Node Localization Based on Distributed Constrained Optimization using Jacobi's Method

Henrique Ferraz, Amr Alanwar, Mani Srivastava, and João P. Hespanha

Abstract— **We consider the spatial localization of nodes in a network, based on measurements of their relative position with respect to their neighbors. These measurements include the nodes' relative positions in a global coordinate system, their distances, or their pseudo ranges. We show that the maximum likelihood estimator associated with these localization problems can be viewed as a constrained optimization problem with a specific structure and provide a distributed algorithm to solve it. Under appropriate assumptions, it is shown that the maximum likelihood estimates are locally asymptotically stable equilibrium points of the proposed algorithm. As a case study, we consider a range-based localization problem and present simulation results to evaluate the proposed algorithm.**

I. Introduction

With the increasing popularity of low-cost and low-powered devices, many applications of network systems such as environmental monitoring, road traffic monitoring, search and rescue operations, and smart structures are experiencing dramatic changes on their sensing and controlling capabilities. For most of these applications, determining the precise location of the sensor position is of utmost importance for the collected data to be meaningful [18].

In some cases, because of limitations in the environment and constraints on power and size, many of the sensors (nodes) in the network do not have access to information about their absolute position in space (from GPS, for example) and have to rely on inter-sensor measurements such as, angle-of-arrival, received signal strength (RSS), time-of-arrival, time-difference-of-arrival, and range to solve the localization problem [13]. In this context, the main challenge is to combine the available measurements to design a scalable algorithm that accurately estimates the nodes' positions efficiently in terms of energy and network bandwidth usage.

Localization algorithms can be divided in two groups: centralized and distributed. Centralized algorithms [24], [3] are impractical in large-scale networks where nodes have limited power and communication capabilities. Also, these algorithms are vulnerable to single-point failure and require the transmission of every sensor measurement to a fusion center, where all the computation is performed. One way to overcome these issues is by using distributed algorithms, where the computation is carried out locally at each node, using only information and measurements available from its neighbors.

We propose a distributed algorithm based on constrained optimization to localize nodes in a network, inspired by Jacobi's method as described in [4]. We show that the maximum likelihood estimation formulation of several localization problems reduces to a constrained optimization with a specific structure. The constraints arise from the need to impose a coordinate system that avoids ambiguities arising from a global rotation and translation.

The distributed algorithm iteratively computes the optimal solution using only the most recent measurements and estimates from the neighboring nodes. Additionally, at each node only local variables need to be stored and transmitted, greatly reducing the complexity of the computation required locally.

By regarding the iterative algorithm as a dynamical systems and linearizing it around the optimal solution, we show that the algorithm converges to the maximum likelihood estimate, provided that it starts sufficiently close to it. While our stability results are local, simulations generally converge to the optimal solution, regardless of how the algorithm is initialized.

Related work

Both centralized and distributed approaches to localization have been extensively investigated by the wireless sensor network community. Evaluation of accuracy and computation of performance bounds for unbiased location estimators are treated in [19]. A centralized convex optimization scheme is proposed in [6], where the communication network is modeled as a set of geometric constraints and the global solution is obtained by solving the resulting convex optimization problem. Distributed localization techniques can be based on multidimensional scaling and coordinate alignment techniques [10], multilateration [21], [14], and graph-theoretical methods [2]. In [17] the authors propose a hop by hop connectivity-based algorithm using trilateration. For range-based and range-free localization problems, [22] presents a convex formulation obtained by relaxation techniques that is solved with a sequential greedy optimization algorithm. Our approach builds upon [4], where it is proposed an algorithm that combines the maximum likelihood estimate and the Jacobi method for localization with relative position measurements. The algorithm analyzed here is close to the one evaluated experimentally in [1] for simultaneous node localization and time synchronization, using custom ultra-wideband wireless transceivers.

H. Ferraz and J. Hespanha are with Center for Control, Dynamical-systems and Computation (CCDC), University of California, Santa Barbara, CA. {henrique,hespanha}@ece.ucsb.edu Research supported by a CAPES BEX 1111-13-2 grant and by the National Science Foundation under Grant No. CNS-1329650.

A. Alanwar and M. Srivastava are with the Department of Electrical Engineering, University of California, Los Angeles, CA. {alanwar,mbs}@ucla.edu

In the context of optimization, the seminal work by Tsitsiklis [25] sets the foundations for the analysis of distributed algorithms. Since then, many works have focused on designing discrete-time algorithms to find the solution of optimization problems, where the cost function is a sum of convex functions, see e.g [20], [11], [15], [5]. For problems with constraints, a distributed primal-dual subgradient method was proposed in [26], where the global constraint set is the intersection of local constraints set. A similar problem setup was studied in [16] for networks with time-varying connectivity, for which a consensus-based distributed algorithm was presented. More recently, algorithms that deal with distributed continuous-time strategies were investigated in [8] and in [12], where a more sophisticated combination of a local continuous-time and discrete-time dynamics for communication with the neighbors was proposed.

Structure: The paper is organized as follows. In Section II, we formulate an optimization problem that generalizes several problems related to node localization. Section III introduces the distributed algorithm and presents conditions for local asymptotic convergence. A case study for a range-based localization is presented in Section IV and numerical simulations illustrate the proposed algorithm. In Section V we conclude the paper and present future research directions.

