Distributed Estimation and Control for Robotic Sensor Networks

Charles Freundlich, Soomin Lee, and Michael M. Zavlanos

Abstract— In this paper, we address the problem of controlling a network of mobile sensors to estimate a collection of hidden states to a user-specified accuracy. The mobile sensors simultaneously take measurements of the hidden states, fuse them in a distributed filter, and plan their next views in order to minimize expected uncertainty. This formulation leads to an optimization problem with as many coupled LMI constraints as the number of hidden states. The network solves this problem by means of a new distributed random approximate projections method that is robust to the state disagreement errors that will exist among the robots as the distributed filter fuses the collected measurements and is computationally light enough to handle large numbers of hidden states.

I. Introduction

In this paper, we control a robotic sensor network to estimate a large collection of hidden states so that desired accuracy thresholds and confidence levels are satisfied. We require that estimation and control are completely distributed, i.e., beliefs and control actions are decided upon locally through pairwise communication among robots in the network. A common assumption in problems like the one discussed herein is that observations depend linearly on the state and are corrupted by Gaussian noise as well as that the sensors are self localized [1]–[5]. Under these assumptions, a typical approach to estimate a set of hidden states is to use an Information Filter (IF). The role of the IF is to fuse new linear-Gaussian observations of the hidden states with a prior Gaussian distribution to create a minimum-variance posterior Gaussian distribution. Typically, the linear-Gaussian observations are subject to error covariance that is a function of, e.g., viewing position, angle of incidence, and the sensing modality. Several such models have been proposed in the robotics and sensing literature, cf. [1]–[5], but regardless of the specific model used for the measurement covariance, the IF combines the data among the nodes in the network in an additive way, so that it is readily decentralized; see, e.g., Information Consensus Filtering (ICF) [6].

The dependence of the measurement model on the sensor state has been widely used to obtain planning algorithms that, given a history of measurements, determine the next best set of observations of a hidden state [1]–[5]. The objective is that this new set of observations optimizes an information-theoretic criterion of interest. In this paper, we choose to maximize the minimum eigenvalue of the posterior information matrix, which minimizes the directional variance of every hidden state, as opposed to, e.g., the trace [1], [2], determinant [4], or a coverage-related objective [5]. This allows us to set a threshold for accuracy and control the sensor network to drive the error below that threshold. We express this requirement by reformulating the problem as a Linear Matrix Inequality (LMI) constrained optimization problem. Specifically, we introduce as many LMI constraints as the number of hidden states, that are coupled with respect to the sensors via an ICF that estimates the uncertainty of each hidden state using the sensor measurements. To solve this problem, we propose a new distributed optimization algorithm, which we call *random approximate projections*, that is robust to the state disagreement errors that exist among the robots as the ICF fuses the collected measurements. Our method, that extends our previous work in [7] and [8], requires only weak connectivity of the communication graph, is computationally light, and can handle large systems of coupled LMIs under disagreement errors, for which no distributed methods currently exist in the literature.

Relevant work in the area of distributed sensor planning and estimation include SLAM [9], localization [3], [10], coverage [5], mobile target tracking [1], [2], [4], [11], and classification [12]. In the general case, sampling is needed in order to reason over possible posteriors [10]–[12]. If self localization uncertainty is also taken into account, a decentralized Partially Observable Markov Decision Process (dec-POMDP) can be formulated for sensor planning [13]. Dec-POMDPs typically require a centralized coordinator and do not scale well with the number of targets and agents. In this paper, and in much of the relevant literature [1]–[4], it is assumed that the posteriors for the hidden states are Gaussian, and that the sensor is self localized. To our knowledge, the ability to control worst-case estimation uncertainty for large numbers of states in a distributed fashion is new.

The paper is organized as follows. Section II formulates the problem. Then, Section III develops the proposed distributed method to maximize the minimum eigenvalue of the information matrix. Section IV states assumptions and main convergence results and shows that the proposed algorithm converges to the optimal point. Section V provides simulations of a landmark localization and mapping problem.

II. Problem Formulation

Consider the problem of estimating m stationary hidden states $\{\mathbf{x}_i \in \mathbb{R}^p\}_{i \in \mathcal{I}}$ using noisy observations from n mobile sensors, where p is the dimension of the hidden states and $\mathcal{I} = \{1, \ldots, m\}$. Denote the locations of the mobile sensors at time t by $\{\mathbf{r}_s(t) \in \mathbb{R}^q\}_{s \in \mathcal{S}}$, where q is the dimension of the sensors' configuration space and $\mathcal{S} = \{1, \ldots, n\}$. The

observation of state i by sensor s at time t is given by

$$\mathbf{y}_{i,s}(t) = \mathbf{x}_i + \boldsymbol{\zeta}_{i,s}(t). \tag{1}$$

We assume that the measurement errors are Normally distributed so that $\boldsymbol{\zeta}_{i,s}(t) \sim N(\mathbf{0}, Q(\mathbf{r}_s(t), \mathbf{x}_i))$, where $Q \colon \mathbb{R}^q \times \mathbb{R}^p \to \mathrm{Sym}_+(p, \mathbb{R})$ denotes the measurement precision matrix, or information matrix (not the covariance), and $\mathrm{Sym}_+(p, \mathbb{R})$ is the set of $p \times p$ symmetric positive semidefinite real matrices. We also assume that if signal $\mathbf{x}_i$ is out of range of sensor s, then the function Q returns the information matrix $\mathbf{0}_{p \times p}$, corresponding to infinite variance. Denote the set of all observations at time t by $\mathcal{O}(t) \triangleq \{(\mathbf{y}_{i,s}(t), Q(\mathbf{r}_s(t), \mathbf{x}_i))\}_{i \in \mathcal{I}, s \in \mathcal{S}}$.

Moreover, let $\hat{\mathbf{x}}_i(t)$ denote the estimate of $\mathbf{x}_i$ at time t and $S_i(t)$ denote the information matrix corresponding to the estimate $\hat{\mathbf{x}}_i(t)$. We assume that initial values for the pair $(\hat{\mathbf{x}}_i(0), S_i(0))$ are available *a priori*. Let also $\hat{\mathbf{x}}(t) \in \mathbb{R}^{pm}$ denote the stack vector of all estimates $\hat{\mathbf{x}}_i(t)$ and $S(t) \in \mathrm{Sym}_+(p, \mathbb{R}) \times \cdots \times \mathrm{Sym}_+(p, \mathbb{R}) \triangleq (\mathrm{Sym}_+(p, \mathbb{R}))^m$ denote the collection of all information matrices $S_i(t)$ at time t. Then, the global data at time t can be captured by the set $\mathcal{D}(t) \triangleq (\hat{\mathbf{x}}(t), S(t))$.

