Optimal Energy Allocation and Storage Control for Distributed Estimation with Sensor Collaboration

Sijia Liu, Yanzhi Wang, Makan Fardad, Pramod K. Varshney

*Department of Electrical Engineering and Computer Science, Syracuse University, NY, 13244 USA
{sliu17, ywang393, makan, varshney}@syr.edu

***Abstract*—In wireless sensor networks with energy harvesting nodes, we study the problem of energy allocation and storage control for distributed estimation with sensor collaboration, where collaboration refers to the act of sharing measurements with neighboring sensors prior to transmission to a fusion center. To jointly design energy allocation and storage control polices, we formulate a nonconvex optimization problem in which the estimation distortion is minimized subject to energy harvesting and storage constraints. We show that the resulting optimization problem contains two special types of nonconvexities: cardinality function and difference of convex functions. By exploiting the problem structure, locally optimal solutions are found via an ℓ_1 relaxation and a convex-concave procedure. Numerical experiments are provided to show the effectiveness of our approach.**

***Index Terms*—Distributed estimation, sensor collaboration, energy harvesting and storage, convex optimization, sparsity.**

I. Introduction

Design of optimal energy allocation schemes for distributed estimation has been widely studied in the literature based on different types of parameters to be estimated (static parameter or random process), communication channels (coherent or orthogonal) and cost functions (energy or estimation distortion) [1]–[3]. However, most of the existing literature has focused on the amplify-and-forward transmission model, in which no inter-sensor collaboration is involved.

Recently, the problem of distributed estimation with sensor collaboration has attracted significant attention [4]–[6]. In [4], the optimal energy allocation strategy was found for a fully-connected network, where all the sensors are allowed to collaborate, namely, share their measurements with other sensors. It was shown that sensor collaboration results in significant improvement of estimation performance compared to the conventional amplify-and-forward transmission scheme. In [5], the optimal collaboration strategy was found under an arbitrary network topology. It was observed that even a partially connected network can yield performance close to that of a fully connected network. In [6], the nonzero collaboration cost was taken into account, and a sparsity induced optimization framework was proposed to simultaneously design both the collaboration topology and the energy allocation scheme.

In the aforementioned literature [4]–[6], collaborative estimation was restricted to parameter estimation during one snapshot, and sensors were assumed to be equipped with conventional power-limited batteries. There have been significant advances in the design of energy-harvesting sensors. Thus, it becomes quite attractive to study collaborative estimation problems in sensor networks with energy harvesting nodes, where at each time step sensors can replenish energy from the environment (such as solar and wind) without the need of battery replacement.

Several research efforts have been devoted to the design of optimal energy allocation schemes in energy harvesting networks for distributed estimation without inter-sensor collaboration [7]–[9]. In [7], a single sensor equipped with an energy harvester was used for parameter estimation, where the optimal energy allocation strategy was designed based on both causal and non-causal side information of energy harvesting. In [8], the design of energy allocation was studied over an orthogonal multiple access channel (MAC), where the estimation distortion resulting from the best linear unbiased estimator was minimized subject to energy harvesting constraints. In [9], a stochastic control problem was formulated for power allocation over a finite or an infinite time horizon.

Different from [4]–[9], here we present a unified optimization framework for the joint design of optimal energy allocation and storage control polices while incorporating the cost of sensor collaboration over a finite time horizon. We show that the resulting optimization problem is nonconvex. However, by identifying the special types of nonconvexities, the methods of convex relaxation and convex-concave procedure can be effectively used to find locally optimal solutions. Extensive numerical results are provided to demonstrate the utility of our approach for energy allocation and storage control in collaborative estimation.

II. Problem Formulation

In this paper, the task of the sensor network is to estimate a time-varying parameter $\theta_t \in \mathbb{R}$ over a time horizon of length T. We consider a collaborative estimation system where sensors are able to collaborate, namely, share observations with other neighboring sensors. The obtained collaborative messages are transmitted through a coherent MAC to the fusion center (FC), which produces a global estimate of θ_t for $t = 1, 2, \ldots, T$. In the network, each sensor is equipped with an energy harvesting device and a finite-capacity energy storage device, where the former harvests the renewable energy from the environment, and the latter governs the storage charging/discharging actions across time for sensor collaboration and data transmission. An overview of the collaborative

978-1-4673-9457-4/16/$31.00 ©2016 IEEE

estimation system with energy harvesting and storage is shown in Fig. 1.

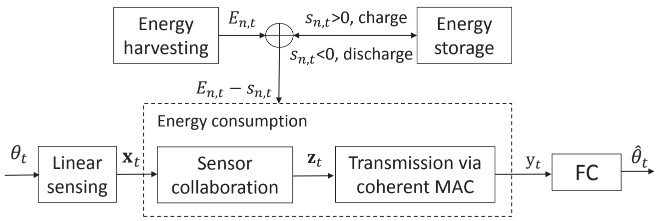

Fig. 1: Collaborative estimation with energy harvesting and storage. θ_t and $\hat{\theta}_t$ denote the parameter and its estimate at time t, respectively. $\mathbf{x}_t$, $\mathbf{z}_t$ and y_t denote the sensor measurements, collaborative signals and transmitted signal, respectively. $E_{n,t}$ and $s_{n,t}$ are the harvested energy (known in advance) and stored energy (to be designed) of the nth sensor at time t.