Notation: For twice continuously differentiable functions $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $\nabla_x f(x) \in R^{m\times n}$ denotes the Jacobian of f, $\nabla^2_{xx} f(x) \in \mathbb{R}^{n\times n}$ is the Hessian of f, and $\nabla^2_{xy} g(x,y) \in \mathbb{R}^{m\times n}$ denotes the second-order derivatives matrix of g. We denote by $0_{m\times n}$ the zero matrix in $\mathbb{R}^{m\times n}$.

II. Problem formulation

Several problems related to the localization of multi-agent systems can be reduced to optimizations of the following form:

$$\min_x \quad \sum_{i\in\mathcal{N}} f_i(x_i) + \sum_{i\in\mathcal{N}} \sum_{j\in\mathcal{N}_i} f_{ij}(x_i, x_j), \tag{1a}$$

$$\text{subject to} \quad h_i(x_i) = 0, \quad \forall i \in \mathcal{N}, \tag{1b}$$

where $x \coloneqq (x_1, x_2, \dots, x_N) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \cdots \times \mathbb{R}^{n_N}$ are the optimization variables, $\mathcal{N} \coloneqq \{1, 2, \dots, N\}$, $\mathcal{N}_i$ is a subset of $\mathcal{N} \setminus \{i\}$ containing the neighbors of node i, and the functions $f_i : \mathbb{R}^{n_i} \to \mathbb{R}$, $f_{ij} : \mathbb{R}^{n_i} \times \mathbb{R}^{n_j} \to \mathbb{R}$, $h_i : \mathbb{R}^{n_i} \times \mathbb{R}^{n_j} \to \mathbb{R}^{m_i}$, $i \in \mathcal{N}, j \in \mathcal{N}_i$ are all twice continuously differentiable.

We consider here four problems, where we want to localize in space a set $\mathcal{N} \coloneqq \{1, 2, \dots, N\}$ of N nodes based on relative measurements of each node $i \in \mathcal{N}$ with respect to its neighbors $\mathcal{N}_i \subset \mathcal{N} \setminus \{i\}$.

A. Relative position measurements

In this scenario, each variable $p_i \in \mathbb{R}^d$, $d = \{2, 3\}$, $i \in \mathcal{N}$ denotes the position of node i in a global coordinate system and the node i has access to noisy measurements $z_{ij} \in \mathbb{R}^d$ of the relative position $p_j - p_i$ of each neighboring node $j \in \mathcal{N}_i$. Specifically,

$$z_{ij} = p_j - p_i + w_{ij}, \quad \forall i \in \mathcal{N}, \ j \in \mathcal{N}_i,$$

where the w_{ij} denote independent zero-mean Gaussian noise with co-variance matrix $\Sigma_{ij} > 0$. For this problem, the symmetric of the log-likelihood of the measurements $\{z_{ij} \in \mathbb{R}^d : i \in \mathcal{N}, \ j \in \mathcal{N}_i\}$ is given by

$$\frac{1}{2} \sum_{i\in\mathcal{N}} \sum_{j\in\mathcal{N}_j} (p_i - p_j + z_{ij})' \Sigma_{ij}^{-1} (p_i - p_j + z_{ij}).$$

Just with relative measurements it is only possible to reconstruct the positions p_i up to a global translation. To avoid this ambiguity one can force the position of one "reference" node (say node $i = 1$) to be the origin of the coordinate system, which corresponds to the constraint

$$p_1 = 0_{n\times 1}.$$

The computation of maximum likelihood estimates for the p_i thus amounts to solving an optimization of the form (1) with

$$\begin{aligned}
&x_i \coloneqq p_i, && f_i(x_i) \coloneqq 0, \\
&f_{ij}(x_i, x_j) \coloneqq \frac{1}{2}(p_i - p_j + z_{ij})' \Sigma_{ij}^{-1} (p_i - p_j + z_{ij}), \\
&h_1(x_1) \coloneqq p_1, && h_i(x_i) \coloneqq 0, \quad \forall i \neq 1.
\end{aligned}$$

When the neighborhoods $\mathcal{N}_i$ induce a graph in which there is a path from the reference node 1 to every other node $i \neq 1$, this optimization is a strictly convex quadratic program [4].

In this formulation and the ones that follow, we ignore any prior information about the positions p_i. When prior distributions for these variables are available, this information could be incorporated into the optimization through the functions $f_i(x_i)$.

B. Range measurements

This scenario is analogous to the previous one, but now the node i has access to noisy measurements $z_{ij} \in \mathbb{R}$ of its distance $\|p_j - p_i\|$ to each of its neighboring nodes $j \in \mathcal{N}_i$. Specifically,

$$z_{ij} = \|p_j - p_i\| + w_{ij}, \quad \forall i \in \mathcal{N}, \ j \in \mathcal{N}_i,$$

where the w_{ij} denote independent zero-mean Gaussian noise with variance $\sigma_{ij} > 0$. For this problem, the symmetric of the log-likelihood of the measurements $\{z_{ij} \in \mathbb{R} : i \in \mathcal{N}, \ j \in \mathcal{N}_i\}$ is given by

$$\frac{1}{2} \sum_{i\in\mathcal{N}} \sum_{j\in\mathcal{N}_j} \frac{(\|p_j - p_i\| - z_{ij})^2}{\sigma_{ij}^2}.$$

