Distributed Projection-Based Algorithms for Source Localization in Wireless Sensor Networks

Yanqiong Zhang, Youcheng Lou, Yiguang Hong, *Senior Member, IEEE*, and Lihua Xie, *Fellow, IEEE*

***Abstract*—In this paper, we investigate source localization for wireless sensor networks based on received signal strength. We first formulate the localization problem as the intersection computation of a group of sensing rings, and then convert this non-convex problem into two weighted convex optimization problems. We next propose a unified distributed alternating projection algorithm to solve the resulting weighted optimization problems, where sensor nodes can communicate only locally with their neighbors over a time-varying jointly-connected topology. We also show that sensor nodes' estimates can achieve consensus on a possible minimizer. Both theoretical analysis and some comparative simulations reveal that the proposed approach has good estimation performance in both the consistent and inconsistent cases.**

***Index Terms*—Source localization, wireless sensor networks, distributed algorithms, intersection computation, constrained convex optimization.**

I. Introduction

In wireless sensor networks, how to locate source nodes is one of the fundamental tasks due to the importance of position information for the network. In recent years, some types of measurement approaches to source localization, including time difference of arrival, angle of arrival and received signal strength, and energy measurement-based approaches have been proposed for wireless sensor networks (WSNs) [1]–[8]. Some classical estimation approaches such as maximum likelihood estimation (MLE) [3] and nonlinear least squares (NLS) [9] estimation derived for the source localization problem are often very complex and maybe trapped by a local optimum, and therefore, numerous suboptimal techniques have been proposed to reduce the complexity in solving the MLE or NLS problems, such as convex relaxation methods [6], [27]–[32].

Source localization algorithms can be categorized into two classes from an implementation point of view: centralized and distributed algorithms. In [10], [11], projection-based methods were studied for source localization, which are usually centralized and implemented in a sequential or parallel manner without rigorous analysis. The main disadvantage of centralized methods is that they require a central unit to gather and process the measurements from the sensor nodes. To overcome this disadvantage, some distributed algorithms were developed, where local communications between neighboring nodes are realized by wireless transmissions. For example, a distributed implementation of the incremental gradient algorithm was proposed to solve the nonlinear least-squares problem in [5], and a weighted direct least-squares formulation was reported in [9] for a tradeoff between performance and computational complexity, while a distributed asynchronous algorithm was proposed to deal with the maximum likelihood convex relaxation in [32].

A fully distributed projection-based approach was also developed for the source localization in [12], where the problem was formulated as a convex feasibility problem or convex intersection problem (CIP), and both the consistent and inconsistent cases were discussed. In fact, distributed optimization to achieve a global optimal objective by local optimization and local information exchange mechanism has received considerable attention [15]–[18], and CIP has also been widely studied in distributed optimization literature. Various distributed projection-based algorithms for CIP were proposed over the past decade [13]–[15], [19], [20].

Clearly, communications among sensor nodes play an important role in the implementation of distributed algorithms. The communication topology of a sensor network may be time-varying because of environmental influence, or link failure, or energy saving. A well-known concept to describe the time-varying connectivity is called uniform joint-strong-connected (UJSC), which has been widely used in the multi-agent literature [16], [17], [20] to demonstrate that their proposed algorithms are robust to the variable communication topology.

Here we give a formulation by treating the source localization problem as an intersection computation problem of a group of sensing rings in two different localization cases: the consistent case (when the intersection of the rings is nonempty) and the inconsistent case (when the intersection is empty). Note that the ring intersection computation is a non-convex constrained optimization problem, whose exact optimal solution is very hard to find. Obviously, the convex problem under fixed topologies discussed in [12] is a very special case of our non-convex problem under switching communication topologies.

Manuscript received March 14, 2014; revised September 2, 2014, January 14, 2015, and February 5, 2015; accepted February 6, 2015. Date of publication February 11, 2015; date of current version June 6, 2015. This work was supported partially by the NNSF of China under Grant 61333001, 973 Program under Grant 2014CB845301/2/3, and National Research Foundation of Singapore under Grant NRF-CRP8-2011-03, NRF2011NRF-CRP001-090, NRF2013EWT-EIRP004-012, NSFC 61120106011. The associate editor coordinating the review of this paper and approving it for publication was A. Giorgetti.

Y. Zhang and Y. Hong are with the Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China (e-mail: yanqiongzh@amss.ac.cn; yghong@iss.ac.cn).

Y. Lou is with the National Center for Mathematics and Interdisciplinary Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China (e-mail: louyoucheng@amss.ac.cn).

L. Xie is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mail: elhxie@ntu.edu.sg).

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/TWC.2015.2402672

In the literature, some convex relaxation methods, based on semidefinite programming [27], [28], second order cone programming [6], [29], sum of squares [30] and Lagrangian dual [22], [23], have been proposed to overcome the difficulty caused by the non-convexity. However, their results can not be applied directly to our problem because, in our problem setup: (i) each sensor can only observe its own sensing ring and share its source localization estimate with its neighbors, without knowing the information of any other sensing rings; (ii) the communication topology defining the neighbor relationship is time-varying and may not be connected at each moment. To solve the problem, we take a convex approximation technique to get a suboptimal estimation. By converting the non-convex problem into a weighted unconstrained convex problem in the inconsistent case and a weighted constrained convex problem in the consistent case, we further propose a unified distributed projection-based algorithm and prove its optimal convergence in these two cases. Notice that the convergence analysis for our problem is much more difficult than that in [12], because we have to handle a constrained convex optimization in a distributed way, rather than the unconstrained optimization, especially when each agent only knows its own constraint set (sensing ring). Moreover, we have to guarantee the convergence under some weaker connectivity conditions, compared with either the same constraint set case or uniform weight assumptions in [16].

This paper is organized as follows. In Section II, we provide a new formulation for the source localization problem along with basic concepts. In Section III, we present an alternating projection algorithm with some preliminary results. Then in Sections IV we give the main results and convergence analysis for both the consistent and inconsistent cases under switching interaction topologies. Following that, we carry out simulation studies to verify our theoretical results in Section V. Finally, we give concluding remarks in Section VI.

II. Problem Formulation

In this section, we give a formulation for source localization based on intersection computation.

Consider a network composed of n (wireless) sensors that can only interact with each other through local time-varying communications in the sensor field denoted by $S \subseteq \mathbb{R}^2$. The problem of interest is to determine the location of an active acoustic source in this sensor network.

A. Topology of Sensor Network

Here we consider the time-varying communication topology for the investigated sensor network.

At first, we review some useful concepts related to graph theory [24]. A directed graph is a pair $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, 2, \cdots, n\}$ is the node set, and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ the edge set. In the paper, we regard sensor i as node i. Node i is a neighbor of j if $(i, j) \in \mathcal{E}$. Let N_i be the set of neighbors of node i. A path of length r from a node i_1 to a node i_{r+1} is a sequence of $r+1$ distinct nodes $i_1, \cdots, i_{r+1}$ such that $(i_q, i_{q+1}) \in \mathcal{E}$ for $q = 1, \cdots, r$. If there is a path between any two nodes $i, j \in \mathcal{V}$, then $\mathcal{G}$ is said to be strongly connected.

However, the communication topologies over sensor networks may be time-varying in practice due to link failure or energy saving. As usual, the topology can be described by a time-varying directed graph $\mathcal{G}_k = (\mathcal{V}, \mathcal{E}(k))$ with $k \geq 0$, where $\mathcal{E}(k)$ represents the set of communication links at time k. For any two sensors $i, j \in \mathcal{V}$, j can get the information from i at time k if and only if $(i, j) \in \mathcal{E}(k)$. Denote $N_i(k) = \{j \in \mathcal{V} | (j, i) \in \mathcal{E}(k)\}$, and we assume $i \in N_i$ in this paper. Let $a_{ij}(k)$ represent the weight of arc (j, i) at time k, $a_{ij}(k) > 0$ if $(j, i) \in \mathcal{E}(k)$; $a_{ij}(k) = 0$ otherwise. $A(k) = [a_{ij}(k)]$ is the corresponding weighted adjacency matrix.

The following two assumptions on the graphs for sensor networks are widely used in the distributed optimization literature [15], [16], [20].

Assumption 1: (Weight Rule) There exists a scalar η with $0 < \eta < 1$ such that $a_{ij}(k) \geq \eta$ if $a_{ij}(k) > 0$ for $i, j \in \mathcal{V}$ and $k \geq 0$. Moreover, $\sum_{i=1}^{n} a_{is}(k) = \sum_{j=1}^{n} a_{sj}(k) = 1$ for $s \in \mathcal{V}$.

Assumption 2: (Connectivity) The communication graph is uniformly jointly strongly connected (UJSC), i.e., there is an integer $T > 0$ such that the directed graph $(\mathcal{V}, \cup_{t=1}^{T} \mathcal{E}(k+t))$ is strongly connected for any $k \geq 0$.

Remark 2.1: Assumption 1 has been widely used in the distributed optimization literature, and in the bidirectional graph case (that is, $(i, j) \in \mathcal{E}(k)$ if and only if $(j, i) \in \mathcal{E}(k)$), it can be easily satisfied by adopting Metropolis weight rule [34]:

$$a_{ij}(k) = \begin{cases} 1/\left(\max\left\{|N_i(k)|, |N_j(k)|\right\} + 1\right), & (i, j) \in \mathcal{E}(k); \\ 1 - \sum_{p \neq i} a_{ip}(k), & j = i; \\ 0, & \text{otherwise.} \end{cases}$$

In the directed graph case, we can also obtain an adjacency matrix to satisfy Assumption 1 starting from an arbitrary adjacency matrix using some algorithms such as the distributed imbalance-correcting algorithm [35].

Denote $\Phi(k, s) = A(k)A(k-1) \cdots A(s)$ for $k \geq s \geq 0$. Then we review some convergence results of transition matrix $\Phi(k, s)$ in [15].

Lemma 2.2: Suppose Assumptions 1 and 2 hold. Then

$$\left| [\Phi(k, s)]_{ij} - \frac{1}{n} \right| \leq \lambda \gamma^{k-s}$$

with $[\Phi(k, s)]_{ij}$ the ij-th entry of transition matrix $\Phi(k, s)$ for $i, j \in \{1, \cdots, n\}$, and $k \geq s \geq 0$, $\lambda = 2(1 + \eta^{(1-n)T})/(1 - \eta^{(n-1)T})$ and $\gamma = (1 - \eta^{(n-1)T})^{1/((n-1)T)}$.

