Decentralized Sparse Signal Recovery for Compressive Sleeping Wireless Sensor Networks

Qing Ling and Zhi Tian, *Senior Member, IEEE*

***Abstract*—This paper develops an optimal decentralized algorithm for sparse signal recovery and demonstrates its application in monitoring localized phenomena using energy-constrained large-scale wireless sensor networks. Capitalizing on the spatial sparsity of localized phenomena, compressive data collection is enforced by turning off a fraction of sensors using a simple random node sleeping strategy, which conserves sensing energy and prolongs network lifetime. In the absence of a fusion center, sparse signal recovery via decentralized in-network processing is developed, based on a consensus optimization formulation and the alternating direction method of multipliers. In the proposed algorithm, each active sensor monitors and recovers its local region only, collaborates with its neighboring active sensors through low-power one-hop communication, and iteratively improves the local estimates until reaching the global optimum. Because each sensor monitors the local region rather than the entire large field, the iterative algorithm converges fast, in addition to being scalable in terms of transmission and computation costs. Further, through collaboration, the sensing performance is globally optimal and attains a high spatial resolution commensurate with the node density of the original network containing both active and inactive sensors. Simulations demonstrate the performance of the proposed approach.**

***Index Terms*—Alternating direction method of multipliers, compressive sensing, consensus optimization, decentralized sparse signal recovery, Wireless sensor networks.**

I. Introduction

RECENT advances in compressive sensing have demonstrated that signals which are sparse in certain domain can be recovered from a small set of measurements [1]–[3]. The appealing reduction in signal acquisition and storage costs has spawned a range of signal processing applications, particularly for imaging and spectral analysis. A niche application of interest in this paper is for monitoring localized phenomena using large-scale wireless sensor networks.

Wireless sensor networks have found increasing applications in important monitoring problems, but the energy efficiency and network robustness issues are still quite perplexing for a large-scale network of battery-powered, low-cost wireless sensors [4], [5]. A particularly challenging scenario is when the physical phenomena under monitoring exhibit localized features that appear sparsely over a large sensing field; that is, the physical field of interest can be described by a sparse signal in the spatial domain. Examples abound in a broad range of monitoring applications, such as tracking multiple sources/targets, sensing the underutilized spectrum in a cognitive radio network, and monitoring civil structural health conditions [6]–[8]. In these cases, sensory measurements do not contribute equally to the monitoring task, because sensors that are far away from source locations may not be able to collect useful measurements for the reconstruction of the physical field. Nevertheless, without prior knowledge of the signal source locations, a large number of sensors need to be densely deployed and always stay on in order to provide adequate spatial resolution for detection and reconstruction of the physical field. Note that the spatial resolution of sensing offered by a uniform network is commensurate to the minimum spatial spacing of sensors; as such, the sensor density becomes an indicator of the spatial sampling rate [9]. When traditional sensing methods are adopted, a wireless sensor network faces the conflicting design objectives of sensing at low energy costs and high spatial resolution.

Recognizing the spatial sparsity of localized phenomena and motivated by the compressive sensing principle, we ask: *is it possible to accurately recover a sparse signal that represents the physical field, at high spatial resolution but using only a fraction of sensory measurements?* Specifically, our idea is to turn off some sensors using a random node sleeping strategy [4], process measurements collected only from active sensors to conserve energy, and recover localized phenomena at a high spatial resolution commensurate to the node density of the original network containing both active and inactive sensors.

The information processing issue raised above is intimately related to the network infrastructure. In a centralized network with a fusion center, the network becomes increasingly energy-consuming and unreliable as the number of sensors increases, due to extensive multi-hop communication between sensors and the fusion center [5]. To improve scalability and robustness of large-scale wireless networks, we focus on decentralized in-network processing in the absence of a fusion center. Under this network structure, active sensors collaboratively recover localized phenomena and seek to reach globally optimal solutions through an iterative in-network procedure, during which each sensor exchanges information only with neighboring active sensors within its one-hop transmission range.

In this paper, we develop a decentralized in-network processing algorithm for recovering spatially sparse signals using a sleeping wireless sensor network. The sensing field under monitoring is represented as a state vector, in which each element

Manuscript received July 20, 2009; accepted March 11, 2010. Date of publication April 08, 2010; date of current version June 16, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Hongbin Li.

Q. Ling is with the Department of Automation, University of Science and Technology of China, Hefei, Anhui, 230027, China (e-mail: qingling@mail.ustc.edu.cn).

Z. Tian is with the Department of Electrical and Computer Engineering, Michigan Technological University, Houghton, MI 49931 USA (e-mail: ztian@mtu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2010.2047721

describes the signal value of the phenomenon occurring at a corresponding sensor point. Because the number of occurring phenomena in the sensing field is smaller than the number of sensors, the state vector is sparse with only a few nonzero elements. The sparsity feature motivates compression during data collection, for which we employ a simple compression scheme by randomly turning off some sensors and using only a fraction of sensory measurements. To guarantee the resolution for monitoring, the key problem is how to make decisions for both active and inactive sensors when inactive sensors are unable to neither collect measurements nor exchange information. This work addresses this problem via formulating the collaborative sparse signal recovery problem as a consensus optimization problem, where an ℓ_1 regularized least squares formulation is adopted to incorporate the sparsity knowledge. An active sensor not only optimizes for itself, but also optimizes for its inactive neighbors; the active sensors finally reach consensus for the inactive sensors. As a result, a phenomenon, regardless of whether occurring at an active or inactive sensor point, can be detected and quantified. This consensus optimization framework is implemented in a decentralized manner using the alternating direction method of multipliers for separable convex programs [10].

The rest of the paper is organized as follows. Section II briefly surveys related work. In Section III, data compression is introduced by a random node sleeping strategy, and the corresponding sparse signal recovery task is formulated as an ℓ_1 regularized least squares problem. A consensus optimization formulation is proposed in Section IV, followed by the development of a decentralized and collaborative in-network processing algorithm that only requires local communication. Extensive simulations are provided in Section V to verify the effectiveness of the decentralized sparse signal recovery algorithm. Section VI summarizes the paper.

II. Related Work

Sparse signal recovery has attracted extensive research interest recently. For centralized sensing systems such as cameras and radars, compressive sensing based on random reduced-dimension projections is discussed for both noise-free and noisy cases in [1] and [2]. The key idea is to solve a mathematical program, which minimizes the ℓ_1 norm of the signal subject to constraints of measurements. Deterministic projections such as (inverse) Fourier transformation, followed by a random selection, can also guarantee the quality of recovery [3]. For wireless compressive sensing in a distributed system, a widely advocated idea for data compression is to transform the sensory measurements via a random matrix and transmit the projected measurements to a fusion center. The fusion center then recovers the original sparse signal in a centralized way [11], [12]. In doing so, all sensors are assumed to be active and the compression comes from the choice of the random projection matrix rather than from the sensing process for data acquisition. This method conserves energy by avoiding sending all measurements directly to the fusion center. However, when in the absence of a fusion center, it involves nontrivial scheduling burden to percolate the randomly projected measurements to distributed sensors, and it is an unsolved technical challenge to collaboratively recover the signal in a decentralized manner. Furthermore, the energy consumption of data acquisition cannot be conserved, since all sensors still need to sense and collect measurements.

In regard of the data collection issue, this paper considers a projection scheme that induces data compression and energy conservation during the sensing process. The sparse signal to be recovered, namely the vector of signal strength of the phenomena occurring in the sensing field, is transformed by a deterministic measurement matrix to the sensory measurement domain. The rows of the measurement matrix are then randomly sampled by randomly activating a fraction of sensors, while the rest sensors stay in a sleeping mode without collecting measurements.

Node sleeping strategies have been investigated for wireless sensor networks where energy efficiency is a primary concern [4]. Selection of the sleeping nodes can be deterministic by optimizing a network utility function, or stochastic by randomly turning off a fraction of sensors [9], [13]. In this paper, we consider a simple random sleeping strategy. In the beginning of a reconfiguration period, a random fraction of sensors are forced to be inactive. At the end of the period, these sensors wake up to wait for a new round of reconfiguration. For the coverage, connectivity, synchronization, and scheduling of sleeping networks, readers are referred to [14], [15], and related literature. This work departs from the networking issue by focusing on information processing for a sleeping network.

In terms of information processing, this paper aims for decentralized in-network processing, which is known for improving the scalability and robustness of large-scale wireless networks [16]. The design objective is to accomplish an otherwise centralized task in a fully decentralized manner, in the absence of a fusion center, using a network where each node exchanges information with its one-hop neighbors only. A well-studied decentralized in-network processing method is consensus averaging [17], [18]. Sensors dynamically exchange current estimates with one-hop neighbors and update their local estimates, until the whole network reaches consensus on an averaged scalar. A more complicated task is to collaboratively optimize an objective function, such as in learning problems [19]. In [20], separable objective functions are optimized based on the decentralized incremental subgradient approach. The distributed event-region detection problem is solved by using hidden Markov random field models in [21] and by a graph-based method in [22]. For constrained optimization problems, a recently developed technique is to construct a consensus optimization formulation [23], [24]. Each sensor holds its own local estimates of all the unknowns, and the estimates of neighboring sensors are forced to consent asymptotically. A powerful implementation tool to solve the constrained consensus optimization problem is the alternating direction method of multipliers, which is basically an augmented Lagrangian method [10]. Taking [23] and [24] as examples, the estimation task is formulated to be with a separable objective function and a set of consensus constraints. By iteratively updating local estimates, the network reaches a consensus which minimizes the estimation error. This idea has been applied in [7], which discusses decentralized spectrum sensing for cognitive radios via exploiting sparsity.