Just with relative measurements, it is only possible to reconstruct the positions $p_i \in \mathbb{R}^d$ up to a global translation and rotation. To avoid this ambiguity one can force the position of one "reference" node, say node 1, to be the origin of the coordinate system and, for $d = 3$, use two other nodes, say nodes 2 and 3, to define the orientation of the coordinate system. Specifically, forcing the 1st axis of the coordinate system to be aligned with the vector from p_1 to p_2 and the second axis to lie in the plane defined by the first 3 nodes (assumed not to be co-linear), corresponds to the constraints

$$p_1 = 0, \qquad e_2' p_2 = e_3' p_2 = 0, \qquad e_3' p_3 = 0,$$

where $e_i \in \mathbb{R}^3$ denotes the ith vector of the canonical basis of $\mathbb{R}^3$. The computation of maximum likelihood estimate for the p_i thus amounts to solving an optimization of the form (1) with

$$x_i := p_i, \qquad f_i(x_i) := 0, \tag{2a}$$

$$f_{ij}(x_i, x_j) := \frac{1}{2}\frac{(\|p_j - p_i\| - z_{ij})^2}{\sigma_{ij}^2}, \tag{2b}$$

$$h_1(x_1) := p_1, \qquad h_2(x_2) := \begin{bmatrix} e_2' \\ e_3' \end{bmatrix} p_2, \tag{2c}$$

$$h_3(x_3) := e_3' p_3, \qquad h_i(x_i) := 0, \quad \forall i > 3. \tag{2d}$$

When $d = 2$, only two nodes are needed to define the coordinate system.

When the neighborhoods $\mathcal{N}_i$ induce a framework that is rigid and the measurements are noiseless, this optimization has an isolated global minima at the true positions of the nodes [3].

C. Pseudo-range measurements

In this scenario, each node $i \in \mathcal{N}$ broadcasts a wireless message to its neighbors $j \in \mathcal{N}_i$ with the value t_i of its local clock at the transmission time and the neighbors record the times of arrival t_{ij} of this message in their local clocks. The clock of each node i has an unknown offset τ_i with respect to the "global" time reference and therefore the actual times at which the message was transmitted by node i and received by node j are given by

$$t_i - \tau_i + w_i \text{ and } t_{ij} - \tau_j + w_{ij},$$

respectively, where w_i and w_{ij} denote time measurement errors. Assuming that messages propagate at a velocity c, we have that

$$t_{ij} - \tau_j + w_{ij} = t_i - \tau_i + w_i + \frac{\|p_j - p_i\|}{c}$$

which can be re-written as

$$t_{ij} - t_i = \tau_j - \tau_i + \frac{\|p_j - p_i\|}{c} + \bar{w}_{ij},$$

where $\bar{w}_{ij} := w_i - w_{ij} \in \mathbb{R}$. Assuming that the $\bar{w}_{ij}$ are independent zero-mean Gaussian random variables with variances $\sigma_{ij} > 0$, the symmetric of the log-likelihood of the measurements $\{z_{ij} \in \mathbb{R} : i \in \mathcal{N},\ j \in \mathcal{N}_i\}$ is given by

$$\frac{1}{2}\sum_{i \in \mathcal{N}}\sum_{j \in \mathcal{N}_j} \frac{(\tau_j - \tau_i + \frac{\|p_j - p_i\|}{c} - t_{ij} + t_i)^2}{\sigma_{ij}^2}.$$

In this problem, we have ambiguity with respect to a global rotation and translation, as in Section II-B, but also with respect to a shift of the time reference, which can be resolved by forcing the clock of node 1 to determine the global time t. This corresponds to the following constraints:

$$p_1 = 0, \quad e_2' p_2 = e_3' p_2 = 0, \quad e_3' p_3 = 0, \quad \tau_1 = 0.$$

The computation of maximum likelihood estimates for the p_i and τ_i thus amounts to solving an optimization of the form (1) with

$$x_i := [p_i, \tau_i], \qquad f_i(x_i) := 0,$$

$$f_{ij}(x_i, x_j) := \frac{1}{2}\frac{(\tau_j - \tau_i + \frac{\|p_j - p_i\|}{c} - t_{ij} + t_i)^2}{\sigma_{ij}^2},$$

$$h_1(x_1) := \begin{bmatrix} p_1 \\ \tau_1 \end{bmatrix}, \qquad h_2(x_2) := \begin{bmatrix} e_2' \\ e_3' \end{bmatrix} p_2,$$

$$h_3(x_3) := e_3' p_3, \qquad h_i(x_i) := 0, \quad \forall i > 3.$$

D. Pseudo-range measurements with clock drift and biases

This scenario is similar to the previous one, but considering clock drifts and biases in the time measurements. In this case, the transmission time measurement t_i for a message sent at time t by node i is given by

$$t_i = \phi_i t + \tau_i + b_i + w_i \quad \Leftrightarrow \quad t = \frac{t_i - \tau_i - b_i - w_i}{\phi_i}$$

and the reception time measurement t_{ij} for a message received at time t by node j is given by

$$t_{ij} = \phi_j t + \tau_j + d_j + w_{ij} \quad \Leftrightarrow \quad t = \frac{t_{ij} - \tau_j - d_j - w_{ij}}{\phi_j},$$

where ϕ_i is the clock drift of node i, b_i a bias in the transmission-time measurement for node i, d_j a bias in the reception-time measurement for node j, and w_i, w_{ij} zero-mean noise.

