DECENTRALIZED SPARSITY-PROMOTING SENSOR SELECTION IN ENERGY HARVESTING WIRELESS SENSOR NETWORKS

Miguel Calvo-Fullana, Javier Matamoros and Carles Antón-Haro

Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Barcelona, Spain

ABSTRACT

This paper considers the problem of sensor selection for the estimation of a stochastic source, being the sensor nodes powered by energy harvesting. Therefore, the interest lies in selecting the subset of most informative sensors that transmit their observations to a fusion center (FC). To that end, we propose to minimize the attained distortion at the FC plus a penalization term that promotes sparsity on the power allocation at the sensors. Then, we propose a decentralized algorithm in which the power allocation (and, thus, the selection policy) and distortion minimization problems can be regarded as separated problems. More specifically, the algorithm consists of: (i) a local computation of the power allocation policy, and (ii) a distortion minimization step. Moreover, for the case where sparsity is promoted via the classical ℓ_1 norm, we show that the resulting local power allocation policy can be readily computed by means of a waterfilling-like algorithm.

Index Terms— Sensor selection, energy harvesting, sparsity, wireless sensor networks.

1. INTRODUCTION

Energy Harvesting (EH) has recently emerged as a technology capable of providing self-sustainable and longer lasting wireless networks. As the name suggests, energy harvesting consists in the scavenging of environmental energy, with common sources ranging from solar, thermal or kinetic energy. One area where this idea has shown considerable promise is in Wireless Sensor Networks (WSNs). These networks consist of inexpensive, small and low-power sensors, making them a prime candidate for being EH-powered.

All of this has lead to a great deal of research interest in energy harvesting (See [1] and references therein for an overview of current advances). Research focus ranges from point-to-point scenarios [2] (with various considerations such as finite battery capacity [3], and source-channel coding [4]) to multi-user scenarios, such as the broadcast channel [5] and the multiple access channel [6].

However, one essential characteristic in wireless sensor networks is the availability of a large number of sensor nodes. In this case, it may not be desirable for all the sensors to transmit (i.e., to be selected) at a same time instant. This is usually the result of constraints such as the available bandwidth or the interference generated between nodes. This *sensor selection problem*—as it is commonly known—is in essence a combinatorial problem (with its inherent computational complexity). However, convex relaxations of the problem have been studied and are commonly used [7]. More recently, a sparsity-aware approach has also been taken. Specifically, the authors in [8] study the minimization of the number of selected sensors subject to a given Mean Square Error (MSE) constraint. Other works have also studied this problem for the case of non-linear measurement models [9]. Also in an sparsity-aware framework, and from an energy efficiency point of view, the authors in [10] used a sparsity-promoting penalty function to discourage the repeated selection of any sensor node in particular.

Likewise in [7–10], this paper considers the problem of sensor selection to estimate a stochastic source. However, unlike the previous works, here sensors are assumed to be powered by energy harvesting. In this context, our previous works [11, 12] have focused on the case where the number of selected sensors is constrained by a prescribed value (i.e. maximum number of communication channels) and no sparsity is promoted. In this setting, we provided in [12] an algorithm achieving a stationary solution of the resulting (nonconvex) sensor selection problem. Instead, in this paper (and more in line with [13]), we resort to an sparsity-promoting framework. Notably, we take an *offline* optimization approach, and, unlike in [13], focus on the derivation of a decentralized sparse sensor selection and power allocation scheme. To do so, we resort to the Alternating Directional Method of Multipliers (ADMM). The proposed procedure consists of: (i) a local power allocation problem, in which the sensor nodes conduct a regularization of their power allocation subject to their energy harvesting constraints; and (ii) a distortion minimization step, in which the FC ensures the local decisions of the sensor nodes minimize the distortion. Interestingly, when sparsity is promoted by a weighted ℓ_1 norm, we show that the power allocation policy can be computed locally by means of a directional waterfilling algorithm. Finally, we assess the per-

This work was partially supported by the Catalan government under grant SGR2014-1567; the Spanish government under grant TEC2013-44591-P (INTENSYV), and grant PCIN-2013-027 (E-CROPS) in the framework of the ERA-NET CHIST-ERA program.

formance of the algorithm (i.e., convergence speed and distortion) by means of simulations.

