SPARSITY-PROMOTING SENSOR SELECTION WITH ENERGY HARVESTING CONSTRAINTS

Miguel Calvo-Fullana[†]*, Javier Matamoros*[†]*, Carles Antón-Haro*[†] *and Sophie M. Fosson*[⋆]

[†]Centre Tecnològic de Telecomunicacions de Catalunya, Barcelona (Spain)
[⋆]Department of Electronics and Telecommunications, Politecnico di Torino (Italy)

ABSTRACT

In this paper, we propose a novel sensor selection scheme for networks equipped with energy harvesting sensing devices. Ultimately, the goal is to minimize the reconstruction distortion at the fusion center by selecting a reduced (i.e., sparse) yet informative enough subset of sensors. The solution must also fulfill the causality constraints associated to the energy harvesting process. For a classical formulation, the optimization problem turns out to be non-convex. To circumvent that, we promote sparsity directly in the power allocation vector by introducing a log-sum penalty term in the cost function. The problem can be iteratively solved by resorting to majorization-minimization procedure leading to a stationary point of the solution. Numerical results reveal that, by using a log-sum penalty term, the sensor selection scheme outperforms others based on the ℓ_1 norm while making an effective use of the harvested energy.

Index Terms— Sensor selection, energy harvesting, sparsity.

1. INTRODUCTION

A Wireless Sensor Network (WSN) is typically composed of spatially distributed sensor nodes acquiring information of a phenomenon of interest and conveying such information to a Fusion Center (FC) for reconstruction. With current technological advances making feasible the deployment of small and inexpensive sensor networks in large numbers, the problem of selecting which subset of them should transmit at a given time naturally arises. This often stems from resource (e.g., bandwidth), interference level or energy consumption constraints, which make massive sensor-to-FC communications barely recommended (or simply not possible). Formally, the problem is that of selecting the best subset of sensors such that a certain optimality criteria is satisfied. In the literature, this is referred to as the *sensor selection problem* and it naturally arises in applications, such as robotics, target tracking, smart grids and others (see [1] and references therein).

While the sensor selection problem is combinatorial in nature, Joshi and Boyd studied in [1] a convex relaxation allowing to (approximately) solve the problem with a reasonable computational cost. More recent approaches have relied in sparsity-promoting techniques for a wide range of scenarios. For instance, the authors in [2] investigate—both from centralized and distributed standpoints—strategies aimed to minimize the number of selected sensors subject to a given Mean Square Error (MSE) target. Non-linear measurement models (such as those in source localization and tracking problems) have been considered in [3], also in a sparsity-promoting framework. Nonetheless, such techniques might lead to situations where only the most informative sensors are being selected. Given the scarce nature of the energy supply of the sensor nodes, such scenarios are undesirable. In order to alleviate this problem, the authors in [4] use an sparsity-promoting penalty function to discourage the repeated selection of the most informative sensor nodes. By doing so, uneven battery drainage can be prevented. Likewise, the same authors propose in [5] a periodic sensor scheduling strategy which limits the number of times that a sensor can be selected in a given period.

Besides, Energy Harvesting (EH) has become a promising technology capable of extending the operational lifetime (or even allowing self-sustainable operation) of WSNs. Instrumental to that, is the design of advanced *power allocation strategies* and sensor scheduling schemes [6–8] making a judicious use of the harvested energy.

In this paper, we study (from an *offline optimization* point of view) the sensor selection problem with energy harvesting constraints. Our objective is the minimization of the reconstruction distortion of a source. As in [2–5] we adopt a sparsity-promoting framework but in addition, we take EH constraints into consideration. Formulating this problem in the classical sensor selection framework leads to a non-convex optimization problem (with its inherent computational complexity) which we studied in [9]. To avoid this, we propose to promote sparsity in the power allocation using a log-sum penalty term, this in turn resulting into sparse sensor selection policies. We provide a majorization-minimization algorithm to find a stationary solution of the problem, which consists in the iterative minimization of a reweighted ℓ_1 penalty function. Finally, we compare the resulting power and sensor selection policies of the reweighted and non-reweighted approaches.