Assuming that messages propagate at a velocity c, we have that

$$\frac{t_{ij} - \tau_j - d_j - w_{ij}}{\phi_j} = \frac{t_i - \tau_i - b_i - w_i}{\phi_i} + \frac{\|p_j - p_i\|}{c}$$

which can be re-written as

$$\phi_j^{-1} t_{ij} - \phi_i^{-1} t_i = \bar{\tau}_j - \bar{\tau}_i + \frac{\|p_j - p_i\|}{c} + \bar{w}_{ij},$$

where $\bar{w}_{ij} := \phi_i^{-1} w_i - \phi_j^{-1} w_{ij} \in \mathbb{R}$, $\bar{\tau}_j = \phi_j^{-1}\tau_j$. Assuming that the $\bar{w}_{ij}$ are independent zero-mean Gaussian random variables with variances $\sigma_{ij} > 0$, the symmetric of the log-likelihood of the measurements $\{z_{ij} \in \mathbb{R} : i \in \mathcal{N},\ j \in \mathcal{N}_i\}$ is given by

$$\frac{1}{2}\sum_{i \in \mathcal{N}}\sum_{j \in \mathcal{N}_j} \frac{(\bar{\tau}_j - \bar{\tau}_i + \frac{\|p_j - p_i\|}{c} - \phi_j^{-1} t_{ij} + \phi_i^{-1} t_i)^2}{\sigma_{ij}^2}.$$

Mapping the computation of the maximum likelihood estimate for p_i, τ_i, ϕ_i, b_i, and d_i with the optimization in (1) is straightforward, so we omit it for the sake of space.

III. Main Results

A. Jacobi Algorithm

We construct a distributed algorithm for solving (1), where each node $i \in \mathcal{N}$ receives estimates x_j of the optimal from its neighbors $j \in \mathcal{N}_i$ and, based on these estimates, computes a value for $x_i \in \mathbb{R}^{n_i}$ that only minimizes the terms in the cost function in (1) that depend on x_i:

$$\min_{x_i} \quad f_i(x_i) + \sum_{j\in\mathcal{N}_i} f_{ij}(x_i,x_j) + \sum_{j\in\mathcal{N}:i\in\mathcal{N}_j} f_{ji}(x_j,x_i), \tag{3a}$$

$$\text{subject to} \quad h_i(x_i) = 0. \tag{3b}$$

In practice, node $i \in \mathcal{N}$ computes values for $x_i \in \mathbb{R}^{n_i}$ and Lagrange multipliers $\lambda_j \in \mathbb{R}^{m_j}$ that satisfy the first-order necessary optimality conditions for (3):

$$\nabla_{x_i} f_i(x_i) + \lambda_i' \nabla_{x_i} h_i(x_i) + \sum_{j\in\mathcal{N}_i} \nabla_{x_i} f_{ij}(x_i,x_j) + \sum_{j\in\mathcal{N}:i\in\mathcal{N}_j} \nabla_{x_i} f_{ji}(x_j,x_i) = 0, \tag{4a}$$

$$h_i(x_i) = 0. \tag{4b}$$

If all nodes succeed in jointly satisfying (4), then the first-order optimality conditions for the optimal (1) are automatically satisfied. This observation motivates the following iterative algorithm, which is inspired by Jacobi's method for solving linear systems of equations as described in [4].

Algorithm 1 (Jacobi):

Step 1. Set $k = 0$.

Step 2. Each node $i \in \mathcal{N}$ initializes its estimates $\hat{x}_i(0) \in \mathbb{R}^{n_i}$ and $\hat{\lambda}_i(0) \in \mathbb{R}^{m_i}$.

Step 3. Each node $i \in \mathcal{N}$ computes vectors $\hat{x}_i(k+1) \in \mathbb{R}^{n_i}$ and Lagrange multiplies $\hat{\lambda}_i(k+1) \in \mathbb{R}^{m_i}$ that solve the first-order necessary optimality conditions (4) for the previous estimates of $\hat{x}_j(k) \in \mathbb{R}^{n_j}$ for the other nodes $j \in \mathcal{N}_j$:

$$\nabla_{x_i} f_i(\hat{x}_i(k+1)) + \hat{\lambda}_i(k+1)' \nabla_{x_i} h_i(\hat{x}_i(k+1)) + \sum_{j\in\mathcal{N}_i} \nabla_{x_i} f_{ij}(\hat{x}_i(k+1), \hat{x}_j(k)) + \sum_{j\in\mathcal{N}:i\in\mathcal{N}_j} \nabla_{x_i} f_{ji}(\hat{x}_j(k), \hat{x}_i(k+1)) = 0, \tag{5a}$$

$$h_i(\hat{x}_i(k+1)) = 0. \tag{5b}$$

Step 4. Increment k by 1 and go back to Step 3 until $\|\hat{x}_i(k+1) - \hat{x}_i(k)\| < \varepsilon_x$, for some user defined tolerance ε_x.

In this work, we restrict our attention to problems, where the first-order necessary optimality conditions (5) uniquely determine $\hat{x}_i(k+1) \in \mathbb{R}^{n_i}$ and $\hat{\lambda}_i(k+1) \in \mathbb{R}^{m_i}$. Lemma 1 below shows that this will happen under are mild assumptions.

B. Local Stability

To study the convergence of Algorithm 1, we view the sequences $x(k) := \big(x_1(k), x_2(k), \dots, x_N(k)\big)$, $\lambda(k) := \big(\lambda_1(k), \lambda_2(k), \dots, \lambda_N(k)\big)$ as the state of a discrete-time dynamical system whose dynamics are defined by (5) and study its local stability around an optimum $x^* := (x_1^*, x_2^*, \dots, x_N^*)$ for (1) and the corresponding Lagrange multiplier $\lambda^* := (\lambda_1^*, \lambda_2^*, \dots, \lambda_N^*)$.

