf{w}}$ ,  $A = \text{diag}(\sqrt{a_t}) \in \mathbb{R}^{n \times n}$  and  $\gamma \triangleq 1/\sigma_w^2$ . Here we have used Sherman-Morrison-Woodbury identity [11]. We note that in order to perform the above MMSE estimation procedure, the fusion center needs to know the source and the noise statistics, including the covariance matrices and  $a_t$ 's. We also note that by adopting a second-order analysis framework and using the optimum linear MMSE filter instead of the MMSE filter at the fusion center, the above error analysis can be also performed under non-Gaussian statistics [10].

### E. Design of Optimal Transmission Strategies

Our goal is to design the optimal transmission strategies in order to minimize the MMSE as follows

$$\min_A \varepsilon(A) \quad (6a)$$

$$\text{s.t.} \sum_{l=1}^t a_l \sigma_{x_l}^2 \leq \sum_{l=1}^t E_l, \quad t = 1, \dots, n-1, \quad (6b)$$

$$\sum_{l=1}^n a_l \sigma_{x_l}^2 = E_{tot}, \quad (6c)$$

where (6b)-(6c) follow from (2), (3) and we have  $E_{tot} \triangleq \sum_{l=1}^n E_l$ ;  $a_t \geq 0, \forall t$ . Since for any optimum strategy all the available energy should be used, (6c) is stated as an equality. Here we consider the scenario where the sensor has access to the energy realization for a look-ahead window of size  $n$ , i.e. off-line optimization [1]. This type of approaches are well-suited for the scenarios where these arrivals can be accurately predicted, such as radio frequency EH systems with dedicated power transfer scheduling. They also serve as benchmarks and provide structural guidelines for the general case.

The objective function of (6) is a convex function since  $\text{tr}[X^{-1}]$  is convex for  $X \succ 0$ . The constraints form convex constraints since they are in the form of linear inequalities and equalities. Hence (6) is a convex formulation and Karush-Kuhn-Tucker (KKT) conditions are necessary and sufficient for optimality under the assumption of a strictly feasible point. Optimal solutions can be found using the standard numerical optimization tools, such as SDPT3, SeDuMi and CVX [12–14]. In Section III and Section IV, we provide analytical solutions that reveal the structure of the optimal power allocations for a number of cases and propose low-complexity policies, respectively. Numerical evaluations are provided in Section V.

## III. OPTIMAL TRANSMISSION POLICIES

Here we discuss the structure of the solutions for a number of illustrative scenarios. These results motivate the low-complexity policies proposed in Section IV.

### A. Uncorrelated Sources

Here we consider the case where the components of  $\mathbf{x}$  are uncorrelated, hence  $K_x = \text{diag}(\sigma_{x_t}^2)$ ,  $\sigma_{x_t}^2 > 0$ . The MMSE can be expressed as follows:

$$\varepsilon(A) = \sum_{t=1}^n \frac{\sigma_{x_t}^2}{1 + \gamma \sigma_{x_t}^2 a_t}. \quad (7)$$

The Lagrangian can be expressed as follows:

$$\mathcal{L} = \sum_{t=1}^n \frac{\sigma_{x_t}^2}{1 + \gamma \sigma_{x_t}^2 a_t} + \sum_{T=1}^{n-1} \eta_T W_T + \nu W_n - \sum_{t=1}^n \mu_t a_t, \quad (8)$$

where  $W_L = \sum_{t=1}^L \sigma_{x_t}^2 a_t - \sum_{t=1}^L E_t$ ,  $1 \leq L \leq n$ . Here  $\eta_T \in \mathbb{R}$ ,  $\eta_T \geq 0$ ,  $1 \leq T \leq n-1$ ,  $\nu \in \mathbb{R}$  and  $\mu_t \in \mathbb{R}$ ,  $\mu_t \geq 0$ ,  $1 \leq t \leq n$  are the Lagrange multipliers. Solving the KKT conditions reveals that the optimal  $a_t$  can be expressed in terms of a water-filling type solution  $a_t = (\sqrt{\frac{1}{\kappa_t \gamma \sigma_{x_t}^2}} - \frac{1}{\gamma \sigma_{x_t}^2})^+$  where  $\kappa_t \triangleq \sum_{T=t}^{n-1} \eta_T + \nu$  is a time-index dependent threshold, which is a typical property of the EH solutions [2]. We now focus on the solution structure in the i.i.d. case:

1) *I.I.D. Sources*: Here  $K_x = \sigma_x^2 I_n$ , hence the MMSE can be expressed as follows:

$$\varepsilon(A) = \sum_{t=1}^n \frac{\sigma_x^2}{1 + \gamma \sigma_x^2 a_t}. \quad (9)$$

In this case, an optimal strategy can be found by adopting the arguments of [3]. More precisely, we note the following:

**Definition 3.1:** [15, Ch.1] Let  $\mathbf{a} = [a_1, \dots, a_n] \in \mathbb{R}^n$  and  $\mathbf{b} = [b_1, \dots, b_n] \in \mathbb{R}^n$ . Then  $\mathbf{a}$  is said to be majorized by

$b$  if the following holds: i)  $\sum_{t=1}^k a_{[t]} \leq \sum_{t=1}^k b_{[t]}$ ,  $k = 1, \dots, n-1$ ; ii)  $\sum_{t=1}^n a_{[t]} = \sum_{t=1}^n b_{[t]}$ . Here  $a_{[t]}$  denotes the components of  $a$  in decreasing order, i.e.  $a_{[1]} \geq \dots \geq a_{[n]}$ . This majorization relationship is denoted by  $a \prec b$ .

Majorization can be interpreted as a measure of how balanced or uniform the distribution of the components of vectors are. In particular, the following relationship holds  $\forall a \in \mathbb{R}^n$ :  $\bar{a} \prec a$ , where  $\bar{a} = (1/n)(\sum_{t=1}^n a_t)[1, \dots, 1] \in \mathbb{R}^n$ . The following is of interest:

**Definition 3.2:** [15, Ch.3] Let us have  $\mathcal{S} \subseteq \mathbb{R}^n$  and  $f(\cdot) : \mathcal{S} \rightarrow \mathbb{R}$ . Then  $f(\cdot)$  is said to be Schur-convex on  $\mathcal{S}$  if  $a \prec b$  on  $\mathcal{S}$  implies  $f(a) \leq f(b)$ .

**Lemma 3.1:** [15, Ch.3] Let  $\mathcal{S} \subseteq \mathbb{R}^n$ , and  $g(\cdot) : \mathcal{S} \rightarrow \mathbb{R}$  be convex. Then  $f(a) = \sum_{t=1}^n g(a_t)$  is Schur-convex.

By Lemma 3.1, (9) is Schur-convex since  $g(a_t) = \frac{\sigma_x^2}{1+\gamma\sigma_x^2 a_t}$  is a convex function of  $a_t \geq 0$ . Hence an optimal solution is given by  $a_t$  that is majorized by all feasible power allocations, i.e. the strategy as balanced/uniform as possible. Characterization of such solutions have been studied in relation to maximizing the rate function in [3]:

**Lemma 3.2:** [3, Thm.3] The power allocation that is majorized by all feasible solutions of (6b), (6c), can be characterized as follows:

$$\bar{a}_r = \frac{\bar{E}_{\tau_k} - \bar{E}_{\tau_{k-1}}}{\tau_k - \tau_{k-1}}, \quad r = \tau_{k-1} + 1, \dots, \tau_k \quad (10)$$

$$\tau_k = \arg \min_{r \in \{\tau_{k-1}+1, \dots, \bar{\tau}\}} \frac{\bar{E}_r - \bar{E}_{\tau_{k-1}}}{r - \tau_{k-1}}, \quad k = 2, \dots, K \quad (11)$$

where  $1 \leq r \leq n$ ,  $\tau_1 = 0$  and  $\bar{\tau} = \tau_{K+1} = n$ , and  $1 \leq K \leq n$  is the number of constant power sections.