2. SYSTEM MODEL AND PROBLEM STATEMENT

Consider a wireless sensor network composed of M energy harvesting sensor nodes (with index set $\mathcal{M} \triangleq \{1,\ldots,M\}$) 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 $\mathcal{T} \triangleq \{1,\ldots,T\}$ of duration T_s. In time slot t, the stationary source $\mathbf{x}$ generates an independent and identically 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 [7], source samples and sensor measurements are related through the following linear model:

$$y_i^{(k)}[t] = \mathbf{a}_i^T\mathbf{x}^{(k)}[t] + w_i^{(k)}[t], \qquad k = 1,\ldots,n \quad (1)$$

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. 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 $\{u_i^{(k)}[t]\}_{k=1}^n$ [14]:

$$u_i^{(k)}[t] = \mathbf{a}_i^T\mathbf{x}^{(k)}[t] + w_i^{(k)}[t] + q_i^{(k)}[t], \quad k = 1,\ldots,n \quad (2)$$

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 [15], that is,

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

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

$$\sigma_{q_i}^2[t] = \frac{\mathbf{a}_i^T\mathbf{\Sigma}_x\mathbf{a}_i + \sigma_w^2}{h_i[t]p_i[t]}. \quad (4)$$

Finally, by means of a Minimum Mean Square Error (MMSE) estimator [16] the FC reconstructs $\{\mathbf{x}^{(k)}[t]\}_{k=1}^n$ from the received codewords $\{u_i^{(k)}[t]\}_{k=1}^n$. The average (MSE) distortion in time slot $t \in \mathcal{T}$ is given by [16]

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

where $\mathrm{tr}(\cdot)$ 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 write the average distortion over all time slots as

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

We consider sensor nodes powered by energy harvesting. Assuming a discrete energy arrival model, at the beginning of time slot t, the i-th sensor harvests $E_i[t]$ Joules of energy. For simplicity, we consider the sensor nodes to be equipped with batteries of infinite capacity and take into account only the transmit power in the energy consumption model. Under these considerations, the causal[2] constraints on the transmit power are given by the convex sets

$$\Omega_i[t] \triangleq \left\{T_s\sum_{l=1}^t p_i[l] \leq \sum_{l=1}^t E_i[l], \quad p_i[t] \geq 0\right\} \quad (7)$$

for all $t \in \mathcal{T}$ and $i \in \mathcal{M}$. Our goal is then to minimize the average distortion (6), subject to the energy harvesting constraints (7). Due to bandwidth and signaling constraints[3], we attempt to minimize (6) by selecting a reduced subset of sensors. Since sensor activity is determined by its transmit power, (i.e., a sensor is active if $p_i[t] > 0$ and idle otherwise), we promote sparsity in the power allocation variable $p_i[t]$. Hence, the regularized optimization problem is given by

$$\begin{aligned} \min_{p_i[t]\in\Omega_i[t]} \quad & \frac{1}{T}\sum_{t=1}^T \mathrm{tr}\left(\frac{1}{\sigma_w^2}\sum_{i=1}^M \frac{p_i[t]}{p_i[t] + \xi_i[t]}\mathbf{a}_i\mathbf{a}_i^T + \mathbf{\Sigma}_x^{-1}\right)^{-1} \\ & + \lambda\frac{1}{T}\sum_{t=1}^T\sum_{i=1}^M f\left(p_i[t]\right) \quad (8a) \end{aligned}$$

where $f : \mathbb{R} \to \mathbb{R}$ is a sparsity-promoting penalty function, and λ the corresponding sparsity parameter. Several penalty functions have been studied in the literature [17, 18], with the most common ones being shown in Figure 1. The commonly called ℓ_0 norm corresponds to the cardinality function (in our case, it equals one for $p_i[t] > 0$ and zero otherwise). This function results in a combinatorial problem, which is generally difficult to solve. To alleviate this, the ℓ_0 norm is typically replaced by its convex envelope, i.e. the ℓ_1 norm, leading to affordable optimization problems. However, other non-convex regularizers such as the logarithmic penalty function

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

[2] Recall that we consider the design of offline power allocation policies.