This paper is organized as follows. In Section 2, we introduce the system model. In Section 3, we formulate the sensor selection problem with energy harvesting constraints. Next, an sparsity-promoting approach to the problem is developed in Section 4. In Section 5, some numerical results are provided. Finally, Section 6 provides some concluding remarks regarding this work.

2. SYSTEM MODEL

Consider a wireless sensor network composed of M energy harvesting sensor nodes (with index set $\mathcal{M} \triangleq \{1, \dots, M\}$) and one fusion center deployed to estimate an underlying source $\mathbf{x} \in \mathbb{R}^m$, with $\mathbf{x} \sim \mathcal{N}(0, \mathbf{\Sigma}_x)$. We consider a time-slotted system with T time slots indexed by the set $\mathcal{T} \triangleq \{1, \dots, T\}$ of duration T_s. In time slot t, the stationary source $\mathbf{x}$ generates an independent and identi-

This work was supported by the Catalan government under grant SGR2014-1567; the Spanish government under grants PCIN-2013-027 (E-CROPS), and TEC2013-44591-P (INTENSYV); and the European Commission under grant agreement n.318306 (NEWCOM#). This work was also partially supported by the European Research Council under grant agreement n.279848 (CRISP).

cally distributed (i.i.d.) large sequence of n samples $\{\mathbf{x}^{(k)}[t]\}_{k=1}^{n} = \left\{\mathbf{x}^{(1)}[t], \ldots, \mathbf{x}^{(n)}[t]\right\}$. As in [1], source samples and sensor measurements are related through the following linear model:

$$y_i^{(k)}[t] = \mathbf{a}_i^T \mathbf{x}^{(k)}[t] + w_i^{(k)}[t], \qquad \begin{array}{l} k = 1, \ldots, n \\ i \in \mathcal{Z}_t, \end{array} \tag{1}$$

where $\{w_i^{(k)}[t]\}_{k=1}^{n}$ stands for i.i.d., zero-mean Gaussian observation noise of variance σ_w^2; vector $\mathbf{a}_i$ gathers the *known* coefficients of the linear model at the i-th sensor; and $\mathcal{Z}_t \subseteq \mathcal{M}$ denotes the subset of active (selected) sensors in time slot t, with cardinality $|\mathcal{Z}_t|$. The ultimate goal is to reconstruct at the FC the sequence $\{\mathbf{x}^{(k)}[t]\}_{k=1}^{n}$ in each time slot.

In the sequel, we assume separability of source and channel coding. As far as *source* coding is concerned, we adopt a rate-distortion optimal encoder. Assuming a quadratic distortion measure at the FC, the encoded measurements at the sensor nodes can be modeled as a sequence of auxiliary random variables $\{u_i^{(k)}[t]\}_{k=1}^{n}$ [10]:

$$u_i^{(k)}[t] = \mathbf{a}_i^T \mathbf{x}^{(k)}[t] + w_i^{(k)}[t] + q_i^{(k)}[t], \qquad \begin{array}{l} k = 1, \ldots, n \\ i \in \mathcal{Z}_t, \end{array} \tag{2}$$

with $q_i^{(k)}[t] \sim \mathcal{N}\left(0, \sigma_{q_i}^2[t]\right)$ modeling the i.i.d. encoding noise. The average encoding rate per sample $R_i[t]$ must satisfy the rate-distortion theorem [11], that is,

$$\begin{aligned} R_i[t] &\geq I(y_i[t]; u_i[t]) = h(u_i[t]) - h(u_i[t]|y_i[t]), \\ &= \frac{1}{2} \log\left(1 + \frac{\mathbf{a}_i^T \mathbf{\Sigma}_x \mathbf{a}_i + \sigma_w^2}{\sigma_{q_i}^2[t]}\right) \end{aligned} \tag{3}$$

for all $i \in \mathcal{Z}_t$. Further, we assume that each *active* sensor encodes its observations at the maximum *channel* rate which is given by the Shannon capacity formula[1]. Hence we have $R_i[t] = \frac{1}{2}\log(1 + h_i[t]p_i[t])$, where $p_i[t]$ and $h_i[t]$ stand for the average transmit power and channel gain, respectively. From this and (3), the variance of the encoding noise reads

$$\sigma_{q_i}^2[t] = \frac{\mathbf{a}_i^T \mathbf{\Sigma}_x \mathbf{a}_i + \sigma_w^2}{h_i[t] p_i[t]}, \quad i \in \mathcal{Z}_t. \tag{4}$$

