Automatica 171 (2025) 111903

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Performance optimization for continuous network localization ☆

Liangming Chen [a,1], Chenyang Liang [b,1], Yibei Li [c], Jie Mei [b,*], Lihua Xie [c]

[a] School of System Design and Intelligent Manufacturing and Shenzhen Key Laboratory of Control Theory and Intelligent Systems, Southern University of Science and Technology, China
[b] School of Mechanical Engineering and Automation, Harbin Institute of Technology, Shenzhen, China
[c] School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore

A R T I C L E I N F O

Article history:
Received 9 October 2023
Received in revised form 21 April 2024
Accepted 29 June 2024
Available online xxxx

Keywords:
Distributed network localization
Angle rigidity
Performance optimization
Sensor networks
Angle measurements
Distributed algorithm

A B S T R A C T

Recent advances in linear localization of sensor networks allow sensors to localize themselves by using inter-sensor measurements, such as distances, bearings and interior angles. According to earlier works, linear localization algorithms' performance is relatively poor, which, however, has not been adequately addressed in the existing literature. The aim of this paper is to improve the performance of linear and continuous localization algorithms. More specifically, we focus on improving three key aspects of linear localization algorithms' performance, i.e., the stability margin, convergence rate and robustness against measurement noises. Firstly, we propose a unified description for networks' linear localization algorithms, given different types of measurements, and show that the stability margin, convergence rate and robustness of linear localization algorithms are *commonly determined by one parameter*, namely, *the minimum eigenvalue of the network's localization matrix*. Secondly, by carefully choosing the decision variable, we formulate the performance optimization problem as an eigenvalue optimization problem, and show the non-differentiability of the eigenvalue optimization problem. Thirdly, we propose a *distributed* optimization algorithm, which guarantees the convergence to an optimal solution of the eigenvalue optimization problem. Finally, simulation examples validate the effectiveness of the proposed distributed optimization algorithm.

1. Introduction

Sensor networks have been widely used in many engineering scenarios, such as industrial manufacturing and multi-robot coordination (Kleiner & Dornhege, 2007; Zhao, 2018). For a static network consisting of anchor sensors and free sensors, the aim of network localization is to determine/estimate the positions of the free sensors using their sensor measurements with respect to their neighbors and the positions of the anchor sensors, which, from a mathematical point of view, is a typical problem of solving nonlinear equations. To avoid multiple equilibria existing in the nonlinear equations, linear and distributed localization algorithms have been proposed recently in Chen (2022), Diao, Lin, and Fu (2014), Lin, Han, Zheng, and Fu (2016) and Zhao and Zelazo (2016), in which the only equilibrium of the estimating dynamics is the desired solution.

The existing linear localization algorithms can be mainly categorized into three classes according to the available sensor measurements among the sensors: distance-based localization (Diao et al., 2014; Xia, Yu, & He, 2022), bearing-based localization (Bishop, Anderson, Fidan, Pathirana, & Mao, 2009; Cao, Han, Lin, & Xie, 2021; Li, Luo and Zhao, 2019; Zhao & Zelazo, 2016), and angle-based localization (Chen, 2022; Chen, Cao, Xie, Li and Feroskhan, 2022; Fang, Li, & Xie, 2020; Jing, Wan, & Dai, 2021). According to Cao et al. (2021), Chen (2022), Diao et al. (2014) and Lin, Fu, and Diao (2015), under the designed linear localization algorithms, the estimated positions of the free sensors globally and exponentially converge to the true positions. Besides the convergence of linear localization algorithms, other indices of these algorithms' performance, such as robustness, are also worthy of investigation since they play an important role in engineering practices. From existing works (Chen, 2022, 2022; Fang et al., 2020; Lin et al., 2016; Zhao & Zelazo, 2016), the stability margin, convergence rate and robustness are three important performance indices of linear localization algorithms. Firstly, the stability margin not only verifies linear localization algorithms' stability, but also quantifies how much system uncertainties can be tolerated before system instability occurs. According to Chen

☆ The material in this paper was partially presented at the 62nd IEEE Conference on Decision and Control (CDC), December 13–15, 2023, Marina Bay Sands, Singapore. This paper was recommended for publication in revised form by Associate Editor Luca Schenato under the direction of Editor Christos G. Cassandras.