The vector of measurements from N sensors at time t is

$$\mathbf{x}_t = \mathbf{h}_t \theta_t + \boldsymbol{\epsilon}_t, \quad t \in [T], \tag{1}$$

where $[T]$ denotes the integer set $\{1, 2, \ldots, T\}$, $\mathbf{x}_t = [x_{t,1}, \ldots, x_{t,N}]^T$ is the vector of measurements, $\mathbf{h}_t = [h_{t,1}, \ldots, h_{t,N}]^T$ is the vector of observation gains, θ_t is the parameter of interest which has zero mean and variance σ_θ^2, $\boldsymbol{\epsilon}_t = [\epsilon_{t,1}, \ldots, \epsilon_{t,N}]^T$ is the noise vector with i.i.d. Gaussian variables $\epsilon_{t,i} \sim \mathcal{N}(0, \sigma_\epsilon^2)$ for $t \in [T]$ and $i \in [N]$.

After linear sensing, each sensor may pass its observation to another sensor for collaboration prior to transmission to the FC. With a relabelling of the sensors, we assume that the first M sensors (out of a total of N sensor nodes) communicate with the FC. The process of sensor collaboration at time t is

$$\mathbf{z}_t = \mathbf{W}_t \mathbf{x}_t, \quad t \in [T] \tag{2}$$

where $\mathbf{z}_t = [z_{t,1}, z_{t,2}, \ldots, z_{t,M}]^T$ is the vector of collaborative signals, and $\mathbf{W}_t \in \mathbb{R}^{M \times N}$ is the collaboration matrix that contains collaboration weights (based on the energy allocated) used to combine sensor measurements.

After sensor collaboration, the message $\mathbf{z}_t$ is transmitted through a coherent MAC so that the received signal y_t at the FC is a coherent sum [2]

$$y_t = \mathbf{g}_t^T \mathbf{z}_t + \varsigma_t, \ \ t \in [T], \tag{3}$$

where $\mathbf{g}_t = [g_{t,1}, , g_{t,2}, \ldots, g_{t,M}]^T$ is the vector of channel gains, and ς_t is temporally white Gaussian noise with zero mean and variance σ_ς^2.

A. Best linear unbiased estimator (BLUE) at the FC

We assume that the FC knows the observation gains, channel gains, and variances of the parameters of interest and additive noises. We employ the best linear unbiased estimator (BLUE) [10, Theorem 6.1] to estimate the parameters θ_t for $t \in [T]$. The use of BLUE is spurred by the lack of prior knowledge about the temporal correlation of time-varying parameters $\{\theta_t\}$ [11]. As demonstrated in Appendix A, the estimation error resulting from BLUE is given by a sum of rational functions with respect to collaboration matrices

$$f(\mathbf{W}_1, \ldots, \mathbf{W}_T) = \sum_{t=1}^{T} \frac{\sigma_\varsigma^2 + \sigma_\epsilon^2 \operatorname{tr}\left(\mathbf{W}_t^T \mathbf{g}_t \mathbf{g}_t^T \mathbf{W}_t\right)}{\operatorname{tr}\left(\mathbf{W}_t^T \mathbf{g}_t \mathbf{g}_t^T \mathbf{W}_t \mathbf{h}_t \mathbf{h}_t^T\right)}. \tag{4}$$

B. Collaboration and transmission costs

It is clear from (2) that the sparsity structure of $\mathbf{W}_t$ characterizes the collaboration topology at time t. For instance, $[\mathbf{W}_t]_{mn} = 0$ indicates the absence of a collaboration link from the nth sensor to the mth sensor, where $[\mathbf{W}_t]_{mn}$ is the (m, n)th entry of $\mathbf{W}_t$. To account for an active collaboration link, we use the cardinality function

$$\operatorname{card}([\mathbf{W}_t]_{mn}) = \begin{cases} 0 & [\mathbf{W}_t]_{mn} = 0 \\ 1 & [\mathbf{W}_t]_{mn} \neq 0. \end{cases} \tag{5}$$

The *collaboration cost* of each sensor is then given by

$$Q_n(\mathbf{W}_t) = \sum_{m=1}^{M} C_{mn} \operatorname{card}([\mathbf{W}_t]_{mn}), \ n \in [N], \ t \in [T], \tag{6}$$

where we assume that sharing of an observation is realized through a reliable (noiseless) communication link that consumes a known power C_{mn}. Note that $C_{mm} = 0$ since each node can collaborate with itself at no cost.

The *transmission cost* of the mth sensor at time t refers to the energy consumption of transmitting the collaborative message $z_{t,m}$ to the FC. Namely,

$$\begin{aligned} T_m(\mathbf{W}_t) &= \mathbb{E}_{\theta_t, \boldsymbol{\epsilon}_t}[z_{t,m}^2] \\ &= \operatorname{tr}\left[\mathbf{W}_t^T \mathbf{e}_m \mathbf{e}_m^T \mathbf{W}_t (\sigma_\theta^2 \mathbf{h}_t \mathbf{h}_t^T + \sigma_\epsilon^2 \mathbf{I})\right], \end{aligned} \tag{7}$$

for $m \in [M]$ and $t \in [T]$, where $\mathbf{e}_m \in \mathbb{R}^M$ is a basis vector with 1 at the mth coordinate and 0s elsewhere. Since only the first M sensors are used to communicate with the FC, we define $T_n(\mathbf{W}_t) := 0$ for $n > M$,

C. Energy harvesting and storage constraints

We assume that knowledge of the harvested energy is available in advance (also known as full side information [12]). Let $E_{n,t}$ denote the harvested energy of the nth sensor at time t, which is used for storage or sensor collaboration and data transmission. We introduce a variable $s_{n,t}$ to represent the charging/discharging operation of the storage device at the nth sensor at time t, with the following sign convention for $s_{n,t}$

$$\begin{cases} s_{n,t} \geq 0 & \text{charging with amount of energy } s_{n,t}, \\ s_{n,t} < 0 & \text{discharging with amount of energy } -s_{n,t}. \end{cases} \tag{8}$$

Given the harvested energy $E_{n,t}$ and the operating mode of the storage device, the energy consumption for sensor collaboration and data transmission satisfies the constraint

$$Q_n(\mathbf{W}_t) + T_n(\mathbf{W}_t) \leq E_{n,t} - s_{n,t}, \ n \in [N], \ t \in [T] \tag{9}$$

where $Q_n(\mathbf{W}_t)$ is the collaboration cost in (6), $T_n(\mathbf{W}_t)$ is the transmission cost in (7), and the quantity $E_{n,t} - s_{n,t}$ stands for the amount of energy available at the nth sensor at time t.