[3] Typically, the number of communication channels to the FC is limited.

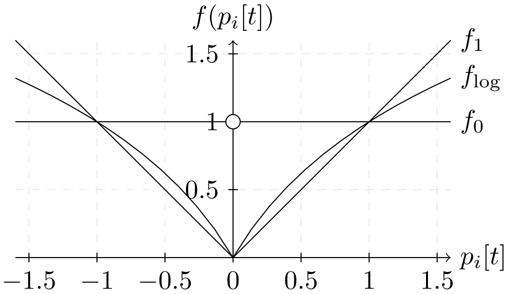


Fig. 1. Sparsity-promoting penalty functions.

might be preferable in some cases. The latter is typically used in combination with Majorization-Minimization (MM) methods allowing us to solve the problem by means of a sequence of surrogate (convex) functionals. Therefore, in the sequel, we assume f to be a convex function.

3. DECENTRALIZED ALGORITHM

In practice, the sensors should implement their selection and power allocation policies locally and, the FC should ultimately ensure that these local decisions attain the minimum distortion. Bearing this in mind, we decouple the optimization problem into those two tasks by introducing consensus variables in the power allocation variables. Namely, we introduce the global variables $\{q_i[t]\}$ in (6) and force consensus with $\{p_i[t]\}$ by introducing the equality constraints $p_i[t] = q_i[t]$. Then, the optimization problem is rewritten as

$$\min_{\substack{p_i[t]\in\Omega_i[t],\\ q_i[t]}} \quad \frac{1}{T}\sum_{t=1}^{T} \mathrm{tr}\left(\frac{1}{\sigma_w^2}\sum_{i=1}^{M}\frac{q_i[t]}{q_i[t]+\xi_i[t]}\mathbf{a}_i\mathbf{a}_i^T + \mathbf{\Sigma}_x^{-1}\right)^{-1} + \lambda\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{M} f\left(p_i[t]\right) \tag{9a}$$

$$\text{s.t.} \quad p_i[t] = q_i[t], \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{M} \tag{9b}$$

Since (9) is a convex optimization problem, we can find a global minimizer [19]. To exploit the separate structure of this problem, we resort to the ADMM algorithm [20]. The augmented Lagrangian of this problem is given by

$$\mathcal{L} = \frac{1}{T}\sum_{t=1}^{T} \mathrm{tr}\left(\frac{1}{\sigma_w^2}\sum_{i=1}^{M}\frac{q_i[t]}{q_i[t]+\xi_i[t]}\mathbf{a}_i\mathbf{a}_i^T + \mathbf{\Sigma}_x^{-1}\right)^{-1} + \lambda\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{M} f\left(p_i[t]\right) + \sum_{t=1}^{T}\sum_{i=1}^{M}\frac{\rho}{2}\left(p_i[t]-q_i[t]\right)^2 + \sum_{t=1}^{T}\sum_{i=1}^{M}\phi_i[t]\left(p_i[t]-q_i[t]\right), \tag{10}$$

where ρ is the penalty parameter of the augmented Lagrangian and $\{\phi_i[t]\}$ stands for the Lagrange multipliers

Algorithm 1 Decentralized algorithm.

Step 1: Sensor computation.

$$p_i^{(k+1)}[t] := \underset{p_i[t]\in\Omega_i[t]}{\arg\min}\left(\frac{\rho}{2}\left(p_i[t]-q_i^{(k)}[t]+\psi_i^{(k)}[t]\right)^2 + \frac{\lambda}{T}f\left(p_i[t]\right)\right)$$

Step 2: FC computation.

$$q_i^{(k+1)}[t] := \underset{q_i[t]}{\arg\min}\left(\sum_{i=1}^{M}\frac{\rho}{2}\left(p_i^{(k+1)}[t]-q_i[t]+\psi_i^{(k)}[t]\right)^2 + \frac{1}{T}\mathrm{tr}\left(\frac{1}{\sigma_w^2}\sum_{i=1}^{M}\frac{q_i[t]}{q_i[t]+\xi_i[t]}\mathbf{a}_i\mathbf{a}_i^T+\mathbf{\Sigma}_x^{-1}\right)^{-1}\right)$$