Here we have adopted the notation  $\bar{E}_r = \sum_{t=1}^r E_t/\sigma_x^2$ ,  $\bar{a}_r = a_t$  for later notational convenience. Due to Schur-convexity of (9), Lemma 3.2 also provides an optimal strategy for the minimization of the MSE in (9). In the subsequent sections, we will utilize this characterization to provide optimal solutions in scenarios even when the source is not i.i.d.

### B. Parameter Estimation

We refer to the scenario where  $K_x$  is of rank 1, hence there is effectively only one random variable to be estimated, as the parameter estimation scenario. Hence (5) can be expressed as

$$\varepsilon(A) = \frac{1}{1/P_x + \gamma \sum_{t=1}^n |[U]_{tk}|^2 a_t} = \frac{1}{1 + \gamma \sum_{t=1}^n \sigma_{x_t}^2 a_t} P_x,$$

where we have used  $|[U]_{tk}|^2 P_x = \sigma_{x_t}^2$ . Since  $\sum_{t=1}^n \sigma_{x_t}^2 a_t = E_{tot}$  by (6c), any feasible strategy is an optimum strategy including the most uniform strategy given by (10)-(11). The optimum error value is given by  $(1 + \gamma E_{tot})^{-1} P_x$ .

### C. A Lower Bound

We will now consider a lower bound on the performance and then utilize this lower bound to propose optimal policies. We consider the following setting:

$$\varepsilon_{LB} = \min_A \varepsilon(A) \quad (12)$$

subject to (6c). Compared to (6), here only the total energy constraint is imposed. Hence (12) forms a relaxation of (6) and the optimum value of (12) provides a lower bound for the optimum value of (6).

We focus on the case where  $\Lambda_{x,s}$  is of the form  $\Lambda_{x,s} = \frac{P_x}{s} I_s$  i.e. the non-zero eigenvalues are all equal. This type of models have been used to represent signal families with a low degree of freedom in various signal applications, for instance as a sparse signal model in compressive sensing literature [16]. We obtain the following analytical expression for  $\varepsilon_{LB}$ :

**Lemma 3.3:** Let  $\Lambda_{x,s} = (P_x/s)I_s$ . Then  $a_t = E_{tot}/P_x$ ,  $\forall t$  is an optimum strategy for (12). The optimal value is given by  $\varepsilon_{LB} = \frac{1}{1+\gamma E_{tot}/s} P_x$ .

The proof is presented in Section VII-A. Hence whenever  $a_t = E_{tot}/P_x$  is a feasible allocation for (6), it is also an optimal strategy. Hence we obtain the following result:

**Corollary 3.1:** Let  $\Lambda_{x,s} = (P_x/s)I_s$ . If  $\frac{1}{P_x} \sum_{l=1}^t \sigma_{x_l}^2 \leq \frac{1}{E_{tot}} \sum_{l=1}^t E_l$ ,  $\forall t$ , then  $a_t = E_{tot}/P_x$  is an optimum strategy for (6) with the optimal value  $\frac{1}{1+\gamma E_{tot}/s} P_x$ .

We note that the conditions of Corollary 3.1 are always satisfied for c.w.s.s. signals with  $\Lambda_{x,s} = (P_x/s)I_s$  under a constant energy arrival scheme,  $E_t = E$ ,  $\forall t$ . Hence the lower bound presented in Lemma 3.3 is achieved even under the energy causality constraints in such scenarios.