This paper also uses the idea of consensus optimization and the tool of the alternating direction method of multipliers. However, a key differentiating feature of our problem formulation is the number of decision variables updated per node in each iteration. Note that an important property of our problem is, the size of the state vector to be recovered is equal to the number of sensors, which is quite large for a large-scale network. Therefore, it can be too costly for each sensor to hold a local estimate of the entire unknown vector as in [23], [24], and [7]. Alternatively, we let an active sensor hold a scalar estimate for itself, and several scalar estimates for its inactive neighbors. Neighboring active sensors then reach consensus for their common inactive neighbors. This dimension reduction scheme considerably reduces the computation and communication costs per sensor, and improves the scalability of the algorithm for a large-scale network. Further, by imposing restrictions on the measurement matrix, we formulate the problem as a separable convex program, which has a neat decentralized solution from the alternating direction method of multipliers.

Finally, we emphasize the difference between *decentralized sparse signal recovery* and *distributed compressed sensing* [25], [26]. The merit of distributed compressed sensing is to recover signals collected from distributed sensing sources via exploiting the joint sparsity; the recovery, however, is generally done at a fusion center in a centralized way. In this paper we also utilize the sparsity of signals observed by distributively located sensors, but focus on decentralized optimization in a large-scale sleeping wireless sensor network. In the absence of a fusion center, we have to deal with the challenge that each node only has its own measurement, not all measurements from all sensors, in the signal recovery process.

III. Problem Formulation

Consider a dense wireless sensor network deployed in a two-dimensional field. The network has a set of L sensors at locations $\{v_i\}$, indexed by $i \in \mathcal{L} = \{1, \ldots, L\} = \mathcal{W} \cup \mathcal{S}$, in which $\mathcal{W}$ is the subset of sensor indices for active sensors with cardinality $W = |\mathcal{W}|$ and $\mathcal{S}$ is the subset of inactive sensors in the sleeping mode for energy conservation with cardinality $S = |\mathcal{S}|$, where $|\cdot|$ denotes cardinality. Transmitting at low power, each sensor can only communicate locally with its one-hop neighbors within the communication range r_C, which is assumed to be the same for all sensors. The network is connected if there is at least one undirected path between any pair of sensors. To enable decentralized decision-making over the entire multi-hop network, we make a basic assumption on the network connectivity: **(A1)** Given the communication range r_C, the original network composed of all sensors in $\mathcal{L}$ and the active network composed of all sensors in $\mathcal{W}$ are both connected. For any inactive sensor, the subnetwork consisting of its active neighboring nodes is also connected.

At each sampling time, multiple phenomena (a.k.a. signal sources) may occur in the sensing field. When sensors are densely deployed to provide adequate spatial resolution, the locations of these phenomena can be well approximated to coincide with some sensor locations. Let $\mathbf{c} = [c_1, \ldots, c_L]^T$ denote the signal source vector of interest, where c_j corresponds to be source value at $v_j, \forall j \in \mathcal{L}$. As depicted in Fig. 1, a unit-intensity phenomenon originating at a sensor point v_j may influence its neighboring area $\mathcal{R}_j$ through an influence function $f_j(v)$, which is non-zero only for locations $v \in \mathcal{R}_j$

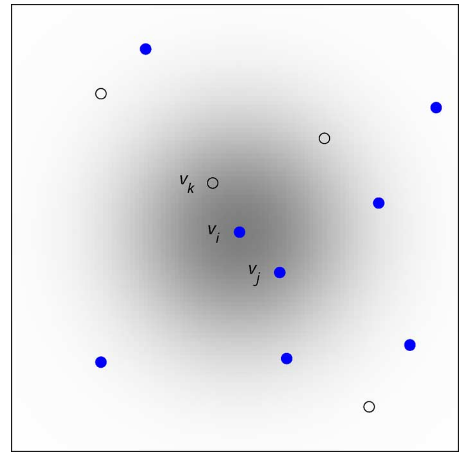


Fig. 1. A sleeping network of both active sensors (solid circles) and inactive sensors (unfilled circles). Only a small region of a large sensing field is depicted to highlight one phenomenon at the sensor point v_i. This phenomenon influences the shaded area centered at v_i, including both active neighbors $j \in \mathcal{W}$ and inactive neighbors $k \in \mathcal{S}$.

and is normalized to obey $f_j(v_j) = 1$. Meanwhile, the field measurement b_i at a sensor point v_i can be described by the superposition of the influence of all phenomena on v_i. This paper focuses on the scenario where phenomena occur sparsely in the large spatial domain. As aforementioned in Section I, this scenario not only appears widely in many practical sensor network applications, but also entails distinct challenges and opportunities for energy-efficient information processing. As such, we have the following assumption on the field signals and measurements: **(A2)** The measurement b_i of sensor i can be represented as

$$b_i = \sum_{j \in \mathcal{L}} f_{ji} c_j + e_i$$

where $c_j \geq 0$ is the amplitude of the phenomenon occurring at sensor point v_j, $f_{ji} c_j = f_j(v_i) c_j$ is the influence of this phenomenon on sensor point v_i, and e_i is the random measurement noise of zero mean. The signal vector $\mathbf{c}$ is sparse, namely, the number of phenomena is much smaller than the number of sensors in the dense large network.

Here we have assumed $\mathbf{c} \geq \mathbf{0}$ for exposition clarity, without loss of generality. An illustrative example is acoustic source localization, in which the networked measurements are generally modeled as linear superposition of acoustic intensities of multiple sources [27]. When the signal sources are independent random variables with zero means, each sensory measurement of the composite signal power is the summation of individual powers after being attenuated by the propagating environment. Hence, c_j refers to the non-negative signal power intensity, and f_{ji} indicates the distance-dependent power attenuation. In other applications where c_j is negative, we can re-write it as $c_j = c_j^+ - c_j^-$ where c_j^+ and c_j^- are both non-negative [28]. Replacing c_j by $[c_j^+, c_j^-]^T$ in the vector $\mathbf{c}$, the assumption $\mathbf{c} \geq \mathbf{0}$ still holds.

It is worth clarifying that the signal source vector $\mathbf{c}$ is sparse, but through the influence functions $\{f_{ji}\}$, the sensor readings $\{b_i\}$ can be mostly nonzero or non-sparse. A source at location

v_i is declared not when $b_i \neq 0$, but only when $c_i \neq 0$. Note that the occurrence and strength of the phenomena, reflected by the locations and values of those nonzero entries in $\mathbf{c}$, are unknown and hence to be monitored; yet the influence function values, described by $\{f_{ji}\}_{j,i\in\mathcal{L}}$, can be learned during the network deployment stage. This is done either through on-site calibration to directly measure $\{f_{ji}\}_{j,i\in\mathcal{L}}$, or by modeling the influence functions $\{f_j(v)\}_{j\in\mathcal{L}}$ by (a few) parameterized basis functions and then learning those parameters during network initialization. This paper focuses on the task of online monitoring, assuming that $\{f_{ji}\}_{j,i\in\mathcal{L}}$ have been acquired.

Further, to facilitate decentralized decision-making via local one-hop communication in a large-scale network, we properly adjust the transmission power of sensors such that the communication range r_C is larger than the radius of each influence area $\mathcal{R}_j, \forall j$. Specifically, **(A3)** For any sensor i, the influence function $f_{ji} = 0$ if the distance from j to i is larger than the communication range $r_C, \forall j$.

The assumption **(A3)** holds for a wide range of sensing problems where the phenomena under monitoring are local events compared with the large sensing field. A phenomenon hardly influences the measurements at a faraway sensor point (cf., Fig. 1). As a result, when r_C is large enough, f_{ji} is close to 0 for two non-neighboring sensors at v_i and v_j; otherwise this assumption results in a truncation error. We will address the issue of truncation errors in the simulations.

Summarizing the assumptions **(A2)** and **(A3)**, and denoting $\mathcal{N}_i$ as the set of neighboring sensors of sensor i, the measurement b_i becomes:

$$b_i = c_i + \sum_{j\in\mathcal{N}_i} f_{ji}c_j + e_i, i \in \mathcal{W}. \tag{1}$$

or in a matrix form:

$$\mathbf{b}_W = \mathbf{\Phi}\mathbf{F}\mathbf{c} + \mathbf{e}_W. \tag{2}$$

where $\mathbf{F}$ is the $L \times L$ matrix whose (i,j)-th element is f_{ji}, $\mathbf{\Phi}$ is the $W \times L$ selection matrix which selects the W rows of $\mathbf{F}$ corresponding to the active sensors, and $\mathbf{b}_W$ and $\mathbf{e}_W$ are the $W \times 1$ measurement vector and noise vector respectively.

Given $\{f_{ji}\}_{j,i\in\mathcal{L}}$, the goal of this paper is to recover the $L\times 1$ sparse signal vector $\mathbf{c}$ from the W measurements $\{b_i\}_{i\in\mathcal{W}}$ collected from active sensors. In particular, we aim to perform the sparse signal recovery in a decentralized manner in the absence of a fusion center.

To solve the inverse problem for the under-determined linear system in (1), the prior knowledge of $\mathbf{c}$ being sparse needs to be utilized. A sparsity metric for $\mathbf{c}$ is its ℓ_1 norm, which reduces to $||\mathbf{c}||_1 = \sum_{i\in\mathcal{L}} c_i$ for $\mathbf{c} \geq \mathbf{0}$. Accordingly, we formulate the following ℓ_1 regularized least squares problem [28], [29], with additional nonnegative constraints:

$$\begin{aligned} \min_{\mathbf{c}} \quad & \frac{\lambda}{2}\|\mathbf{\Phi}\mathbf{F}\mathbf{c} - \mathbf{b}_W\|_2^2 + \|\mathbf{c}\|_1 \\ & = \frac{\lambda}{2}\sum_{i\in\mathcal{W}}\left(b_i - c_i - \sum_{j\in\mathcal{N}_i} f_{ji}c_j\right)^2 + \sum_{j\in\mathcal{L}} c_j \\ \text{s.t.} \quad & \mathbf{c} \geq 0. \end{aligned} \tag{3}$$

Here the objective function consists of a least-squares ℓ_2 norm term and a sparsity-enforcing ℓ_1 norm term, with a nonnegative weighting coefficient λ reflecting the tradeoff between these two terms. This formulation is a generalized form of the well-known basis pursuit de-noising (BPDN) [30] and the least absolute shrinkage and selection operator (LASSO) [31], which are conventionally solved by convex programming in a centralized manner.