* Corresponding author.
E-mail addresses: chenlm6@sustech.edu.cn (L. Chen), liangchenyang@stu.hit.edu.cn (C. Liang), yibei.li@ntu.edu.sg (Y. Li), jmei@hit.edu.cn (J. Mei), ELHXIE@ntu.edu.sg (L. Xie).

[1] Liangming Chen and Chenyang Liang contributed equally to this work.

https://doi.org/10.1016/j.automatica.2024.111903

(2022), Fang et al. (2020) and Zhao and Zelazo (2016), all the eigenvalues of the defined network's localization matrix determining the localization system's stability are real, which indicates that the value of the stability margin equals the minimum eigenvalue of the localization matrix. Secondly, it is expected that the localization error could converge to zero as fast as possible, which is commonly determined by the minimum eigenvalue of the localization matrix (Chen, 2022; Fang et al., 2020; Zhao & Zelazo, 2016). Thirdly, when considering the existence of measurement noises, the linear localization systems are still exponentially stable, provided that the norm of a defined error matrix is less than the minimum eigenvalue of the localization matrix (Chen, 2022; Fang et al., 2020). This indicates that the robustness of linear localization algorithms against measurement noises also depends on the minimum eigenvalue of the localization matrix. From the aforementioned introduction, linear localization algorithms' stability margin, convergence rate and robustness against measurement noises, are commonly determined by the minimum eigenvalue of the localization matrix. However, the minimum eigenvalue of the localization matrix is associated with the network's topology and sensors' positions, which are fixed. According to the simulation examples in our earlier work (Chen, Xie, Li, Fang and Feroskhan, 2022), the minimum eigenvalue of the angle-based network's localization matrix approximately equals 0.03, which indicates that the proposed angle-based localization algorithm is of poor robustness and low convergence rate, that is, the position estimation error needs more than 300 seconds to decay to 10% of the initial error. Therefore, it is crucial to investigate how to increase the minimum eigenvalue of the localization matrix.

Obviously, increasing the minimum eigenvalue of the localization matrix is an eigenvalue optimization problem. Due to the nonlinearity from the decision variables, such as sensors' positions, to the minimum eigenvalue of the localization matrix, how to increase the minimum eigenvalue of the localization matrix is a challenging problem. Another challenge is the non-differentiability of the eigenvalue optimization problem when the eigenvalue's algebraic multiplicity is greater than one (Overton, 1988, 1992). In our earlier work (Liang, Chen, Li, Mei, & Xie, 2023), we propose a centralized approach to address this problem. Based on it, in this paper we aim to develop a distributed optimization algorithm to solve this problem, which is more applicable to large-scale networks. The main difficulties of developing a distributed algorithm are in two aspects. Firstly, there exist three types of constraints in this eigenvalue optimization problem, i.e., the local feasible set constraint, the linear matrix inequality (LMI) constraint and the vector equality constraint. The presence of multiple constraints increases the optimization problem's complexity, making the algorithm design more challenging. Secondly, the coupling of the problem's cost function and constraints means that sensors must update their decision variables based on both local information and the information related to other sensors, which is challenging for sensor networks without central agents.

Related works on improving specific localization performance of Multi-Agent Systems (MAS) are presented in Schoof, Chapman, and Mesbahi (2017), Sun, Yu, and Anderson (2015), Trinh, Van Tran, and Ahn (2019) and Zhu and Hu (2009). A centralized algorithm is proposed in Trinh et al. (2019) to ensure minimal bearing rigidity of MAS. The concept of stiffness matrix is introduced in Zhu and Hu (2009) to characterize the formation rigidity. A novel Gramian matrix-based approach is proposed in Sun et al. (2015) to enhance the worst-case rigidity level of MAS, using a decentralized algorithm. In Schoof et al. (2017), edge-weighted optimization techniques are proposed, which significantly contributes to ensuring the overall formation performance of the MAS. In contrast to the existing literature that primarily aims to enhance the formation rigidity, our focus in this paper is on improving the linear localization algorithms' performance in terms of convergence rate, stability margin and robustness against noise.