### D. Circularly Wide-Sense Stationary Signals

We now focus on the c.w.s.s. signals, which constitute a finite dimensional analog of wide-sense stationary signals [17], [18]. By definition, the covariance matrix is circulant and the unitary matrix  $U$  in the s.v.d. of  $K_x$  is given by the DFT matrix [17], [18]. Let  $F^n$  denote the DFT matrix of size  $n \times n$ , i.e.  $[F^n]_{tk} = (1/\sqrt{n}) \exp(-j\frac{2\pi}{n}(t-1)(k-1))$ ,  $1 \leq t, k \leq n$ . Hence the reduced s.v.d. of  $K_x$  is given by  $K_x = F_\Omega^n \Lambda_{x,s} F_\Omega^{n\dagger}$ , where  $\Lambda_{x,s} = \text{diag}(\lambda_k) \in \mathbb{R}^{s \times s}$ ,  $F_\Omega^n \in \mathbb{C}^{n \times s}$  is the matrix that consists of  $s$  columns of  $F^n$  corresponding to non-zero eigenvalues. Here due to circulant structure,  $\sigma_{x_t}^2 = \sigma_x^2 = P_x/n$ ,  $\forall t$ .

Let  $e_j \in \mathbb{R}^n$ ,  $1 \leq j \leq n$  denote the  $j^{\text{th}}$  unit vector. We obtain the following result, which we will utilize later:

**Lemma 3.4:** Let the s.v.d. of  $K_x$  be given by  $K_x = F^n \Lambda_x F^{n\dagger}$  with  $\Lambda_x = \beta I_n + \alpha e_j e_j^\dagger$  with  $-\beta < \alpha$ ,  $\beta > 0$ ,  $\alpha, \beta \in \mathbb{R}$ . Then (10)-(11) is an optimal strategy for (6).

The proof is given in Section VII-B.

1) *Almost I.I.D. Sources:* When  $x_t$  is i.i.d. distributed, we have  $K_x = \sigma_x^2 I_n$ . Hence the s.v.d. of  $K_x$  is given by  $K_x = U \Lambda_x U^\dagger$  with  $\Lambda_x = \sigma_x^2 I_n$  where  $U$  is an arbitrary unitary matrix. Motivated by this, we refer to the case where  $\Lambda_x \propto I_n - \epsilon e_j e_j^\dagger$ ,  $0 < \epsilon < 1$  as an *almost i.i.d.* source. We obtain the following:

**Corollary 3.2:** Let  $\mathbf{x}$  be almost i.i.d. with  $K_x = F^n \Lambda_x F^{n\dagger}$ ,  $\Lambda_x = I_n - \epsilon e_j e_j^\dagger$ ,  $0 < \epsilon < 1$ . Then (10)-(11) is an optimal strategy for (6).

The result follows from Lemma 3.4. When the source is not exactly i.i.d. but only close to being i.i.d. as defined above, the most uniform feasible allocation is still an optimal solution.

2) *Static Correlation Coefficient:* We now consider signals whose covariance matrix has the following form

$$K(\rho) = \frac{P_x}{n} \begin{bmatrix} 1 & \rho & \dots & \rho \\ \dots & \dots & \dots & \dots \\ \rho & \dots & \dots & 1 \end{bmatrix} \quad (13)$$

where  $K(\rho) \in \mathbb{R}^n$ ,  $0 \leq |\rho| \leq 1$ ,  $\rho \in \mathbb{R}$ . Hence the correlation coefficient between  $x_i$  and  $x_j$ ,  $i \neq j$  does not depend on  $i, j$ . We obtain the following result:

**Lemma 3.5:** *Let  $K_x = K(\rho)$ . Then (10)-(11) is an optimal strategy for (6).*

*Proof:* Let  $v$  be the first row of  $K_x$ , i.e.  $v = (P_x/n)[1, \rho \dots \rho] \in \mathbb{C}^n$  and  $z = [\lambda_1, \dots, \lambda_n] \in \mathbb{R}^n$ . Then  $z = \sqrt{n}F_n v$  [18]. Hence we obtain  $z_1 = (P_x/n)(\rho(n-1)+1)$  and  $z_i = (P_x/n)(1-\rho)$ ,  $i \neq 1$ . Hence Lemma 3.4 applies.  $\square$

Hence regardless of the value of  $\rho$ , i.e. the level of statistical dependency of the signal components, the strategy that allocates the power as uniform as possible is an optimal strategy. Nevertheless, we recall that in the limiting case of parameter estimation, any strategy that spends all the energy is an optimal strategy, hence correlation can be used to compensate for the unreliable nature of the energy arrival process.