Given prior data $\mathcal{D}(t-1)$ known by all sensors, and observations $\mathcal{O}(t-1)$, the current set of data $\mathcal{D}(t)$ can be determined in a distributed way using an Information Consensus Filter (ICF) [6] as[1]

$$(\mathcal{D}(t-1), \mathcal{O}(t-1)) \overset{\text{ICF}}{\longmapsto} \mathcal{D}(t). \tag{2}$$

Then, the goal is to control the next set of observations $\mathcal{O}(t)$ in order to reduce the future uncertainty in $\mathcal{D}(t+1)$. We can achieve this goal by controlling the mobile sensors and, therefore, the expected information $\{Q(\mathbf{r}_s(t), \hat{\mathbf{x}}_i(t))\}_{i \in \mathcal{I}, s \in \mathcal{S}}$ associated with the observations $\{\mathbf{y}_{i,s}(t)\}_{i \in \mathcal{I}, s \in \mathcal{S}}$. Specifically, let $\mathbf{r}_s(t) = \mathbf{r}_s(t-1) + \mathbf{u}_s$ denote the motion model of sensor s, where $\mathbf{u}_s \in \mathbb{R}^q$ is a candidate control input used to drive sensor s to a new position $\mathbf{r}_s(t)$. Then, our goal is to maximize the minimum eigenvalues of the information matrices in $\mathcal{D}(t+1)$ that, for each hidden state i, are given by

$$S_i(t+1) = \overbrace{S_i(t)}^{\text{from } \mathcal{D}(t)} + \underbrace{\sum_{s \in \mathcal{S}} Q(\mathbf{r}_s(t-1) + \mathbf{u}_s, \overbrace{\hat{\mathbf{x}}_i(t)}^{\text{from } \mathcal{D}(t)})}_{\text{from } \mathcal{O}(t)}. \tag{3}$$

We assume that the set of feasible controls $\mathcal{U}_s \ni \mathbf{0}$ is compact and has diameter less than δ for some $\delta > 0$ and that $\mathbf{u}_s \in \mathcal{U}_s$. Define $\mathcal{U} \triangleq \prod_{s \in \mathcal{S}} \mathcal{U}_s$ and let $\mathbf{u} \in \mathbb{R}^{qn}$ denote the stack of all controllers $\mathbf{u}_s$. Additionally, define the set $\Gamma \triangleq \prod_{i \in \mathcal{I}} [0, \tau_i]$, where τ_i represent the given user-specified estimation thresholds for each hidden state i, and define $\boldsymbol{\gamma} \triangleq (\gamma_1, \ldots, \gamma_m)^\top \in \mathbb{R}^m$. With the new notation, the problem we consider in this paper can be defined as

$$\max_{(\boldsymbol{\gamma}, \mathbf{u}) \in \Gamma \times \mathcal{U}} \sum_{i \in \mathcal{I}} \gamma_i \tag{4a}$$

$$\text{s.t} \quad \gamma_i I \preceq S_i + \sum_{s \in \mathcal{S}} Q(\mathbf{r}_s + \mathbf{u}_s, \hat{\mathbf{x}}_i) \ \forall i \in \mathcal{I}, \tag{4b}$$

[1]Explicit details on the mapping (2) are provided in Section III.

where we have dropped dependence on the time t for notational convenience. For the explicit references to time in (4b) see (3). The symbol $\preceq$ denotes an ordering on the negative semidefinite cone, i.e., $A \preceq B$ if and only if $A - B$ is negative semidefinite.

III. Concurrent Estimation and Sensor Planning

To simplify the nonlinear semidefinite program in problem (4), we use the first order approximation to $Q(\cdot, \hat{\mathbf{x}}_i)$ about the input $\mathbf{u}_s = 0$. Specifically, we define

$$h \colon \mathbb{R}^{m+qn} \times \left(\mathbb{R}^{pm} \times \left(\mathrm{Sym}_+(p, \mathbb{R}) \right)^m \right) \times \mathcal{I} \to \mathrm{Sym}(p, \mathbb{R}),$$

where $\mathrm{Sym}(p, \mathbb{R})$ denotes the set of symmetric $p \times p$ dimensional matrices, by

$$h\left((\boldsymbol{\gamma}, \mathbf{u}); \mathcal{D}, i \right) \triangleq \gamma_i I - S_i \tag{5}$$
$$- \sum_{s \in \mathcal{S}} \left[Q(\mathbf{r}_s, \hat{\mathbf{x}}_i) + \sum_{j=1}^q \frac{\partial Q}{\partial [\mathbf{r}]_j} (\mathbf{r}_s, \hat{\mathbf{x}}_i) [\mathbf{u}_s]_j \right].$$

The matrix-valued function h is negative semidefinite at a point $((\boldsymbol{\gamma}, \mathbf{u}); \mathcal{D}, i)$ if and only if the linearized version of the constraint (4b) is satisfied for γ_i using the data $\mathcal{D} = (\hat{\mathbf{x}}, S)$.

To simplify notation, collect all decision variables in the vector $\mathbf{z} = (\boldsymbol{\gamma}, \mathbf{u}) \in \mathbb{R}^{m+qn}$, and to decompose the linearized version of (4) over the set of sensors, let $\mathbf{z}_s$ denote a copy of $\mathbf{z}$ that is local to sensor $s \in \mathcal{S}$. Additionally, define the local functions $f_s(\mathbf{z}_s) = -\sum_{i \in \mathcal{I}} \gamma_{i,s}$ and the global objective $f(\mathbf{z}_1, \ldots, \mathbf{z}_n) = \sum_{s \in \mathcal{S}} f_s(\mathbf{z}_s)$. Then, problem (4) can be expressed in a distributed way as

$$\min_{\{\mathbf{z}_s\}_{s \in \mathcal{S}} \subset \mathcal{X}_0} \sum_{s \in \mathcal{S}} f_s(\mathbf{z}_s) \tag{6a}$$

$$\text{s.t} \quad h(\mathbf{z}_s; \mathcal{D}, i) \preceq \mathbf{0} \quad \forall i \in \mathcal{I}, \forall s \in \mathcal{S}, \tag{6b}$$

where $\mathcal{X}_0 = \Gamma \times \mathcal{U}$. In problem (6), the objective function (6a) is separable among the sensors and the constraints (6b) are linear in $\mathbf{z}_s$ and local to every sensor but coupled by $\mathcal{D}$. In this form, problem (6) can be decomposed and solved in a distributed way. A necessary requirement is that, at the solution, the local variables $\mathbf{z}_s$ need to agree for all sensors s, i.e., $\mathbf{z}_s = \mathbf{z}_v$ for all $s, v \in \mathcal{S}$.

The remainder of this section discusses an iterative approach to solving problem (6) that takes place during the time window $[t-1, t)$ to determine what the control input $\mathbf{u}$ will be at time t. Let $\mathcal{D}_{s,k}$ denote the data and $\mathbf{z}_{s,k}$ denote the decision variables of the sensor s at iteration k of this decision process. The first step is to broadcast $\mathbf{z}_{s,k-1}$ and a summary of $\mathcal{D}_{s,k-1}$. Following the ICF [6], this summary is captured by the variables $\boldsymbol{\Xi}_{s,k}$ and $\boldsymbol{\xi}_{s,k}$, defined in lines 3 and 4 of Algorithm 1. It is shown in [6] that for all $s \in \mathcal{S}$, $\left\{ \left([n\boldsymbol{\Xi}_{s,k}]^{-1} \boldsymbol{\xi}_{s,k}, [n\boldsymbol{\Xi}_{s,k}]^{-1} \right) \right\}_{k \in \mathbb{N}} \to \mathcal{D} \triangleq \mathcal{D}(t)$, where $\mathbb{N} = \{1, 2, 3, \ldots\}$ denotes the natural numbers.