Finally, by means of a Minimum Mean Square Error (MMSE) estimator [12] the FC reconstructs $\{\mathbf{x}^{(k)}[t]\}_{k=1}^{n}$ from the received codewords $\{u_i^{(k)}[t]\}_{k=1}^{n}$, $i \in \mathcal{Z}_t$. The average (MSE) distortion in time slot $t \in \mathcal{T}$ is given by [12]

$$D[t] = \operatorname{tr}\left(\sum_{i=1}^{M} \frac{1}{\sigma_w^2 + \sigma_{q_i}^2[t]} \mathbf{a}_i \mathbf{a}_i^T + \mathbf{\Sigma}_x^{-1}\right)^{-1}, \tag{5}$$

where $\operatorname{tr}(\cdot)$ denotes the trace operator. By substituting expression (4) in (5) and defining $\xi_i[t] \triangleq \left(\frac{\mathbf{a}_i^T \mathbf{\Sigma}_x \mathbf{a}_i / \sigma_w^2 + 1}{h_i[t]}\right)$, we can rewrite the distortion expression (5) as

$$D[t] = \operatorname{tr}\left(\frac{1}{\sigma_w^2} \sum_{i=1}^{M} \frac{p_i[t]}{p_i[t] + \xi_i[t]} \mathbf{a}_i \mathbf{a}_i^T + \mathbf{\Sigma}_x^{-1}\right)^{-1}. \tag{6}$$

[1] For simplicity, we let the number of channel uses per sensor be equal to the number of samples in a time slot.

3. SENSOR SELECTION WITH ENERGY HARVESTING CONSTRAINTS

Since sensor nodes are capable of harvesting energy from the environment, the transmit power is necessarily constrained by the scavenged energy. Thus, at time slot $t \in \mathcal{T}$ we have

$$T_s \sum_{l=1}^{t} p_i[l] \leq \sum_{l=1}^{t} E_i[l], \quad t \in \mathcal{T}, i \in \mathcal{M}. \tag{7}$$

where $E_i[t]$ denotes the energy harvested by the i-th sensor node at time slot t. We aim to minimize the sum distortion over all time slots (6) subject to the energy harvesting constraints (7). Accordingly, the optimization problem reads

$$\begin{aligned} \underset{\mathbf{p}[t]}{\text{minimize}} \quad & \sum_{t=1}^{T} \operatorname{tr}\left(\frac{1}{\sigma_w^2} \sum_{i=1}^{M} \frac{p_i[t]}{p_i[t] + \xi_i[t]} \mathbf{a}_i \mathbf{a}_i^T + \mathbf{\Sigma}_x^{-1}\right)^{-1} && \text{(8a)} \\ \text{subject to} \quad & T_s \sum_{l=1}^{t} p_i[l] \leq \sum_{l=1}^{t} E_i[l], \forall t \in \mathcal{T}, \forall i \in \mathcal{M} && \text{(8b)} \\ & \mathbf{p}[t] \geq \mathbf{0}, \quad \forall t \in \mathcal{T} && \text{(8c)} \end{aligned}$$

where $\mathbf{p}[t] = [p_1[t], \ldots, p_M[t]]^T$ stands for the power allocation vector in a given time slot; $\mathbf{0}$ denotes the all-zeros vector (of appropriate dimension); and vector inequality (8c) is defined elementwise. By, introducing the the auxiliary vector $\mathbf{s}[t] = [s_1[t], \ldots, s_M[t]]^T$, the problem above is equivalent to