B. Sensing Area

We now discuss the sensing area of each sensor to locate a source. We start with some preliminaries about convex analysis [26]. A set $K \subseteq \mathbb{R}^m$ is convex if $ax + (1-a)y \in K$ for any $x, y \in K$ and $0 \leq a \leq 1$. For any $x \in \mathbb{R}^m$, there is a unique element $P_K(x) \in K$ satisfying $\|x - P_K(x)\| = \inf_{y \in K} \|x - y\|$, which is denoted by $|x|_K$, where $K \subseteq \mathbb{R}^m$ is a nonempty closed convex set and P_K denotes the projection operator onto K. Denote $|x|_K = 0$ for any x if $K = \emptyset$, for convenience. A function $f(\cdot) : \mathbb{R}^m \to \mathbb{R}$ is said to be convex if $f(ax + (1-a)y) \leq af(x) + (1-a)f(y)$ for any $x, y \in \mathbb{R}^m$ and $0 \leq a \leq 1$. Let

$f(\cdot) : \mathbb{R}^m \to \mathbb{R}$ be a convex function, vector $\zeta \in \mathbb{R}^m$ is a subgradient of f at $y \in \mathbb{R}^m$ if

$$f(x) - f(y) \geq \zeta^T(x - y), \qquad \forall\, x \in \mathbb{R}^m, \tag{1}$$

where ζ^T denotes the transpose of vector ζ. The following are some useful results [16], [21] on the properties of projection operator which are instrumental for the analysis later.

Lemma 2.3: Let K be a closed convex set in $\mathbb{R}^m$. Then

(i) $\|P_K(x) - y\|^2 \leq \|x - y\|^2 - \|P_K(x) - x\|^2$ for any $y \in K$ and x;
(ii) $||x|_K - |y|_K| \leq \|x - y\|$ for any x and y;
(iii) $|x|_K$ is a convex function and $|x|_K^2$ is continuously differentiable with $\nabla |x|_K^2 = 2(x - P_K(x))$, where ∇ stands for the gradient.

Let the source be located at an unknown coordinate pair $\rho = [\rho_1, \rho_2]^T$ in a sensor field $S \subseteq \mathbb{R}^2$ which emits a signal with power level P. The considered n sensors perform sensing based on energy detection. At sensor i with its known coordinate s_i, $i = 1, \cdots, n$, the received signal power can be written as follows [3], [4], [8]:

$$P_{s_i} = g_i \frac{P}{D_{is}^{\nu}} + \epsilon_i, \tag{2}$$

where g_i is the gain factor of sensor i, $D_{is} := \|\rho - s_i\|$, ν is the power-loss factor, which depends on the propagation environment and varies from 2 to 5, and ϵ_i is the receiver noise at the i-th sensor. For simplicity, we assume that $\epsilon_i (i = 1, \cdots, n)$ are zero-mean uncorrelated Gaussian processes with variances $\{\sigma_i^2\}$, i.e., $\epsilon_i \sim N(0, \sigma_i^2)$.

In this paper, we assume that the source power P is known, and the only parameter to be estimated is the location vector ρ of the source. The maximum likelihood estimator (MLE) is found by solving the nonlinear least-squares problem when the noise is Gaussian, that is,

$$\rho_{ML}^* = \arg\min_{\rho} \sum_{i=1}^{n} \frac{\left(P_{s_i} - \frac{g_i P}{D_{is}^{\nu}}\right)^2}{\sigma_i^2} = \arg\min_{\rho} \sum_{i=1}^{n} \phi_i(\rho),$$

where $\phi_i(\rho) := \frac{\left(P_{s_i} - \frac{g_i P}{D_{is}^{\nu}}\right)^2}{\sigma_i^2}$. Clearly, $\phi_i(\rho)$ attains its minimum 0 on the circle defined as follows:

$$\mathcal{C}_i = \left\{\rho \in \mathbb{R}^2 : \|\rho - s_i\| = \left(\frac{g_i P}{P_{s_i}}\right)^{1/\nu}\right\}.$$

However, due to the observation noise, the source may not appear exactly on the circles $\{\mathcal{C}_i\}_{i=1}^n$ but be contained in a group of sensing areas described by rings in the following forms:

$$R_i = \left\{\rho : \left(\frac{g_i P}{P_{s_i} + \xi_i \sigma_i}\right)^{\frac{1}{\nu}} < \|\rho - s_i\| < \left(\frac{g_i P}{P_{s_i} - \xi_i \sigma_i}\right)^{\frac{1}{\nu}}\right\}$$

where $\xi_i \sigma_i$ indicates the area of noise distribution for sensor i with a given trust parameter $0 < \xi_i < \frac{P_{s_i}}{\sigma_i}$. Obviously, as ξ_i increases, the possibility of the source located in the ring R_i increases, which will be 99.7% if $\xi_i = 3$. In this paper, we take $\xi_i = \min\left\{3, \frac{P_{s_i}}{\sigma_i}\right\}$.

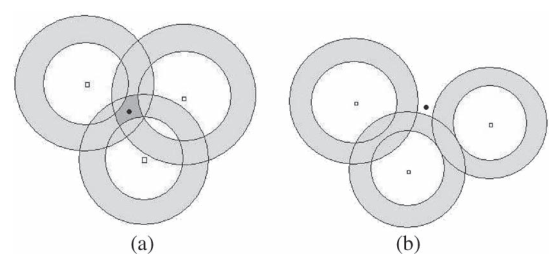


Fig. 1. (a) Consistent case; (b) inconsistent case.

Denote the outer open disk of ring R_i by

$$X_i^0 = \left\{\rho \in \mathbb{R}^2 : \|\rho - s_i\| < \left(\frac{g_i P}{P_{s_i} - \xi_i \sigma_i}\right)^{1/\nu}\right\},$$

and the inner open disk by

$$Y_i^0 = \left\{\rho \in \mathbb{R}^2 : \|\rho - s_i\| < \left(\frac{g_i P}{P_{s_i} + \xi_i \sigma_i}\right)^{1/\nu}\right\}.$$

Let X_i and Y_i be the closures of X_i^0 and Y_i^0, respectively. $R_i = X_i^0 \setminus Y_i = \{x : x \in X_i^0, x \notin Y_i\}$ is an open ring for $i = 1, \cdots, n$. Since the ring R_i is the sensing area of sensor i, it is natural to assume that sensor i can get the information about the sets X_i and Y_i, though it does not know anything about R_j when $j \neq i$.

C. Formulation

The source localization problem can be formulated as finding a point in the intersection set of the sensing rings as its location estimate. However, for any given ξ_i, there is still some possibility for the source to be located outside of the ring R_i. Hence, there are two cases for the source localization as a ring intersection computation problem: i) Consistent case: $R_0 := \bigcap_{i=1}^n R_i \neq \emptyset$; ii) Inconsistent case: $R_0 = \emptyset$ (as shown in Fig. 1, the black dot and small squares denote the source and the sensors, respectively).

Remark 2.4: In the consistent case, since $\bigcap_{i=1}^n X_i^0 \subseteq X_0 := \bigcap_{i=1}^n X_i$ is an open and nonempty set, there exists at least one interior point in the intersection set X_0.

Generally speaking, the intersection computation of rings, which are not convex, is still very complicated. One of the reasons is that it is difficult to ensure that the optimal point is located outside of all the inner disks $\{Y_i\}_{i=1}^n$ in the non-convex problem. Therefore, we formulate the source localization problem as the following two convex problems:

i) Inconsistent case: the source can not be in R_i for all $i = 1, \cdots, n$, and we give a trade-off solution for the localization problem by finding a point $x^* \in \mathbb{R}^2$ to minimize the weighted sum of the squares of the distances to the outer disks $\{X_i\}_{i=1}^n$ and inner disks $\{Y_i\}_{i=1}^n$, i.e.,

$$\min \sum_{i=1}^{n} \left(|x|_{X_i}^2 + b_i |x|_{Y_i}^2\right); \tag{3}$$

ii) Consistent case: the source is inside the intersection set X_0, and then the source localization problem can be solved by finding a point $x^* \in X_0$ to minimize the weighted sum function (3), that is,

$$\min \sum_{i=1}^{n} \left(|x|_{X_i}^2 + b_i |x|_{Y_i}^2 \right), \quad \text{subject to } x \in X_0, \tag{4}$$

or equivalently,

$$\min \sum_{i=1}^{n} b_i |x|_{Y_i}^2, \quad \text{subject to } x \in X_0, \tag{5}$$

where $b_1, \cdots, b_n$ are the positive scalar parameters.

Roughly speaking, $\{b_i\}_{i=1}^n$ can be viewed as the weighted parameters to make our considered problem become a "virtual" convex intersection problem for a group of certain convex sets $\{Q_i\}_{i=1}^n$ with $Y_i \subseteq Q_i \subseteq X_i$.

Denote by Z^* and X^* the optimal solution sets of the unconstrained problem (3) and constrained problem (4), respectively. Then we give our localization formulation as follows:

Definition 2.5: The distributed source localization problem is solved if we can find a distributed algorithm such that, for any initial condition $x_i(0) \in \mathbb{R}^2$, $i \in \mathcal{V}$, there exists an optimal solution $x^* \in Z^*$ with $\lim_{k\to\infty} x_i(k) = x^*$, $i \in \mathcal{V}$ if problem is inconsistent; there exists an optimal solution $x^* \in X^*$ with $\lim_{k\to\infty} x_i(k) = x^*$ for $i \in \mathcal{V}$ otherwise. In both cases, x^* is termed as the optimal estimate of the source location.

Remark 2.6: Our formulation of the source localization is more general than or different from those given in the literature such as [12], [27]–[29]. First of all, we consider jointly-connected switching communication topologies, while fixed topologies were basically discussed in [12], [27]–[29]. Next, some existing results formulate the localization problem as an easily-solved convex optimization problem in [12], but ours is a constrained problem. Moreover, different from the problem setups of the rings' intersection investigated in some literature [27]–[29], each sensor in our problem can only access the information about its own sensing ring without knowing any other rings, and the only available information from its time-varying neighbor sensors is their estimates about the source. Certainly, if the information about the sensing rings can be shared among all the sensors, the semidefinite programming (SDP) and linear cone programming (LCP) methods can be easily applied.