IV. Decentralized Sparse Signal Recovery

In this section, we reformulate the sparse signal recovery problem in (3) to an equivalent consensus optimization problem. An essential difference from a conventional consensus optimization formulation is that we let each sensor make decisions for both itself and its inactive neighbors, but not to seek recovery of the entire field vector $\mathbf{c}$. Decisions on all active and inactive nodes eventually reach network-wide consensus. This strategy effectively reduces the number of decision variables for each active sensor, and in turn alleviates the computational costs and expedites convergence during iterative consensus optimization.

A. *Consensus Optimization Formulation*

During online monitoring, the goal is to decide the signal strength c_i at each sensor point $v_i, i \in \mathcal{L}$. The main challenge in designing a decentralized algorithm for (3) is that inactive sensors are unable to decide for themselves. Our approach to tackle this challenge is to let each active sensor $i \in \mathcal{W}$ decide the signal c_i occurring at its own location as well as the signals c_k occurring at its inactive neighboring sensors $\forall k \in \mathcal{S} \cap \mathcal{N}_i$. To do so, we let each active sensor i keep local copies of its decisions on c_i and $\{c_k\}_{k\in\mathcal{S}\cap\mathcal{N}_i}$ as $c_i^{(i)}$ and $\{c_k^{(i)}\}_{k\in\mathcal{S}\cap\mathcal{N}_i}$, respectively. The decisions on each inactive sensor are forced to reach consensus among all its neighboring active sensors, such that the network eventually consents on all estimates that are globally optimal. Based on this idea, we reformulate (3) into an equivalent consensus optimization problem as follows:

$$\min \quad \frac{\lambda}{2}\sum_{i\in\mathcal{W}}\left(b_i - c_i^{(i)} - \sum_{k\in\mathcal{S}\cap\mathcal{N}_i} f_{ki}c_k^{(i)} - \sum_{j\in\mathcal{W}\cap\mathcal{N}_i} f_{ji}c_j^{(j)}\right)^2$$

$$+ \sum_{i\in\mathcal{W}} c_i^{(i)} + \sum_{k\in\mathcal{S}} w_k \sum_{j\in\mathcal{W}\cap\mathcal{N}_k} c_k^{(j)} \tag{4a}$$

$$\text{s.t.} c_k^{(i)} = c_k^{(j)}, \forall i \in \mathcal{W}, j \in \mathcal{W}\cap\mathcal{N}_i, \quad k \in \mathcal{S}\cap\mathcal{N}_i\cap\mathcal{N}_j \tag{4b}$$

$$c_i^{(i)} \geq 0, \forall i \in \mathcal{W} \tag{4c}$$

$$c_k^{(i)} \geq 0, \forall i \in \mathcal{W}, \forall k \in \mathcal{S}\cap\mathcal{N}_i. \tag{4d}$$

Here, w_k^{-1} denotes the number of active neighbors of a sleeping node at v_k, and the corresponding signal value c_k inside the ℓ_1 norm term in (3) is replaced in (4a) by averaging the local copies of all its active neighboring sensors $j \in \mathcal{W}\cap\mathcal{N}_k$. The constraint (4b) enforces consensus on the decisions for each inactive sensor among all its active neighboring sensors. The non-negativity constraints in (4c) and (4d) are imposed for all decision variables.

We have the following proposition for the equivalence of (3) and (4):

TABLE I
DECENTRALIZED SPARSE SIGNAL RECOVERY ALGORITHM

Step 1:	*Initialization.* Each active sensor sets the decision variable, slack variable, and multipliers as 0.
Step 2:	*Exchanging Decision Variables.* To each of its active neighbors, an active sensor i transmits its current decision variables for itself and the common inactive neighbors.
Step 3:	*Updating Slack Variables and Lagrange multipliers.* An active sensor i locally updates its slack variable s_i, multipliers $\{z_{ki}\}$, $\{p_{ki}\}$ and $\{\lambda_{jik}\}$ according to (6), (7), (8), and (9).
Step 4:	*Exchanging Slack Variables and Lagrange multipliers.* To each active neighbor, an active sensor v_i transmits its current slack variable and corresponding Lagrange multipliers.
Step 5:	*Updating Decision Variables.* An active sensor i updates its own decision variable $c_i^{(i)}$, and the decision variables $\{c_k^{(i)}\}_{k\in\mathcal{N}_i\cap\mathcal{S}}$ of its inactive neighbors according to (10) and (11).
Step 6:	*Iteration.* All active sensors simultaneously execute Steps 2-5 or take turns to do so, and iteratively repeat these steps until convergence.

Proposition 1: Under the assumption **(A1)**, the formulations in (3) and (4) are equivalent.

Proof: From **(A1)**, for any inactive sensor k, the subnetwork containing its active neighbors is connected. Hence for any $i,j\in\mathcal{W}\cap\mathcal{N}_k$, $c_k^{(i)}$ and $c_k^{(j)}$ are forced to be equal according to the consensus constraints (4b). Simply rewriting $c_k=c_k^{(i)}=c_k^{(j)},\forall k\in\mathcal{S},\forall i,j\in\mathcal{W}\cap\mathcal{N}_k$ and $c_i=c_i^{(i)},\forall i\in\mathcal{W}$, it follows immediately that (3) and (4) are equivalent.

B. Decentralized Algorithm Design

To facilitate decentralized processing, we further reformulate (4) as a separable convex program, and solve it using the alternating direction method of multipliers [10]. To do so, let us introduce a set of slack variables $\{s_i\}_{i\in\mathcal{W}}$ to indicate the measurement errors. Then (4) can be equivalently rewritten as

$$\min\ \frac{\lambda}{2}\sum_{i\in\mathcal{W}}s_i^2+\sum_{i\in\mathcal{W}}c_i^{(i)}+\sum_{k\in\mathcal{S}}w_k\sum_{j\in\mathcal{W}\cap\mathcal{N}_k}c_k^{(j)} \tag{5a}$$

$$\text{s.t.}\ c_i^{(i)}+\sum_{k\in\mathcal{S}\cap\mathcal{N}_i}f_{ki}c_k^{(i)}+\sum_{j\in\mathcal{W}\cap\mathcal{N}_i}f_{ji}c_j^{(j)}=s_i+b_i,\ \forall i\in\mathcal{W} \tag{5b}$$

$$c_k^{(i)}=c_k^{(j)},\ \forall i\in\mathcal{W},j\in\mathcal{W}\cap\mathcal{N}_i,k\in\mathcal{S}\cap\mathcal{N}_i\cap\mathcal{N}_j \tag{5c}$$

$$c_i^{(i)}\geq 0,\quad\forall i\in\mathcal{W} \tag{5d}$$

$$c_k^{(i)}\geq 0,\quad\forall i\in\mathcal{W},\forall k\in\mathcal{S}\cap\mathcal{N}_i. \tag{5e}$$

The alternating direction method of multipliers forms a constrained augmented Lagrangian function from (5), and then iteratively optimizes it based on the block coordinate descent algorithm. During each iteration, each active sensor $i\in\mathcal{W}$ minimizes the constrained augmented Lagrangian function over its own decision variable $c_i^{(i)}$, slack variable s_i, and decision variables $\{c_k^{(i)}\}_{k\in\mathcal{S}\cap\mathcal{N}_i}$ for its sleeping neighbors. Meanwhile, the multipliers are updated and exchanged among neighboring active sensors. Two neighboring active sensors also exchange the estimates on their common neighboring inactive sensors in order to enforce the consensus constraints. The optimal solution to (5) is an iterative one, which is derived in detail in the Appendix and summarized in Proposition 2. Accordingly, we propose a decentralized sparse signal recovery algorithm presented in Table I.

Proposition 2: Let $[\cdot]^+$ denote the projection operator $\max\{\cdot,0\}, m_i=|\mathcal{N}_i|+1$ denote the number of neighbors of any active sensor i plus 1, $\forall i$, and β be a constant coefficient in the augmented Lagrangian method. The iterative steps in (6)–(11) converge to the globally optimal solution of (5).