3) *Low-Pass Signals:* Let  $n/s \in \mathbb{Z}$ . Let  $\lambda_k$  denote the eigenvalue that corresponds to the eigenvector in the  $k^{\text{th}}$  column of  $F^n$ , where  $F^n$  is as defined above. Here we consider low-pass signals, i.e. signals for which  $\lambda_1, \dots, \lambda_s = P_x/s$ , and the rest are zero. Similar to their deterministic counterparts, given  $\sigma_w^2 = 0$ , these signals can be recovered from their equidistant samples with zero mean-square error when the number of samples is larger than  $s$ , or equivalently the spacing between the samples satisfies  $\Delta \leq n/s$  [16]. Motivated by this, we consider communication strategies that send one out of every  $\Delta$  samples, i.e. strategies in the form of

$$a_t = \begin{cases} \geq 0 & \text{if } t = \Delta r + t_d + 1, \quad 0 \leq r \leq m-1 \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

where  $m = n/\Delta$ , and  $t_d \in 0, \dots, \Delta - 1$ , the initial delay before sending the first data, is fixed.

**Lemma 3.6:** *Let  $\Delta = n/s$ . An optimal strategy for (6) under the setting in (14) is given by (10)-(11) with  $\bar{a}_r \triangleq a_{\Delta r + t_d + 1}$ ,  $\bar{E}_r = \sum_{t=1}^{\Delta r + t_d + 1} E_t / \sigma_x^2$  and  $\tau_1 = 0$ ,  $\bar{\tau} = \tau_{K+1} = s$ , and  $1 \leq K \leq s$ .*

The proof is provided in Section VII-C. This strategy allocates the power as uniform as possible among the  $s$  samples sent. Hence for a sampled low-pass c.w.s.s. signal, the most balanced feasible power allocation is an optimum strategy.

In general, there may be more than one optimal strategy for (6) for c.w.s.s. signals with  $s < n$ . An example is the case of low-pass signals under  $E_t = E \forall t$ . In this scenario, both the uniform power allocation over all components, i.e.  $a_t = E/\sigma_x^2, \forall t$ , and the uniform allocation over the equidistant samples with  $\bar{a}_r = nE/(s\sigma_x^2)$ ,  $\Delta = n/s$ ,  $t_d = \Delta - 1$  are optimal strategies since they both achieve the lower bound in Lemma 3.3. (Here the performance of  $\bar{a}_r$  can be evaluated, for instance, by adopting the arguments in Section VII-C.)

#### IV. LOW-COMPLEXITY TRANSMISSION POLICIES

We now propose a number of heuristic schemes. These schemes provide possibly sub-optimal but nevertheless low-complexity schemes. We consider the following upper bound which avoids the matrix inverse in (5) in the optimization formulation

$$\varepsilon(A) \leq \sum_{t=1}^n \frac{\sigma_{x_t}^2}{1 + \gamma \sigma_{x_t}^2 a_t}, \quad (15)$$