In what follows, we will construct a unified algorithm for both cases and then provide its convergence proofs in the two cases, respectively.

III. Distributed Algorithm and Preliminaries

In this section, we propose a distributed alternating projection algorithm (DAPA) for the source localization problem with jointly-connected switching topologies and then give some preliminary results.

A. Algorithm

The main idea of DAPA is to allow sensor $i (i \in \mathcal{V})$ with an initial estimate $x_i(0) \in \mathbb{R}^2$ to update its estimate by combining the estimates received from its neighbors, and meanwhile, to employ the projections on its inner disk Y_i and its outer closed disk X_i, respectively. To be precise, sensor i is allowed to update its estimate according to the following rule:

$$\begin{cases} v_i(k) = \sum_{j=1}^n a_{ij}(k) x_j(k), \\ w_i(k) = v_i(k) - \alpha_k \nabla f_i(v_i(k)), \\ x_i(k+1) = w_i(k) - \beta_k \nabla g_i(w_i(k)), \end{cases} \tag{6}$$

where the weights $a_{ij}(k)$ are nonnegative, $f_i(x) = \frac{b_i}{2}|x|_{Y_i}^2$ and $g_i(x) = \frac{1}{2}|x|_{X_i}^2$ with respective gradients $\nabla f_i(v_i(k)) = b_i(v_i(k) - P_{Y_i}(v_i(k)))$ at $v_i(k)$ and $\nabla g_i(w_i(k)) = w_i(k) - P_{X_i}(w_i(k))$ at $w_i(k)$, and $\alpha_k \in [0, \min\{1, 1/b_0\}]$ with $b_0 = \max\{b_1, \cdots, b_n\}$ and $\beta_k \in [0,1]$ for $k = 0, 1, \cdots$ are the step-sizes. Clearly, our algorithm is a distributed alternating gradient-based algorithm. In fact, $w_i(k) = (1 - \alpha_k b_i) v_i(k) + \alpha_k b_i P_{Y_i}(v_i(k))$ is the approximate projection of point $v_i(k)$ onto Y_i, and $x_i(k+1) = (1-\beta_k) w_i(k) + \beta_k P_{X_i}(w_i(k))$ is the approximate projection of point $w_i(k)$ onto X_i. Here α_k and β_k can be viewed as projection accuracies.

Remark 3.1: Note that the ring intersection computation problem is very complicated. In the above, we made a first attempt to derive a more accurate estimator for source localization. In fact, its complexity can be significantly reduced if it becomes the widely-studied convex set intersection computation problem [12], [16], [19], [20] by taking $Y_i = \emptyset$ for $i = 1, \cdots, n$. In this way, the non-convex ring intersection problem can be reduced to the following unconstrained optimization problem:

$$\min_{x \in \mathbb{R}^2} \sum_{i=1}^{n} |x|_{X_i}^2 \tag{7}$$

In this case, $w_i(k) = v_i(k)$ in the algorithm (6). Then the corresponding distributed algorithm of problem (7) is obtained in the following form:

$$\begin{cases} v_i(k) = \sum_{i=1}^n a_{ij}(k) x_j(k), \\ x_i(k+1) = v_i(k) - \beta_k \left(v_i(k) - P_{X_i}(v_i(k)) \right). \end{cases} \tag{8}$$

Our results obtained in the next section are directly applicable to the convex optimization problem in the consistent $(\bigcap_{i=1}^n X_i \neq \emptyset)$ and inconsistent $(\bigcap_{i=1}^n X_i = \emptyset)$ cases discussed in [12].

B. Preliminary Results

To prove the convergence of DAPA, we introduce two useful lemmas obtained in [16] and [25], respectively.

Lemma 3.2: Let $0 < \mu < 1$ and $\{\phi_k\}_{k=0}^\infty$ be a positive scalar sequence. If $\lim_{k\to\infty} \phi_k = 0$, then $\lim_{k\to\infty} \sum_{s=0}^k \mu^{k-s} \phi_s = 0$. Moreover, if $\sum_{k=0}^\infty \phi_k < \infty$, then $\sum_{k=0}^\infty \sum_{s=0}^k \mu^{k-s} \phi_s < \infty$.

Lemma 3.3: Let $\{a_k\}_{k=0}^\infty$ and $\{\bar{a}_k\}_{k=0}^\infty$ be two non-negative sequences with $\sum_{k=0}^\infty \bar{a}_k < \infty$. If $a_{k+1} \leq a_k + \bar{a}_k, \ \forall k \geq 0$, then $\lim_{k\to\infty} a_k$ is a finite number.

Let $X_{co} = co\{X_1, \cdots, X_n\}$ be the convex hull consisting of all finite convex combinations of points in X_i, $i = 1, \cdots, n$. Due to the boundedness of X_i $(1 \leq i \leq n)$, X_{co} is bounded and $d_0 := \sup_{x,y\in X_{co}} \|x - y\| < \infty$. Denote $d_i(k) = b_i(v_i(k) - P_{Y_i}(v_i(k)))$ and $\delta_i(k) = w_i(k) - P_{X_i}(w_i(k))$. Then we establish a lemma for the boundedness of sequences $\{d_i(k)\}$ and $\{\delta_i(k)\}$ obtained by our algorithm (6).

Lemma 3.4: With Assumption 1, there exists a positive constant L such that

$$\|d_i(k)\| \leq L \quad \text{and} \quad \|\delta_i(k)\| \leq L \tag{9}$$

for $i = 1, \cdots, n$ and $k = 0, 1, \cdots$.

Denoting $e_i(k) = x_i(k+1) - w_i(k)$, it follows from (9) that

$$\|e_i(k)\| \leq L\beta_k. \tag{10}$$

With the transition matrix $\Phi(k, s)$, we have

$$\begin{aligned} x_i(k+1) = &\sum_{j=1}^{n} [\Phi(k,0)]_{ij}\, x_j(0) - \alpha_k d_i(k) + e_i(k) \\ &- \sum_{r=1}^{k} \left(\sum_{j=1}^{n} [\Phi(k,r)]_{ij}\, \alpha_{r-1} d_j(r-1) \right) \\ &+ \sum_{r=1}^{k} \left(\sum_{j=1}^{n} [\Phi(k,r)]_{ij}\, e_j(r-1) \right). \end{aligned}$$

Due to $\hat{x}(k) = \frac{1}{n}\sum_{j=1}^{n} x_j(k)$, (9), (10), and Lemma 2.2, we obtain

$$\begin{aligned} &\|x_i(k+1) - \hat{x}(k+1)\| \\ &\quad \leq \lambda\gamma^k \sum_{j=1}^{n} \|x_j(0)\| + nL\lambda \sum_{r=1}^{k} \gamma^{k-r}\alpha_{r-1} + 2L\alpha_k \\ &\qquad + 2L\beta_k + nL\lambda \sum_{r=1}^{k} \gamma^{k-r}\beta_{r-1}. \end{aligned} \tag{11}$$

From Lemma 3.2 and (11), we establish the following result.

Lemma 3.5: Under Assumptions 1 and 2, if the step-sizes $\{\alpha_k\}$ and $\{\beta_k\}$ satisfy $\sum_{k=0}^{\infty} \alpha_k^2 < \infty$ and $\sum_{k=0}^{\infty} \beta_k^2 < \infty$, then, for any $i \in \{1, \cdots, n\}$, we have $\lim_{k\to\infty} \|x_i(k) - \hat{x}(k)\| = 0$, $\sum_{k=0}^{\infty} \alpha_k \|x_i(k) - \hat{x}(k)\| < \infty$, and $\sum_{k=0}^{\infty} \beta_k \|x_i(k) - \hat{x}(k)\| < \infty$.

IV. Intersection Computation for Localization

In this section, we give the proofs for the unified DAPA in the inconsistent case and the consistent case with switching topologies, respectively.

In the inconsistent case, we can verify that $\sum_{i=1}^{n} b_i|x|_{Y_i}^2 + |x|_{X_i}^2$ is a convex function from Lemma 2.3 (iii). With the boundedness of sets $\{X_i, Y_i\}_{i=1}^{n}$, we get that $\sum_{i=1}^{n} b_i|x|_{Y_i}^2 + |x|_{X_i}^2$ is coercive (i.e., $\lim_{\|x\|\to\infty} \sum_{i=1}^{n} b_i|x|_{Y_i}^2 + |x|_{X_i}^2 = \infty$). Then by Proposition 2.1.1 in [26], we obtain the optimal solution set Z^* of problem (3) is nonempty. We have the following convergence theorem.

Theorem 4.1: Under Assumptions 1 and 2, the sequences $\{x_i(k)\}_{i=1}^{n}$ generated by algorithm (6) converge to the optimal solution of problem (3), i.e.,

$$\lim_{k\to\infty} x_i(k) = p^* \text{ for some } p^* \in Z^* \text{ and all } i,$$

if $\sum_{k=0}^{\infty} \alpha_k = \infty$, $\sum_{k=0}^{\infty} \beta_k = \infty$, $\sum_{k=0}^{\infty} \alpha_k^2 < \infty$, $\sum_{k=0}^{\infty} \beta_k^2 < \infty$, and $\sum_{k=0}^{\infty} |\alpha_k - \beta_k| < \infty$.

Proof: See Appendix A. ■

Theorem 4.1 shows that the algorithm (6) can achieve the optimal point p^* to minimize the unconstrained optimization problem (3) in a distributed way. This subsection can also be viewed as dealing with the convergence of the proposed algorithm in the inconsistent case. This means that the obtained p^* is an estimate of the source location in the inconsistent case. Moreover, the conditions in Theorem 4.1 can be easily satisfied by taking $\alpha_k = \frac{1}{k+2}$ and $\beta_k = \frac{1}{k+1}$ for $k = 0, 1, \cdots$.

In the consistent case, we need further to show that the optimal estimate belongs to set $X_0 = \bigcap_{i=1}^{n} X_i$. In this case, it follows from Lemma 2.3 (iii) that $|\cdot|_{Y_i}^2$, $i = 1, \cdots, n$ are convex and continuously differentiable functions. According to the compactness of X_0 and the Weierstrass' Theorem, the optimal solution set X^* of problem (5) is nonempty.