$$\begin{aligned} \underset{\mathbf{s}[t], \mathbf{p}[t]}{\text{minimize}} \quad & \sum_{t=1}^{T} \operatorname{tr}\left(\sum_{i=1}^{M} \frac{s_i[t]}{\sigma_w^2} \mathbf{a}_i \mathbf{a}_i^T + \mathbf{\Sigma}_x^{-1}\right)^{-1} && \text{(9a)} \\ \text{subject to} \quad & s_i[t] \leq \frac{p_i[t]}{p_i[t] + \xi_i[t]}, \forall t \in \mathcal{T}, \forall i \in \mathcal{M} && \text{(9b)} \\ & T_s \sum_{l=1}^{t} p_i[l] \leq \sum_{l=1}^{t} E_i[l], \forall t \in \mathcal{T}, \forall i \in \mathcal{M} && \text{(9c)} \\ & \mathbf{p}[t] \geq \mathbf{0}, \quad \forall t \in \mathcal{T} && \text{(9d)} \\ & \mathbf{s}[t] \geq \mathbf{0}, \quad \forall t \in \mathcal{T}. && \text{(9e)} \end{aligned}$$

This is a convex program. Hence, this problem has a global minimizer [13] and can be efficiently solved by optimization packages such as CVX [14]. However, since there are no restrictions regarding the selection of sensors, the optimal solution tends to select all the available sensors available at each time slot so as to minimize the resulting distortion.

Usually (due to constraints such as bandwidth, signaling, interference or complexity), one should attempt to select only a small yet informative subset of sensors at each time slot. In a classical formulation of the sensor selection problem (without energy harvesting), this is done by first introducing a boolean selection variable $z_i[t]$ for each sensor and then relaxing the problem (to render it convex) by letting $z_i[t]$ take values in the real interval $[0, 1]$ [1].

4. SPARSITY-PROMOTING SENSOR SELECTION

In (9), due to the dependence of the distortion on the power allocation $p_i[t]$, introducing a selection variable $z_i[t]$ would lead to the bilinear form $p_i[t]z_i[t]$ in constraint (9b), thus turning (9) into a non-convex optimization problem. To circumvent that, we promote sparsity in the power allocation vector $\mathbf{p}[t]$ itself. The resulting optimization problem thus reads

$$\underset{\mathbf{s}[t],\mathbf{p}[t]}{\text{minimize}} \quad \sum_{t=1}^{T} \operatorname{tr}\left(\sum_{i=1}^{M} \frac{s_i[t]}{\sigma_w^2}\mathbf{a}_i\mathbf{a}_i^T + \mathbf{\Sigma}_x^{-1}\right)^{-1} + \lambda \sum_{t=1}^{T} f(\mathbf{p}[t]) \tag{10a}$$

$$\text{subject to} \quad s_i[t] \leq \frac{p_i[t]}{p_i[t]+\xi_i[t]}, \forall t \in \mathcal{T}, \forall i \in \mathcal{M} \tag{10b}$$

$$T_s \sum_{l=1}^{t} p_i[l] \leq \sum_{l=1}^{t} E_i[l], \forall t \in \mathcal{T}, \forall i \in \mathcal{M} \tag{10c}$$

$$\mathbf{p}[t] \geq \mathbf{0}, \quad \forall t \in \mathcal{T} \tag{10d}$$

$$\mathbf{s}[t] \geq \mathbf{0}, \quad \forall t \in \mathcal{T}, \tag{10e}$$

where $f : \mathbb{R}^M \to \mathbb{R}$ is a sparsity-inducing penalty function and λ is the corresponding sparsity parameter. For the ease of notation, in the sequel we will denote the constraints (10b)-(10e) by the convex set $\mathcal{C}$. Three common penalty functions are illustrated in Figure 1 for the scalar case. Namely, the ℓ_0 norm, the ℓ_1 norm and the log-sum function. Function $f_0 = \|\mathbf{p}[t]\|_0$ merely counts the non-zero elements of the input vector $\mathbf{p}[t]$, which leads to an optimization problem which is combinatorial in nature (and thus intractable). The most common convex (and thus tractable) approximation of f_0 is given by the ℓ_1 norm $f_1(\mathbf{p}[t]) = \|\mathbf{p}[t]\|_1$, which has been shown to provide good performance [15]. In our scenario, however, this results in an *homogeneous* penalization of the allocated power. That is, an increase in power allocation in an already selected sensor will be penalized the same way as an increase in power allocation in a non-selected sensor. This leads to scenarios where only a small subset of the most informative sensors are repeatedly selected without using their total available energy. To circumvent that, we need a better approximation of f_0. In particular, we adopt the well known log-sum penalty function $f_{\log}(\mathbf{p}[t]) = \sum_{i=1}^{M} \log(|p_i[t]| + \epsilon)$, which promotes sparsity more efficiently than the ℓ_1 norm (see e.g., [16, 17]).