Let $\mathcal{G}_k = (\mathcal{S}, \mathcal{E}_k)$ represent a communication graph such that $(s, v) \in \mathcal{E}_k$ if and only if sensors s and v can communicate at iteration k. This defines the *neighbor set*, $\mathcal{N}_{s,k} \triangleq \{s\} \cup \{v \in \mathcal{S} \mid (v, s) \in \mathcal{E}_k\}$. Sensor s receives $\{\mathbf{z}_{v,k-1}, \boldsymbol{\xi}_{v,k}, \boldsymbol{\Xi}_{v,k} \mid v \in \mathcal{N}_{s,k}\}$. With this new information, the sensor updates its own optimization variable and data summary via using weighted averaging. In particular, let W_k

Algorithm 1 Concurrent Filtering and Optimization for $s \in \mathcal{S}$ during the interval $[t-1, t]$

Require: $\mathcal{D}_s(t-1)$, $\{\mathbf{y}_{i,s}(t-1)\}_{i \in \mathcal{I}}$, feasible $\mathbf{z}_{s,0}$
1: $(\hat{\mathbf{x}}, S) \leftarrow \mathcal{D}_s(t-1)$
2: $\mathbf{y} \leftarrow (\mathbf{y}_{1,s}(t-1), \dots, \mathbf{y}_{m,s}(t-1))^\top$
3: $\Xi_{s,0} = \frac{1}{n} S + \text{blkdiag}(Q(\mathbf{r}_s, \hat{\mathbf{x}}_1), \dots, Q(\mathbf{r}_s, \hat{\mathbf{x}}_m))$
4: $\boldsymbol{\xi}_{s,0} = \frac{1}{n} S + \text{blkdiag}(Q(\mathbf{r}_s, \hat{\mathbf{x}}_1), \dots, Q(\mathbf{r}_s, \hat{\mathbf{x}}_m))\mathbf{y}$
5: **for** $k = 1, 2, \dots$ **do**
6:     Broadcast $\mathbf{z}_{s,k-1}, \Xi_{s,k-1}$, and $\boldsymbol{\xi}_{s,k-1}$ to $\mathcal{N}_{s,k}$ and collect incoming data according to

$$(\mathbf{p}_{s,k}, \Xi_{s,k}, \boldsymbol{\xi}_{s,k}) \leftarrow \sum_{j \in \mathcal{N}_{s,k}} [W_k]_{sj} (\mathbf{p}_{j,k-1}, \Xi_{j,k}, \boldsymbol{\xi}_{j,k})$$

7:     $\mathcal{D}_{s,k} = \left((n\Xi_{s,k})^{-1} \boldsymbol{\xi}_{s,k}, n\Xi_{s,k}\right)$
8:     $\mathbf{v}_{s,k} = \Pi_{\mathcal{X}_0}(\mathbf{p}_{s,k} - \alpha_k f'_s(\mathbf{p}_{s,k}))$
9:     Choose $\omega_{s,k} \in \mathcal{I}$, compute $h_+(\mathbf{v}_{s,k}; \mathcal{D}_{s,k}, \omega_{s,k})$ and $h'_+(\mathbf{v}_{s,k}; \mathcal{D}_{s,k}, \omega_{s,k})$
10:     $\beta_{s,k} = h_+(\mathbf{v}_{s,k}; \mathcal{D}_{s,k}, \omega_{s,k}) / \|h'_+(\mathbf{v}_{s,k}; \mathcal{D}_{s,k}, \omega_{s,k})\|^2$
11:     $\mathbf{z}_{s,k} = \Pi_{\mathcal{X}_0}(\mathbf{v}_{s,k} - \beta_{s,k} h'_+(\mathbf{v}_{s,k}; \mathcal{D}_{s,k}, \omega_{s,k}))$
12: **end for**
13: **return** $\mathcal{D}(t) \leftarrow \mathcal{D}_{s,k}$ and $(\mathbf{u}, \boldsymbol{\gamma}) \leftarrow \mathbf{z}_{s,k}$

---

be an $n \times n$ row stochastic matrix such that $[W_k]_{sv} > 0 \iff v \in \mathcal{N}_s$. The exact assumptions for W_k will be introduced later, when they are necessary for the proof of our main result. The agent can compute the updated data $\mathcal{D}_{s,k}$ as shown in line 7, concluding the ICF portion of iteration k. Note that, given initial conditions $\{\mathcal{D}_{s,0}\}_{s \in \mathcal{S}}$, the data for any sensor s at time k is bounded by the initial conditions as $\|\mathcal{D}_{s,k}\| \leq \max_{v \in \mathcal{S}} \|(\mathcal{D}_{v,0})\|$. Moreover, all sensors' data converge to a consensus, i.e., $\{\|\mathcal{D}_{s,k} - \mathcal{D}\|\}_{k \in \mathbb{N}} \to 0$ for all $s \in \mathcal{S}$; cf. [6] for an in-depth discussion of ICF.

The *random approximate projection* part of the algorithm begins in line 8 by computing a negative subgradient of the local objective function f_s, which we denote by $-f'_s(\mathbf{z}_{s,k})$, taking a gradient descent step, and projecting the iterate back to the simple constraint set $\mathcal{X}_0$. The positive step size α_k is diminishing, non-summable and square-summable. The last step of our method is a subgradient step in the direction that minimizes the violation of a *single* LMI constraint. Specifically, line 9 randomly assigns $\omega_{s,k} \in \mathcal{I}$, effectively choosing to activate $h(\cdot; \cdot, \omega_{s,k})$. The idea behind line 11 is to minimize a scalar metric that measures the amount of violation of $h(\cdot; \cdot, \omega_{s,k})$, which we denote by $h_+(\cdot; \cdot, \omega_{s,k})$, while remaining feasible. Let $\Pi_+ : \text{Sym}(p, \mathbb{R}) \to \text{Sym}_+(p, \mathbb{R})$ denote the projection operator onto the set of positive semidefinite matrices, given by $\Pi_+ X = E \text{diag}[(\lambda_1)_+, \dots, (\lambda_p)_+] E^\top$, where $(\cdot)_+ \triangleq \max\{\cdot, 0\}$ and $\lambda_1, \dots, \lambda_p$ are the eigenvalues and E is a matrix with columns $\mathbf{e}_1, \dots, \mathbf{e}_p$ that are the corresponding eigenvectors of X. A measure of the "positive definiteness" of X is thus $\|\Pi_+ X\|_F = \sqrt{(\lambda_1)_+^2 + \dots + (\lambda_p)_+^2}$. This defines our measure of constraint violation,

$$h_+(\mathbf{z}; \mathcal{D}, i) \triangleq \|\Pi_+ h(\mathbf{z}; \mathcal{D}, i)\|_F. \tag{7}$$

The step size of the gradient step in line 10 is

$$\beta_{s,k} = \frac{h_+(\mathbf{v}_{s,k}; \mathcal{D}_{s,k}, \omega_{s,k})}{\|h'_+(\mathbf{v}_{s,k}; \mathcal{D}_{s,k}, \omega_{s,k})\|^2}, \tag{8}$$

where h'_+ denotes the gradient of h_+, is a variant of the Polyak step size [14]. Note that this is a well defined step size as we let $h'_+(\mathbf{v}_{s,k}; \mathcal{D}_{s,k}, \omega_{s,k}) = d\mathbf{1}_{m+qn}$ whenever $h_+(\mathbf{v}_{s,k}; \mathcal{D}_{s,k}, \omega_{s,k}) = 0$, and h'_+ is nonzero elsewhere.