Step 3: Dual update.

$$\psi_i^{(k+1)}[t] := \psi_i^{(k)}[t] + p_i^{(k+1)}[t] - q_i^{(k+1)}[t]$$

Step 4: Go to Step 1 until convergence.

associated to the equality constraints (9b). Then, the resulting optimization procedure is summarized[4] in Algorithm 1, where for convenience we have introduced the scaled dual variable $\psi_i[t] = (1/\rho)\phi_i[t]$. Algorithm 1 consists on a local step and a consensus step. The local step is computed at each sensor node and it is constrained by each sensor's own energy harvesting process. Then, sensors transmit their *locally* computed power allocation $p_i[t]$ to the FC. During the consensus step, the FC[5] finds new values of $\{q_i[t]\}$ which minimize the distortion plus the regularization term. These new $\{q_i[t]\}$ values are then transmitted to all the sensor nodes, which locally update the dual variables $\{\psi_i[t]\}$. The role of dual variables is to enforce consensus between the FC variables and the power allocation $\{p_i[t]\}$ computed by the sensor nodes. The process is repeated until convergence.

3.1. Local step for a weighted ℓ_1 norm

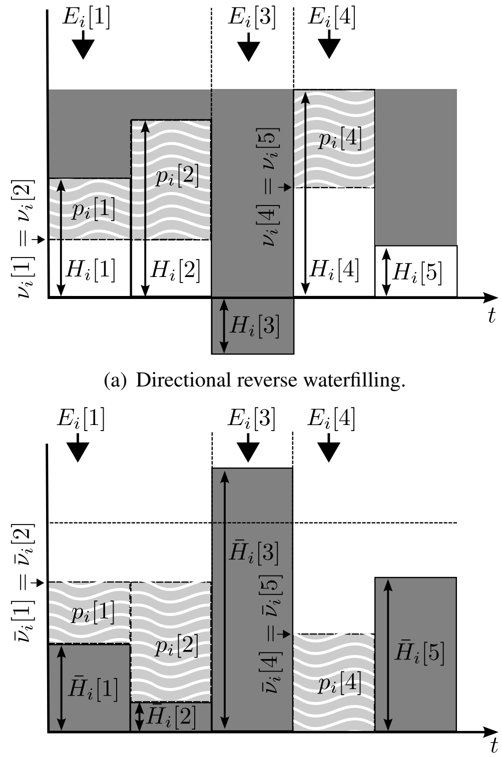
A frequent sparsity-promoting function that arises in a variety of scenarios is the weighted ℓ_1 norm [18], (in our case, $f\left(p_i[t]\right) = w_i[t]p_i[t]$, where $w_i[t]$ is a nonnegative scalar weight). In this case, the resulting local optimization problem can be interpreted as a slight variation of the classical waterfilling algorithm. First, including the energy harvesting constraints given by (7), we form the Lagrangian of the local step in Algorithm 1

$$\mathcal{L}_i[t] = \frac{\rho}{2}\left(p_i[t]-q_i^{(k)}[t]+\psi_i^{(k)}[t]\right)^2 + \frac{\lambda}{T}w_i[t]p_i[t] + \beta_i[t]\left(T_s\sum_{l=1}^{t}p_i[l] - \sum_{l=1}^{t}E_i[l]\right) + \eta_i[t]p_i[t] \tag{11}$$

with $\beta_i[t]$, and $\eta_i[t]$ standing for the Lagrange multipliers of the constraints defined in (7). Taking the derivative of the

[4]This scheme can also be implemented in the case that there is no fusion center and the sensor nodes are supported on a fully connected graph.

[5]We assume that the coefficients $\mathbf{a}_i$ are known at the fusion center.



(a) Directional reverse waterfilling.

(b) Directional waterfilling.

Fig. 2. Local step interpretation for a weighted ℓ_1 norm.