However, the log-sum penalty function $f_{\log}$ is concave (see Fig. 1), thus turning the objective function of the optimization problem (10) into a difference of convex functions. Though a global minimizer of this problem cannot be expected to be found without resorting to an exhaustive search, we can find a local minimum of the problem by resorting to a Majorization-Minimization (MM) algorithm [16]. In doing so, we can converge to a stationary solution of the problem (10) by iteratively minimizing a surrogate optimization problem in which, at iteration k, we approximate $f(\mathbf{p}[t])$ in (10a) by its linearization around $\mathbf{p}^{(k-1)}[t]$, that is,

$$\hat{f}_{\log}^{(k)}(\mathbf{p}[t]) = f_{\log}(\mathbf{p}^{(k-1)}[t]) + \nabla f_{\log}(\mathbf{p}^{(k-1)}[t])^T \left(\mathbf{p}[t] - \mathbf{p}^{(k-1)}[t]\right) \tag{11}$$

with $f_{\log}(\mathbf{p}[t]) = \sum_{i=1}^{M} \log(|p_i[t]| + \epsilon)$ and removing the constant terms

$$\hat{f}_{\log}^{(k)}(\mathbf{p}[t]) = \sum_{i=1}^{M} \frac{|p_i[t]|}{|p_i^{(k-1)}[t]| + \epsilon}. \tag{12}$$

By defining the set of weights as

$$w_i^{(k)}[t] = \frac{1}{|p_i^{(k-1)}[t]| + \epsilon}, \tag{13}$$

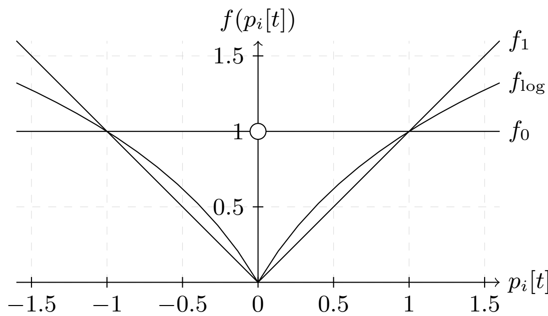

Fig. 1. Sparsity-promoting penalty functions.

Algorithm 1 Reweighted ℓ_1 minimization algorithm.

1: **Initialize:** $\{w_i[t]\} := 1$, set λ and ϵ.
2: **Step 1:** Solve reweighted ℓ_1 problem:
3: $(\mathbf{s}^{(k)}[t], \mathbf{p}^{(k)}[t]) := \underset{\mathbf{s}[t],\mathbf{p}[t]\in\mathcal{C}}{\arg\min} \left\{\sum_{t=1}^{T} D[t] + \lambda\|\mathbf{W}^{(k)}[t]\mathbf{p}[t]\|_1\right\}$
4: **Step 2:** Update weights:
5: $w_i^{(k+1)}[t] := \frac{1}{|p_i^{(k)}[t]|+\epsilon}$
6: **Step 3:** Go to Step 1 until convergence.

which can be more conveniently expressed by the diagonal matrix $\mathbf{W}^{(k)}[t] = \operatorname{diag}(w_1^{(k)}[t], \ldots, w_M^{(k)}[t])$, we can then interpret this as the following reweighted ℓ_1 penalty function

$$\hat{f}_{\log}^{(k)}(\mathbf{p}[t]) = \|\mathbf{W}^{(k)}[t]\mathbf{p}[t]\|_1, \tag{14}$$

where at each iteration, we solve the optimization problem (10), with the penalty function given by (14), and the weights are updated after each iteration according to (13). This procedure is summarized in Algorithm 1.