By using Remark 2.4, we can obtain the following lemma in [16].

Lemma 4.2: For any $x_i \in X_i$, $i = 1, \cdots, n$, let $\hat{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $\bar{x} \in \cap_{i=1}^{n} X_i^0$. Then $s = \frac{\epsilon}{\epsilon+\delta}\bar{x} + \frac{\delta}{\epsilon+\delta}\hat{x} \in X_0$ and

$$\|\hat{x} - s\| \leq \frac{1}{\delta n} \left(\sum_{j=1}^{n} \|x_j - \bar{x}\| \right) \left(\sum_{j=1}^{n} |\hat{x}|_{X_j} \right),$$

where $\epsilon = \sum_{j=1}^{n} |\hat{x}|_{X_j}$ and δ is a positive number satisfying $\{y \,|\, \|y - \bar{x}\| \leq \delta\} \subseteq X_0$.

In fact, when $v_i(k) \in X_i$, $w_i(k) \in X_i$ and then $|w_i(k)|_{X_i} = 0$. When $v_i(k) \notin X_i$, $|w_i(k)|_{X_i} = 0$ if $w_i(k) \in X_i$; otherwise, it is not hard to get

$$|w_i(k)|_{X_i} = |v_i(k)|_{X_i} - \alpha_k b_i\, |v_i(k)|_{Y_i}.$$

Next, we demonstrate how to select the sequence $\{\alpha_k\}$. For any $k \geq 0$ and $i \in \mathcal{N}_{out} := \{i | v_i(k) \notin X_i\}$, α_k must satisfy at least one of the following two conditions:

- $|v_i(k)|_{X_i} - \alpha_k b_i |v_i(k)|_{Y_i} \leq \alpha_k \tilde{L}$;
- $v_i(k) - \alpha_k d_i(k) \in X_i$,

where $\tilde{L}$ is a given large enough positive constant. We term this property as the *alternative property*. From the alternative property, we have

$$|w_i(k)|_{X_i} \leq \alpha_k \tilde{L}, \qquad \forall i \in \mathcal{V}. \tag{12}$$

Remark 4.3: Once $v_i(k) \in X_i$ for sensor i, the above condition is not affected by α_k. Hence, we only need to consider sensor i $(i \in \mathcal{N}_{out})$ to determine α_k. In this case, the alternative property means that α_k should satisfy that the approximate projection of $v_i(k)$ on its inner ring Y_i is sufficiently close to X_i or belongs to X_i. Our algorithm objective is to find a point in the set $X_0 = \bigcap_{i=1}^{n} X_i$ to minimize the sum function $\sum_{i=1}^{n} b_i|\cdot|_{Y_i}^2$. It is a natural assumption that makes its state close to its own constrained set X_i as much as possible when sensor i optimizes its own function $b_i|\cdot|^2_{Y_i}$. Therefore, the alternative property assumption is reasonable. In other words, we first make all the sensors' estimates converge to the intersection set X_0 with the alternative property assumption, and then optimize the sum function $\sum_{i=1}^n b_i|\cdot|^2_{Y_i}$ to achieve a consensus point.

The next theorem presents the convergence result for the distributed algorithm (6) in the consistent case.

Theorem 4.4: Under Assumptions 1 and 2, the sequences $\{x_i(k)\}_{i=1}^n$ produced by the algorithm (6) converge to the optimal solution of problem (5), i.e.,

$$\lim_{k\to\infty} x_i(k) = x^* \text{ for some } x^* \in X^* \text{ and all } i,$$

if the sequence $\{\alpha_k\}$ satisfies the alternative property along with $\sum_{k=0}^{\infty} \alpha_k = \infty$, $\sum_{k=0}^{\infty} \beta_k = \infty$, $\sum_{k=0}^{\infty} \alpha_k^2 < \infty$, and $\sum_{k=0}^{\infty} \beta_k^2 < \infty$.

Proof: See Appendix B. ■

Remark 4.5: The condition on α_k in Theorem 4.4 can be easily satisfied by setting $\alpha_k = \frac{1}{k+2}$ and $\beta_k = \frac{1}{k+1}$, for instance. Moreover, the alternative property can also be guaranteed with a sufficiently large constant $\tilde{L} > 0$. In the next section, we can see that the value of $\tilde{L}$ is not necessarily too large. In addition, the optimal solution of constrained optimization problem (5) is also the global optimal solution of unconstrained optimization problem (3) if $\sum_{k=0}^{\infty} |\alpha_k - \beta_k| < \infty$ holds.

Remark 4.6: Note that [12], [15] focused on unconstrained convex problem, but we consider a constrained convex problem (in the consistent case). Clearly, the constrained convex optimization is much more difficult than the optimization without constraints, especially when each agent can only know its own constraint. In [16], all the constraint sets are assumed to be identical or the communication topologies are assumed to be fixed and fully connected ($a_{ij} = \frac{1}{n}$ for all i, j). In contrast, Theorem 4.4 gives a complete analysis for the case of different constraints under switching topologies. Moreover, we also try our best to provide a unified algorithm to handle two different cases (both consistent and inconsistent cases).

V. Simulation and Discussion

In this section, we give simulation results for the consistent and inconsistent cases to illustrate the convergence performance of the distributed alternating projection algorithm proposed in this paper. Then we give a comparative study about the estimation performance of our algorithm, the distributed projection algorithm (DPA) in [12], the grid search method for MLE in [3], and the convex relaxation method for linear cone programming (LCP) in [12], as well as the Cramer-Rao lower bound (CRLB) by a series of Monte Carlo simulations.

Consider a group of five sensor nodes with the switching interaction topologies randomly placed in a 10 m × 10 m field for illustration. The measurement of the source energy at each sensor is generated by (2). Suppose that the gain factors g_i equal 1 for all sensors and the power-loss factor is set as $\nu = 2$. The source is located at $\rho = [5.5, 5.5]^T$ and emits a signal with P set to 20 mW. An AWGN channel is assumed with $\epsilon_i \sim N(0, \sigma_i^2)$, where $0 \le \sigma_i \le 0.3$, $i = 1, \cdots, 5$. The actual

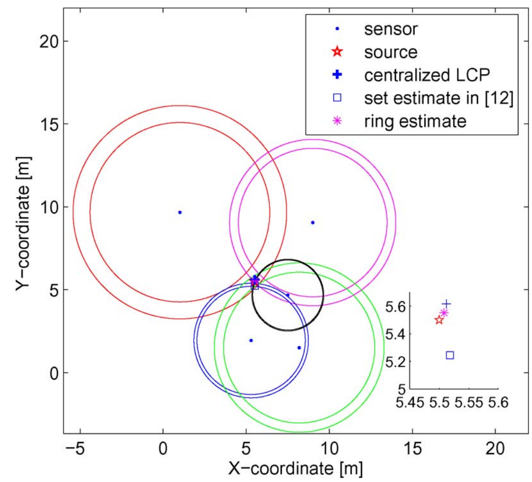


Fig. 2. Rings determined by sensors.

receiver SNR (signal to noise ratio) at different sensors depends on the distance to the source. For instance, if the variance of noise is 0.01 mW, the receiver SNR of a sensor 5 m away from the source is $10 \times \log_{10} \frac{(20/5^2)}{0.01} = 19$ dB. To simplify the simulation, here we take $b_1 = \cdots = b_5 = b$ considering the inner disks have the same priority. The standard deviation of background noise for each sensor takes the same value σ, and then the trust parameters $\xi_1 = \cdots = \xi_5 = \xi$.

The time-varying interaction topology $\mathcal{G}$ is periodically switched between two graphs $\mathcal{G}_i,\ i = 1, 2$, where $\mathcal{G}(2k) = \mathcal{G}_1$ and $\mathcal{G}(2k+1) = \mathcal{G}_2$ for $k = 0, 1, \cdots$. The corresponding adjacency matrices are described by $A_1 = [1, 0, 0, 0, 0; 0, 0.6, 0.2, 0, 0.2; 0, 0, 0.6, 0.4, 0; 0, 0.4, 0, 0.6, 0; 0, 0, 0.2, 0, 0.8]$ and $A_2 = [0.6, 0.4, 0, 0, 0; 0, 0.6, 0, 0.4, 0; 0, 0, 0.8, 0, 0.2; 0.4, 0, 0, 0.6, 0; 0, 0, 0.2, 0, 0.8]$.

A. Convergence Performance

In this subsection, we give a comparative study about the estimation performance of our proposed algorithm (6), the DPA for convex set intersection (without caring about $\{Y_i\}_{i=1}^n$) in [12], and the LCP method in [29].

At first, we take $b = 1$, $\sigma = 0.1$, and $\xi = 1 < 3$, and the intersection problem is inconsistent most of time in our simulation, which means the intersection of the corresponding rings determined by sensors is empty, i.e., $R_0 = \bigcap_{i=1}^5 R_i = \emptyset$. Fig. 2 shows a source localization scenario and the corresponding sensing rings, determined by the sensors, with dots standing for the sensors and $\star$ for the source.

Take the step-sizes (projection accuracies) with $\alpha_k = \frac{1}{k+2}$ and $\beta_k = \frac{1}{k+1}$, which satisfy the conditions in Theorem 4.1. The estimation errors, expressed as the logarithm of the Euclidean distance between the estimated and true source location (i.e., $\log_{10} \|x_i - \rho\|$, $i = 1, \cdots, 5$), are shown, respectively with our algorithm (6) and the DPA in Fig. 3. It is shown that the corresponding average estimation value $\hat{x} = \frac{1}{5}\sum_{i=1}^5 x_i$ of DPA converges to $[5.518, 5.245]^T$ (marked with □ in Fig. 2) after about 10000 iteration steps, the average estimation $\hat{x}$ of our algorithm converges to $[5.508, 5.551]^T$ (marked with ∗ in Fig. 2) after 10000 iterations, which is very close to the centralized

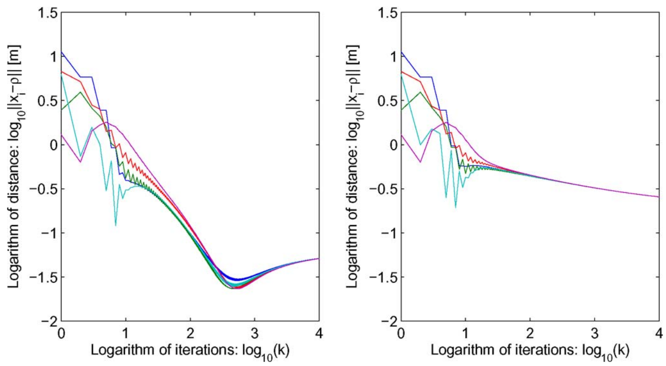


Fig. 3. Estimation errors by proposed algorithm (6) (left) and the DPA in [12] (right) in the inconsistent case.