IV. MAIN RESULTS

We begin by introducing the following notation regarding the optimal value $f^* = \min_{\mathcal{X}} f(\mathbf{z})$ and optimal set $\mathcal{X}^* = \{\mathbf{z} \in \mathcal{X} \mid f(\mathbf{z}) = f^*\}$ of (6), where $\mathcal{X}$ denotes the feasible set, i.e., $\mathcal{X} = \mathcal{X}_0 \cap \{\mathbf{z} \mid h(\mathbf{z}; \mathcal{D}, i) \preceq 0, \ \forall i \in \mathcal{I}\}$. Note that $\mathbf{0} \in \Gamma \times \mathcal{U}$ and $h(\mathbf{0}; \mathcal{D}, i) \preceq 0$ for all possible $\mathcal{D}$ and i, thus there is always a feasible point. Denote the disagreement error in the constraint violation by

$$\nu_{s,k} = h_+(\mathbf{z}_{s,k}; \mathcal{D}_{s,k}, \omega_{s,k}) - h_+(\mathbf{z}_{s,k}; \mathcal{D}, \omega_{s,k}) \tag{9}$$

and the disagreement error in the gradient of the constraint violation by

$$\boldsymbol{\delta}_{s,k} = h'_+(\mathbf{z}_{s,k}; \mathcal{D}_{s,k}, \omega_{s,k}) - h'_+(\mathbf{z}_{s,k}; \mathcal{D}, \omega_{s,k}). \tag{10}$$

The norms of $\nu_{s,k}$ and $\boldsymbol{\delta}_{s,k}$ are bounded, i.e., for some scalars $D, N > 0$, for all $k \in \mathbb{N}$ and $s \in \mathcal{S}$, it holds *a.s.* that $|\nu_{s,k}| \leq N$, and $\|\boldsymbol{\delta}_{s,k}\| \leq D$. Due to space limitations, we omit a direct calculation of D and N.

The set $\mathcal{X}_0 = \Gamma \times \mathcal{U}$ is convex and compact. Therefore, there exists a constant $C_\mathbf{z} > 0$ such that for any $\mathbf{z}_1, \mathbf{z}_2 \in \mathcal{X}_0$ it holds that $\|\mathbf{z}_1 - \mathbf{z}_2\| \leq C_\mathbf{z}$. Also, as defined herein, the functions $f_s(\cdot)$ and $h(\mathbf{z}; \mathcal{D}, i)$ for $i \in \mathcal{I}$ are convex (not necessarily differentiable) over some open set that contains $\mathcal{X}_0$. A direct consequence of this and the compactness of $\mathcal{X}_0$ is that the subdifferentials $\partial f_s(\mathbf{z})$ and $\partial h_+(\mathbf{z}; \mathcal{D}, i)$ are nonempty over $\mathcal{X}_0$. It also implies that the subgradients $f'_s(\mathbf{z}) \in \partial f_s(\mathbf{z})$ and $h'_+(\mathbf{z}; \mathcal{D}, i) \in \partial h_+(\mathbf{z}; \mathcal{D}, i)$ are uniformly bounded over the set $\mathcal{X}_0$. That is, there is a scalar L_f such that for all $f'_s(\mathbf{z}) \in \partial f_s(\mathbf{z})$ and $\mathbf{z} \in \mathcal{X}_0$, it holds that $\|f'_s(\mathbf{z})\| \leq L_f$, and for any $\mathbf{z}_1, \mathbf{z}_2 \in \mathcal{X}_0$, it holds that $|f_s(\mathbf{z}_1) - f_s(\mathbf{z}_2)| \leq L_f \|\mathbf{z}_1 - \mathbf{z}_2\|$. Note that for the objective function f_s in (6), we have $L_f = 1$. Also, there is a scalar L_h such that for all $h'_+(\mathbf{z}; \mathcal{D}, i) \in \partial h_+(\mathbf{z}; \mathcal{D}, i)$, $\mathbf{z} \in \mathcal{X}_0$, $i \in \mathcal{I}$, it holds that $\|h'_+(\mathbf{z}; \mathcal{D}, i)\| \leq L_h$.

From relation (10) we have that for any $i \in \mathcal{I}$, $s \in \mathcal{S}$, $k \in \mathbb{N}$, and $\mathbf{z} \in \mathcal{X}_0$, it holds that

$$\|h'_+(\mathbf{z}; \mathcal{D}_{s,k}, i)\| = \|h'_+(\mathbf{z}; \mathcal{D}, i) + \boldsymbol{\delta}_{s,k}\| \leq L_h + D,$$

and for any $\mathbf{z}_1, \mathbf{z}_2 \in \mathcal{X}_0$ that

$$|h_+(\mathbf{z}_1; \mathcal{D}_{s,k}, i) - h_+(\mathbf{z}_2; \mathcal{D}_{s,k}, i)| \leq (L_h + D) \|\mathbf{z}_1 - \mathbf{z}_2\|.$$

The compactness of $\mathcal{X}_0$, the boundedness of the data sequences $\mathcal{D}_{s,k}$, and the continuity of the functions $h_+(\mathbf{z}; \mathcal{D}, i)$ for $i \in \mathcal{I}$ imply that there exist a constant $C_h > 0$ such that for any $s \in \mathcal{S}$, $k \in \mathbb{N}$, $i \in \mathcal{I}$ and $\mathbf{z} \in \mathcal{X}_0$, it holds that $|h_+(\mathbf{z}; \mathcal{D}, i)| \leq C_h$ and

$$|h_+(\mathbf{z}; \mathcal{D}_{s,k}, i)| = |h_+(\mathbf{z}; \mathcal{D}, i) + \nu_{s,k}| \leq C_h + N, \tag{11}$$

which follows from (9). Furthermore, when $h_+(\mathbf{z}; \mathcal{D}_{s,k}, i) \neq 0$, we have $h'_+(\mathbf{z}; \mathcal{D}_{s,k}, i) \neq \mathbf{0}$. Therefore, there exists a constant $c_h > 0$ such that $\|h'_+(\mathbf{z}; \mathcal{D}_{s,k}, i)\| \geq c_h$ for all $\mathbf{z} \in \mathcal{X}_0$, $s \in \mathcal{S}$, $k \in \mathbb{N}$, and $i \in \mathcal{I}$. This and relation (11) imply that

$$\beta_{s,k} = \frac{h_+(\mathbf{z}; \mathcal{D}_{s,k}, \omega_{s,k})}{\|h'_+(\mathbf{z}; \mathcal{D}_{s,k}, \omega_{s,k})\|^2} \leq \frac{C_h + N}{c_h^2}. \quad (12)$$

Note that when $h_+(\mathbf{z}; \mathcal{D}_{s,k}, \omega_{s,k}) = 0$, the above bound holds trivially. Next, for our algorithm to converge, we require the following assumptions to hold.

Assumption 4.1: The information function Q (a) is bounded, (b) is twice subdifferentiable (c) has bounded subdifferentials up to the second order, and (d) has relatively few critical points, i.e., the sets of critical points of Q and its partial derivatives up to the second order are measure zero.

Assumption 4.2: In the k-th iteration of the inner-loop, each element i of $\mathcal{I}$ is generated with nonzero probability, i.e., for any $s \in \mathcal{S}$ and $k \in \mathbb{N}$, the probability that i is selected $\pi_i \triangleq \Pr\{\omega_{s,k} = i\} > 0$.

Assumption 4.3: For all $s \in \mathcal{S}$ and $k \in \mathbb{N}$, there exists a constant $\kappa > 0$ such that for all $\mathbf{z} \in \mathcal{X}_0$, it holds that

$$\text{dist}^2(\mathbf{z}, \mathcal{X}) \leq \kappa \mathsf{E}\left[h_+^2(\mathbf{z}; \mathcal{D}_{s,k}, \omega_{s,k})\right].$$

Note that this assumption is known to hold true for a class of functions including the definition of h_+ in (7); cf. [15].

Assumption 4.4: For all $k \in \mathbb{N}$, it holds for the weighted graphs $\mathcal{G}_k = (\mathcal{S}, \mathcal{E}_k, W_k)$ that (a) there exists a scalar $\eta \in (0, 1)$ such that $[W_k]_{sj} \geq \eta$ if $j \in \mathcal{N}_{s,k}$. Otherwise, $[W_k]_{sj} = 0$, (b) the weight matrix W_k is doubly stochastic, i.e., $\sum_{s \in \mathcal{S}} [W_k]_{sj} = 1$ for all $j \in \mathcal{S}$ and $\sum_{j \in \mathcal{S}} [W_k]_{sj} = 1$ for all $s \in \mathcal{S}$, and (c) there exists a scalar $B > 0$ such that the graph $(\mathcal{S}, \cup_{\ell=0,\dots,B-1} \mathcal{E}_{t+\ell})$ is strongly connected for any $t \in \mathbb{N}$.