The following result is a direct consequence of the Implicit Function Theorem [9] applied to (5):

Lemma 1: Let $x^* := (x_1^*, x_2^*, \dots, x_N^*)$ be an optimum for (1) and $\lambda^* := (\lambda_1^*, \lambda_2^*, \dots, \lambda_N^*)$ the corresponding Lagrange multiplier. Assume that

A1 all the functions f_i, f_{ij}, h_i, $i \in \mathcal{N}$, $j \in \mathcal{N}_j$ are twice continuous differentiable in an open neighborhood of (x^*, λ^*); and

A2 the following Jacobian matrix is invertible at $x = x^*$ and $\lambda = \lambda^*$

$$\begin{bmatrix} L_i & D_i' \\ D_i & 0 \end{bmatrix}_{(n_i+m_i)\times(n_i+m_i)},$$

where $L_i := \nabla^2_{x_i x_i} f_i(x_i) + \lambda_i' \nabla^2_{x_i x_i} h_i(x_i) + \sum_{j\in\mathcal{N}_i} \nabla^2_{x_i x_i} f_{ij}(x_i,x_j) + \sum_{j\in\mathcal{N}:i\in\mathcal{N}_j} \nabla^2_{x_i x_i} f_{ji}(x_j,x_i)$ is a symmetric matrix and $D_i := \nabla_{x_i} h_i(x_i)$. Then there exists an open neighborhood of (x^*, λ^*) such that if $(\hat{x}(k), \hat{\lambda}(k))$ belong to this neighborhood, $x_i^+ := \hat{x}_i(k+1)$ and $\lambda^+ := \hat{\lambda}_i(k+1)$ are uniquely defined by (5) and we have that

$$\nabla_{x_\ell} \begin{bmatrix} x_i^+ \\ \lambda_i^+ \end{bmatrix} = \begin{bmatrix} L_i & D_i' \\ D_i & 0 \end{bmatrix}^{-1} \begin{bmatrix} S_{i\ell} \\ 0 \end{bmatrix}_{(n_i+m_i)\times n_\ell} \tag{6a}$$

$$\nabla_{\lambda_\ell} \begin{bmatrix} x_i^+ \\ \lambda_i^+ \end{bmatrix} = 0_{(n_i+m_i)\times m_\ell}, \tag{6b}$$

where $S_{i\ell} \in \mathbb{R}^{n_i \times n_\ell}$,

$$S_{i\ell} := \begin{cases} -\nabla^2_{x_i x_\ell} f_{i\ell}(x_i, x_\ell) \\ \qquad -\nabla^2_{x_i x_\ell} f_{\ell i}(x_\ell, x_i) & \ell \in \mathcal{N}_i, i \in \mathcal{N}_\ell \\ -\nabla^2_{x_i x_\ell} f_{i\ell}(x_i, x_\ell) & \ell \in \mathcal{N}_i, i \notin \mathcal{N}_\ell \\ -\nabla^2_{x_i x_\ell} f_{\ell i}(x_\ell, x_i) & \ell \notin \mathcal{N}_i, i \in \mathcal{N}_\ell \\ 0 & \text{otherwise.} \end{cases}$$

■

Remark 1 ([7]): For the problems discussed in Section II, it can be shown that Assumption A1 of Lemma 1 holds as long as no two nodes are at the same position. Assumption A2 has simple geometric interpretations for the several localization problems. To express these conditions, we denote by $\bar{\mathcal{N}}_i$ the union of the set $\mathcal{N}_i$ of neighbors of i together with the set of nodes $j \in \mathcal{N}$ to which i is a neighbor, i.e.,

$$\bar{\mathcal{N}}_i := \mathcal{N}_i \cap \{j \in \mathcal{N} : i \in \mathcal{N}_j\}.$$

For points p_i for which there are no constraints these conditions are as follows:

- For the relative measurements problem in Section II-A, Assumption A2 holds provided that $\bar{\mathcal{N}}_i$ is not empty.
- For the range measurements problem in Section II-B with points in $\mathbb{R}^2$ and in the absence of noise, Assumption A2 holds provided that $\bar{\mathcal{N}}_i$ contains at least two points that are not co-linear with p_i.
- For the range measurements problem in Section II-B with points in $\mathbb{R}^3$ and in the absence of noise, Assumption A2 holds provided that $\bar{\mathcal{N}}_i$ contains at least three points that are not co-planar with p_i.
- For the pseudo-range measurements problem in Section II-C with points in $\mathbb{R}^3$ and in the absence of noise, Assumption A2 holds as long as $\bar{\mathcal{N}}_i$ contains at least four non co-planar points. □

Lemma 1 enables us to compute the local linearization of the discrete-time dynamical system whose dynamics are defined by (5) around an optimum $x^* := (x_1^*, x_2^*, \ldots, x_N^*)$ for (1) with Lagrange multipliers $\lambda^* := (\lambda_1^*, \lambda_2^*, \ldots, \lambda_N^*)$. Denoting by $\delta x := (\delta x_1, \delta x_2, \ldots, \delta x_N)$, $\delta\lambda := (\delta\lambda_1, \delta\lambda_2, \ldots, \delta\lambda_N)$ the perturbations of the state with respect to the equilibrium point (x^*, λ^*), under the assumptions of Lemma 1 we conclude from (6a) that the next-state vector $(\delta x^+, \delta\lambda^+)$ is uniquely defined by the following system of equations on the unknowns δx_i^+ and $\delta\lambda_i^+$.