The stored energy at each sensor per time satisfies

$$-\sum_{k=0}^{t-1} s_{n,k} \leq s_{n,t} \leq E_{n,t}, \ n \in [N], \ t \in [T], \tag{10}$$

where $s_{n,0} \geq 0$ is the known amount of initially stored energy at the nth sensor, $\sum_{k=0}^{t-1} s_{n,k}$ is the amount of stored energy at time $t-1$ which might be discharged from the storage device at the next time, and $E_{n,t}$ is the amount of harvested energy at time t which can be used to charge the storage device.

There is a capacity limit of the storage device [13], denoted by $\hat{S}$, so that

$$\sum_{k=0}^{t} s_{n,k} \leq \hat{S},\ n \in [N],\ t \in [T]. \tag{11}$$

To increase the network lifetime, it is also desirable to have some remaining stored energy, denoted by $\check{S}$, at the end of time horizon. This gives the following constraint [14]

$$\sum_{k=0}^{T} s_{n,k} \geq \check{S},\ n \in [N]. \tag{12}$$

D. Sensor collaboration with energy storage management

In order to seek the optimal collaboration schemes $\{\mathbf{W}_t\}$ and storage control policy $\{s_{n,t}\}$, we formulate the optimization problem, in which the estimation error (4) is minimized subject to energy constraints (9) – (12),

$$\begin{array}{ll} \text{minimize} & f(\mathbf{W}_1, \ldots, \mathbf{W}_T) \\ \text{subject to} & \text{constraints (9)} - \text{(12)}, \end{array} \tag{13}$$

where the optimization variables are $\mathbf{W}_t \in \mathbb{R}^{M \times N}$ and $\mathbf{s}_t = [s_{1,t}, s_{2,t}, \ldots, s_{N,t}]^T$ for $t \in [T]$. As will be evident later, problem (13) is not convex due to its objective function and the constraint (9) that involves the cardinality function. We will elaborate on the problem structure in the next section.

III. Equivalent Optimization Problem

In this section, we first simplify problem (13) by converting quadratic matrix functions into quadratic vector functions. We then show that problem (13) is equivalent to a nonconvex optimization problem with special types of nonconvexities.

It is clear from (4) and (7) that both the estimation distortion and the transmission cost contain a quadratic matrix function[1]. For simplicity of representation, we transform a quadratic matrix function into a quadratic vector function by concatenating the entries of a matrix into a column vector.

Specifically, a columnwise vector $\mathbf{w}_t$ of the collaboration matrix $\mathbf{W}_t$ can be written as $\mathbf{w}_t = [w_{t,1}, w_{t,2}, \ldots, w_{t,L}]^T$, where $L = MN$, $w_{t,l}$ is the lth entry of $\mathbf{w}_t$, $w_{t,l} = [\mathbf{W}_t]_{m_l n_l}$, $m_l = l - (\lceil \frac{l}{M} \rceil - 1)M$, $n_l = \lceil \frac{l}{M} \rceil$, and $\lceil x \rceil$ is the ceiling function that yields the smallest integer not less than x.

Using the relationship between the Kronecker product and the vectorization of a matrix[2], we can rewrite the matrix quadratic functions in (4) and (7) as

$$f(\mathbf{w}_1, \ldots, \mathbf{w}_T) = \sum_{t=1}^{T} \frac{\sigma_\varsigma^2 + \sigma_\epsilon^2 \mathbf{w}_t^T \boldsymbol{\Omega}_{\mathrm{PN},t} \mathbf{w}_t}{\mathbf{w}_t^T \boldsymbol{\Omega}_{\mathrm{PD},t}}, \text{ and} \tag{14}$$

$$T_n(\mathbf{w}_t) = \begin{cases} \mathbf{w}_t^T \boldsymbol{\Gamma}_{m,t} \mathbf{w}_t & n \in [M], \\ 0 & n = m+1, \ldots, N, \end{cases} \tag{15}$$

[1]A quadratic matrix function is a function $g : \mathbb{R}^{m \times n} \to \mathbb{R}$ of the form $g(\mathbf{X}) = \operatorname{tr}(\mathbf{X}^T \mathbf{A} \mathbf{X}) + 2 \operatorname{tr}(\mathbf{B}^T \mathbf{X}) + c$, where $\mathbf{A} \in \mathbb{R}^{m \times m}$, $\mathbf{B} \in \mathbb{R}^{m \times n}$ and $c \in \mathbb{R}$ are given coefficients.

[2]For appropriate matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$, we obtain $\operatorname{tr}(\mathbf{A}^T \mathbf{B} \mathbf{C} \mathbf{D}^T) = \operatorname{vec}(\mathbf{A})^T (\mathbf{D} \otimes \mathbf{B}) \operatorname{vec}(\mathbf{C})$, where $\otimes$ denotes the Kronecker product, and $\operatorname{vec}(\mathbf{Z})$ is the columnwise vector of $\mathbf{Z}$.

where with an abuse of notation, we used f and T_n introduced in (4) and (7) to represent the estimation distortion and tranmission cost with respect to $\{\mathbf{w}_t\}$, $\boldsymbol{\Omega}_{\mathrm{PN},t} = \mathbf{I}_N \otimes (\mathbf{g}_t \mathbf{g}_t^T)$, $\boldsymbol{\Omega}_{\mathrm{PD},t} = (\mathbf{h}_t \mathbf{h}_t^T) \otimes (\mathbf{g}_t \mathbf{g}_t^T)$, and $\boldsymbol{\Gamma}_{m,t} = (\sigma_\theta^2 \mathbf{h}_t \mathbf{h}_t^T + \sigma_\epsilon^2 \mathbf{I}) \otimes (\mathbf{e}_m \mathbf{e}_m^T)$. Clearly, the matrices $\boldsymbol{\Omega}_{\mathrm{PN},t}$, $\boldsymbol{\Omega}_{\mathrm{PD},t}$ and $\boldsymbol{\Gamma}_{m,t}$ are positive semidefinite, and in particular $\boldsymbol{\Omega}_{\mathrm{PD},t}$ is of rank one.