Theorem 4.5 (Convergence Under Disagreement Errors): Let Assumptions 4.1 - 4.4 hold. Assume that $\{\alpha_k\}_{k \in \mathbb{N}}$ is square summable but not absolutely summable. Then, there exists $\mathbf{z}^* \in \mathcal{X}^*$ such that $\{\mathbf{z}_{s,k}\}_{k \in \mathbb{N}} \to \mathbf{z}^*$ *a.s.* for all $s \in \mathcal{S}$.

To prove Theorem 4.5, we need some intermediate results.

Lemma 4.6 ([16]): Let $\mathcal{X} \subseteq \mathbb{R}^d$ be a nonempty closed convex set. The function $\Pi_{\mathcal{X}} : \mathbb{R}^d \to \mathcal{X}$ is nonexpansive, i.e., $\|\Pi_{\mathcal{X}}[\mathbf{a}] - \Pi_{\mathcal{X}}[\mathbf{b}]\| \leq \|\mathbf{a} - \mathbf{b}\|$ for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$.

Lemma 4.7 (Lemma 10-11, p. 49-50 in [17]): Let $\mathcal{F}_k \triangleq \{v_\ell, u_\ell, a_\ell, b_\ell\}_{\ell=0}^k$ denote a collection of nonnegative real random variables for $k \in \mathbb{N} \cup \{\infty\}$ such that $\mathsf{E}[v_{k+1}|\mathcal{F}_k] \leq (1+a_k)v_k - u_k + b_k$ for all $k \in \{0\} \cup \mathbb{N}$ *a.s.* Assume further that $\{a_k\}$ and $\{b_k\}$ are *a.s.* summable. Then, we have *a.s.* that (i) $\{u_k\}$ is summable and (ii) there exists a nonnegative random variable v such that $\{v_k\} \to v$.

The next lemma relates $\mathbf{p}_{s,k}$ and $\mathbf{z}_{s,k}$ in Line 6 and 11 of Algorithm 1. The proof is omitted due to space limitations.

Lemma 4.8: For any convex function $g : \mathbb{R}^n \to \mathbb{R}$, we have $\sum_{s \in \mathcal{S}} g(\mathbf{p}_{s,k}) \leq \sum_{s \in \mathcal{S}} g(\mathbf{z}_{s,k-1})$.

Lastly, we state a result that ensures successful consensus in the presence of a well-behaved disturbance sequence.

Lemma 4.9 ([18]): Let Assumption 4.4 hold and define $\mathbf{e}_{s,k} \triangleq \mathbf{z}_{s,k} - \mathbf{p}_{s,k}$. Consider the iterates generated by

$\boldsymbol{\theta}_{s,k} = \sum_{v \in \mathcal{S}} [W_k]_{sv} \boldsymbol{\theta}_{v,k-1} + \mathbf{e}_{s,k}$. Suppose there exists a nonnegative nonincreasing scalar sequence $\{\alpha_k\}$ such that $\{\alpha_k \|\mathbf{e}_{s,k}\|\}_{k \in \mathbb{N}}$ is summable for all $s \in \mathcal{S}$. Then, for all $s, v \in \mathcal{S}$, $\{\alpha_k \|\boldsymbol{\theta}_{s,k} - \boldsymbol{\theta}_{v,k}\|\}_{k \in \mathbb{N}}$ is summable.

In addition to the well-known results of Lemmas 4.6-4.9, we need the following intermediate results. The proofs are omitted due to space limitations.

Lemma 4.10: For *a.e.* bounded sequence $\{\mathbf{z}_{s,k}\}_{k \in \mathbb{N}}$, the resulting sequence $\{\nu_{s,k}\}_{k \in \mathbb{N}}$, defined in (9), is *a.s.* absolutely summable.

Lemma 4.11: For *a.e.* bounded sequence $\{\mathbf{z}_{s,k}\}_{k \in \mathbb{N}}$, the resulting sequence $\{\boldsymbol{\delta}_{s,k}\}_{k \in \mathbb{N}}$, defined in (10), is *a.s.* absolutely summable.

Lemma 4.12 (Basic Iterate Relation): Consider the sequences $\{\mathbf{z}_{s,k}\}_{k=0}^\infty$ and $\{\mathbf{p}_{s,k}\}_{k=0}^\infty$ for $s \in \mathcal{S}$ generated by Algorithm 1. Then, for any $\mathbf{z}, \check{\mathbf{z}} \in \mathcal{X}$, $s \in \mathcal{S}$ and $k \in \mathbb{N}$, we have *a.s.* that

$$\|\mathbf{z}_{s,k} - \mathbf{z}\|^2 \leq \|\mathbf{p}_{s,k} - \mathbf{z}\|^2 - 2\alpha_k(f_s(\check{\mathbf{z}}) - f_s(\mathbf{z}))$$
$$+ 2\frac{C_\mathbf{z}(C_h + N)}{c_h^2}\|\boldsymbol{\delta}_{s,k}\| + 2\frac{C_h + N}{c_h^2}|\nu_{s,k}| + \frac{1}{4\eta}\|\mathbf{p}_{s,k} - \check{\mathbf{z}}\|^2$$
$$- \frac{\tau - 1}{\tau(L_h + D)^2} h_+^2(\mathbf{p}_{s,k}; \mathcal{D}_{s,k}, \omega_{s,k}) + A_{\eta,\tau}\alpha_k^2 L_f^2,$$

where $A_{\eta,\tau} = 1 + 4\eta + \tau$ and $\eta, \tau > 0$ are arbitrary.

Note that, because we only minimize the scalar violation of a single constraint at each $k \in \mathbb{N}$ in Algorithm 1, we cannot guarantee the feasibility of the full sequences $\mathbf{p}_{s,k}$ and $\mathbf{z}_{s,k}$. The next lemma states that each $\mathbf{p}_{s,k}$ for all $s \in \mathcal{S}$ asymptotically achieves feasibility.

Lemma 4.13: Let Assumption 4.2 and 4.3 hold. Assume α_k is square summable. Consider $\{\mathbf{p}_{s,k}\}_{k=0}^\infty$ for $s \in \mathcal{S}$, generated by Algorithm 1. Then, for any $s \in \mathcal{S}$, we have *a.s.* that $\{\text{dist}(\mathbf{p}_{s,k}, \mathcal{X})\}_{k \in \mathbb{N}}$ is square summable.

Lemma 4.14: Let Assumptions 4.2-4.4 hold. Assume $\{\alpha_k\}$ is nonnegative and square summable. Define $\mathbf{e}_{s,k} \triangleq \mathbf{z}_{s,k} - \mathbf{p}_{s,k}$. Then, we have *a.s.* that (i) $\{\mathbf{e}_{s,k}\}_{k \in \mathbb{N}}$ is absolutely square summable for all $s \in \mathcal{S}$ and (ii) $\{\alpha_k \|\check{\mathbf{z}}_{s,k} - \bar{\check{\mathbf{z}}}_k\|\}_{k \in \mathbb{N}}$ is summable for all $s \in \mathcal{S}$, where $\check{\mathbf{z}}_{s,k} \triangleq \Pi_{\mathcal{X}}[\mathbf{p}_{s,k}]$ and $\bar{\check{\mathbf{z}}}_k \triangleq \frac{1}{n}\sum_{s \in \mathcal{S}} \check{\mathbf{z}}_{s,k}$.