Lagrangian (11) with respect to $p_i[t]$, and solving for $p_i[t]$ (while satisfying the KKT conditions) we have

$$p_i[t] = \frac{T_s}{\rho}\left[\frac{\rho}{T_s}\left(q_i^{(k)}[t] - \psi_i^{(k)}[t] - \frac{\lambda}{\rho T} w_i[t]\right) - \sum_{l=t}^{T} \beta_i[l]\right]^+$$

where $[\cdot]^+ = \max\{\cdot, 0\}$. We define the heights $H_i[t] \triangleq \rho\big(q_i^{(k)}[t] - \psi_i^{(k)}[t] - \frac{\lambda}{\rho T} w_i[t]\big)/T_s$ and waterlevels $\nu_i[t] \triangleq \sum_{l=t}^{T} \beta_i[l]$. Then, we can interpret the solution as a directional reverse waterfilling, scaled by the widths T_s/ρ. This is shown in Figure 2(a). For any sensor $i \in \mathcal{M}$, each time slot t is associated to a hollow rectangle of height $H_i[t]$. Right-permeable walls are placed at the beginning of each time slot with an energy arrival. Then, water flows from the top down inside those rectangles, until a waterlevel $\nu_i[t]$ is reached. Then, the power allocation corresponds to the filled area of water above zero.

An equivalent interpretation as the more common directional waterfilling can be derived by defining the new heights $\bar{H}_i[t] = H_i^{\max} - H_i[t]$, where $H_i^{\max} \triangleq \max\{\{H_i[t]\}_{t\in\mathcal{T}}\}$. We illustrate this in Figure 2(b). This equivalent problem corresponds to the *mirroring* of the original problem with respect to the axis defined by the taller rectangle $H_i^{\max}$. In this equivalent problem, the rectangles $\bar{H}_i[t]$ correspond to a solid material over which water is then poured to a waterlevel $\bar{\nu}_i[t] = H_i^{\max} - \nu_i[t]$. In this case there is a ceiling $H_i^{\max}$ up to which power is allocated. Thus, any water flowing over this ceiling will not be allocated as power.

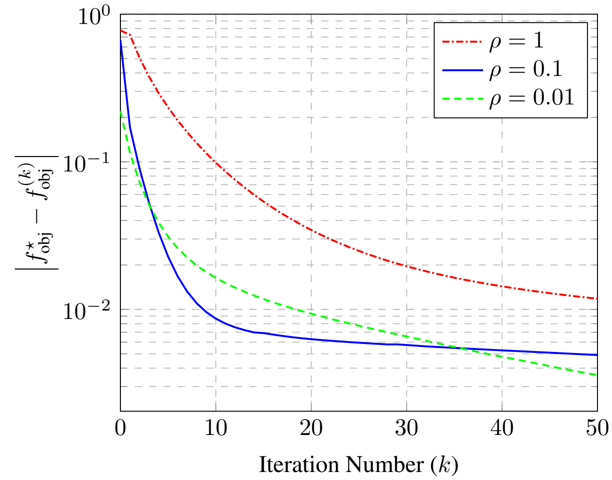


Fig. 3. Convergence of the proposed algorithm.

4. NUMERICAL RESULTS

In this section, we study the performance of the proposed algorithm. We consider a WSN 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 3, we show the numerical convergence of the proposed algorithm when the sparsity-promoting function is the ℓ_1 norm. In particular, we show on the y-axis the error between the optimal value $f^\star_{\text{obj}}$ of the objective function (9a) and its value at the k-th iteration $f^{(k)}_{\text{obj}}$. We plot the error for several values of the penalty parameter ρ. As shown, an appropriate choice of ρ allows convergence to an acceptable error of 10^{-2} in just 10 iterations, which ensures a low communication overhead between the sensors and the FC.