Moreover, we express the collaboration cost (6) as

$$Q_n(\mathbf{w}_t) = \sum_{m=1}^{M} c_{m+(n-1)M} \operatorname{card}(w_{t,m+(n-1)M}) \tag{16}$$

for $n \in [N]$ and $t \in [T]$, where c_i denotes the ith entry of the columnwise vector of the collaboration cost matrix $\mathbf{C} \in \mathbb{R}^{M \times N}$, whose (m,n)th entry is given by C_{mn} in (6).

Based on (15) and (16), the energy consumption constraint (9) can be cast as a quadratic inequality that involves the cardinality function with respect to $\mathbf{w}_t$,

$$Q_n(\mathbf{w}_t) + T_n(\mathbf{w}_t) \leq E_{n,t} - s_{n,t},\ n \in [N],\ t \in [T]. \tag{17}$$

Problem (13) is further written as

$$\begin{array}{ll} \text{minimize} & f(\mathbf{w}_1, \ldots, \mathbf{w}_T) \\ \text{subject to} & \text{storage constraints (10)} - \text{(12)} \\ & \text{energy consumption constraint (17)}, \end{array} \tag{18}$$

where $\mathbf{w}_t \in \mathbb{R}^L$ and $\mathbf{s}_t \in \mathbb{R}^N$ are optimization variables for $t \in [T]$, and we recall that $L = MN$.

It is a nontrivial exercise to solve problem (18) due to its objective function in terms of minimizing a sum of quadratic ratios [15]. However, we will show in Proposition 1 that problem (18) can be transformed to a special nonconvex optimization problem, where the resulting problem structure helps us in developing an efficient optimization approach.

Proposition 1: Problem (18) is equivalent to

$$\text{minimize}\ \mathbf{1}^T \mathbf{u} \tag{19a}$$
$$\text{subject to storage constraints (10)} - \text{(12)} \tag{19b}$$
$$\text{energy consumption constraint (17)} \tag{19c}$$
$$\sigma_\varsigma^2 + \sigma_\epsilon^2 \mathbf{w}_t^T \boldsymbol{\Omega}_{\mathrm{PN},t} \mathbf{w}_t \leq p_t,\ q_t > 0, t \in [T] \tag{19d}$$
$$2p_t + (q_t^2 + u_t^2) - (q_t + u_t)^2 \leq 0,\ t \in [T] \tag{19e}$$
$$q_t - \mathbf{w}_t^T \boldsymbol{\Omega}_{\mathrm{PD},t} \mathbf{w}_t \leq 0,\ t \in [T], \tag{19f}$$

where $\mathbf{u} \in \mathbb{R}^T$, $\mathbf{p} \in \mathbb{R}^T$, $\mathbf{q} \in \mathbb{R}^T$, $\mathbf{w}_t \in \mathbb{R}^L$ and $\mathbf{s}_t \in \mathbb{R}^N$ are optimization variables for $t \in [T]$, and $\mathbf{1}$ is the vector of all ones.

Proof: See Appendix B. ■

It is clear from (19) that there exist two sources of nonconvexity: a) the cardinality function in the collaboration cost (16), b) the difference of convex (DC) quadratic functions (19e) and (19f). In what follows, we will show that the difficulties caused by the cardinality function and the nonconvex function of DC type can be efficiently handled via certain convexifications.

IV. Optimization Approach

A. ℓ_1 relaxation

Problem (19) is combinatorial in nature due to the presence of the cardinality function (also known as ℓ_0 norm). In the sparsity-promoting optimization context [16], an ℓ_0 norm is commonly relaxed to an ℓ_1 norm. Based on that, the collaboration cost (16) is approximated as

$$\tilde{Q}_n(\mathbf{w}_t) := \|\mathbf{C}_n \mathbf{w}_t\|_1, \ n \in [N], \ t \in [T], \tag{20}$$

where $\mathbf{C}_n = [c_{1+(n-1)M}\mathbf{e}_{1+(n-1)M}, \ldots, c_{nM}\mathbf{e}_{nM}]^T \in \mathbb{R}^{M \times L}$, and $\mathbf{e}_i \in \mathbb{R}^L$ is a basis vector. Accordingly, an ℓ_1 proxy for the energy consumption constraint (17) is given by

$$\tilde{Q}_n(\mathbf{w}_t) + T_n(\mathbf{w}_t) \le E_{n,t} - s_{n,t}, \ n \in [N], \ t \in [T]. \tag{21}$$

We remark that the ℓ_1 norm given by (20) can be eliminated by introducing additional optimization variables so that the ℓ_1 constraint (21) is cast as a smooth convex constraint together with a linear inequality constraint [17]

$$\begin{aligned} &\mathbf{1}^T \mathbf{r}_{n,t} + T_n(\mathbf{w}_t) \le E_{n,t} - s_{n,t}, \ \text{and} \\ &-\mathbf{r}_{n,t} \le \mathbf{C}_n \mathbf{w}_t \le \mathbf{r}_{n,t}, \ n \in [N], \ t \in [T], \end{aligned} \tag{22}$$

where $\mathbf{r}_{n,t} \in \mathbb{R}^M$ is the newly introduced optimization variable for $n \in [N]$ and $t \in [T]$, and the linear inequality is defined in an elementwise fashion.