Note that the sequences $\{\mathbf{z}_{s,k}\}_{k \in \mathbb{N}}$ for $s \in \mathcal{S}$ generated by Algorithm 1 can be represented as in Lemma 4.9. That is, $\mathbf{z}_{s,k} = \sum_{s \in \mathcal{S}} [W_k]_{sj} \mathbf{z}_{j,k-1} + \mathbf{e}_{s,k}$, $\forall s \in \mathcal{S}$, where $\mathbf{e}_{s,k}$ is as defined in Lemma 4.14. Therefore, from Lemma 4.14(a), we have that $\sum_{k \in \mathbb{N}} \alpha_k \|\mathbf{e}_{s,k}\| = \sum_{k \in \mathbb{N}} \frac{1}{2}\left(\alpha_k^2 + \|\mathbf{e}_{s,k}\|^2\right) < \infty$. Invoking Lemma 4.9, consensus among $\mathbf{z}_{s,k}$ for $s \in \mathcal{S}$ follows.

Proof: [Theorem 4.5] We invoke Lemma 4.12 with $\check{\mathbf{z}} = \check{\mathbf{z}}_{s,k}$ (Note that $\check{\mathbf{z}}_{s,k}$ is defined as $\Pi_{\mathcal{X}}[\mathbf{p}_{s,k}]$ in Lemma 4.14), $\tau = 4$ and $\eta = \kappa(L_h + D)^2$. We also let $\mathbf{z} = \mathbf{z}^*$ for an arbitrary $\mathbf{z}^* \in \mathcal{X}^*$. Therefore, for any $\mathbf{z}^* \in \mathcal{X}^*$, $s \in \mathcal{S}$ and $k \in \mathbb{N}$, we have *a.s.* that

$$\|\mathbf{z}_{s,k} - \mathbf{z}^*\|^2 \leq \|\mathbf{p}_{s,k} - \mathbf{z}^*\|^2 - 2\alpha_k(f_s(\check{\mathbf{z}}_{s,k}) - f_s(\mathbf{z}^*))$$
$$+ 2\frac{C_\mathbf{z}(C_h + N)}{c_h^2}\|\boldsymbol{\delta}_{s,k}\| + 2\frac{C_h + N}{c_h^2}|\nu_{s,k}|$$
$$+ \frac{1}{4\kappa(L_h + D)^2}\text{dist}^2(\mathbf{p}_{s,k}, \mathcal{X}) + A\alpha_k^2 L_f^2$$

$$-3h_+^2(\mathbf{p}_{s,k};\mathcal{D}_{s,k},\omega_{s,k})/(4(L_h+D)^2), \qquad (13)$$

where $A = 5 + 4\kappa(L_h+D)^2$. Denote by $\mathcal{F}_k$ the sigma-field induced by the history of the algorithm up to time k, i.e., $\mathcal{F}_k = \{\mathbf{z}_{s,0}, \ (\omega_{s,\ell}, \ 1 \le \ell \le k), \ s \in \mathcal{S}\}$ for $k \in \mathbb{N}$, and $\mathcal{F}_0 = \{\mathbf{z}_{s,0}, \ s \in \mathcal{S}\}$. From Assumption 4.3, we know that

$$\mathrm{dist}^2(\mathbf{p}_{s,k},\mathcal{X}) \le \kappa \mathsf{E}\left[h_+^2(\mathbf{p}_{s,k};\mathcal{D}_{s,k},\omega_{s,k}) \mid \mathcal{F}_{k-1}\right].$$

Taking the expectation conditioned on $\mathcal{F}_{k-1}$ in relation (13), summing this over $s \in \mathcal{S}$, and using the above relation, we obtain

$$\sum_{s\in\mathcal{S}} \mathsf{E}\left[\|\mathbf{z}_{s,k}-\mathbf{z}^*\|^2 \mid \mathcal{F}_{k-1}\right] \qquad (14)$$
$$\le \sum_{s\in\mathcal{S}} \|\mathbf{z}_{s,k-1}-\mathbf{z}^*\|^2 - 2\alpha_k \sum_{s\in\mathcal{S}}(f_s(\check{\mathbf{z}}_{s,k})-f_s(\mathbf{z}^*))$$
$$+ 2\frac{C_{\mathbf{z}}(C_h+N)}{c_h^2}\sum_{s\in\mathcal{S}}\mathsf{E}[\|\boldsymbol{\delta}_{s,k}\| \mid \mathcal{F}_{k-1}]$$
$$+ 2\frac{C_h+N}{c_h^2}\sum_{s\in\mathcal{S}}\mathsf{E}[|\nu_{s,k}| \mid \mathcal{F}_{k-1}]$$
$$- \frac{1}{2\kappa_2(L_h+D)^2}\sum_{s\in\mathcal{S}}\mathrm{dist}^2(\mathbf{p}_{s,k},\mathcal{X}) + A\alpha_k^2 n L_f^2,$$

where we used Lemma 4.8 for the first term on the right-hand side. Recall that $f(\mathbf{z}) = \sum_{s\in\mathcal{S}} f_s(\mathbf{z})$ and $\bar{\check{\mathbf{z}}}_k \triangleq \frac{1}{n}\sum_{\ell\in\mathcal{S}}\check{\mathbf{z}}_{\ell,k}$. Using $\bar{\check{\mathbf{z}}}_k$ and f, we can rewrite the term $\sum_{s\in\mathcal{S}}(f_s(\check{\mathbf{z}}_{s,k}) - f_s(\mathbf{z}^*))$ as follows:

$$\sum_{s\in\mathcal{S}}(f_s(\check{\mathbf{z}}_{s,k})-f_s(\mathbf{z}^*))$$
$$= \sum_{s\in\mathcal{S}}(f_s(\check{\mathbf{z}}_{s,k})-f_s(\bar{\check{\mathbf{z}}}_k)) + (f(\bar{\check{\mathbf{z}}}_k)-f(\mathbf{z}^*)) \quad (15a)$$
$$\le \sum_{s\in\mathcal{S}}\langle f_s'(\bar{\check{\mathbf{z}}}_k), \bar{\check{\mathbf{z}}}_k - \check{\mathbf{z}}_{s,k}\rangle + (f(\bar{\check{\mathbf{z}}}_k)-f(\mathbf{z}^*)) \quad (15b)$$
$$\le \sum_{s\in\mathcal{S}}\|f_s'(\bar{\check{\mathbf{z}}}_k)\| \, \|\check{\mathbf{z}}_{s,k}-\bar{\check{\mathbf{z}}}_k\| + (f(\bar{\check{\mathbf{z}}}_k)-f(\mathbf{z}^*)) (15c)$$
$$\le L_f \sum_{s\in\mathcal{S}}\|\check{\mathbf{z}}_{s,k}-\bar{\check{\mathbf{z}}}_k\| + (f(\bar{\check{\mathbf{z}}}_k)-f(\mathbf{z}^*)), \qquad (15d)$$

where (15a) follows from adding and subtracting $f_s(\bar{\check{\mathbf{z}}}_k)$; (15b) follows from the convexity of the function f_s; (15c) follows from the Schwarz inequality; and (15d) follows from the fact that $\|f_s'(\mathbf{z})\| \le L_f$ and $\bar{\check{\mathbf{z}}}_k \in \mathcal{X} \subseteq \mathcal{X}_0$. Combining (15d) with (14), we obtain