5. NUMERICAL RESULTS

In this section we assess the performance of our proposed sensor selection scheme. For this purpose, we consider a wireless sensor network composed of $M = 100$ sensors measuring an uncorrelated source (i.e., $\mathbf{\Sigma}_x = \mathbf{I}$) of length $m = 5$. We have $T = 20$ time slots of duration $T_s = 1$ each. The linear combination coefficients are given by $\mathbf{a}_i \sim \mathcal{N}(0, \mathbf{I}/\sqrt{m})$ and the variance of the measurement noise by $\sigma_w^2 = 0.01$. The harvested energies $E_i[t]$ are modeled by means of Poisson processes of common intensity rate $\mu = 1$.Further, we assume non-fading communication channels.

In Figure 2, we compare the resulting distortion of our proposed scheme when using the ℓ_1 norm as a penalty function (f_1) and the reweighted ℓ_1 penalty function ($f_{\log}$). Two different scenarios are compared, consisting of energy arrivals of low and high harvested energy E. As the sparsity parameter λ is not comparable between the two algorithms, we solve the optimization problem (10) for different values of λ and map the resulting distortion to the average number of selected sensors. As expected, the distortion monotonically decreases as the average number of selected sensors increases. More importantly, the reweighted ℓ_1 penalty function clearly outperforms the non-reweighted ℓ_1 norm. Also, note that the gap between the solutions of the two penalty functions becomes broader for scenarios with larger amounts of harvested energy. This is due to the

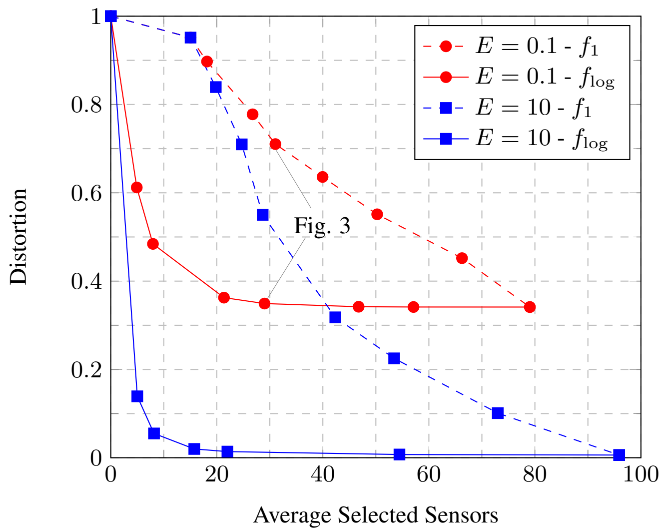

Fig. 2. Distortion vs. Average number of selected sensor for high and low energy scenarios.

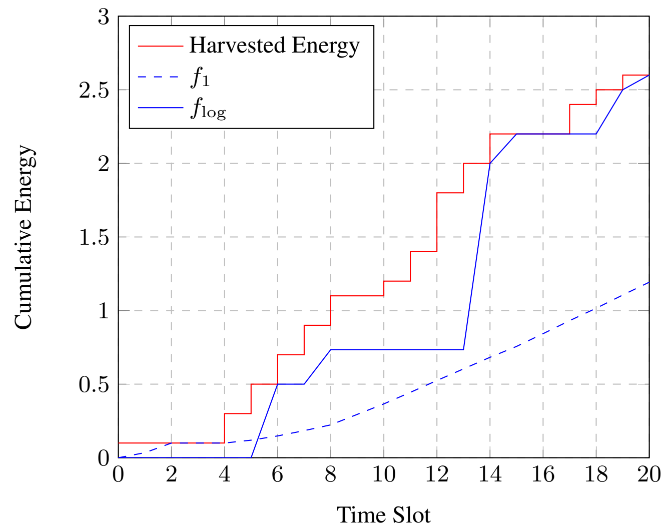

Fig. 3. Power allocation for a single sensor.

linear nature of the ℓ_1 norm and its impact on the resource allocation, as discussed in the next paragraph.

In Figure 3, we depict the power allocation of an individual sensor. This sensor is taken from the selection of approximately 30 sensors and $E = 0.1$ (this scenario is marked in Fig. 2). We observe that, in the solution obtained when using the ℓ_1 penalty function, the sensor node still has a considerable amount of unused energy at the end of the observation period ($t = 20$). A solution leading to lower distortion and the same sensor selection schedule can be found by simply increasing the transmit power during time slots 5 to 20 so as to consume all the available energy by the deadline. On that account, we confirm that the ℓ_1 norm does not lead to good solutions. Also, note how the sensor is selected during most of the time, with the exception of time slots 3 and 4, which is not a very sparse schedule. In contrast, when using the reweighted ℓ_1 penalty function, the sensor allocates all of its available energy by the last time slot and exhibits a more sparse sensor selection schedule, being selected only 6 out of the 20 total time slots.