$$\begin{bmatrix} L_i^* & D_i^{*\prime} \\ D_i^* & 0 \end{bmatrix} \begin{bmatrix} \delta x_i^+ \\ \delta\lambda_i^+ \end{bmatrix} = \begin{bmatrix} \sum_{\ell \in \mathcal{N}_i} S_{i\ell}^* \delta x_\ell \\ 0 \end{bmatrix}, \quad \forall i \in \mathcal{N},$$

where

$$\begin{aligned}
L_i^* &:= \nabla^2_{x_i x_i} f_i(x_i^*) + \lambda_i^{*\prime} \nabla^2_{x_i x_i} h_i(x_i^*) \\
&\quad + \sum_{j \in \mathcal{N}_i} \nabla^2_{x_i x_i} f_{ij}(x_i^*, x_j^*) + \sum_{j \in \mathcal{N} : i \in \mathcal{N}_j} \nabla^2_{x_i x_i} f_{ji}(x_j^*, x_i^*), \\
D_i^* &:= \nabla_{x_i} h_i(x_i^*), \\
S_{i\ell}^* &:= \begin{cases} -\nabla^2_{x_i x_\ell} f_{i\ell}(x_i^*, x_\ell^*) & \\ \qquad\qquad -\nabla^2_{x_i x_\ell} f_{\ell i}(x_\ell^*, x_i^*) & \ell \in \mathcal{N}_i, i \in \mathcal{N}_\ell \\ -\nabla^2_{x_i x_\ell} f_{i\ell}(x_i^*, x_\ell^*) & \ell \in \mathcal{N}_i, i \notin \mathcal{N}_\ell \\ -\nabla^2_{x_i x_\ell} f_{\ell i}(x_\ell^*, x_i^*) & \ell \notin \mathcal{N}_i, i \in \mathcal{N}_\ell \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

In vector form, these equations can be expressed as

$$\begin{bmatrix} L^* & D^{*\prime} \\ D^* & 0 \end{bmatrix} \begin{bmatrix} \delta x^+ \\ \delta\lambda^+ \end{bmatrix} = \begin{bmatrix} S^* \delta x \\ 0 \end{bmatrix},$$

where

$$\begin{aligned}
L^* &:= \operatorname{diag}(L_1^*, L_2^*, \ldots, L_N^*) \in \mathbb{R}^{n \times n}, \qquad n := \sum_{i \in \mathcal{N}} n_i, \\
D^* &= \operatorname{diag}(D_1^*, D_2^*, \ldots, D_N^*) \in \mathbb{R}^{m \times m}, \quad m := \sum_{i \in \mathcal{N}} m_i, \\
S^* &:= \left[S_{i\ell}^*\right]_{i \in \mathcal{N}, \ell \in \mathcal{N}}, \in \mathbb{R}^{n \times n}.
\end{aligned}$$

The next result provides a sufficient condition for the local stability of this system.

Theorem 1: Suppose that the assumptions of Lemma 1 hold and that there exists a scalar $\sigma \in \mathbb{R}$ such that,

$$\begin{aligned}
L^* + \sigma D^{*\prime} D^* - \frac{1}{2}(S^{*\prime} + S^*) > 0, \\
L^* + \sigma D^{*\prime} D^* + \frac{1}{2}(S^{*\prime} + S^*)^* > 0.
\end{aligned}$$

Then the optimum (x^*, δ^*) for (1) is a locally asymptotically stable equilibrium point of the discrete-time dynamical system whose dynamics are defined by (5). □

This result is a direct consequence of Lemma 1 and the following lemma.

Lemma 2: Consider a discrete-time linear time-invariant system

$$z^+ = Az, \quad z \in \mathbb{R}^n, \tag{7}$$

where, for every $z \in \mathbb{C}^n$, the next state $z^+ \in \mathbb{C}$ can be uniquely determined by the following equation

$$\exists u \in \mathbb{C}^m : \begin{bmatrix} L & D' \\ D & 0_{m \times m} \end{bmatrix} \begin{bmatrix} z^+ \\ u \end{bmatrix} = \begin{bmatrix} Sz \\ 0_{m \times 1} \end{bmatrix}$$

for appropriate matrices $L = L', S \in \mathbb{R}^{n \times n}$, $D \in \mathbb{R}^{m \times n}$ such that there exists a scalar $\sigma \in \mathbb{R}$

$$L + \sigma D'D - \frac{1}{2}(S' + S) > 0, \tag{8a}$$

$$L + \sigma D'D + \frac{1}{2}(S' + S) > 0. \tag{8b}$$

Then A is Schur and the original system (7) is asymptotically stable. □

Proof of Lemma 2. Consider an eigenvalue $\lambda \in \mathbb{C} \backslash \{0\}$ and the corresponding eigenvector $v \in \mathbb{R}^n$ of A. Under the lemma's hypothesis,

$$\begin{aligned}
Av = \lambda v \quad &\Leftrightarrow \quad \exists u \in \mathbb{C}^m : \begin{bmatrix} L & D' \\ D & 0_{m \times m} \end{bmatrix} \begin{bmatrix} \lambda v \\ u \end{bmatrix} = \begin{bmatrix} Sv \\ 0_{m \times 1} \end{bmatrix} \\
&\Leftrightarrow \quad \exists u \in \mathbb{C}^m : \lambda L v + D'u = Sv, \ \lambda D v = 0. \\
&\Leftrightarrow \quad \exists u \in \mathbb{C}^m : \lambda L v + \lambda\sigma D'Dv + D'u = Sv, \ \lambda D v = 0.
\end{aligned}$$