$$\sum_{s\in\mathcal{S}} \mathsf{E}\left[\|\mathbf{z}_{s,k}-\mathbf{z}^*\|^2 \mid \mathcal{F}_{k-1}\right]$$
$$\le \sum_{s\in\mathcal{S}} \|\mathbf{z}_{s,k-1}-\mathbf{z}^*\|^2 - 2\alpha_k(f(\bar{\check{\mathbf{z}}}_k)-f(\mathbf{z}^*))$$
$$+ 2\frac{C_{\mathbf{z}}(C_h+N)}{c_h^2}\sum_{s\in\mathcal{S}}\mathsf{E}[\|\boldsymbol{\delta}_{s,k}\| \mid \mathcal{F}_{k-1}]$$
$$+ 2\frac{C_h+N}{c_h^2}\sum_{s\in\mathcal{S}}\mathsf{E}[|\nu_{s,k}| \mid \mathcal{F}_{k-1}]$$
$$+ L_f \sum_{s\in\mathcal{S}}\alpha_k\|\check{\mathbf{z}}_{s,k}-\bar{\check{\mathbf{z}}}_k\| + A\alpha_k^2 n L_f^2,$$

where we omitted the negative term. Since $\bar{\check{\mathbf{z}}}_k \in \mathcal{X}$, we have $f(\bar{\check{\mathbf{z}}}_k) - f(\mathbf{z}^*) \ge 0$. Thus, under the assumption that $\{\alpha_k\}$ is square summable and Lemma 4.14(b), the above relation satisfies all the conditions of Lemma 4.7. We have that
Result 1 $\{\|\mathbf{z}_{s,k}-\mathbf{z}^*\|\} \to 0$ *a.s.* for $\mathbf{z}^* \in \mathcal{X}^*$ and all $s \in \mathcal{S}$.
Result 2 $\{\alpha_k(f(\bar{\check{\mathbf{z}}}_k)-f^*)\}$ is *a.s.* summable.
As a direct consequence of *Result 1*, we know that the

sequences $\{\|\mathbf{p}_{s,k}-\mathbf{z}^*\|\}_{k\in\mathbb{N}}$ and $\{\|\check{\mathbf{z}}_{s,k}-\mathbf{z}^*\|\}_{k\in\mathbb{N}}$ also converge *a.s.* This is straightforward from Line 6 of Algorithm 1 and Lemma 4.13, with the knowledge that $\mathrm{dist}^2(\mathbf{p}_{s,k},\mathcal{X}) = \|\mathbf{p}_{s,k}-\check{\mathbf{z}}_{s,k}\|^2$. Since $\|\bar{\check{\mathbf{z}}}_k - \mathbf{z}^*\| \le \frac{1}{n}\sum_{s\in\mathcal{S}}\|\check{\mathbf{z}}_{s,k}-\mathbf{z}^*\|$, we further know that the sequence $\{\|\bar{\check{\mathbf{z}}}_k-\mathbf{z}^*\|\}$ is also convergent *a.s.* As a direct consequence of *Result 2* and the fact that $\{\alpha_k\}$ is not summable, we know that $\liminf_{k\to\infty}(f(\bar{\check{\mathbf{z}}}_k) - f(\mathbf{z}^*)) = 0$ *a.s.* From this relation and the continuity of f, it follows that the sequence $\{\bar{\check{\mathbf{z}}}_k\}$ must have at least one accumulation point in the set $\mathcal{X}^*$ *a.s.* This and the fact that $\{\|\bar{\check{\mathbf{z}}}_k - \mathbf{z}^*\|\} \to 0$ *a.s.* imply that $\{\bar{\check{\mathbf{z}}}_k\} \to \mathbf{z}^*$ *a.s.* Also, by Lemma 4.14(b), we have $\liminf_{k\to\infty}\|\check{\mathbf{z}}_{s,k} - \bar{\check{\mathbf{z}}}_k\| = 0$ for all $s \in \mathcal{S}$ *a.s.* The fact that $\{\|\check{\mathbf{z}}_{s,k} - \mathbf{z}^*\|\}$ is convergent *a.s.* for all $s \in \mathcal{S}$ thus implies that $\{\check{\mathbf{z}}_{s,k}\} \to \mathbf{z}^*$ for all $s \in \mathcal{S}$ *a.s.* Since $\|\mathbf{z}_{s,k} - \check{\mathbf{z}}_{s,k}\| \le \|\mathbf{z}_{s,k} - \mathbf{p}_{s,k}\| + \|\mathbf{p}_{s,k}-\check{\mathbf{z}}_{s,k}\|$, by invoking Lemma 4.13 and 4.14(a), we have $\{\|\mathbf{z}_{s,k} - \check{\mathbf{z}}_{s,k}\|\} \to 0$ for all $s \in \mathcal{S}$ *a.s.* It follows that $\{\mathbf{z}_{s,k}\} \to \mathbf{z}^\star$ for all $s \in \mathcal{S}$ *a.s.* $\blacksquare$

V. NUMERICAL EXPERIMENTS

We illustrate Algorithm 1 for a sparse landmark localization problem. To maintain network connectivity, we use a distributed connectivity maintenance controller [19]. We assume that the landmarks and the robots live in $\mathbb{R}^2$. We use a generic model for the measurement covariance matrix for sparse landmark localization in $\mathbb{R}^2$, which can be found in e.g., [1], [4], [5]. These models have two important characteristics that apply to a wide variety of localization sensors: (i) Measurement quality is inversely proportional to viewing distance, and (ii) The direction of maximum uncertainty is the viewing direction, i.e., the sensors are more proficient at sensing bearing than range. The general idea is to use the vector $\hat{\mathbf{x}}_i(t) - \mathbf{r}_s(t)$ to construct a diagonal matrix in a coordinate frame relative to the sensor, then rotate the matrix to a global coordinate frame so that it can be compared to other observations.