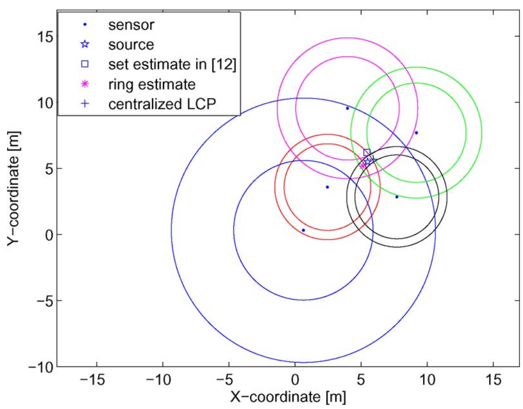


Fig. 4. Rings determined by sensors.

solution $[5.512, 5.554]^T$. Additionally, the source estimate by using the centralized LCP method is $[5.512, 5.616]^T$ (marked with + in Fig. 2), and the corresponding estimation errors are 0.2558, 0.0515, 0.1166 by these three methods, which implies that our algorithm has a better estimation performance in this case.

Next, when $\xi = 3$, the intersection problem becomes consistent with a high probability. Fig. 4 shows a source localization scenario, and the corresponding sensing rings determined by the sensors, with dots and $\star$ representing the sensors and the source, respectively. From Fig. 4, we clearly see that the ring intersection is nonempty (i.e., $R_0 \neq \emptyset$).

Take $\tilde{L} = 30$ and the algorithm step-sizes $\alpha_k = \frac{1}{k+2}$ and $\beta_k = \frac{1}{k+1}$, which satisfy the conditions in Theorem 4.4. The estimation errors and convergence results are presented with our algorithm (6) and DPA in [12] in Fig. 5, respectively. DPA cannot converge to X_0 after 10000 steps and the corresponding average estimation value $\hat{x} = \frac{1}{5}\sum_{i=1}^{5} x_i = [5.447, 6.216]^T \notin X_0$ (marked with $\square$ in Fig. 4). On the other hand, the average estimation value $\hat{x} = [5.123, 5.198]^T \in X_0$ (marked with $\star$ in Fig. 4) after 10000 steps with our method, which is very close to the centralized solution $[5.133, 5.187]^T$. Additionally, the source estimation by using centralized LCP method in [29] is $[5.937, 5.708]^T$ (marked with + in Fig. 4). Thus, the corresponding estimation errors are 0.7184, 0.4830, 0.4845 for these three methods, respectively. This simulation results show that our algorithm has better performance than DPA and the centralized LCP methods.

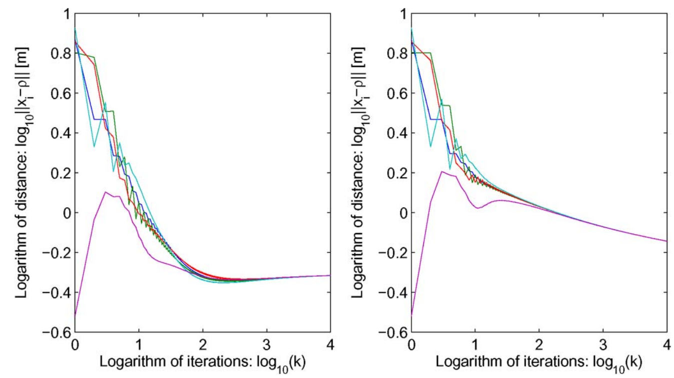


Fig. 5. Estimation errors by proposed algorithm (6) (left) and the DPA in [12] (right) in the consistent case.

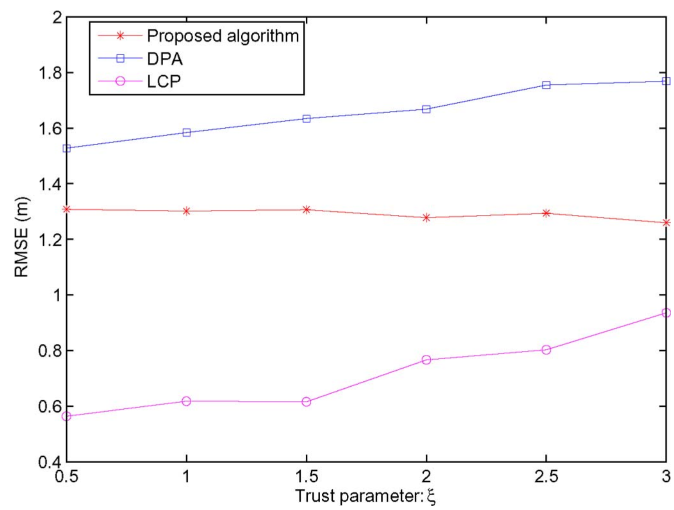


Fig. 6. Estimate error as trust parameter ξ changes.

Although the selection of trust parameter ξ determines if the problem is consistent or inconsistent, the above simulation results of the proposed algorithm (6) show good convergence performance in both consistent and inconsistent cases without knowing which case it is beforehand.

Then we investigate the convergence performance of the proposed algorithm with an analysis on ξ. In this simulation, sensors are randomly placed in each run with $\sigma = 0.1$, $b = 1$. Next, we compare the estimation accuracy between our proposed algorithm (6), DPA and LCP methods with changing ξ, based on the well-known localization accuracy metric, root mean square error (RMSE),

$$RMSE = \sqrt{\frac{1}{T_r}\sum_{i=1}^{T_r} \|\rho_i^* - \rho\|^2},$$

where T_r denotes the number of trials and ρ_i^* denotes the estimate of source location at the i-th trail for $i \in \{1, \cdots, T_r\}$. Fig. 6 shows the RMSE as ξ changes under 1000 trials, where the estimation error for DPA and LCP roughly increases as ξ increases. In fact, the possibility of the source located in R_i increases as ξ increases, and therefore, the intersection set of these rings becomes larger, which may yield larger estimation

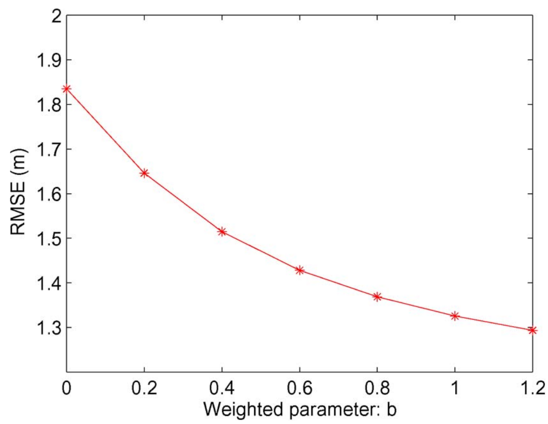


Fig. 7. Estimation errors as weighted parameter b changes.

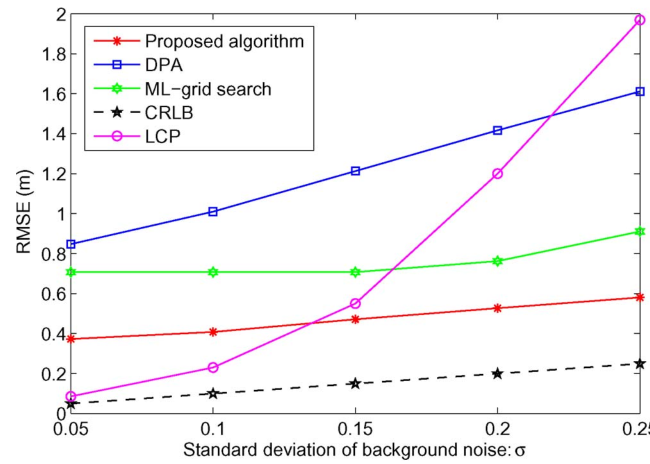


Fig. 8. Comparisons in accuracy as noise standard deviation changes through 5000 trails for a fixed sensor configuration.

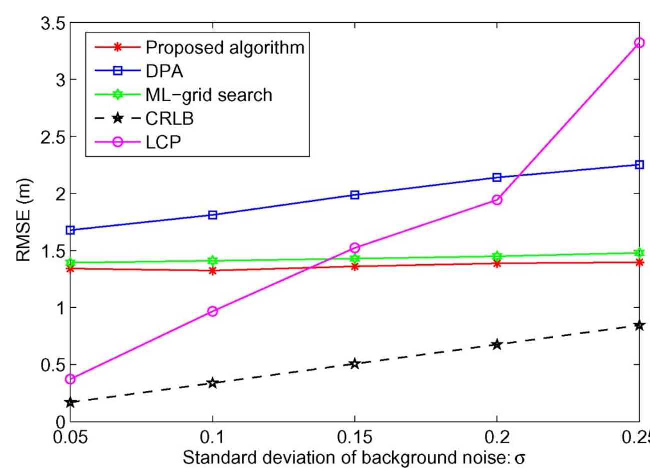


Fig. 9. Comparisons in accuracy as noise standard deviation changes through 8000 trails for randomly placed sensors.

errors. However, our proposed algorithm actually solves the "constrained" problem

$$\min \sum_{i=1}^{n} \left(|x|_{X_i}^2 + |x|_{Y_i}^2 \right),$$
$$s.t.,\ x \in \mathbb{R}^2 \text{ (inconsistent) or } x \in X_0 \text{ (consistent)}$$

and ξ has a less impact on this objective function $\sum_{i=1}^{n}(|x|_{X_i}^2 + |x|_{Y_i}^2)$ than the term $\sum_{i=1}^{n}|x|_{X_i}^2$, which leads to better performance in the statistical sense.

Moreover, the stepsize conditions in Theorems 4.1 and 4.4 can be easily satisfied by setting $\alpha_k = \frac{1}{k+2}$ and $\beta_k = \frac{1}{k+1}$, for instance. Furthermore, the alternative property in the consistent case can also be guaranteed with a sufficiently large constant $\tilde{L} > 0$. In the above simulation example, we can see that the value of $\tilde{L}$ is not necessarily too large.