After ℓ_1 approximation, problem (19) becomes

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T \mathbf{u} \\ \text{subject to} & \text{storage constraints (10)–(12)} \\ & \text{energy consumption constraints (22)} \\ & \text{quadratic constraints (19d)} \\ & \textit{nonconvex} \text{ quadratic constraints (19e)–(19f),} \end{array} \tag{23}$$

where $\mathbf{u} \in \mathbb{R}^T$, $\mathbf{p} \in \mathbb{R}^T$, $\mathbf{q} \in \mathbb{R}^T$, $\mathbf{w}_t \in \mathbb{R}^L$, $\mathbf{s}_t \in \mathbb{R}^N$, and $\mathbf{r}_{n,t} \in \mathbb{R}^M$ are optimization variables for $n \in [N]$ and $t \in [T]$.

Problem (23) is convex except for the last nonconvex quadratic constraints (19e) and (19f). However, we will show that such a nonconvex optimization problem can be efficiently solved via a convex-concave procedure.

B. Convex-concave procedure

In problem (23), we recall that the nonconvex quadratic constraints (19e)–(19f) have the DC type,

$$\phi(\mathbf{v}) - \psi(\mathbf{v}) \le 0, \tag{24}$$

where both $\phi(\cdot)$ and $\psi(\cdot)$ are convex functions. In (19e), we have $\phi(p_t, q_t, u_t) = 2p_t + (q_t^2 + u_t^2)$, and $\psi(p_t, q_t, u_t) = (q_t + u_t)^2$. In (19f), $\phi(q_t) = q_t$, and $\psi(\mathbf{w}_t) = \mathbf{w}_t^T \mathbf{\Omega}_{\mathrm{PD},t} \mathbf{w}_t$.

We convexify (24) by linearizing the convex function $\psi(\cdot)$ around a feasible point $\hat{\mathbf{v}}$

$$\phi(\mathbf{v}) - \hat{\psi}(\mathbf{v}) \le 0, \tag{25}$$

where $\hat{\psi}(\mathbf{v}) := \psi(\hat{\mathbf{v}}) + (\mathbf{v} - \hat{\mathbf{v}})^T \frac{\partial \psi(\hat{\mathbf{v}})}{\partial \mathbf{v}}$, and $\frac{\partial \psi(\hat{\mathbf{v}})}{\partial \mathbf{v}}$ denotes the first-order derivative of $\psi(\cdot)$ at the point $\hat{\mathbf{v}}$. We emphasize that since $\hat{\psi}(\mathbf{v})$ is an affine lower bound on the convex function $\psi(\mathbf{v})$, the set of $\mathbf{v}$ that satisfy (25) is a strict subset of the set of $\mathbf{v}$ that satisfy (24). This implies that a solution of the optimization problem with the linearized constraint (25) is locally optimal to the problem with the original nonconvex constraint of DC type.

We eventually obtain a 'restricted' convex version of problem (23) by linearizing the nonconvex quadratic constraints (19e)–(19f). We then solve a sequence of convex problems based on iteratively updated linearization points. The use of linearization to convexify nonconvex problems with DC type functions is known as a convex-concave procedure (CCP) [18], [19]. At each iteration of CCP, we solve the convexified problem of (23)

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T \mathbf{u} \\ \text{subject to} & \text{storage constraints (10)–(12)} \\ & \text{energy consumption constraints (22)} \\ & \text{quadratic constraints (19d)} \\ & \text{linearized quadratic constraints (19e)–(19f):} \\ & 2p_t + (q_t^2 + u_t^2) \le \hat{\psi}_1(q_t, u_t), \ \text{and} \\ & q_t \le \hat{\psi}_2(\mathbf{w}_t), \ t \in [T], \end{array} \tag{26}$$

where $\mathbf{u} \in \mathbb{R}^T$, $\mathbf{p} \in \mathbb{R}^T$, $\mathbf{q} \in \mathbb{R}^T$, $\mathbf{w}_t \in \mathbb{R}^L$, $\mathbf{s}_t \in \mathbb{R}^N$, and $\mathbf{r}_{n,t} \in \mathbb{R}^M$ are optimization variables for $n \in [N]$ and $t \in [T]$, $\psi_1(q_t, u_t) := 2(u_t + q_t)(\hat{u}_t + \hat{q}_t) - (\hat{u}_t + \hat{q}_t)^2$, $(\hat{u}_t, \hat{q}_t)$ is the linearization point used in $\hat{\psi}_1$, $\hat{\psi}_2(\mathbf{w}_t) = 2\mathbf{w}_t^T \mathbf{\Omega}_{\mathrm{PD},t} \hat{\mathbf{w}}_t - \hat{\mathbf{w}}_t^T \mathbf{\Omega}_{\mathrm{PD},t} \hat{\mathbf{w}}_t$, and $\hat{\mathbf{w}}_t$ is the linearization point used in $\hat{\psi}_2$. We summarize CCP for solving problem (23) in Algorithm 1.