Figure 1 shows an example of the resulting active sensing trajectories for a network of four sensors that cooperatively localize a uniformly random distribution of 100 landmarks in a square $100\,\mathrm{m}^2$ workspace to a desired accuracy of $1\,\mathrm{m}$. This corresponds to an eigenvalue tolerance of $\tau_i = 9\,\mathrm{m}^{-2}$. for all landmarks. The initial mean estimate and information matrix were set to the initial observations $\{\mathbf{y}_{i,s}(0) \mid i \in \mathcal{I}, s \in \mathcal{S}\}$ and the corresponding information matrices were set to $\{Q(\mathbf{r}_s(0), \mathbf{y}_{i,s}(0))\}_{i\in\mathcal{I},s\in\mathcal{S}}$. The feasible control set $\mathcal{U}$ was set to $\mathcal{U} = \{\mathbf{u}_s \in \mathbb{R}^2 \mid \|\mathbf{u}_s\| \le 1\,\mathrm{m}\}$. In the figure, note that between $t = 194$ and $t = 242$, the agent in the bottom right moves north in order to maintain connectivity while the other agents move to finish the localization task. Fig. 2 shows the evolution of the minimum eigenvalues of the information matrices $S_i(t)$ from the simulation shown in Fig. 1. Even after the minimum eigenvalue associated with a landmark has surpassed τ, more information is still collected, but it does not affect the control signal. The maximal error in the bottom panel does not go below the desired threshold, even though the minimal eigenvalue does go above the threshold in the top
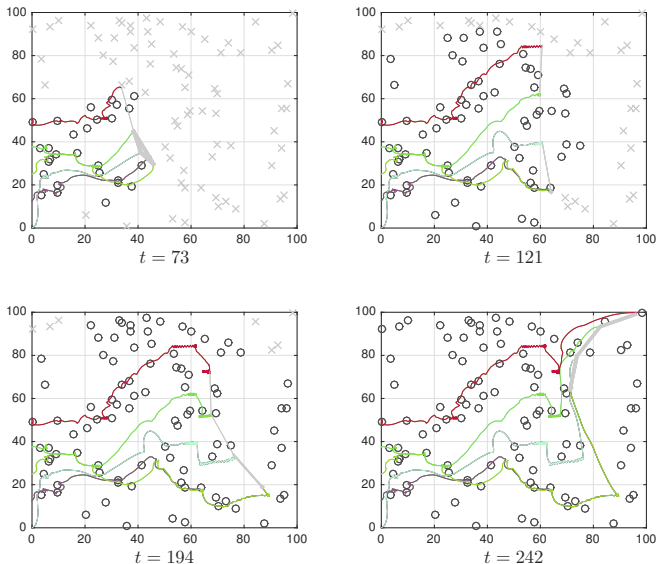
Fig. 1. Four sensors localize 100 landmarks. Gray ×'s indicate landmarks with uncertainty above the threshold, and black ∘'s indicate localized landmarks. Gray lines connecting the agents indicate the communication links.
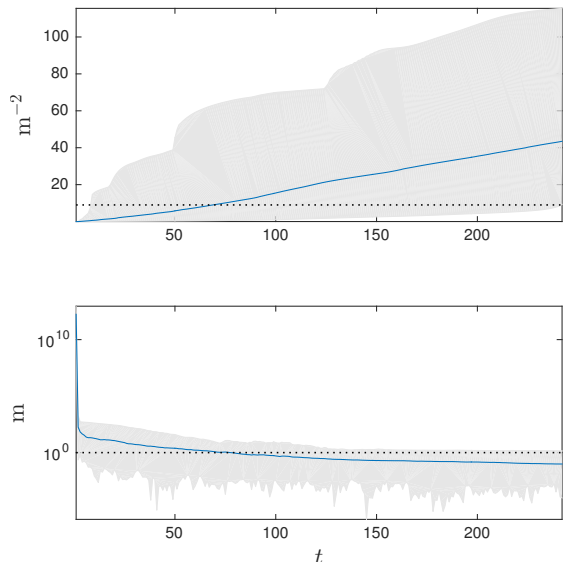

Fig. 2. Top: The blue curve is the mean eigenvalue among all 100 targets, with the grey block extending to the lowest and largest eigenvalues at each time instant. The dotted line is at the tolerance $\tau = 9\,\mathrm{m}^{-2}$. Bottom: The blue curve is the mean localization error among all 100 targets, with the grey block extending to the minimum and maximum errors at each time instant. The dotted line is at the desired accuracy of $1\,\mathrm{m}$.

panel. The scale on vertical axis of the bottom panel is large because the initial estimates are extremely poor. We terminated Algorithm 1 when both $\|\mathcal{D}_{s,k} - \mathcal{D}_{s,k-1}\|$ and $\|\mathbf{z}_{s,k} - \mathbf{z}_{s,k-1}\|$ changed by less than 10%. Typically, this converges in about 30 iterations.

We also measured the amount of constraint violation of the original nonlinear constraints in (4b). For the $\mathcal{U}$ and the parameters for the function Q used in these simulations, the constraint violation was very small, and driven to zero typically for $t > 10$. We tested larger feasible sets $\mathcal{U}$ and (predictably) observed increases in the constraint violation,

although we still observed reasonable robot behavior.

VI. Conclusion

In this paper, we addressed the problem of controlling a network of mobile sensors so that a set of hidden states are estimated up to a user-specified accuracy. Our approach interleaves filtering and optimization. We showed that the network converges to the optimal solution despite disagreement errors and the the fact that only a single constraint is considered in each step of the optimization. The numerical technique is robust toward network disagreements, is computationally light, and can handle large systems of coupled LMIs.

References

[1] P. Jalalkamali and R. Olfati-Saber, "Information-driven self-deployment and dynamic sensor coverage for mobile sensor networks," in *IEEE American Control Conf. (ACC)*, 2012, pp. 4933–4938.

[2] C. Freundlich, P. Mordohai, and M. M. Zavlanos, "Hybrid control for mobile target localization with stereo vision," in *IEEE Conf. on Decision and Control (CDC)*, 2013, pp. 2635–2640.

[3] J. Vander Hook, P. Tokekar, and V. Isler, "Algorithms for cooperative active localization of static targets with mobile bearing sensors under communication constraints," *IEEE Trans. on Robotics*, vol. 31, no. 4, pp. 864–876, 2015.

[4] T. H. Chung, J. W. Burdick, and R. M. Murray, "A decentralized motion coordination strategy for dynamic target tracking," in *IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2006, pp. 2416–2422.

[5] A. Simonetto and T. Keviczky, "Distributed multi-target tracking via mobile robotic networks: a localized non-iterative sdp approach," in *IEEE Conf. on Decision and Control (CDC)*, 2011, pp. 4226–4231.

[6] A. T. Kamal, J. Farrell, and A. K. Roy-Chowdhury, "Information weighted consensus," in *IEEE Conf. on Decision and Control (CDC)*, 2012, pp. 2732–2737.

[7] S. Lee and A. Nedić, "Distributed random projection algorithm for convex optimization," *IEEE J. Selected Topics in Signal Processing*, vol. 7, pp. 221–229, April 2013.

[8] S. Lee and M. Zavlanos, "Approximate projections for decentralized optimization with SDP constraints," 2015, http://arxiv.org/abs/1509.08007.

[9] E. D. Nerurkar, K. J. Wu, and S. I. Roumeliotis, "C-KLAM: Constrained keyframe-based localization and mapping," in *IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2014, pp. 3638–3643.

[10] M. T. Spaan, T. S. Veiga, and P. U. Lima, "Decision-theoretic planning under uncertainty with information rewards for active cooperative perception," *Int. Found. for Autonomous Agents and Multiagent Sys. (IFAAMAS)*, pp. 1–29, 2014.

[11] G. Huang, K. Zhou, N. Trawny, and S. I. Roumeliotis, "A bank of maximum a posteriori (MAP) estimators for target tracking," *IEEE Trans. on Robotics*, vol. 31, no. 1, pp. 85–103, 2015.

[12] B. J. Julian, M. Angermann, M. Schwager, and D. Rus, "Distributed robotic sensor networks: An information-theoretic approach," *Int. J. Robotics Research*, vol. 31, no. 10, pp. 1134–1154, 2012.

[13] S. Omidshafiei, A.-a. Agha-mohammadi, C. Amato, and J. P. How, "Decentralized control of partially observable markov decision processes using belief space macro-actions," in *IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2015.

[14] B. Polyak, "Minimization of unsmooth functionals," *USSR Computational Mathematics and Mathematical Physics*, vol. 9, no. 3, pp. 14–29, 1969.

[15] F. Facchinei and J.-S. Pang, *Finite-dimensional Variational Inequalities and Complementarity Problems*. Springer, New York, 2003.

[16] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar, *Convex analysis and optimization*. Athena Scientific, 2003.

[17] B. Polyak, *Intro. to optimization*. Optimization software, Inc., Publications division, New York, 1987.

[18] S. S. Ram, A. Nedić, and V. V. Veeravalli, "A new class of distributed optimization algorithms: application to regression of distributed data," *Optimization Methods and Software*, vol. 27, no. 1, pp. 71–88, 2012.

[19] B. M. M. Zavlanos, M. B. Egerstedt, and G. J. Pappas, "Graph-theoretic connectivity control of mobile robot networks," *Proceedings of the IEEE*, vol. 99, no. 9, pp. 1525–1540, 2011.