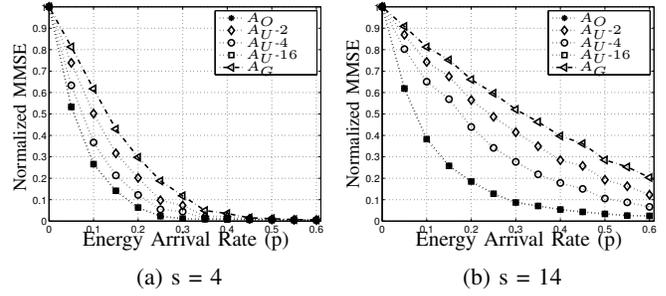


Fig. 2: Normalized MMSE versus energy arrival rate

where the inequality follows from the fact that the right hand side of (15) is the error of the scheme where the possible correlation between the field values are ignored, i.e. the covariance matrix of  $\mathbf{x}$  is assumed to be in the form of  $\text{diag}(\sigma_{x_t}^2)$ . We propose a sliding window approach based on the minimization of this upper bound. Let  $1 \leq l_w \leq n \in \mathbb{Z}$  with  $n/l_w \in \mathbb{Z}$  be the look-ahead window size. Let  $t_i = (i-1)l_w + 1$ . At time index  $t_i$ ,  $i = 1, \dots, n/l_w$ , the sensor looks ahead  $l_w$  time steps and designs the following strategy:

$$\min_{a_{t_i}, \dots, a_{t_{i+1}-1}} \sum_{t=t_i}^{t_{i+1}-1} \frac{\sigma_{x_t}^2}{1 + \gamma \sigma_{x_t}^2 a_t} \quad (16a)$$

$$\text{s.t.} \quad \sum_{l=t_i}^t a_l \sigma_{x_l}^2 \leq \sum_{l=t_i}^t E_l, \quad t = t_i, \dots, t_{i+1} - 2, \quad (16b)$$

$$\sum_{l=t_i}^{t_{i+1}-1} a_l \sigma_{x_l}^2 = \sum_{l=t_i}^{t_{i+1}-1} E_l, \quad (16c)$$

The overall strategy  $a_t, \forall t$  is obtained by solving (16) over  $n/l_w$  non-overlapping windows. We note that using (15) as a performance metric is consistent with the c.w.s.s. signal scenario with the static correlation coefficient, where a balanced power allocation (which is optimal for the uncorrelated case) is an optimal strategy regardless of the correlation of level.

#### V. NUMERICAL RESULTS

We now present the numerical evaluations. Let  $n = 16$ ,  $s = 4, 14$ ,  $P_x = n$ ,  $\gamma = 20\text{dB}$ ,  $\Lambda_{x,s} = \frac{P_x}{\text{tr}[\Lambda]} \Lambda$ ,  $\Lambda = \text{diag}(\alpha_k)$ ,  $\alpha_k = 0.8^k$ ,  $0 \leq k \leq s-1$ . The unitary matrix  $U$  is drawn from the uniform (Haar) unitary matrix distribution [19] and fixed throughout the experiments. The energy arrivals are generated with  $E_t = \delta_t E_0$ ,  $E_0 = 1$  where  $\delta_t$ 's are i.i.d. Bernoulli with probability of success  $p$ ,  $0 \leq p \leq 1$ . Average error over  $N = 200$  energy arrival realizations are reported. The error is normalized as  $\varepsilon/P_x$ .

The error versus energy arrival rate curves are presented in Fig 2. Here  $A_O$  and  $A_{U-l_w}$  refers to the solution of (6) and the solutions provided by (16), respectively. The greedy approach where the energy is spent as soon as it arrives is denoted by  $A_G$ . As expected, due to the low degree of freedom of the signal ( $s = 4$ ) and the possible high correlation between the field values it is possible to obtain lower error values in Fig 2a. In Fig 2a, the gap between the performance of the optimal and the sub-optimal schemes are relatively small compared to Fig 2b. This is consistent with the low degree of freedom of the signal in Fig 2a and the relative insensitivity

of the performance to the energy allocation as suggested by the parameter estimation case. In both scenarios, the low-complexity scheme with the look-ahead window of  $l_w = n$ ,  $A_U - n$ , is remarkably successful so that the performance of  $A_O$  and  $A_U - n$  are not distinguishable from each other in the plots. In the case of Fig 2a, this is again consistent with the insensitivity of the performance to the energy allocation for parameter estimation and the fact that the correlation may have limited effect on the optimal strategies as illustrated by the static correlation coefficient case. In the case of Fig 2b, the close performance of  $A_O$  and  $A_U - l_w$  is supported by the relative closeness of the source to an uncorrelated source due to the relatively high degree of freedom provided by  $s = 14$ .