$$\begin{aligned}s_i(t+1)=\frac{\beta}{\lambda m_i^2+\beta}\Bigg(&\sum_{k\in(\mathcal{N}_i\cap\mathcal{W})\cup i}\Big(f_{ki}c_k^{(k)}(t)-z_{ki}(t)+\frac{1}{\beta}p_{ki}(t)\Big)\\&+\sum_{k\in\mathcal{N}_i\cap\mathcal{S}}\Big(f_{ki}c_k^{(i)}(t)-z_{ki}(t)+\frac{1}{\beta}p_{ki}(t)\Big)\Bigg),\quad\forall i\in\mathcal{W}\end{aligned} \tag{6}$$

$$\begin{aligned}z_{ki}(t+1)=&f_{ki}c_k^{(k)}(t)+\frac{1}{m_i}\Bigg(b_i-\sum_{j\in(\mathcal{N}_i\cap\mathcal{W})\cup i}f_{ji}c_j^{(j)}(t)-\sum_{j\in\mathcal{N}_i\cap\mathcal{S}}f_{ji}c_j^{(i)}(t)\Bigg)\\&+\frac{1}{\beta}\Bigg(p_{ki}(t)-\frac{1}{m_i}\sum_{j\in\mathcal{N}_i\cup i}p_{ji}(t)\Bigg)\\&\forall i\in\mathcal{W},k\in(\mathcal{N}_i\cap\mathcal{W})\cup i;\end{aligned} \tag{7}$$

$$\begin{aligned}z_{ki}(t+1)=&f_{ki}c_k^{(i)}(t)+\frac{1}{m_i}\Bigg(b_i-\sum_{j\in(\mathcal{N}_i\cap\mathcal{W})\cup i}f_{ji}c_j^{(j)}(t)-\sum_{j\in\mathcal{N}_i\cap\mathcal{S}}f_{ji}c_j^{(i)}(t)\Bigg)\\&+\frac{1}{\beta}\Bigg(p_{ki}(t)-\frac{1}{m_i}\sum_{j\in\mathcal{N}_i\cup i}p_{ji}(t)\Bigg),\\&\forall i\in\mathcal{W},k\in\mathcal{N}_i\cap\mathcal{S};\\p_{ki}(t+1)=&p_{ki}(t)+\beta\Big(f_{ki}c_k^{(k)}(t)-\frac{1}{m_i}s_i(t+1)-z_{ki}(t+1)\Big)\\&\forall i\in\mathcal{W},k\in\mathcal{N}_i\cap\mathcal{W}\cup i,\end{aligned} \tag{8}$$

$$p_{ki}(t+1) = p_{ki}(t) + \beta\left(f_{ki}c_k^{(i)}(t) - \frac{1}{m_i}s_i(t+1) - z_{ki}(t+1)\right) \quad \forall i \in \mathcal{W}, k \in \mathcal{N}_i \cap \mathcal{S}$$

$$\lambda_{ijk}(t+1) = \lambda_{ijk}(t) + \beta\left(c_k^{(i)}(t) - c_k^{(j)}(t)\right) \quad \forall i \in \mathcal{W}, j \in \mathcal{W} \cap \mathcal{N}_i, \; k \in \mathcal{S} \cap \mathcal{N}_i \cap \mathcal{N}_j, \tag{9}$$

$$c_i^{(i)}(t+1) = \left[\frac{m_i^{(i)}(t+1)}{n_i^{(i)}(t+1)}\right]^+ \quad \forall i \in \mathcal{W} \tag{10}$$

where

$$m_i^{(i)}(t+1) = \sum_{k \in (\mathcal{N}_i \cap \mathcal{W}) \cup i} \beta f_{ik} \times \left(\frac{1}{m_k}s_k(t+1) + z_{ik}(t+1) - \frac{1}{\beta}p_{ik}(t+1)\right) - 1,$$

$$n_i^{(i)}(t+1) = \sum_{k \in (\mathcal{N}_i \cap \mathcal{W}) \cup i} \beta f_{ik}^2,$$

$$c_k^{(i)}(t+1) = \left[\frac{m_k^{(i)}(t+1)}{n_k^{(i)}(t+1)}\right]^+, \quad \forall i \in \mathcal{W}, k \in \mathcal{S} \cap \mathcal{N}_i, \tag{11}$$

where

$$m_k^{(i)}(t+1) = \sum_{j \in \mathcal{N}_i \cap \mathcal{N}_k \cap \mathcal{W}} \left(2\beta c_k^{(j)}(t) - \lambda_{ijk}(t+1) + \lambda_{jik}(t+1)\right) + \beta f_{ki}\left(\frac{1}{m_i}s_i(t+1) + z_{ki}(t+1) - \frac{1}{\beta}p_{ki}(t+1)\right) - w_k,$$

$$n_k^{(i)}(t+1) = \beta f_{ki}^2 + 2\sum_{j \in \mathcal{W} \cap \mathcal{N}_i \cap \mathcal{N}_k} \beta.$$

For an active sensor i, s_i in (6) is the slack variable that helps to construct the separable convex program in (5); z_{ki} and p_{ki} in (7) and (8) are two intermediate variables held by sensor i for its neighbor k, playing the role of Lagrange multipliers associated with the equality constraint (5b). For any two neighboring active sensors i and j and their common inactive neighbor k, λ_{jik} in (9) is the Lagrange multiplier held by i, associated with the consensus constraint (5b). Finally, each active sensor iteratively updates its own decision variable and the decision variables for its inactive neighbors according to (10) and (11). Evidently, the operations of each active sensor simply boils down to summations and multiplications, which are manageable for low-cost sensor nodes.

In a practical sleeping network, sensors change their sleeping mode based on a predefined mechanism. Thereafter the network reconfigures to collect the parameters $\{w_i\}_{i \in \mathcal{S}}$ and $\{m_i\}_{i \in \mathcal{W}}$. Each sensor scheduled to sleep needs to count the number of its neighboring sensors that are scheduled to be active, and broadcast to its active neighboring sensors, while each sensor scheduled to be active needs to count the number of its neighboring sensors. Given the network configuration, the decentralized sparse signal recovery algorithm is executed upon new measurements. After the algorithm converges, each active sensor holds the signal estimates for itself and its inactive neighbors. Such decentralized in-network processing is performed for energy-efficient online monitoring, and sensors may alarm or communicate with a central console to report the estimated phenomena when needed.

C. Discussions

This section discusses several important application-related issues for the proposed decentralized sparse signal recovery algorithm, including communication load, recovery accuracy, and resolution for detection. These issues guide the choice of system parameters, such as communication range, influence function, sensor density, the weighting coefficient λ, and the fraction of inactive sensors in the network.

Communication Load: Communication consumes a significant portion of the energy in a wireless sensor network, and hence communication load is a major concern for decentralized algorithm design. In the proposed algorithm, active sensors need to exchange intermediate decision variables, slack variables, and Lagrange multipliers in each iteration. This can be done via local broadcasting such that all its one-hop active sensors can acquire the data. Fig. 2 depicts the information flow for an illustrative small network containing two connected active sensors at v_1 and v_3 and one common neighboring inactive sensor at v_2. Each active sensor at v_i transmits one decision variable, one slack variable, and two multipliers to its active neighbor $j \in \mathcal{W} \cap \mathcal{N}_i$. Also, it transmits one decision variable and one multiplier λ_{jik} to its active neighbor j, if there is an inactive sensor $k \in \mathcal{S} \cap \mathcal{N}_i \cap \mathcal{N}_j$ within one-hop from both i and j. Therefore, the total number of messages transmitted from all sensors for updating all local estimates during each iteration is given by $N_{\text{com}} = 4\sum_{i \in \mathcal{W}} |\mathcal{N}_i \cap \mathcal{W}| + 2\sum_{i,j \in \mathcal{W}} |\mathcal{N}_i \cap \mathcal{N}_j \cap \mathcal{S}|$. This quantity can be further reduced if each sensor i simply broadcasts its messages to all neighbors instead of talking to each neighbor one by one.

The communication overhead analyzed above is to be calibrated under the context of large-scale dense networks. Let us consider a large circular sensing field whose area is $A = 2\pi R^2$ with R being the radius. A large number of L sensors are randomly deployed at a density $\rho = L/A$. The ratio of the one-hop communication region to the entire sensing field is indicative of the network locality, which we define as $\eta = r_C^2/R^2 \ll 1$. Note that we have selected r_C to be comparable to the radii of the influence areas of phenomena, as in the assumption **(A3)**. Further, we denote the fraction of active sensors as $\mu = W/L \in (0, 1]$. Given r_C and ρ, the average number of sensors within each one-hop region is $L_{\text{lo}} = \rho \cdot 2\pi r_C^2$. When L and A are large, and assuming uniform distribution of sensor points, the average

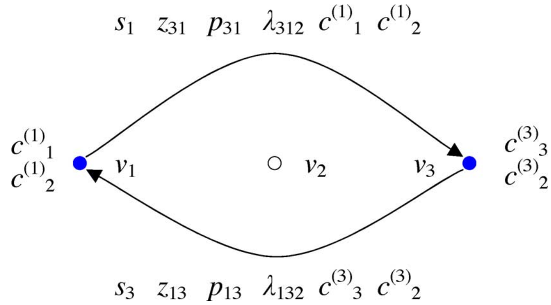


Fig. 2. Information flow for a small network containing two connected active sensors at v_1 and v_3 and one common inactive neighboring sensor at v_2. During each iteration, sensor 1 sends one slack variable s_1, three multipliers z_{31}, p_{31}, and λ_{312}, and two decision variables $c_1^{(1)}$ and $c_2^{(1)}$ to 3.

number of messages exchanged for all nodes to update their decisions in one iteration is approximately $N_{\rm com} = 4\mu^2 L_{\rm lo} L + 2(1-\mu)\mu^2 L_{\rm lo}^2 L = \alpha_{\rm com} L$, which is linear in the network size L for some scalar $\alpha_{\rm com}$ of localized scale.

Let us assume the free-space model in which power falls off proportionally to square-distance. Accordingly, the power cost in communication is on the order of $\mathcal{O}(\alpha_{\rm com} L r_C^2 T)$, where T is the number of iterations for convergence. In contrast, in a centralized network, if each active node raises its transmit power to send one message (measurement) to a fusion center in the center of the sensing field, then the power cost is $\mathcal{O}(\mu L R^2/4)$, which can be quite large for a large network with a large R. Alternatively, each sensor can send its message to the fusion center via multiple hops, where the number of hops is inversely proportional to $\sqrt{\eta^{-1}} = R/r_C$. At the expense of routing cost, this strategy reduces the communication power cost to be on the order of $\mathcal{O}(\mu L r_C R)$, which is still quite high for a large R. Overall, our decentralized optimal solution entails an energy saving on the order of at least r_C/R, which can be quite significant for a large-scale network.