6. CONCLUSIONS

In this paper, we have investigated the sensor selection problem with energy harvesting. Due to the non-convexity of the classical formulation in this scenario, we have adopted an sparsity-promoting approach to solving this problem. This has been accomplished by introducing a regularization term that promotes sparsity in the power allocation, which in turn leads to sparse sensor selection schedules. Further, we have found that strictly concave penalty functions are desirable in order to ensure the proper consumption of the harvested energy. Specifically, we have proposed the use of a log-sum penalty function, which can be interpreted as a reweighted ℓ_1 norm. Numerical results show that the proposed solution rapidly approaches the asymptotic distortion, by just selecting 20% of the available sensors in average.

7. REFERENCES

[1] Siddharth Joshi and Stephen Boyd, “Sensor selection via convex optimization,” *Signal Processing, IEEE Transactions on*, vol. 57, no. 2, pp. 451–462, 2009.

[2] Hadi Jamali-Rad, Andrea Simonetto, and Geert Leus, “Sparsity-aware sensor selection: Centralized and distributed algorithms,” *Signal Processing Letters, IEEE*, vol. 21, no. 2, pp. 217–220, 2014.

[3] Sundeep Prabhakar Chepuri and Geert Leus, “Sparsity-promoting sensor selection for non-linear measurement models,” *Signal Processing, IEEE Transactions on*, vol. 63, no. 3, pp. 684–698, 2015.

[4] Sijia Liu, Aditya Vempaty, Makan Fardad, Engin Masazade, and Pramod K Varshney, “Energy-aware sensor selection in field reconstruction,” *Signal Processing Letters, IEEE*, vol. 21, no. 12, pp. 1476–1480, 2014.

[5] Sijia Liu, Mohammad Fardad, Pramod K Varshney, and Engin Masazade, “Optimal periodic sensor scheduling in networks of dynamical systems,” *Signal Processing, IEEE Transactions on*, vol. 62, no. 12, pp. 3055–3068, 2014.

[6] Jing Yang and Sennur Ulukus, “Optimal packet scheduling in an energy harvesting communication system,” *IEEE Transactions on Communications*, vol. 60, no. 1, pp. 220–230, 2012.

[7] Kaya Tutuncuoglu and Aylin Yener, “Optimum transmission policies for battery limited energy harvesting nodes,” *IEEE Transactions on Wireless Communications*, vol. 11, no. 3, pp. 1180–1189, 2012.

[8] Omur Ozel, Kaya Tutuncuoglu, Jing Yang, Sennur Ulukus, and Aylin Yener, “Transmission with energy harvesting nodes in fading wireless channels: Optimal policies,” *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 8, pp. 1732–1743, 2011.

[9] Miguel Calvo-Fullana, Javier Matamoros, and Carles Antón-Haro, “Sensor selection in energy harvesting wireless sensor networks,” in *Signal and Information Processing (GlobalSIP), 2015 IEEE Global Conference on*, December 2015.

[10] Prakash Ishwar, Rohit Puri, Kannan Ramchandran, and S Sandeep Pradhan, "On rate-constrained distributed estimation in unreliable sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 4, pp. 765–775, 2005.

[11] Thomas M Cover and Joy A Thomas, *Elements of Information Theory*, John Wiley & Sons, 2012.

[12] Steven M Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*, Prentice Hall, 1993.

[13] Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*, Cambridge University Press, 2009.

[14] Michael Grant and Stephen Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," `http://cvxr.com/cvx`, Mar. 2014.

[15] Robert Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.

[16] Emmanuel J Candes, Michael B Wakin, and Stephen P Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *Journal of Fourier analysis and applications*, vol. 14, no. 5-6, pp. 877–905, 2008.

[17] Hui Zou, "The adaptive lasso and its oracle properties," *Journal of the American statistical association*, vol. 101, no. 476, pp. 1418–1429, 2006.