Algorithm 1 CCP for solving problem (23)

Require: $\epsilon_{\mathrm{ccp}} > 0$ and feasible initial points $\{\hat{\mathbf{w}}_t\}$, $\hat{\mathbf{q}}$ and $\hat{\mathbf{u}}$

1: **for** iteration $i = 1, 2, \ldots$ **do**
2: set $(\{\mathbf{w}_t^i\}, \mathbf{q}^i, \mathbf{u}^i)$ as the solution of problem (26)
3: update the linearizing point: $\hat{\mathbf{w}}_t = \mathbf{w}_t^i$, $\hat{\mathbf{q}} = \mathbf{q}^i$, $\hat{\mathbf{u}} = \mathbf{u}^i$
4: **until** $|\mathbf{1}^T \mathbf{u}^i - \mathbf{1}^T \mathbf{u}^{i-1}| \le \epsilon_{\mathrm{ccp}}$ for $i \ge 2$
5: **end for**

For simplicity, we randomly choose the initial points (drawn from the standard uniform distribution) that are scaled to satisfy the constraints of problem (23). It is worth mentioning that reference [19] has proposed a penalized version of CCP which allows the use of infeasible initial points. Our numerical examples show that Algorithm 1 works effectively and converges to almost the same optimal value for different initial points; see Fig. 2. The convergence of Algorithm 1 is also guaranteed, since at each iteration, we solve a linearized convex problem with a smaller feasible set which contains the linearization point (i.e., the solution at the previous iteration) [19].

The computation cost of Algorithm 1 is dominated by solving the solution of a convex program with quadratic constraints at Step 2, which takes the computational complexity $O(T^{3.5}L^{3.5})$ via the interior-point algorithm [20, Sec. 10], where recalling that T is the length of time horizon, and $L = NM$ is the maxinum number of collaboration links.

V. Numerical Results

In this section, we empirically show the effectiveness of the proposed energy allocation and storage algorithm. To specify the system model in Fig. 1, we assume that $N = 5$ sensors are deployed in a unit square region, and let $T = 10$, $M = 5$, and $\sigma_\theta^2 = \sigma_\epsilon^2 = \sigma_\varsigma^2 = 1$. The values of observation gains $\mathbf{h}_t$ and channel gains $\mathbf{g}_t$ are drawn from the uniform distribution $U(0.1, 1)$. The amount of energy harvesting $E_{n,t}$ satisfies a Poisson process [8], namely, $\mathrm{prob}\{E_{n,t} = m\mu_0\} = \frac{\lambda^m e^{-\lambda}}{m!}$ for $n \in [N]$ and $t \in [T]$, where $\mu_0 = 1$ is the average energy harvesting rate, and $\lambda = 5$. The initial energy in the storage device is set by $s_{n,0} = 0$ for $n \in [N]$. The capacity limit and the desired residual energy in the storage device are given by $\hat{S} = 100$ and $\check{S} = 0.2 \sum_{t=1}^{T} \sum_{n=1}^{N} \frac{E_{n,t}}{TN}$, respectively. The cost of establishing a collaboration link is modeled by $C_{mn} = \alpha_c \|\mathbf{a}_m - \mathbf{a}_n\|_2$ for $m, n \in [N]$, where $\mathbf{a}_i$ is the location of the ith sensor, and $\alpha_c = 1$ is the collaboration cost parameter. In algorithm 1, we select $\epsilon_{\mathrm{ccp}} = 10^{-3}$.

In Fig. 2, we present the convergence trajectory of Algorithm 1 as a function of iteration index for 5 different initial points. As we can see, much of the benefit of using Algorithm 1 is gained from its first few iterations. Moreover, Algorithm 1 converges to almost the same objective value for different initial points. That is because the linearization of nonconvex quadratic functions, which have rank-one coefficient matrices, e.g, $\mathbf{\Omega}_{\mathrm{PD},t}$ in (26), leads to good proxies of DC type nonconvex functions.

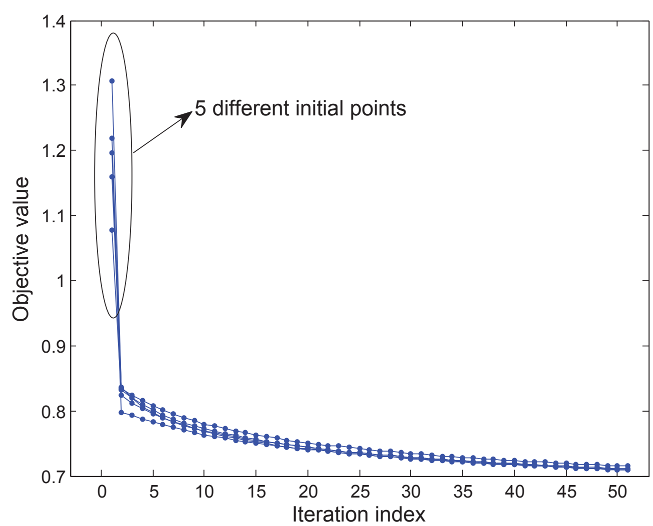

Fig. 2: Convergence of Algorithm 1 for 5 different initial points.

In Fig. 3, we show the obtained energy allocation and storage schemes at the first sensor by solving problem (19). We observe that the harvested energy is not wasted, namely, the sum of consumed energy and stored energy is equal to the amount of the harvested energy. We also note that the storage device operates in a charging mode when there is a large amount of harvested energy, or the stored energy is low. By contrast, the storage device operates in a discharging mode when the harvested energy is low, or there is a large amount of stored energy. We finally remark that the energy consumption for sensor collaboration and data transmission varies in time due to the time-varying sensor network.

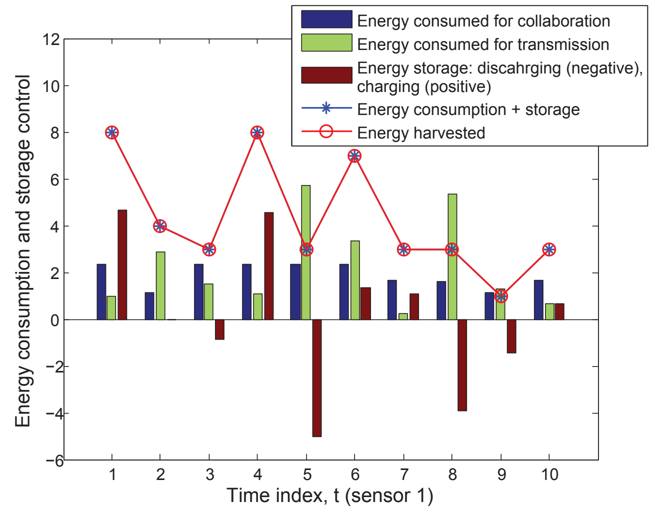

Fig. 3: Energy consumption and storage at the first sensor.