Recovery Accuracy: Depending on the practical applications of interest, it may be unnecessary to accurately recover the strength of phenomena; rather, the main concerns can be to avoid false alarms or missed detection. Hence, the recovery accuracy shall be assessed based on the network goal. In general, there are several major performance-determining factors.

Firstly, the accuracy of field recovery is decided by the choice of the influence functions. Besides a parameterized approach that we will illustrate through simulations, a general approach is to assume a set of common basis functions and then generate the individual influence functions as linear combinations of them, where the combining coefficients are acquired through online learning [8]. In this approach, the basis functions must be carefully selected in order to accurately describe the propagation of phenomena in the sensing area.

Secondly, the collected data shall contain adequate information in order to guarantee successful recovery. If the fraction of active sensors μ is too small, then the optimization based on (3) may lead to incorrect result. This problem is analogue to the choice of compression ratio in compressive sensing.

Thirdly, the communication range affects recovery accuracy. As discussed in the assumption **(A2)**, the influence of a phenomenon on a sensor is assumed to be negligible, if the phenomenon occurs outside the communication range of the sensor. Hence the communication range should be large enough such that the truncation error is negligible or small. A properly selected weight λ in (3) helps to improve the robustness of recovery accuracy against truncation errors, as we will illustrate via simulations.

Lastly, we briefly discuss the role of the weight λ in (3) that trades off the importance of the least squares ℓ_2 norm term and the sparsity-enforcing ℓ_1 norm term in the design objective. The limiting behaviors as $\lambda \to \infty$ and $\lambda \to 0$ have been discussed in [29]. As $\lambda \to \infty$, the limiting point of the optimal solution to $\mathbf{c}$ has the smallest ℓ_1 norm among all non-negative points that satisfy $\mathbf{F}^T \mathbf{\Phi}^T (\mathbf{\Phi}\mathbf{F}\mathbf{c} - \mathbf{b}_W) = 0$, if these points exist. And there is a constant $\lambda_{\min} = 1/||\mathbf{F}^T \mathbf{\Phi}^T \mathbf{b}_W||_\infty$, such that if $\lambda \le \lambda_{\min}$, the optimal solution is 0. Suppose that all sensors are active, namely $\mathbf{\Phi}$ is an identity matrix, and the measurement matrix $\mathbf{\Phi}\mathbf{F}$ is invertible. In the noise-free case, $\lambda \to \infty$ leads to exact recovery. However, in the presence of measurement noise or modelling errors, $\lambda \to \infty$ tends to produce a non-sparse solution. On the other hand, a small λ results in a sparse solution; it tends to correctly identify the nonzero support of the signal $\mathbf{c}$ and hence the locations of occurring phenomena, but the estimation errors of those nonzero elements can be large. Nevertheless, λ should be larger than $\lambda_{\min}$ in order to avoid the trivial solution of all zeros and hence alleviate missed detection.

Resolution for Detection: In **(A2)**, it is assumed that a source only occurs at a sensor point. The modeling error is small when sensors are densely deployed. To bypass this assumption, it is possible to assume that sources can occur at any points, whose positions can be set as optimization variables. However, the resulting nonlinear formulation is computationally intractable in a practical sensor network. This nonlinear inverse problem can be avoided by assuming a virtual grid in the sensing field and confining the sources to appear sparsely on the grid points [7]. The resolution of detection is improved by adopting a fine-scale grid, at the expense of greatly increased computational and communication costs. Specifically, the number of decision variables to be solved at each active sensor is the same as the number of grid points, and active sensors need to reach consensus for all grid points. In contrast, in this paper the tentative decisions of each sensor only involve itself and neighboring sensors. Each sensor does not seek global awareness of the entire field, in exchange for lowered communication load and improved network convergence.

Because of the modeling simplicity in our current setting, the resolution for detection depends on the sensor density. For a non-sleeping network, the spatial resolution is $\rho = L/A$, at a power cost proportional to L. For a regular sleeping network that ignores inactive sensors during decision making, the spatial resolution is reduced to $\mu\rho$, at a reduced power cost proportional to μL. Our proposed algorithm performs sparse signal recovery to make decisions for both active and inactive sensors, which retains the resolution at ρ. Meanwhile, the power cost is reduced to be proportional to μL. This is a notable advantage, as we optimize both the spatial resolution and energy conservation by coupling the sleeping strategy with sparse signal recovery.

V. Simulation Results

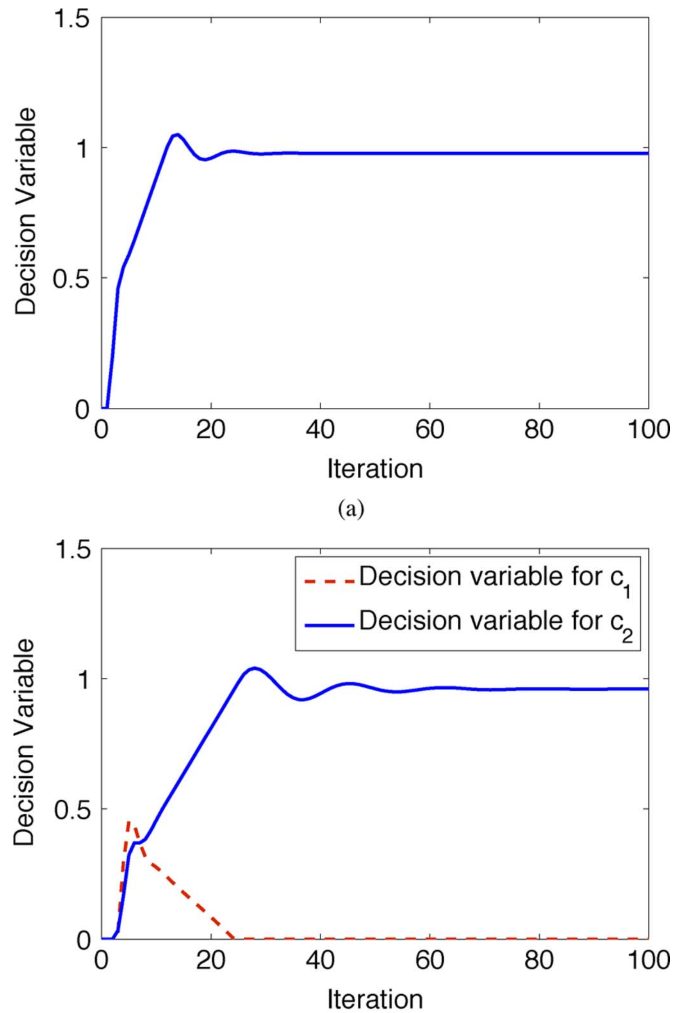
This section provides extensive simulation results to validate the effectiveness of proposed sparse signal recovery algorithm. First, small networks are used to demonstrate the basic properties of the proposed algorithm. Then, large networks are simulated to demonstrate the scalability of the decentralized algorithm. In all simulations, the constant coefficient in the augmented Lagrangian method is set to be $\beta = 1$.

A. Small Networks

We firstly consider applications in small networks to demonstrate optimality, noise resilience, recovery accuracy, and detection resolution of the proposed algorithm. Without loss of generality, we focus on the scenario in which: 1) sensors are evenly deployed in a one-dimensional space; 2) only one phenomenon occurs in the sensing area; and 3) sensors can directly communicate with each other, namely, the communication range r_C is infinite. Taking a parameterized approach for modeling, the influence function is supposed to be a Gaussian shape with a tunable width σ that can be learned. Specifically, if a phenomenon occurs at sensor point v_i, then the noise-free output at sensor point v_j is $c_i f_i(v_j) = c_i \exp(-d_{ij}^2/\sigma^2)$, where d_{ij} is the distance between v_i and v_j, and σ is known after learning.

Suppose that three sensors are deployed at $v_1 = (0,0)$, $v_2 = (20,0)$, and $v_3 = (40,0)$. A phenomenon occurs at v_2, with $\sigma = 40$ and $c_2 = 1$. Hence the optimal recovery is $\mathbf{c} = [0,1,0]^T$. In the absence of measurement noise, we set the weight $\lambda = 10$ in the proposed algorithm. Fig. 3 depicts the optimization results of $c_2^{(2)}$ when all sensors are active, and $c_2^{(1)}$ and $c_1^{(1)}$ when the sensor at v_2 is inactive, respectively. When the three sensors are all active, each sensor makes decision for itself. Decision variables $c_1^{(1)}$ and $c_3^{(3)}$ remain to be 0, whereas $c_2^{(2)}$ converges to the optimal solution 0.9781, which is near to the true value 1. When sensor 2 is inactive, sensors at v_1 and v_3 need to make decisions for themselves, and further reach a consensus for the sensor at v_2. The decision variable $c_1^{(1)}$ converges to 0 after a transient state, and $c_2^{(1)}$ also converges to its optimal solution 0.9603, which is also near to the true value 1. Convergence of $c_3^{(3)}$ and $c_2^{(3)}$ is similar to that of $c_1^{(1)}$ and $c_2^{(1)}$. Note that in the first case $\lambda_{\min} = 0.5423$ and in the second case $\lambda_{\min} = 0.4519$; both of them are smaller than the chosen weight $\lambda = 10$. The small bias in the estimate is due to the ℓ_1 norm term in (3), which enforces a sparse solution in order to reduce the estimation mean-square error and the false alarm rate in the noisy case [30], [31].

Now, suppose that the measurements are polluted by Gaussian random noise with zero mean. When the sensor at v_2 is inactive, Fig. 4(a) shows the mean values of $c_2^{(1)}$ under different standard variances of noise and weights for 100 random realizations. A larger λ leads to more accurate recovery of $c_2^{(1)}$ in the mean square-error sense, but also shows weaker noise resilience in terms of the sparsity of solution, as shown in Fig. 4(b). In the noisy case, false alarms increase as λ increases; when $\lambda = 0$, which is smaller than $\lambda_{\min}$, the solution is sparsest but with worst accuracy, namely all missed detections.