Next, we discuss the effectiveness of augmentation term $\sum_{i=1}^{n} b_i|x|_{Y_i}^2$. With the noise environment $\sigma = 0.1$ and trust parameter $\xi = \min\left\{3, \frac{P_{s_i}}{\sigma}, i = 1, \ldots, 5\right\}$, Fig. 7 shows the effect of weighted parameter b on the RMSE under the 1000 different selections of sensors' locations and initial estimate values by our proposed algorithm (6). In addition, in most of our simulations, we find that, when $b \le 1$, the estimation performance of our algorithm is better than that of the algorithm (8), where $b = 0$.

B. Estimation Performance Comparison

Let us study the estimation performance of our proposed algorithm compared with the DPA [12], the grid search method for MLE in [3], the convex relaxation method for linear cone programming (LCP) in [29], as well as the Cramer-Rao lower bound (CRLB) through thousands of Monte Carlo simulations. The CRLB is computed as follows:

$$CRLB = \sqrt{\frac{1}{T_r}\sum_{i=1}^{T_r} CRLB_i}$$

where $CRLB_i$ denotes the CRLB at the i-th trail.

Here the grid search resolution is set to 1 m × 1 m and the LCP method can be solved by using the SeDuMi toolbox [33]. Note that both grid search and LCP methods are operated in a centralized way. With $b = 1$ and $\xi = \min\left\{3, \frac{p_{s_i}}{\sigma}, i = 1, \ldots, 5\right\}$, Figs. 8 and 9 demonstrate the localization accuracy metric RMSE of these four methods as a function of the standard deviation of the background noise through 5000 trials for a fixed sensor configuration [7, 9;3, 5;5, 2;7, 2;9, 3] and 8000 trials for random sensor configuration in each run, respectively. The simulation results in Figs. 8 and 9 exhibit that our algorithm outperforms the ML-grid search and DPA methods. With the increasing standard deviation of the background noise, the performance of our algorithm becomes better than the centralized convex relaxation method LCP.

At last, we investigate the scalability with the number of sensors for our algorithm. For simplicity, the communication topology graph is switching between two different 3-regular connected undirected graphs with Metropolis weights. With $\sigma = 0.1$, $b = 1$ and $\xi = \min\left\{3, \frac{p_{s_i}}{\sigma}, i = 1, \ldots, 5\right\}$, Fig. 10 demonstrates the localization accuracy metric RMSE of four different number of sensors through 1000 trials for randomly placed sensors in each run. It can be observed from Fig. 10 that the convergence speeds for various scales of network are about the same. Also, the result shows that the estimation error

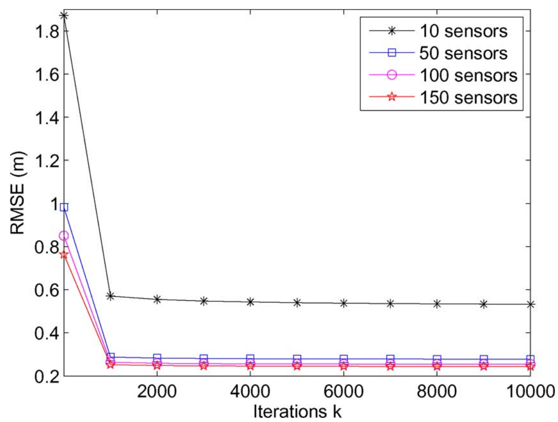


Fig. 10. Estimation errors by proposed algorithm (6) for four different number of sensors.

indeed decreases with the increasing number of sensors, but the estimation accuracy will not be effectively improved after the number of the sensors reaches a certain level. In addition, we find that the estimation accuracy of our algorithm keeps almost the same after a certain number of iterations in a large scale sensor network: Fig. 10 shows that our localization algorithm (6) will almost converge (that is, the estimation error keeps almost unchanged) at 2000 iterations no matter how many sensors are in the sensor network.

VI. Conclusion

In this paper, the source localization problem was considered by cooperatively computing the intersection of the sensing rings of the sensors. A distributed alternating projection algorithm was proposed to deal with this complicated non-convex optimization problem. With alternating computing the projection on the inner and outer disks of each sensor, sufficient convergence conditions were obtained for the distributed localization algorithm in both the consistent and inconsistent cases. Simulation results show that the performance of our algorithm perform good in a comparison with some existing algorithms, especially convex relaxation based on convex intersection computation, semidefinite and cone programming. In fact, many interesting and challenging problems still remain to be addressed, including how to extend this distributed non-convex ring intersection algorithm from single source to multiple sources and how to balance the localization performance and the computational complexity in solving the non-convex optimization.

Appendix A
Proof of Theorem 4.1

Proof: For any $q^* \in Z^*$, it follows from Lemma 2.3 (iii) and (1) that

$$\begin{aligned}\|x_i(k+1) - q^*\|^2 &\le \|w_i(k) - q^*\|^2 + L^2\beta_k^2 - 2\beta_k \delta_i^T(k)\left(w_i(k) - q^*\right)\\ &\le \beta_k\left(|q^*|^2_{X_i} - |w_i(k)|^2_{X_i}\right) + \|w_i(k) - q^*\|^2 + L^2\beta_k^2.\end{aligned}$$

With the boundedness of sequences $\{x_i(k)\}$, $\{v_i(k)\}$ and $\{w_i(k)\}$, we obtain that

$$M := \max_{i,k}\left\{|w_i(k)|_{X_i}, |v_i(k)|_{Y_i}, |q^*|_{X_i}, \|\hat{x}(k) - q^*\|, |\hat{x}(k)|_{X_i}\right\}$$

is a finite number. It follows from Lemma 2.3 (ii) that

$$\begin{aligned}&\beta_k\left(|q^*|^2_{X_i} - |w_i(k)|^2_{X_i}\right)\\ &\qquad \le \beta_k\left(|q^*|^2_{X_i} - |\hat{x}(k)|^2_{X_i}\right) + 2M\beta_k \|\hat{x}(k) - w_i(k)\|.\end{aligned}$$

Since

$$\begin{aligned}\|\hat{x}(k) - w_i(k)\| &\le \|\hat{x}(k) - v_i(k)\| + \|v_i(k) - w_i(k)\|\\ &\le \sum_{j=1}^n a_{ij}(k)\|\hat{x}(k) - x_j(k)\| + L\alpha_k\end{aligned}$$

we have

$$\begin{aligned}&\|x_i(k+1) - q^*\|^2\\ &\quad \le \|w_i(k) - q^*\|^2 + \beta_k\left(|q^*|^2_{X_i} - |\hat{x}(k)|^2_{X_i}\right) + 2ML\alpha_k\beta_k\\ &\qquad + 2M\sum_{j=1}^n a_{ij}(k)\beta_k\|\hat{x}(k) - x_j(k)\| + L^2\beta_k^2. \qquad (13)\end{aligned}$$

With a similar expanding of $\|x_i(k+1) - q^*\|^2$ given in (13), we have

$$\begin{aligned}&\|w_i(k) - q^*\|^2\\ &\le \sum_{j=1}^n a_{ij}(k)\|x_j(k) - q^*\|^2 + \alpha_k b_i\left(|q^*|^2_{Y_i} - |\hat{x}(k)|^2_{Y_i}\right)\\ &\quad + L^2\alpha_k^2 + b_0(d_0 + 2M)\sum_{j=1}^n a_{ij}(k)\alpha_k\|\hat{x}(k) - x_j(k)\|. \qquad (14)\end{aligned}$$

Since

$$\begin{aligned}&\alpha_k\sum_{i=1}^n b_i\left(|q^*|^2_{Y_i} - |\hat{x}(k)|^2_{Y_i}\right) + \beta_k\sum_{i=1}^n\left(|q^*|^2_{X_i} - |\hat{x}(k)|^2_{X_i}\right)\\ &\quad \le \alpha_k\sum_{i=1}^n\left(\left(b_i|q^*|^2_{Y_i} + |q^*|^2_{X_i}\right) - \left(b_i|\hat{x}(k)|^2_{Y_i} + |\hat{x}(k)|^2_{X_i}\right)\right)\\ &\qquad + 2nM^2|\beta_k - \alpha_k|\end{aligned}$$

it follows from (13) and (14) that

$$\begin{aligned}&\sum_{i=1}^n\|x_i(k+1) - q^*\|^2\\ &\quad \le \sum_{i=1}^n\|x_i(k) - q^*\|^2 + b_0(d_0 + 2M)\sum_{j=1}^n\alpha_k\|\hat{x}(k) - x_j(k)\|\\ &\qquad + \alpha_k\sum_{i=1}^n\left(\left(b_i|q^*|^2_{Y_i} + |q^*|^2_{X_i}\right) - \left(b_i|\hat{x}(k)|^2_{Y_i} + |\hat{x}(k)|^2_{X_i}\right)\right)\\ &\qquad + nML\left(\alpha_k^2 + \beta_k^2\right) + nL^2\beta_k^2 + nL^2\alpha_k^2\\ &\qquad + 2M\sum_{j=1}^n\beta_k\|\hat{x}(k) - x_j(k)\| + 2nM^2|\beta_k - \alpha_k|. \qquad (15)\end{aligned}$$

Due to $q^* \in Z^*$, we have

$$\sum_{i=1}^{n}\left(\left(b_i|q^*|_{Y_i}^2+|q^*|_{X_i}^2\right)-\left(b_i\left|\hat{x}(k)\right|_{Y_i}^2+\left|\hat{x}(k)\right|_{X_i}^2\right)\right)\leq 0.$$

With the given conditions $\sum_{k=0}^{\infty}\alpha_k^2<\infty$, $\sum_{k=0}^{\infty}\beta_k^2<\infty$, $\sum_{k=0}^{\infty}|\alpha_k-\beta_k|<\infty$ and Lemma 3.5, it is easy to find that the sequence $\{\sum_{i=1}^{n}\|x_i(k)-q^*\|^2\}$ is convergent for every $q^*\in Z^*$ from Lemma 3.3. Consequently, for any i, the sequence $\{\|x_i(k)-q^*\|\}$ is also convergent for every $q^*\in Z^*$. Notice that $\lim_{k\to\infty}\|x_i(k)-\hat{x}(k)\|=0$ for $i\in\mathcal{V}$, which implies the sequence $\{\|\hat{x}(k)-q^*\|\}$ is convergent for any $q^*\in Z^*$.