We thus conclude that

$$\begin{aligned}
\lambda v^\dagger (L + \sigma D'D) v + v^\dagger D' u = v^\dagger S v, \ u^\dagger D v = 0 \\
\Rightarrow \lambda v^\dagger (L + \sigma D'D) v = v^\dagger S v,
\end{aligned} \tag{9}$$

where $v^\dagger$ and $u^\dagger$ denote the complex conjugate transpose of v and u, respectively. On the other hand, from (8) we have that

$$v^\dagger (L + \sigma D'D) v > \frac{1}{2} v^\dagger (S' + S) v = v^\dagger S v, \tag{10a}$$

$$v^\dagger (L + \sigma D'D) v > -\frac{1}{2} v^\dagger (S' + S) v = -v^\dagger S v, \tag{10b}$$

which implies that $v^\dagger (L + \sigma D'D) v > 0$. This allow us to conclude from (9) and (10) that

$$\begin{aligned}
\lambda v^\dagger (L + \sigma D'D) v = v^\dagger S v \quad &\Rightarrow \quad \lambda = \frac{v^\dagger S v}{v^\dagger (L + \sigma D'D) v} \\
v^\dagger (L + \sigma D'D) v > v^\dagger S v \quad &\Rightarrow \quad \frac{v^\dagger S v}{v^\dagger (L + \sigma D'D) v} < 1 \\
v^\dagger (L + \sigma D'D) v > -v^\dagger S v \quad &\Rightarrow \quad \frac{v^\dagger S v}{v^\dagger (L + \sigma D'D) v} > -1.
\end{aligned}$$

Therefore λ must be a real number in the (open) interval $(-1, 1)$. ■

IV. Numerical Example

To illustrate the proposed algorithm and the theoretical results obtained in the previous section, we present an example of node localization for range-based measurements in the plane, as described in Section II-B.

Inspired by the Henneberg construction of rigid frameworks [23], we generate random networks by successively adding a node at a random position to an existing rigid framework. The new node is connected using bidirectional edges with two existing nodes such that these three nodes are not co-linear.

We have generated a large number of rigid frameworks using the procedure described above and verified that the corresponding matrices L^*, D^*, and S^* verified the conditions in Theorem 1 for $\sigma = 1$. We thus conjecture that these conditions hold generically (perhaps excluding singular configurations) for rigid frameworks. We are currently working towards formally proving this conjecture.

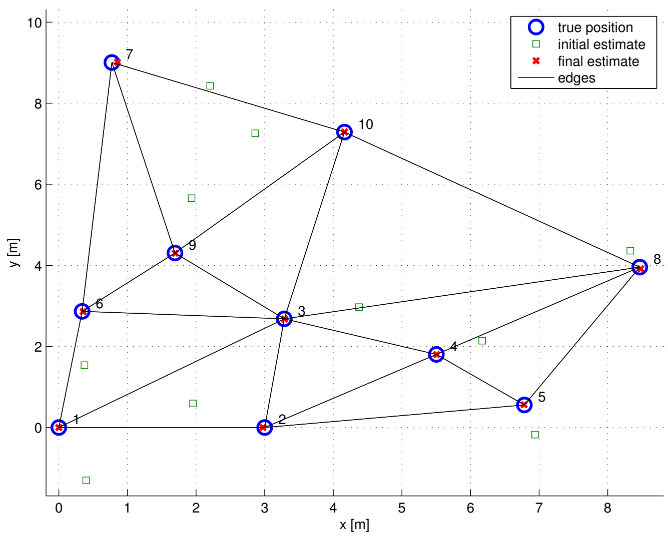

Fig. 1. Sensor network with 10 randomly distribute nodes. The nodes share information about their current estimates and the noisy range measurements with a limited number of neighbors, represented by the edge connection. The initial estimates are random.

One such network consisting of 10 geographically randomly distributed nodes is shown in Figure 1. Figure 2 shows a typical evolution of the local cost function and the estimation error for two nodes, as a function of the iteration number of the Jacobi algorithm described in Section III-A starting with a random initialization for the node estimate within a ball centered at their true positions. The interior-point method was used to solve each optimization step. We observe that node 6, that is one hop away from the reference node 1, converges faster than node 10, that is 2 hops away from the reference node 2. This is to be expected, because the convergence of the references nodes occur faster than the rest of the network, improving the speed of convergence of its direct neighbors.

V. Conclusion

We presented a distributed algorithm that iteratively computes the optimal solution to constrained optimizations that arise on node localization, using only locally available measurements. Sufficient conditions for local asymptotic stability were derived by linearizing the system around the optimal solution. A range-based localization problem was used to illustrate the proposed algorithm and we showed that the nodes estimated position converge to their true position in the presence of noisy and with random initial estimates. The localization of mobile nodes and estimation under adversarial noise are topics for future research.

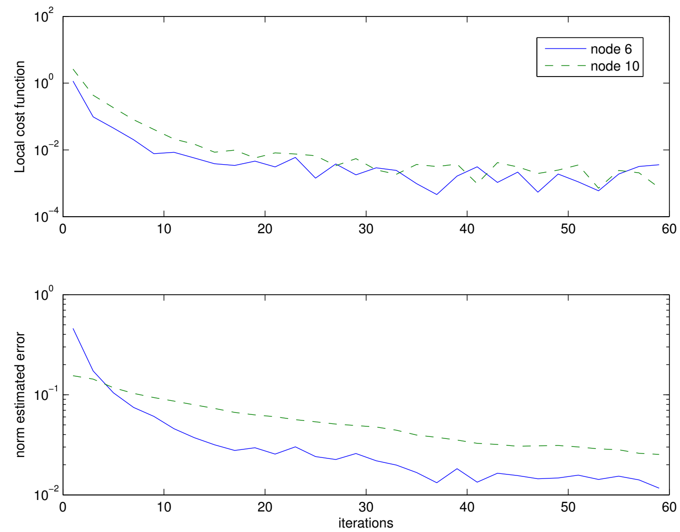

Fig. 2. Cost function and error evolution for two nodes. The dashed line represents a node that is more distant to the reference nodes than the node represented by a solid line.