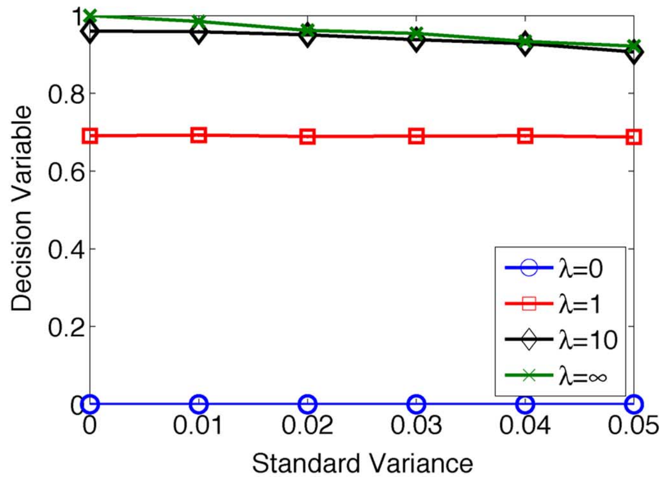
(a)

(b)

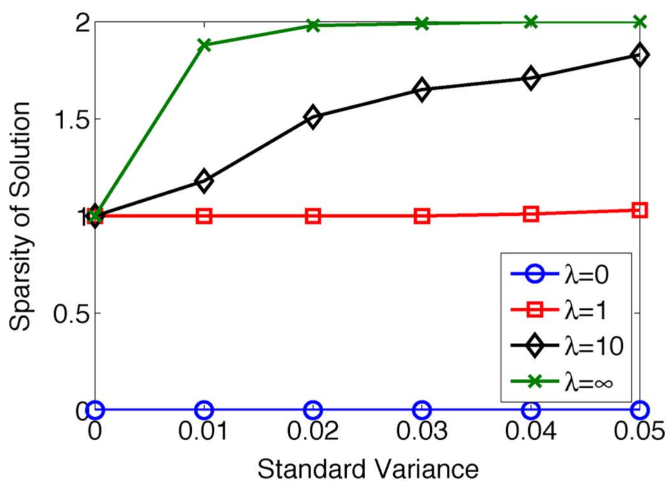
Fig. 3. Decision variables: (a) $c_2^{(2)}$ when all sensors are active; and (b) $c_2^{(1)}$ and $c_1^{(1)}$ when the sensor at v_2 is inactive.

We now discuss the model mismatch issue in the noise-free case. In practical applications, prior knowledge of the influence function can be biased. We use different values of $\sigma \in [35, 45]$ in the Gaussian shape to generate the actual influence function of the phenomenon, but set $\sigma = 40$ in the signal recovery algorithm for all cases. Again, the sensor at v_2 is inactive, and the phenomenon occurs at v_2 with $c_2 = 1$. As shown in Fig. 5, estimation of $c_2^{(1)}$ is resilient to model mismatch.

The resolution for detection in the sparse signal recovery algorithm depends on the sensor density. To demonstrate this dependence, we deploy only two active sensors at $(0,0)$ and $(40,0)$, and then evenly deploy inactive sensors along the line between $(0,0)$ and $(40,0)$. The phenomenon occurs at $(16,0)$, with $\sigma = 40$ and noise-free. We set $\lambda = 10$, collect nonzero elements in the estimates for different node density, and mark them in Fig. 6. Because we have assumed in **(A2)** that the phenomenon occurs at a sensor point, two inactive sensors beside the phenomenon will have nonzero decision variables if **(A2)** does not hold. As the node density increases, we are able to reach accurate position estimation. That is, the proposed



(a)



(b)

Fig. 4. (a) Mean values of $c_2^{(1)}$; and (b) average sparsity of solutions vs. the standard variance of the noise, for various values of the weight λ.

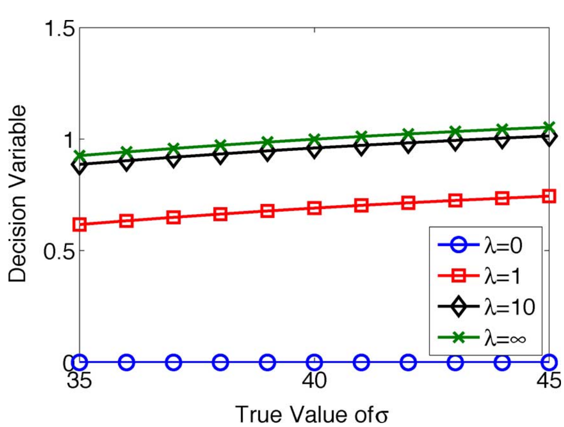


Fig. 5. Decision variables $c_2^{(1)}$ for different true values of σ and weights.

algorithm is able to improve the resolution, even when the node density is increased by adding inactive sensors only.

B. Large Networks

Now we consider large wireless sensor networks, in which sensors may not directly communicate with each other. In this case, if the communication range r_C is smaller than the radius of influence function, then the assumption **(A3)** will result in truncation errors. To address this issue, we demonstrate the relationship between the communication range and the radius of influence function via an illustrative simulation. Suppose that 11 active sensors are evenly deployed on the line from $(0,0)$ to $(200,0)$, with the sensor spacing being 20. There is one phenomenon occurring at sensor position $(100,0)$ with amplitude 1. The influence function has a Gaussian shape $\exp(-d_{ij}^2/\sigma^2)$ and the measurements are noise-free. Fig. 7 shows the impact of the communication range r_C on the sparsity of solution. The communication range r_C is chosen from 10 to 110, such that one sensor can communicate with 0 to 5 neighboring sensors, respectively. Smaller λ values result in sparser solutions when the communication range limits the global information exchange. From this point of view, truncation errors induced by the limited communication range can be treated as a kind of systematic measurement errors. Fig. 7 also suggests that when the influence function shows a long-tailed property, r_C can be properly chosen such that the assumption **(A3)** approximately holds. For example, if $f_{ij} = \exp(-d_{ij}^2/\sigma^2)$, then $r_C > 2\sigma$ is a proper value since $f_{ij} = \exp(-r_C^2/\sigma^2) < 0.0183 \simeq 0$ such that the truncation error is small enough.

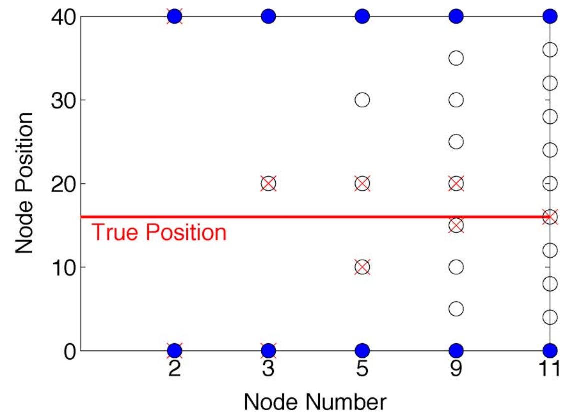


Fig. 6. Positions of sensors which are determined to have nonzero decision variables versus different node densities. Active sensors are marked as solid circles and inactive sensors are marked as hollow circles. The sensor points which have nonzero estimates are marked as cross symbols.

Knowing how to handle the truncation error, we consider a large random network to check the validity of the assumption **(A1)** and to illustrate the algorithm performance when the assumptions **(A2)** and **(A3)** are slightly violated. In a 200×200 area, 100 sensors are uniformly randomly deployed, among which 50% sensors are randomly set to be inactive. There are four phenomena occurring in the sensing field, denoted as $P1, P2, P3$, and $P4$, respectively. All phenomena are of unit amplitude, and the influence functions are Gaussian-shaped with $\sigma = 40$. Phenomena $P1$ and $P2$ occur at position (50,50) and (150,150), which do not coincide with any sensor points. Phenomena $P3$ and $P4$ occur at an inactive sensor point and an active sensor point, respectively. The measurements are supposed to be noise-free.

To simulate a practical network, we need to decide the communication range r_C according to the phenomena and the network connectivity. According to the discussion above,

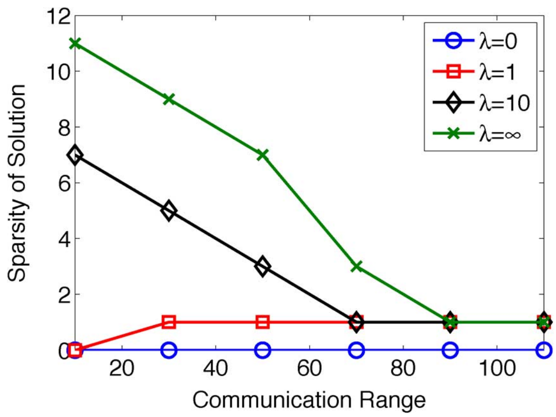


Fig. 7. Impact of the communication range r_C on the sparsity of solution.

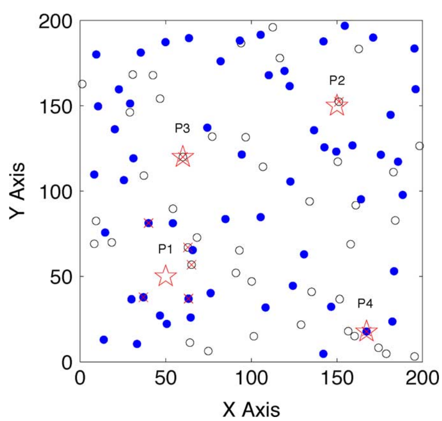


Fig. 8. Simulation results for a random network, which contains 100 sensors with 50% sensors being inactive, in a 200×200 area. Four phenomena, denoted as pentagrams, occur in the sensing area. Active sensors are marked as solid circles and inactive sensors are marked as hollow circles. The sensor points which have nonzero estimates are marked as cross symbols.

we choose $r_C = 80$, such that the influence of a unit phenomenon on the edge of the communication range is $\exp(-80^2/40^2) = 0.0183$. The weight is set as $\lambda = 10$. Then we check the network connectivity based on [15]. It is proven that the assumption **(A1)** (i.e., connectivity of the network and subnetworks) satisfies, with high probability, when $r_C = 80$.