By summing (15) over k, we have

$$\begin{aligned}&\sum_{r=0}^{k}\alpha_r\sum_{i=1}^{n}\left(\left(b_i\left|\hat{x}(r)\right|_{Y_i}^2+\left|\hat{x}(r)\right|_{X_i}^2\right)-\left(b_i|q^*|_{Y_i}^2+|q^*|_{X_i}^2\right)\right)\\&\leq\sum_{i=1}^{n}\|x_i(0)-q^*\|^2+(nML+nL^2)\sum_{r=0}^{k}\left(\alpha_r^2+\beta_r^2\right)\\&\quad+2nM^2\sum_{r=0}^{k}|\beta_r-\alpha_r|+2M\sum_{j=1}^{n}\sum_{r=0}^{k}\beta_r\left\|\hat{x}(r)-x_j(r)\right\|\\&\quad+b_0(d_0+2M)\sum_{j=1}^{n}\sum_{r=0}^{k}\alpha_r\left\|\hat{x}(r)-x_j(r)\right\|\\&<\infty.\end{aligned}$$

Since $\sum_{r=0}^{\infty}\alpha_r=\infty$, we have

$$\liminf_{r\to\infty}\sum_{i=1}^{n}\left(\left(b_i\left|\hat{x}(r)\right|_{Y_i}^2+\left|\hat{x}(r)\right|_{X_i}^2\right)-\left(b_i|q^*|_{Y_i}^2+|q^*|_{X_i}^2\right)\right)=0.$$

Hence, there exists a limit point p^* of $\{\hat{x}(k)\}$ such that $p^*\in Z^*$. Similarly, we get that the sequence $\{\|\hat{x}(k)-p^*\|\}$ is convergent. Therefore, $\lim_{k\to\infty}\|x_i(k)-p^*\|=0$ for all i; namely,

$$\lim_{k\to\infty}x_i(k)=p^*,\quad i=1,\cdots,n.$$

Thus, the conclusion follows. ■

Appendix B
Proof of Theorem 4.4

Proof: For any $z^*\in X^*$, it follows from $X^*\subseteq X_0\subseteq X_i$ and Lemma 2.3 (i) that

$$\begin{aligned}&\|x_i(k+1)-z^*\|^2\\&\leq(1-\beta_k)\|w_i(k)-z^*\|^2+\beta_k\left\|P_{X_i}(w_i(k))-z^*\right\|^2\\&\leq(1-\beta_k)\|w_i(k)-z^*\|^2+\beta_k\|w_i(k)-z^*\|^2-\beta_k|w_i(k)|_{X_i}^2\\&\leq\|v_i(k)-z^*\|^2+L^2\alpha_k^2+2\alpha_k d_i^T(k)\left(z^*-v_i(k)\right)\\&\leq\sum_{j=1}^{n}a_{ij}(k)\|x_j(k)-z^*\|^2+\alpha_k b_i\left(|z^*|_{Y_i}^2-|v_i(k)|_{Y_i}^2\right)\\&\quad+L^2\alpha_k^2.\end{aligned}\tag{16}$$

From Remark 2.4, there are a vector $\bar{x}\in\cap_{i=1}^{n}X_i^0$ and a number $\delta>0$ such that $\{y|\|y-\bar{x}\|\leq\delta\}\subseteq X_0$. Define $s(k)=\frac{\epsilon}{\epsilon+\delta}\bar{x}+\frac{\delta}{\epsilon+\delta}\hat{x}(k)$, where $\epsilon=\sum_{j=1}^{n}|\hat{x}(k)|_{X_j}$. According to Lemma 4.2, $s(k)\in X_0$. Note that $\alpha_k b_i(|z^*|_{Y_i}^2-|v_i(k)|_{Y_i}^2)=\alpha_k b_i(|z^*|_{Y_i}^2-|s(k)|_{Y_i}^2)+\alpha_k b_i(|s(k)|_{Y_i}+|v_i(k)|_{Y_i})(|s(k)|_{Y_i}-|v_i(k)|_{Y_i})$. With the boundedness of sequences $\{x_i(k)\}$, $\{v_i(k)\}$ and $\{w_i(k)\}$, we obtain

$$W:=\max_{i,k}\left\{|w_i(k)|_{X_i},|v_i(k)|_{Y_i},\|z^*-v_i(k)\|\right\}<\infty.$$

It follows from $|s(k)|_{Y_i}\leq d_0$ and Lemma 2.3 (ii) that

$$\begin{aligned}&\alpha_k b_i\left(|z^*|_{Y_i}^2-|v_i(k)|_{Y_i}^2\right)\\&=\alpha_k b_i\left(|z^*|_{Y_i}^2-|s(k)|_{Y_i}^2\right)+\alpha_k b_i(d_0+W)\|s(k)-v_i(k)\|\\&\leq\alpha_k b_i\left(|z^*|_{Y_i}^2-|s(k)|_{Y_i}^2\right)+\alpha_k b_0(d_0+W)\|s(k)-\hat{x}(k)\|\\&\quad+\alpha_k b_0(d_0+W)\|\hat{x}(k)-v_i(k)\|.\end{aligned}$$

From (16), we have

$$\begin{aligned}&\sum_{i=1}^{n}\|x_i(k+1)-z^*\|^2\\&\leq\sum_{i=1}^{n}\|x_i(k)-z^*\|^2+\sum_{i=1}^{n}\alpha_k b_i\left(|z^*|_{Y_i}^2-|s(k)|_{Y_i}^2\right)\\&\quad+b_0(d_0+W)\sum_{j=1}^{n}\alpha_k\|\hat{x}(k)-x_j(k)\|+nL^2\alpha_k^2\\&\quad+nb_0(d_0+W)\alpha_k\|s(k)-\hat{x}(k)\|.\end{aligned}\tag{17}$$

Since $\{x_i(k)\}$ is bounded and $\bar{x}\in X_0$, $\{x_i(k)-\bar{x}\}$ is bounded for any i. Without loss of generality, we assume $\|x_i(k)-\bar{x}\|\leq W$. By Lemma 4.2,

$$\begin{aligned}\|\hat{x}(k)-s(k)\|&\leq\frac{1}{\delta n}\left(\sum_{j=1}^{n}\|x_j(k)-\bar{x}\|\right)\left(\sum_{j=1}^{n}|\hat{x}(k)|_{X_j}\right)\\&\leq\frac{W}{\delta}\sum_{j=1}^{n}\left\|\hat{x}(k)-P_{X_j}(w_j(k))\right\|.\end{aligned}$$

As a result,

$$\begin{aligned}&\|\hat{x}(k)-s(k)\|\\&\quad\leq\frac{W}{\delta}\sum_{j=1}^{n}\|\hat{x}(k)-x_j(k)\|+\frac{W}{\delta}(1-\beta_k)\sum_{j=1}^{n}|w_j(k)|_{X_j}\\&\quad+\frac{nLW}{\delta}\beta_k+\frac{nLW}{\delta}\alpha_k.\end{aligned}\tag{18}$$

Recalling (12), it follows from (18), Lemma 3.5, $\sum_{k=0}^{\infty}\alpha_k^2 < \infty$, and $\sum_{k=0}^{\infty}\beta_k^2 < \infty$ that

$$\sum_{k=0}^{\infty}\alpha_k \|\hat{x}(k) - s(k)\| \le \frac{W}{\delta}\sum_{j=1}^{n}\sum_{k=0}^{\infty}\alpha_k \|\hat{x}(k) - x_j(k)\| + \frac{nLW}{\delta}\sum_{k=0}^{\infty}\left(2\alpha_k^2 + \beta_k^2\right) + \frac{nW}{\delta}\tilde{L}\sum_{k=0}^{\infty}\alpha_k^2 < \infty.$$

Since $s(k) \in X_0$ and $z^* \in X^*$, we obtain

$$\sum_{i=1}^{n} b_i \left(|z^*|_{Y_i}^2 - |s(k)|_{Y_i}^2\right) \le 0.$$

Then it follows from Lemma 3.3 and (17) that the sequence $\{\sum_{i=1}^{n}\|x_i(k) - z^*\|^2\}$ is convergent for every $z^* \in X^*$. Consequently, the sequence $\{\|x_i(k) - z^*\|\}$ is also convergent for any i and $z^* \in X^*$.

By Lemma 3.5, we obtain $\lim_{k\to\infty}\|x_i(k) - \hat{x}(k)\| = 0$ for all i. It can be verified from (12) and (18) that $\lim_{k\to\infty}\|\hat{x}(k) - s(k)\| = 0$. Therefore, $\lim_{k\to\infty}\|x_i(k) - s(k)\| = 0$, which implies that the sequence $\{\|s(k) - z^*\|\}$ is convergent for every $z^* \in X^*$.

By summing (17) over k, we have

$$\sum_{i=1}^{n}\|x_i(k+1) - z^*\|^2 + \sum_{r=0}^{k}\sum_{i=1}^{n}\alpha_r b_i\left(|s(r)|_{Y_i}^2 - |z^*|_{Y_i}^2\right) \le \sum_{i=1}^{n}\|x_i(0) - z^*\|^2 + nb_0(d_0 + W)\sum_{r=0}^{k}\alpha_r\|s(r) - \hat{x}(r)\| + nL^2\sum_{r=0}^{k}\alpha_r^2 + b_0(d_0 + W)\sum_{j=1}^{n}\sum_{r=0}^{k}\alpha_r\|\hat{x}(r) - x_j(r)\|.$$

Thus,

$$\sum_{r=0}^{\infty}\alpha_r\sum_{i=1}^{n}b_i\left(|s(r)|_{Y_i}^2 - |z^*|_{Y_i}^2\right) < \infty.$$

Since $\sum_{r=0}^{\infty}\alpha_r = \infty$, we obtain

$$\liminf_{r\to\infty}\sum_{i=1}^{n}b_i\left(|s(r)|_{Y_i}^2 - |z^*|_{Y_i}^2\right) = 0.$$

As a result, there exists a limit point x^* of $\{s(k)\}$ such that $x^* \in X^*$. Then with a similar analysis we get that the sequence $\{\|s(k) - x^*\|\}$ is convergent, which leads to $\lim_{k\to\infty}\|s(k) - x^*\| = 0$. Therefore, $\lim_{k\to\infty}\|x_i(k) - x^*\| = 0$, which completes the proof. ■

References

[1] J. Bachrach and C. Taylor, *Handbook of Sensor Networks*. Hoboken, NJ, USA: Wiley, 2005.