References

[1] A. Alanwar et al. "D-SLATS: Distributed Simultaneous Localization and Time Synchronization", *Submitted to conference publication*, June 2017.

[2] B. D. Anderson, B. D., Belhumeur, P. N., Eren, T., Goldenberg, D. K., Morse, A. S., Whiteley, W., and Yang, Y. R. "Graphical properties of easily localizable sensor networks." *Wireless Networks* no. 15, vol.2, pp 177-191, 2009.

[3] J. Aspnes et al., "A Theory of Network Localization," in *IEEE Transactions on Mobile Computing*, vol. 5, no. 12, pp. 1663-1678, Dec. 2006.

[4] P. Barooah and J. P. Hespanha, "Estimation on graphs from relative measurements," in *IEEE Control Systems*, vol. 27, no. 4, pp. 57-74, Aug. 2007.

[5] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers", *Found. Trends Mach. Learn.*, vol. 3, pp. 1-122, 2011.

[6] L. Doherty, K. S. J. Pister and L. El Ghaoui, "Convex position estimation in wireless sensor networks," *Proceedings IEEE INFOCOM 2001. Conference on Computer Communications*. vol.3, pp. 1655-1663, 2001.

[7] H. Ferraz and J. P. Hespanha, "Node Localization Based on Distributed Constrained Optimization using Jacobi's Method", tech. rep., Univ. California, Santa Barbara, 2017. `http://www.ece.ucsb.edu/~hespanha/techrep.html`.

[8] B. Gharesifard and J. Cortés, "Distributed Continuous-Time Convex Optimization on Weight-Balanced Digraphs," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 781-786, March 2014.

[9] J. K Hale, *Ordinary Differential Equations*, Krieger Publishing Company, 1980.

[10] X. Ji and H. Zha, "Sensor positioning in wireless ad-hoc sensor networks using multidimensional scaling," *IEEE INFOCOM 2004*, vol.4, pp. 2652-2661, 2004,

[11] B. Johansson, M. Rabi, and M. Johansson, "A randomized incremental subgradient method for distributed optimization in networked systems, *SIAM Journal on Control and Optimization*, vol. 20, no. 3, pp. 11571170, 2009.

[12] S. S. Kia, J. Cortés, S. Martínez, "Distributed convex optimization via continuous-time coordination algorithms with discrete-time communication", *Automatica*, vol. 55, pp. 254-264, May 2015,

[13] G. Mao, B. Fidan, B.D.O. Anderson, "Wireless sensor network localization techniques", *Computer Networks*, vol. 51, issue 10, 11, pp. 2529-2553, July 2007.

[14] R. Nagpal,H. Shrobe, and J. Bachrach "Organizing a global coordinate system from local information on an ad hoc sensor network". *In*

Information Processing in Sensor Networks, pp. 333-348, Springer Berlin Heidelberg, 2003.

[15] A. Nedíc and A. Ozdaglar, "Distributed subgradient methods for multiagent optimization, *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 4861, 2009.

[16] A. Nedic, A. Ozdaglar and P. A. Parrilo, "Constrained Consensus and Optimization in Multi-Agent Networks," in *IEEE Transactions on Automatic Control*, vol. 55, no. 4, pp. 922-938, April 2010.

[17] D. Niculescu and B. Nath, "Ad hoc positioning system (APS)," *Global Telecommunications Conference (GLOBECOM)*, vol. 5, pp. 2926-2931, 2001.

[18] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses and N. S. Correal, "Locating the nodes: cooperative localization in wireless sensor networks," in *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 54-69, July 2005.

[19] N. Patwari, A. O. Hero, M. Perkins, N. S. Correal and R. J. O'Dea, "Relative location estimation in wireless sensor networks," in *IEEE Transactions on Signal Processing*, vol. 51, no. 8, pp. 2137-2148, Aug. 2003.

[20] M. Rabbat and R. Nowak, "Distributed optimization in sensor networks," *Third International Symposium on Information Processing in Sensor Networks, 2004. IPSN 2004*, pp. 20-27, 2004

[21] A. Savvides, H. Park, and M, B. Srivastava. "The bits and flops of the n-hop multilateration primitive for node localization problems." *Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*. pp. 112-121, 2002.

[22] Q. Shi, C. He, H. Chen and L. Jiang, "Distributed Wireless Sensor Network Localization Via Sequential Greedy Optimization Algorithm," in *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3328-3340, June 2010.

[23] T. S. Tay, W. Whiteley, "Generating isostatic frameworks", *Structural Topol.* no. 11, pp. 2169, 1985.

[24] F. Thomas and L. Ros, "Revisiting trilateration for robot localization," in *IEEE Transactions on Robotics*, vol. 21, no. 1, pp. 93-101, Feb. 2005.

[25] J. N. Tsitsiklis, "Problems in Decentralized Decision Making and Computation, Ph.D. dissertation", Dept. Elect. Eng. Comp. Sci., Massachusetts Institute of Technology, Cambridge, 1984.

[26] M. Zhu and S. Martinez, "On Distributed Convex Optimization Under Inequality and Equality Constraints," in *IEEE Transactions on Automatic Control*, vol. 57, no. 1, pp. 151-164, Jan. 2012.