The simulation results are depicted in Fig. 8. The active sensors are marked as solid circles; the inactive sensors are marked as hollow circles; the phenomena are denoted as pentagrams; and the sensor points which have nonzero estimates are marked as cross symbols. Both $P3$ and $P4$ are correctly estimated, no matter whether they occur at an active or inactive sensor point. The phenomenon $P2$ is also identified as occurring in a nearby inactive sensor. The phenomenon $P1$ can not be identified since it has no nearby sensors. Instead, multiple neighboring sensors are identified to be with nonzero amplitudes. It is hence possible to infer the location of $P1$ by averaging the locations of these neighboring sensors with nonzero estimates. This result is encouraging, because even when the assumption **(A2)** (i.e., the phenomena being occurring at the sensor points) is invalid, we are still able to improve the detection resolution via some clustering technique.

Fig. 8 demonstrates the effectiveness of the proposed sparse signal recovery algorithm and the random sleeping strategy. By exploiting sparsity of the signal, we are able to successfully recover it with compressed sensing data without loss of resolution. In this example, the resolution provided by 50 active sensors is equal to that of 100 active sensors; thus nontrivial energy consumption is saved for the whole network. On the other hand, the decentralized in-network processing scheme improves the robustness and scalability of the network, comparing with its centralized counterpart.

Finally we further address the issue of compression ratio, namely the impact of the percentage of inactive sensors and the number of phenomena on the probabilities of missed detection and false alarm. There are 100 sensors deployed uniformly randomly in a 200×200 area, and multiple unit-amplitude phenomena occur at sensor points. The measurements are supposed to be noise-free; the influence function, communication range, weight λ are all set as in the previous simulation.

Fig. 9 depicts the mean values and error bars for the numbers of missed detection and false alarms versus the percentage of inactive sensors, when the number of phenomena varies from 1 to 4, and a total of 100 trials is simulated for each setting. We declare a missed detection when a phenomenon occurs at a sensor point but the corresponding recovered signal value is smaller than 0.01. Conversely, we declare a false alarm when no phenomenon occurs at a sensor point but the corresponding recovered signal value is larger than 0.01. It is shown in Fig. 9 that the number of missed detection remains near to 0 when the compression ratio is smaller than 60%, but increases quickly when the compression ratio increases beyond 60%. Meanwhile, the number of false alarms varies slightly as the compression ratio varies, which suggests that the algorithm tends to provide sparse solutions. This simulation confirms the advantage of the ℓ_1 regularized least squares formulation in (3) over a traditional least squares formulation, because the latter may yield non-sparse solutions and result in a large probability of false alarms in the presence of measurement noise, model mismatch and truncation errors.

VI. Conclusion

This paper investigates the problem of monitoring sparse phenomena using a large-scale and distributed sleeping wireless sensor network. Random node sleeping strategies are adopted for energy conservation, which effect compression during the measurement collection process. A decentralized sparse signal recovery algorithm is developed based on ℓ_1 regularized least squares and consensus optimization. Each active sensor not only optimizes for itself, but also optimizes for its inactive neighbors. Through iterative one-hop information exchange, active sensors are able to reach consensus for inactive sensors. As a result, a phenomenon, no matter whether it occurs at an active sensor point or inactive sensor point, can be detected and quantified. It is theoretically proved that sensors eventually reach globally optimal decisions for their local regions, at scalable computation and communication costs with respect to the network size. Benefiting from the decentralized optimization scheme, the sleeping strategy, and most of all, the recognition of signal sparsity, the proposed decentralized sparse signal recovery algorithm improves the scalability and robustness of large networks,

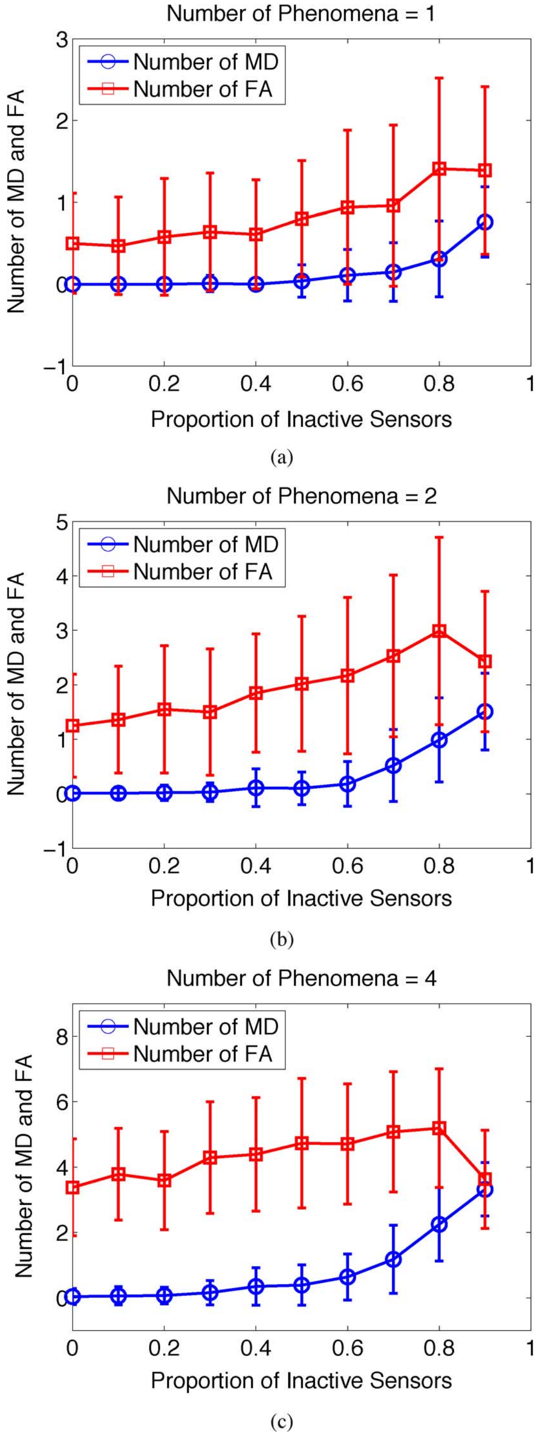


Fig. 9. Mean values and error bars of the numbers of missed detection (MD) and false alarms (FA) versus the percentage of inactive sensors when the number of phenomena is (a) 1; (b) 2; and, (c) 4.

preserves energy of wireless sensors, and at the same time guarantees high spatial resolution for monitoring.

Appendix

This appendix derives a decentralized algorithm for implementing the optimal solution of the consensus optimization problem in (5).

Let m_i denote the number of neighbors of an active sensor i plus 1. The following auxiliary variables are introduced:

$$\begin{aligned} z_{ki} &= f_{ki}c_k^{(k)} - \frac{1}{m_i}s_i, \quad \forall i \in \mathcal{W}, k \in (\mathcal{N}_i \cap \mathcal{W}) \cup i, \\ z_{ki} &= f_{ki}c_k^{(i)} - \frac{1}{m_i}s_i, \quad \forall i \in \mathcal{W}, k \in \mathcal{N}_i \cap \mathcal{S}. \end{aligned} \tag{12}$$

Substituting (12) into (5b), (5) can be rewritten as

$$\begin{aligned} \min \quad & \sum_{i\in\mathcal{W}} c_i^{(i)} + \sum_{k\in\mathcal{S}} \sum_{i\in\mathcal{N}_k\cap\mathcal{W}} w_k c_k^{(i)} + \frac{\lambda}{2}\sum_{i\in\mathcal{W}} s_i^2, \\ \text{s.t.} \quad & z_{ki} = f_{ki}c_k^{(k)} - \frac{1}{m_i}s_i \\ & \qquad \forall i \in \mathcal{W}, k \in (\mathcal{N}_i \cap \mathcal{W}) \cup i \\ & z_{ki} = f_{ki}c_k^{(i)} - \frac{1}{m_i}s_i \\ & \qquad \forall i \in \mathcal{W}, k \in \mathcal{N}_i \cap \mathcal{S}, \\ & \sum_{k\in\mathcal{N}_i\cup i} z_{ki} = b_i \\ & \qquad \forall i \in \mathcal{W} \\ & c_k^{(i)} = c_k^{(j)} \\ & \qquad \forall i \in \mathcal{W}, j \in \mathcal{W} \cap \mathcal{N}_i, k \in \mathcal{S} \cap \mathcal{N}_i \cap \mathcal{N}_j \\ & c_i^{(i)} \geq 0, \quad \forall i \in \mathcal{W} \\ & c_k^{(i)} \geq 0, \quad \forall i \in \mathcal{W}, k \in (\mathcal{N}_i \cap \mathcal{S}) \cup i. \end{aligned} \tag{13}$$