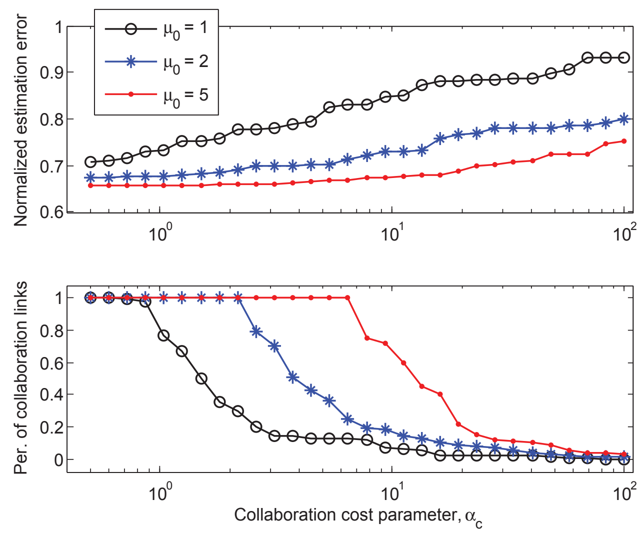

Fig. 4: Estimation error and collaboration links while varying α_c.

In Fig. 4, we present the estimation error and the percentage of active collaboration links as a function of the collaboration cost parameter α_c for $\mu_0 \in \{1, 3, 5\}$. As we can see, by fixing the energy harvesting rate μ_0, both the estimation accuracy and the number of collaboration links decrease as α_c increases. This is expected, since a larger value of α_c corresponds to a larger cost of sensor collaboration. Moreover, if we fix the value of α_c, both the estimation accuracy and the number of collaboration links increase when the energy harvesting rate increases, since more energy would be harvested and can be used for sensor collaboration.

VI. Conclusion

We proposed a tractable optimization framework to design optimal energy allocation and storage control policies for distributed estimation with sensor collaboration. We showed that the problem is nonconvex and combinatorial in nature. By exploiting the problem structure, we convexified the problem by using ℓ_1 relaxation and the convex-concave procedure. The convexified problem rendered a locally optimal solution. Numerical results were provided to show the effectiveness of our approach. In future work, it will be worthwhile to study the problem of sensor collaboration in energy harvesting networks with causal side information or over an infinite time horizon.

ACKNOWLEDGMENT

This work was supported by U.S. Air Force Office of Scientific Research (AFOSR) under Grant FA9550-10-1-0458, and National Science Foundation (NSF) under EAGER Award ECCS-1545270 and Award CNS-1329885.

APPENDIX A
ESTIMATION DISTORTION OF BLUE

According to (1), (2) and (3), the received signals at the FC over T time steps can be written as the succinct form

$$\mathbf{y} = \mathbf{G}_w \mathbf{D}_h \boldsymbol{\theta} + \boldsymbol{\upsilon}, \tag{27}$$

where $\mathbf{y} = [y_1, y_2, \ldots, y_T]^T$, $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_T]^T$, $\boldsymbol{\upsilon} = [\upsilon_1, \upsilon_2, \ldots, \upsilon_T]^T$, $\upsilon_t = \mathbf{g}_t^T \mathbf{W}_t \boldsymbol{\epsilon}_t + \varsigma_t$, $\mathbf{D}_h = \mathrm{blkdiag}\{\mathbf{h}_t\}_{t=1}^T$, $\mathbf{G}_w = \mathrm{blkdiag}\{\mathbf{g}_t^T \mathbf{W}_t\}_{t=1}^T$, and $\mathrm{blkdiag}\{\mathbf{A}_i\}_{i=1}^n$ denotes the block-diagonal matrix with diagonal blocks $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n$.

The estimation distortion resulting from BLUE is given by [10, Theorem 6.1]

$$\begin{aligned} f(\{\mathbf{W}_t\}) &:= \mathrm{tr}\left(\mathbf{D}_h^T \mathbf{G}_w^T (\sigma_\epsilon^2 \mathbf{G}_w \mathbf{G}_w^T + \sigma_\varsigma^2 \mathbf{I})^{-1} \mathbf{G}_w \mathbf{D}_h\right)^{-1} \\ &= \sum_{t=1}^T \frac{\sigma_\varsigma^2 + \sigma_\epsilon^2 \, \mathrm{tr}\left(\mathbf{W}_t^T \mathbf{g}_t \mathbf{g}_t^T \mathbf{W}_t\right)}{\mathrm{tr}\left(\mathbf{W}_t^T \mathbf{g}_t \mathbf{g}_t^T \mathbf{W}_t \mathbf{h}_t \mathbf{h}_t^T\right)}, \end{aligned}$$

where $\mathbf{D}_\upsilon = \mathbb{E}[\boldsymbol{\upsilon}\boldsymbol{\upsilon}^T] = \sigma_\epsilon^2 \mathbf{G}_w \mathbf{G}_w^T + \sigma_\varsigma^2 \mathbf{I}$. ■

APPENDIX B
PROOF OF PROPOSITION 1

We introduce a new vector of optimization variables $\mathbf{u} = [u_1, \ldots, u_T]^T$, and express problem (18) as

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T \mathbf{u} \\ \text{subject to} & \dfrac{\sigma_\varsigma^2 + \sigma_\epsilon^2 \mathbf{w}_t^T \boldsymbol{\Omega}_{\mathrm{PN},t} \mathbf{w}_t}{\mathbf{w}_t^T \boldsymbol{\Omega}_{\mathrm{PD},t} \mathbf{w}_t} \le u_t, \ t \in [T], \\ & \text{constraints of (18)}, \end{array} \tag{28}$$

where $\mathbf{u} \in \mathbb{R}^T$, $\mathbf{w}_t \in \mathbb{R}^L$ and $\mathbf{s}_t \in \mathbb{R}^N$ are optimization variables for $t \in [T]$.