## VI. CONCLUSIONS

We have considered remote estimation of a time-correlated field using an EH sensor. We have provided the optimal power allocation strategies for a number of scenarios and proposed low-complexity policies for the general case.

## VII. APPENDIX

### A. Proof of Lemma 3.3

We have

$$\varepsilon(A) = \sum_{i=1}^s \frac{1}{1 + \gamma \lambda_i(R_A)} \frac{P_x}{s} \quad (17)$$

where  $R_A = \frac{P_x}{s} U_\Omega^\dagger \text{diag}(a_t) U_\Omega$ . We observe that  $\text{tr}[R_A] = \frac{P_x}{s} \text{tr}[U_\Omega^\dagger \text{diag}(a_t) U_\Omega] = \frac{P_x}{s} \text{tr}[\text{diag}(a_t) U_\Omega U_\Omega^\dagger]$  and hence  $\text{tr}[R_A] = \text{tr}[\text{diag}(a_t) K_x] = \sum_{t=1}^n a_t \sigma_{x_t}^2$ . Here we have used  $\text{tr}[AB] = \text{tr}[BA]$  for matrices with appropriate dimensions and  $K_x = \frac{P_x}{s} U_\Omega U_\Omega^\dagger$ . Hence (6c) is equivalent to the condition  $\text{tr}[R_A] = \sum_i \lambda_i(R_A) = E_{tot}$ . Since (17) is a Schur-convex function of  $\lambda_i(R_A)$ , (17) is lower bounded by the performance of a uniform eigenvalue distribution, i.e.  $\lambda_i = E_{tot}/s$ ,  $i = 1, \dots, s$ . Such an eigenvalue distribution, hence the associated performance is achievable by choosing  $a_t = E_{tot}/P_x$ , since  $(P_x/s) U_\Omega^\dagger \text{diag}(a_t) U_\Omega = (E_{tot}/s) I_s$  where  $U_\Omega^\dagger U_\Omega = I_s$ .

### B. Proof of Lemma 3.4

We rewrite  $\varepsilon(A)$  to show it is a symmetric function of  $a_t$ :

$$\varepsilon(A) = \text{tr} \left[ (\bar{\beta} I_n + \bar{\alpha} e_j e_j^\dagger + \gamma F^{n\dagger} \text{diag}(a_t) F^n)^{-1} \right], \quad (18)$$

$$= \text{tr} \left[ R^{-1} - \frac{R^{-1} \bar{\alpha} e_j e_j^\dagger R^{-1}}{1 + \bar{\alpha} e_j^\dagger R^{-1} e_j} \right] \quad (19)$$

$$= \sum_{t=1}^n \theta_t - \frac{\bar{\alpha}}{1 + \bar{\alpha} \frac{1}{n} \sum_{t=1}^n \theta_t} \frac{1}{n} \sum_{t=1}^n \theta_t^2 \quad (20)$$

where  $\bar{\alpha} = 1/(\alpha + \beta) - 1/\beta$ ,  $\bar{\beta} = 1/\beta > 0$  and  $R = \bar{\beta} I_n + \gamma F^{n\dagger} \text{diag}(a_t) F^n = F^{n\dagger} \text{diag}(\bar{\beta} + \gamma a_t) F^n$  and (19) follows from Sherman-Morrison-Woodbury identity with  $1 + \bar{\alpha} e_j^\dagger R^{-1} e_j \neq 0$  [11]. In (20), we have introduced the notation  $\theta_t = 1/(\bar{\beta} + \gamma a_t)$  and we have used  $R^{-1} = F^{n\dagger} \text{diag}(\theta_t) F^n$ ,  $[R^{-1}]_{ii} = \sum_{t=1}^n \theta_t |F^n|_{it}|^2 = (1/n) \sum_{t=1}^n \theta_t$ ,  $[R^{-2}]_{ii} = (1/n) \sum_{t=1}^n \theta_t^2$  and  $\text{tr}[AB] = \text{tr}[BA]$  for matrices with appropriate dimensions. Here (20) reveals that  $\varepsilon(A)$  is a symmetric function of  $a_t$ . Since  $\varepsilon(A)$  is also a convex function of  $a_t$ , (due to, for instance, (18) and the fact that  $\text{tr}[X^{-1}]$  is convex for  $X \succeq 0$ )  $\varepsilon(A)$  is Schur-convex by [15, Ch.3-Prop.C2]. The result follows from Lemma 3.2.