Now, we consider the case in which the sparsity-promoting function is given by a reweighted ℓ_1 norm, i.e., $f(p_i[t]) = w_i[t]p_i[t]$. In this case, we solve Algorithm 1 over ten iterations, with weights at the l-th iteration given by $w_i^{(l)}[t] = 1/(p_i^{(l-1)}[t] + \epsilon)$, and $\epsilon = 0.01$ [18]. In Figure 4, we plot the resulting distortion vs. the average number of selected sensors (the latter being directly related with the regularization parameter λ). We show results for different amounts of the harvested energy, namely, $E_i[t]$. The resulting distortion, as expected, is monotonically decreasing with the average number of selected sensors. More interestingly, by just se-

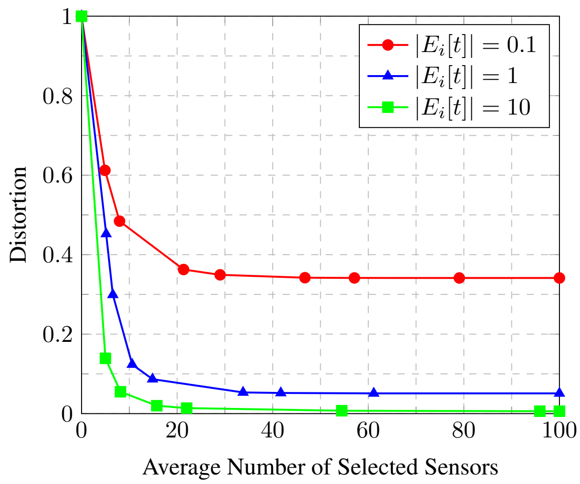


Fig. 4. Distortion vs. average number of active sensors.

lecting 20% to 30% of the sensors, the resulting distortion is identical to that of the whole set. The impact of the amount of harvested energy (or equivalently, the arrival intensity rate) on such percentages is marginal.

5. CONCLUSIONS

In this paper, we have addressed the sensor selection problem for the estimation of a stochastic source when sensors are powered by energy harvesting. In particular, we have proposed to minimize a functional that consists of the attained distortion plus a penalty term. The rationale of this penality term is to promote sparsity on the power allocation and, thus, on the sensor activity as well. Then, by leveraging on the ADMM, we have proposed a decentralized iterative algorithm in which the tasks of selection and reconstruction are decoupled. Moreover, we have shown that when the sparsity function takes the form of a weighted ℓ_1 norm, the power allocation policy can be computed by means of a directional waterfilling algorithm. Finally, numerical results have shown the fast convergence of the proposed scheme.

6. REFERENCES

[1] S. Ulukus, A. Yener, E. Erkip, O. Simeone, M. Zorzi, P. Grover, and K. Huang, "Energy harvesting wireless communications: A review of recent advances," *IEEE Journal on Selected Areas in Communications*, vol. PP, no. 99, pp. 1–1, 2015.

[2] 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.

[3] 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.

[4] O. Orhan, D. Gunduz, and E. Erkip, "Source-channel coding under energy, delay, and buffer constraints," *Wireless Communications, IEEE Transactions on*, vol. 14, no. 7, pp. 3836–3849, July 2015.

[5] Omur Ozel, Jing Yang, and Sennur Ulukus, "Optimal broadcast scheduling for an energy harvesting rechargeable transmitter with a finite capacity battery," *IEEE Transactions on Wireless Communications*, vol. 11, no. 6, pp. 2193–2203, 2012.

[6] Jing Yang and Sennur Ulukus, "Optimal packet scheduling in a multiple access channel with energy harvesting transmitters," *Communications and Networks, Journal of*, vol. 14, no. 2, pp. 140–150, 2012.

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

[8] 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.

[9] S.P. Chepuri and G. Leus, "Sparsity-promoting sensor selection for non-linear measurement models," *Signal Processing, IEEE Transactions on*, vol. 63, no. 3, pp. 684–698, Feb 2015.

[10] 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.

[11] 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.

[12] Miguel Calvo-Fullana, Javier Matamoros, and Carles Antón-Haro, "Sensor selection and power allocation strategies for energy harvesting wireless sensor networks," *submitted*, 2016.

[13] Miguel Calvo-Fullana, Javier Matamoros, Carles Antón-Haro, and Sophie M. Fosson, "Sparsity-promoting sensor selection with energy harvesting constraints," in *Acoustics, Speech and Signal Processing (ICASSP). 2016 IEEE International Conference on*, March 2016.

[14] 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.

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

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

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

[18] 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.

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

[20] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.