By introducing additional variables p_t and q_t for $t \in [T]$, problem (28) can be then rewritten as

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T \mathbf{u} \\ \text{subject to} & \dfrac{p_t}{q_t} \le u_t, \ t \in [T], \\ & \sigma_\varsigma^2 + \sigma_\epsilon^2 \mathbf{w}_t^T \boldsymbol{\Omega}_{\mathrm{PN},t} \mathbf{w}_t \le p_t, \ t \in [T], \\ & \mathbf{w}_t^T \boldsymbol{\Omega}_{\mathrm{PD},t} \mathbf{w}_t \ge q_t, \ q_t > 0, \ t \in [T], \\ & \text{constraints of (18)}, \end{array} \tag{29}$$

where $\mathbf{u} \in \mathbb{R}^T$, $\mathbf{p} = [p_1, \ldots, p_T]^T \in \mathbb{R}^T$, $\mathbf{q} = [q_1, \ldots, q_T]^T \in \mathbb{R}^T$, $\mathbf{w}_t \in \mathbb{R}^L$ and $\mathbf{s}_t \in \mathbb{R}^N$ are optimization variables for $t \in [T]$. Note that the ratio $(p_t/q_t) \le u_t$ can be reformulated as a quadratic constraint given $q_t > 0$,

$$2p_t \le (q_t + u_t)^2 - (q_t^2 + u_t^2). \tag{30}$$

Substituting (30) into (29), we eventually reach the optimization problem (19). ■

REFERENCES

[1] F. Jiang, J. Chen, and A. L. Swindlehurst, "Optimal power allocation for parameter tracking in a distributed amplify-and-forward sensor network," *IEEE Transactions on Signal Processing*, vol. 62, no. 9, pp. 2200–2211, May 2014.

[2] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. J. Goldsmith, "Linear coherent decentralized estimation," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 757–770, 2008.

[3] G. Thatte and U. Mitra, "Sensor selection and power allocation for distributed estimation in sensor networks: Beyond the star topology," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 2649–2661, July 2008.

[4] J. Fang and H. Li, "Power constrained distributed estimation with cluster-based sensor collaboration," *IEEE Transactions on Wireless Communications*, vol. 8, no. 7, pp. 3822–3832, 2009.

[5] S. Kar and P. Varshney, "Linear coherent estimation with spatial collaboration," *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3532–3553, 2013.

[6] S. Liu, S. Kar, M. Fardad, and P. K. Varshney, "Sparsity-aware sensor collaboration for linear coherent estimation," *IEEE Transactions on Signal Processing*, vol. 63, no. 10, pp. 2582–2596, May 2015.

[7] Y. Zhao, B. Chen, and R. Zhang, "Optimal power allocation for an energy harvesting estimation system," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2013, pp. 4549–4553.

[8] C. Huang, Y. Zhou, T. Jiang, P. Zhang, and S. Cui, "Power allocation for joint estimation with energy harvesting constraints," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2013, pp. 4804–4808.

[9] M. Nourian, S. Dey, and A. Ahlen, "Distortion minimization in multi-sensor estimation with energy harvesting," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 3, pp. 524–539, Mar. 2015.

[10] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall, Englewood Cliffs, NJ, 1993.

[11] M. Fanaei, M. C. Valenti, A. Jamalipour, and N. A. Schmid, "Optimal power allocation for distributed blue estimation with linear spatial collaboration," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 5452–5456.

[12] C. K. Ho and R. Zhang, "Optimal energy allocation for wireless communications with energy harvesting constraints," *IEEE Transactions on Signal Processing*, vol. 60, no. 9, pp. 4808–4818, Sept. 2012.

[13] L. Huang, J. Walrand, and K. Ramchandran, "Optimal demand response with energy storage management," in *Proc. IEEE Third International Conference on Smart Grid Communications (SmartGridComm)*, Nov. 2012, pp. 61–66.

[14] Y. Wang, X. Lin, and M. Pedram, "A near-optimal model-based control algorithm for households equipped with residential photovoltaic power generation and energy storage systems," *IEEE Transactions on Sustainable Energy*, vol. PP, no. 99, pp. 1–10, 2015.

[15] F. Bugarin, D. Henrion, and J.-B. Lasserre, "Minimizing the sum of many rational functions," *Mathematical Programming Computation*, pp. 1–29, 2015.

[16] F. Bach, R. Jenatton, J. Mairal, and G. Obozinski, "Optimization with sparsity-inducing penalties," *Found. Trends Mach. Learn.*, vol. 4, no. 1, pp. 1–106, Jan. 2012.

[17] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, 2004.

[18] A. L. Yuille and A. Rangarajan, "The concave-convex procedure," *Neural Computation*, vol. 15, no. 4, pp. 915–936, 2003.

[19] T. Lipp and S. Boyd, "Variations and extensions of the convex-concave procedure," http://web.stanford.edu/~boyd/papers/pdf/cvx_ccv.pdf, 2014.

[20] A. Nemirovski, "Interior point polynomial time methods in convex programming," 2012 [Online], Available: *http://www2.isye.gatech.edu/~nemirovs/Lect_IPM.pdf*.