### C. Proof of Lemma 3.6

Let  $f_n = \exp(-j \frac{2\pi}{n})$ . Here  $F_\Omega^n$  consists of the first  $s$  columns of  $F^n$ . Hence equidistantly row sampled  $F_\Omega^n$  can be associated with the DFT matrix of size  $s$ ,  $F^s$ , as follows  $[F_\Omega^n]_{(n/s)r+t_a+1, k+1} = (1/\sqrt{n}) f_n^{rk} f_n^{t_a k} = \sqrt{s/n} [F^s]_{r+1, k+1} f_n^{t_a k}$ , where  $0 \leq k \leq s-1$ ,  $0 \leq r \leq s-1$ . Let  $D = \text{diag}(d_k)$ ,  $d_k = f_n^{t_a k}$ . Hence

$$\varepsilon(\bar{A}) = \text{tr} \left[ \left( \frac{s}{P_x} I_s + \gamma \frac{s}{n} D F_s^\dagger \bar{A}^\dagger \bar{A} F_s D \right)^{-1} \right], \quad (21)$$

$$= \text{tr} \left[ \left( \frac{s}{P_x} I_s + \gamma \frac{s}{n} \bar{A}^\dagger \bar{A} \right)^{-1} \right] = \sum_{r=0}^{s-1} \frac{1}{\frac{s}{P_x} + \frac{s}{n} \gamma \bar{a}_r}, \quad (22)$$

where  $\bar{A} = \text{diag}(\sqrt{\bar{a}_r}) \in \mathbb{R}^{s \times s}$  and (22) follows from the fact that  $F_s$  and  $D$  are unitary matrices. Due to Lemma 3.1, (22) is a Schur-convex function. The result follows from Lemma 3.2.

## 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, 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.
- [3] O. Ozel and S. Ulukus, "Achieving AWGN capacity under stochastic energy harvesting," *IEEE Trans. Inf. Theory*, vol. 58, pp. 6471–6483, Oct 2012.
- [4] 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.
- [5] S. Knorn, S. Dey, A. Ahlen, 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.
- [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] M. Calvo-Fullana, J. Matamoros, and C. Anton-Haro, "Reconstruction of Correlated Sources with Energy Harvesting Constraints," in *European Wireless Conf. 2015*, pp. 1–6, 2015.
- [9] 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.
- [10] B. D. O. Anderson and J. B. Moore, *Optimal filtering*. Prentice-Hall, 1979.
- [11] H. V. Henderson and S. R. Searle, "On deriving the inverse of a sum of matrices," *SIAM Review*, vol. 23, no. 1, pp. 53–60, 1981.
- [12] R. H. Tütüncü, K. C. Toh, and M. J. Todd, "Solving semidefinite-quadratic-linear programs using SDPT3," *Mathematical Programming*, vol. 95, no. 2, pp. 189–217, 2003.
- [13] J. F. Sturm, "Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11, no. 1–4, pp. 625–653, 1999.
- [14] CVX Research Inc., "CVX: Matlab software for disciplined convex programming 2.0." <http://cvxr.com/cvx>, 2012.
- [15] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*. Academic Press, 1979.
- [16] 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.
- [17] 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.
- [18] R. M. Gray, *Toeplitz and Circulant Matrices: a Review*. Now Publishers Inc., 2006.
- [19] F. Mezzadri, "How to generate random matrices from the classical compact groups," *Notices of the AMS*, vol. 54, pp. 592 – 604, 2007.