[2] H. L. Song, "Automatic vehicle location in cellular communications systems," *IEEE Trans. Veh. Technol.*, vol. 43, no. 4, pp. 902–908, Nov. 1994.

[3] X. Sheng and Y. H. Hu, "Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor network," *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 44–53, Jan. 2005.

[4] W. Meng, W. Xiao, and L. Xie, "An efficient EM algorithm for energy-based sensor network multi-source localization in wireless sensor networks," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 3, pp. 1017–1027, Mar. 2011.

[5] M. Rabbat and R. Nowak, "Decentralized source localization and tracking," in *Proc. ICASSP*, Mar. 2004, vol. 3, pp. 921–924.

[6] S. Srirangarajan, A. Tewfik, and Z.-Q. Luo, "Distributed sensor network localization using SOCP relaxation," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4886–4895, Dec. 2008.

[7] W. Wang, V. Srinivasan, B. Wang, and K. Chua, "Coverage for target localization in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 2, pp. 667–676, Feb. 2008.

[8] K. C. Ho and M. Sun, "An accurate algebraic closed-form solution for energy-based source localization," *IEEE Trans. Audio, Speech, Language Process.*, vol. 15, no. 8, pp. 2542–2550, Nov. 2007.

[9] C. Meesookho, U. Mitra, and S. Narayanan, "On energy-based acoustic source localization for sensor networks," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 365–376, Jan. 2008.

[10] D. Blatt and A. O. Hero, "Energy based sensor network source localization via projection on to convex sets," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3614–3619, Sep. 2006.

[11] P. L. Combettes, "Inconsistent signal feasibility problems: Least squares solutions in a product space," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 2955–2966, Nov. 1994.

[12] W. Meng, W. Xiao, L. Xie, and A. Pandharipande, "Diffusion based projection method for distributed source localization in wireless sensor networks," in *Proc. IEEE Comput. Commun. Workshops*, 2011, pp. 537–542.

[13] F. Deutsch and H. Hundal, "The rate of convergence for the cyclic projections algorithm III: Regularity of convex sets," *J. Approx. Theory*, vol. 155, no. 2, pp. 155–184, Dec. 2008.

[14] H. Bauschke and J. M. Borwein, "On projection algorithms for solving convex feasibility problems," *SIAM Rev.*, vol. 38, no. 3, pp. 367–426, Sep. 1996.

[15] A. Nedic and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Trans. Autom. Control*, vol. 54, no. 1, pp. 48–61, Jan. 2009.

[16] A. Nedic, A. Ozdaglar, and P. A. Parrilo, "Constrained consensus and optimization in multi-agent networks," *IEEE Trans. Autom. Control*, vol. 55, no. 4, pp. 922–938, Apr. 2010.

[17] G. Shi and K. H. Johansson, "Randomized optimal consensus of multi-agent systems," *Automatica*, vol. 48, no. 12, pp. 3018–3030, Dec. 2012.

[18] Z. J. Towfic and A. H. Sayed, "Adaptive penalty-based distributed stochastic convex optimization," *IEEE Trans. Signal Process.*, vol. 62, no. 15, pp. 3924–3938, Aug. 2014.

[19] G. Shi, K. H. Johansson, and Y. Hong, "Reaching optimal consensus: Dynamical systems that compute intersections of convex sets," *IEEE Trans. Autom. Control*, vol. 58, no. 3, pp. 610–622, Mar. 2013.

[20] Y. Lou, G. Shi, K. H. Johansson, and Y. Hong, "Approximate projected consensus for convex intersection computation: Convergence analysis and critical error angle," *IEEE Trans. Autom. Control*, vol. 59, no. 7, pp. 1722–1736, Jul. 2014.

[21] J. Aubin and A. Cellina, *Differential Inclusions*. Berlin, Germany: Springer-Verlag, 1984.

[22] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier system," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, Jul. 2006.

[23] J. Lee, R. Mazumdar, and N. Shroff, "Non-convex optimization and rate control for multi-class services in the Internet," *IEEE/ACM Trans. Networking*, vol. 13, no. 4, pp. 827–840, Aug. 2005.

[24] C. Godsil and G. Royle, *Algebraic Graph Theory*. New York, NY, USA: Springer-Verlag, 2001.

[25] B. T. Polyak, *Introduction to Optimization*. New York, NY, USA: Optimization Software, Inc., 1987.

[26] D. P. Bertsekas, A. Nedic, and A. Ozdaglar, *Convex Analysis and Optimization*. Belmont, MA, USA: Athena Scientific, 2003.

[27] P. Biswas and Y. Ye, "Semidefinite programming for Ad Hoc wireless sensor network localization," in *Proc. 3rd Int. Symp. Inf. Process. Sensor Netw.*, 2004, pp. 46–54.

[28] P. Biswas, T.-C. Lian, T.-C. Wang, and Y. Ye, "Semidefinite programming based algorithms for sensor network localization," *ACM Trans. Sens. Netw.*, vol. 2, no. 2, pp. 188–220, 2006.

[29] H. Chen, Q. Shi, R. Tan, H. V. Poor, and K. Sezaki, "Mobile element assisted cooperative localization for wireless sensor networks with obstacles," *IEEE Trans. Wireless Commun.*, vol. 9, no. 3, pp. 956–963, Mar. 2010.

[30] J. Nie, "Sum of squares methods for sensor network localizaiton," *Comput. Optim. Appl.*, vol. 43, pp. 151–179, 2009.

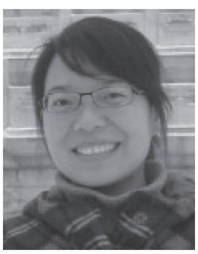
[31] P. Oguz-Ekim, J. Gomes, J. Xavier, and P. Oliveira, "A convex relaxation for approximate maximum-likelihood $2D$ source localization from range measurements," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2010, pp. 2698–2701.

[32] A. Simonetto and G. Leus, "Distributed maximum-likelihood sensor network localization," *IEEE Trans. Signal Process.*, vol. 62, no. 6, pp. 1424–1437, Mar. 2014.

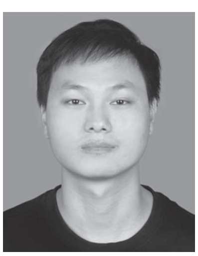
[33] J. F. Sturm, "Using SeDuMi 1.02 a Matlab toolbox for optimization over symmetric cone," *Optim. Methods Softw.*, vol. 11/12, no. 1–4, pp. 625–653, 1999.

[34] L. Xiao and S. Boyd, "A scheme for robust distributed sensor fusion based on average consensus," in *Proc. 4th Int. Symp. Inf. Process. Sensor Netw.*, 2005, pp. 63–70.

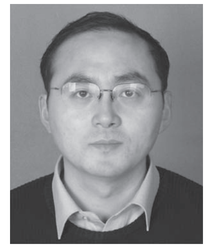
[35] B. Gharesifard and J. Cortes, "Distributed strategies for generating weight-balanced and doubly stochastic digraph," *Eur. J. Control*, vol. 18, no. 6, pp. 539–557, 2012.



Yanqiong Zhang received the bachelor's degree from Wuhan University in 2010. She is pursuing the Ph.D. degree at the Academy of Mathematics and Systems Science, Chinese Academy of Sciences. Her research interest covers distributed optimization, sensor networks, and multi-agent networks.



Youcheng Lou received the B.Sc. degree in mathematics and applied mathematics from Zhengzhou University in 2008, and Ph.D. degree in complex systems and control from Academy of Mathematics and Systems Science, Chinese Academy of Sciences in 2013. He was a visiting scholar at the Chinese University of Hong Kong from October 2014 to December 2014. Currently he is a postdoctoral researcher in the National Center for Mathematics and Interdisciplinary Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences. His research interests include multi-agent networks, distributed optimization and the applications of distributed methods in operation research.



Yiguang Hong (SM'02) received the B.Sc. and M.Sc. degrees from Peking University, Beijing, China, and the Ph.D. degree from the Chinese Academy of Sciences (CAS), Beijing. He is currently a Professor in the Institute of Systems Science, Academy of Mathematics and Systems Science, CAS, and serves as the Director of Key Lab of Systems and Control, CAS and the Director of the Information Technology Division, National Center for Mathematics and Interdisciplinary Sciences, CAS. His research interests include nonlinear dynamics and control, multi-agent systems, distributed optimization, social networks, and software reliability. Prof. Hong serves as Editor-in-Chief of the *Control Theory and Technology*, Deputy Editor-in-Chief of *Acta Automatica Sinca*, and also serves or served as Associate Editors for several journals including the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, *IEEE Control Systems Magazine, Nonlinear Analysis: Hybrid Systems*, and *Kybernetika*. He is a recipient of the Guang Zhaozhi Award at the Chinese Control Conference (in 1997), Young Author Prize of the IFAC World Congress (in 1999), Young Scientist Award of CAS (in 2001), and the Youth Award for Science and Technology of China (in 2006), and the National Natural Science Prize of China (in 2008).

Lihua Xie (F'07) received the B.E. and M.E. degrees in electrical engineering from Nanjing University of Science and Technology in 1983 and 1986, respectively, and the Ph.D. degree in electrical engineering from the University of Newcastle, Australia, in 1992. Since 1992, he has been with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, where he is currently a professor and served as the Head of Division of Control and Instrumentation from July 2011 to June 2014. He held teaching appointments in the Department of Automatic Control, Nanjing University of Science and Technology from 1986 to 1989 and Changjiang Visiting Professorship with South China University of Technology from 2006 to 2011. His research interests include robust control and estimation, networked control systems, multi-agent networks, and sensor networks. In these areas, he has published over 260 journal papers and co-authored two patents and six books. He has served as an editor of IET Book Series in Control and an Associate Editor of a number of journals including IEEE TRANSACTIONS ON AUTOMATIC CONTROL, AUTOMATICA, IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY, and IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS-II. Dr. Xie is a Fellow of IEEE and IFAC.