The augmented Lagrangian function of (13) is given by

$$\begin{aligned} \min \quad & \sum_{i\in\mathcal{W}} c_i^{(i)} + \sum_{k\in\mathcal{S}} \sum_{i\in\mathcal{N}_k\cap\mathcal{W}} w_k c_k^{(i)} + \frac{\lambda}{2}\sum_{i\in\mathcal{W}} s_i^2 \\ & + \sum_{i\in\mathcal{W}} \sum_{k\in(\mathcal{N}_i\cap\mathcal{W})\cup i} p_{ki}\left(f_{ki}c_k^{(k)} - \frac{1}{m_i}s_i - z_{ki}\right) \\ & + \sum_{i\in\mathcal{W}} \sum_{k\in(\mathcal{N}_i\cap\mathcal{W})\cup i} \frac{\beta}{2}\left(f_{ki}c_k^{(k)} - \frac{1}{m_i}s_i - z_{ki}\right)^2 \\ & + \sum_{i\in\mathcal{W}} \sum_{k\in\mathcal{N}_i\cap\mathcal{S}} p_{ki}\left(f_{ki}c_k^{(i)} - \frac{1}{m_i}s_i - z_{ki}\right) \\ & + \sum_{i\in\mathcal{W}} \sum_{k\in\mathcal{N}_i\cap\mathcal{S}} \frac{\beta}{2}\left(f_{ki}c_k^{(i)} - \frac{1}{m_i}s_i - z_{ki}\right)^2 \\ & + \sum_{i\in\mathcal{W}\cap\mathcal{N}_j\cap\mathcal{N}_k} \sum_{j\in\mathcal{W}\cap\mathcal{N}_k} \sum_{k\in\mathcal{S}} \lambda_{ijk}\left(c_k^{(i)} - c_k^{(j)}\right) \\ & + \sum_{i\in\mathcal{W}\cap\mathcal{N}_j\cap\mathcal{N}_k} \sum_{j\in\mathcal{W}\cap\mathcal{N}_k} \sum_{k\in\mathcal{S}} \frac{\beta}{2}\left(c_k^{(i)} - c_k^{(j)}\right)^2 \\ \text{s.t.} \quad & \sum_{v_k\in\mathcal{N}_i\cup v_i} z_{ki} = b_i, \quad \forall v_i \in \mathcal{W}, \\ & c_i^{(i)} \geq 0, \quad \forall i \in \mathcal{W} \\ & c_k^{(i)} \geq 0, \quad \forall i \in \mathcal{W}, k \in (\mathcal{N}_i \cap \mathcal{S}) \cup i. \end{aligned} \tag{14}$$

Here p_{ki} and λ_{ijk} are Lagrange multipliers; β is a weighting factor for the augmented quadratic terms.

The basic idea of the alternating direction method of multipliers is to iteratively optimize the constrained augmented Lagrangian function (14) based on the block coordinate descent algorithm. First, the slack variables and auxiliary variables are optimized under corresponding constraints, which result in the updated $s_i(t+1), \forall i \in \mathcal{W}$, expressed in (6), and $z_{ki}(t+1), \forall i \in \mathcal{W}, k \in (\mathcal{N}_i \cap \mathcal{W}) \cup i$, and $z_{ki}(t+1), \forall i \in \mathcal{W}, k \in \mathcal{N}_i \cap \mathcal{S}$, respectively, as shown in (7).

Second, the multipliers are updated based on subgradient descent, as in (8) for updating $p_{ki}(t+1)$ and (9) for updating $\lambda_{ijk}(t+1)$.

Finally, the decision $c_i^{(i)}(t+1)$ that an active sensor i makes for itself, $\forall i \in \mathcal{W}$, and the decisions $c_k^{(i)}(t+1)$ that it makes for its neighboring sleeping sensors $k, \forall i \in \mathcal{W}, k \in \mathcal{S} \cap \mathcal{N}_i$, are optimized from (14) and described in (10) and (11).

It can be shown straightforwardly from [10, pp. 254–261] that the above iterative steps in (6)–(11) derived from the alternating direction method of multipliers converge to the optimal solution to the original problem formulated in (5) for any positive constant β.

References

[1] D. Donoho, M. Elad, and V. Temlyakov, "Stable recovery of sparse overcomplete representations in the presense of noise," *IEEE Trans. Inf. Theory*, vol. 52, pp. 6–18, 2006.

[2] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, pp. 1289–1306, 2006.

[3] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, pp. 489–509, 2006.

[4] V. Raghunathan, C. Schurgers, S. Park, and M. Srivastava, "Energy-aware wireless microsensor networks," *IEEE Signal Process. Mag.*, vol. 19, pp. 40–50, 2002.

[5] I. Akyildiz and I. Kasimoglu, "Wireless sensor and actuator networks: Research challenges," *Ad Hoc Netw.*, vol. 2, pp. 351–367, 2004.

[6] J. Kotecha, V. Ramachandran, and A. Sayeed, "Distributed multi-target classification in wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, pp. 703–713, 2005.

[7] J. Barzerque and G. Giannakis, "Distributed spectrum sensing for cognitive radios by exploiting sparsity," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1847–1862, Mar. 2008.

[8] Q. Ling, Z. Tian, Y. Yin, and Y. Li, "Localized structural health monitoring in wireless sensor networks," *IEEE Sens. J.*, vol. 9, pp. 1596–1604, 2009.

[9] M. Perillo, Z. Ignjatovic, and W. Heinzelman, "An energy conservation method for wireless sensor networks employing a blue noise spatial sampling technique," in *Proc. IEEE IPSN Conf.*, 2004.

[10] D. Bertsekas and J. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*, 2nd ed. New York: Athena Scientific, 1997.

[11] W. Bajwa, J. Haupt, A. Sayeed, and R. Nowak, "Compressive wireless sensing," in *Proc. IEEE IPSN Conf.*, 2006.

[12] W. Wang, M. Garofalakis, and K. Ramchandran, "Distributed sparse random probjections for refinable approximation," in *Proc. IEEE IPSN Conf.*, 2007.

[13] W. Ye, J. Heidemann, and D. Estrin, "Medium access control with coordinated adaptive sleeping for wireless sensor networks," *IEEE Trans. Netw.*, vol. 12, pp. 493–506, 2004.

[14] S. Kumar, T. Lai, and J. Balogh, "On k-coverage in a mostly sleeping sensor network," in *Proc. ACM MOBICOM Conf.*, 2004.

[15] Q. Ling and Z. Tian, "Minimum node degree and k-connectivity of a wireless multihop network in a bounded area," in *Proc. IEEE GLOBECOM Conf.*, 2007.

[16] J. Xiao, A. Ribeiro, Z. Luo, and G. Giannakis, "Distributed compression-estimation using wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 23, pp. 27–41, 2006.

[17] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Trans. Inf. Theory*, vol. 52, pp. 2508–2530, 2006.

[18] F. Fagnani and S. Zampieri, "Randomized consensus algorithms over large scale networks," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 634–649, 2008.

[19] J. Predd, S. Kulkarni, and V. Poor, "A collaborative training algorithm for distributed learning," *IEEE Trans. Inf. Theory*, vol. 55, pp. 1856–1871, 2009.

[20] M. Rabbat and R. Nowak, "Quantized incremental algorithms for distributed optimization," *IEEE J. Sel. Areas Commun.*, vol. 23, pp. 798–806, 2005.

[21] A. Dogandzic and B. Zhang, "Distributed estimation and detection for sensor networks using hidden Markov random field models," *IEEE Trans. Signal Process.*, vol. 54, pp. 3200–3215, 2006.

[22] J. Fang and H. Li, "Distributed event-region detection in wireless sensor networks," *EURASIP J. Adv. Signal Process., Article ID 287870*, 2008.

[23] I. Schizas, A. Ribeiro, and G. Giannakis, "Consensus in ad hoc WSNs with noisy links. Part I: Distributed estimation of deterministic signals," *IEEE Trans. Signal Process.*, vol. 56, pp. 350–364, 2008.

[24] I. Schizas, G. Mateos, and G. Giannakis, "Distributed LMS for consensus-based in-network adaptive processing," *IEEE Trans. Signal Process.*, vol. 57, pp. 2365–2382, 2009.

[25] M. Duarte, S. Sarvotham, D. Baron, M. Wakin, and R. Baraniuk, "Distributed compressed sensing," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, 2005.

[26] A. Hormati and M. Vetterli, "Distributed compressed sensing: Sparsity models and reconstruction algorithms using annihilating filter," in *Proc. IEEE ICASSP Conf.*, 2008.

[27] X. Sheng and Y. Hu, "Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 53, pp. 44–53, 2005.

[28] M. Figueiredo, R. Nowak, and S. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Sel. Top. Signal Process.*, vol. 1, pp. 586–597, 2007.

[29] S. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An interior-point method for large-scale ℓ_1 regularized least squares," *IEEE J. Sel. Top. Signal Process.*, vol. 1, pp. 606–617, 2007.

[30] S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Scientif. Comput.*, vol. 20, pp. 33–61, 1998.

[31] R. Tibshirani, "Regression shrinkage and selection via the Lasso," *J. Royal Statist. Soc. B*, vol. 58, pp. 267–288, 1996.

Qing Ling received the B.S. degree in automation and the Ph.D. degree in control theory and control engineering from University of Science and Technology of China, Hefei, Anhui, in 2001 and 2006, respectively.

From 2006 to 2009, he was a Postdoctoral Research Fellow with the Department of Electrical and Computer Engineering, Michigan Technological University, Houghton. Since 2009, he has been an Assistant Professor with the Department of Automation, University of Science and Technology of China. His current research focuses on decentralized optimization of wireless sensor networks.

Zhi Tian (M'98–SM'06) received the B.E. degree in electrical engineering from the University of Science and Technology of China, Hefei, in 1994, the M.S. and Ph.D. degrees from George Mason University, Fairfax, VA, in 1998 and 2000, respectively.

Since August 2000, she has been on the faculty of Michigan Technological University, Houghton, where she is currently an Associate Professor. Her general interests are in the areas of signal processing for wireless communications, estimation, and detection theory. Current research focuses on cognitive radio networks and distributed wireless sensor networks.

Dr. Tian served as an Associate Editor for the IEEE Transactions on Wireless Communications and the IEEE Transactions on Signal Processing. She received a CAREER award in 2003 from the U.S. National Science Foundation.