Distributed algorithms for constraint optimization problems have been studied in many works, including (Li, Deng, Zeng and Hong, 2021; Li, Zeng, Hong and Ji, 2021; Nedic, Ozdaglar, & Parrilo, 2010; Van Tran, Sun, Anderson, & Ahn, 2022; Wang, Yang, Guo, Wen, & Huang, 2022). Authors in Van Tran et al. (2022) introduce an innovative approach to formulate the graph matching problem as a convex relaxation problem, replacing the untractable feasible set with a relaxed one, and further present a distributed algorithm to solve it. Refs. Li, Deng et al. (2021) and Li, Zeng et al. (2021) introduce slack variables and consensus constraints to solve optimization problems with coupled LMI constraints. Additionally, Wang et al. (2022) presents an algorithm for optimization problems involving equality constraints. For a comprehensive overview, we refer readers to Yang et al. (2019). However, the algorithms mentioned above, which primarily focus on addressing optimization problems with one or two types of constraints, cannot be directly applied in this paper due to the presence of three types of constraints.

Motivated by the aforementioned discussions, in this paper we propose a distributed continuous-time optimization algorithm to address the eigenvalue optimization problem. Specifically, the projection operator is used to deal with the feasible set constraint. Furthermore, to tackle the coupled LMI and vector equality constraints, we decouple the constraints and then introduce a consensus protocol for the Lagrangian multipliers used in the proposed algorithm. Our approach can tackle a class of optimization problems with different types of constraints, demonstrating its flexibility and applicability. The main contributions of this paper are summarized as follows. Firstly, we propose a unified description for networks' linear localization algorithms, and show that the stability margin, convergence rate and robustness of linear localization algorithms are commonly determined by the minimum eigenvalue of the network's localization matrix. Secondly, by carefully choosing the decision variable, we formulate the performance optimization problem as a constrained eigenvalue optimization problem, and show the non-differentiability of this problem. Thirdly, we propose a distributed optimization algorithm to obtain an optimal solution of the eigenvalue optimization problem.

The remainder of this paper is organized as follows. We formulate the optimization problem in Section 2. Centralized and distributed optimization algorithms are proposed in Section 3 and Section 4, respectively. Section 5 provides simulation examples to validate our algorithms.

2. Problem formulation

2.1. Notations

Consider a two-dimensional and static sensor network consisting of $n_a \in \mathbb{N}$ ($n_a \geq 2$) anchor sensors and $n_f \in \mathbb{N}$ free sensors. The set of anchor sensors is denoted by $\mathcal{V}_a = \{1, 2, \ldots, n_a\}$ with known positions represented by $p_a = [p_1^{\mathrm{T}}, \ldots, p_{n_a}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{2n_a}$. The set of free sensors is represented by $\mathcal{V}_f = \{n_a + 1, \ldots, n\} = \mathcal{V} - \mathcal{V}_a$, where $\mathcal{V}$ denotes the set of all sensors and $|\mathcal{V}_f| = n_f = n - n_a$. The positions of these free sensors, which need to be determined, are denoted by $p_f = [p_{n_a+1}^{\mathrm{T}}, \ldots, p_n^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{2n_f}$. We assume that there are no overlapping points in $p = [p_a^{\mathrm{T}}, p_f^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{2n}$. Define $\sum_g$ as the fixed global coordinate frame, and let each free sensor $i \in \mathcal{V}_f$ hold a fixed and local coordinate frame $\sum_i$ to conduct distance/bearing/angle measurements with respect to its neighbors. Define p_i as the position of sensor i in $\sum_g$ and

p_i^j as the position of sensor i in $\sum_j$. Let I_n denote the n-by-n identity matrix, 1_n the n-by-1 column vector of all ones, $\otimes$ the Kronecker product, $\lambda_{\max}$ the maximum eigenvalue, and $\lambda_{\min}$ the minimum eigenvalue of a real matrix. Denote by $R(\theta)$ the rotation matrix with rotation angle θ. Define $\mathbb{S}^n$ and $\mathbb{S}_+^n$ as the sets of n-by-n symmetric matrices and symmetric and positive semi-definite matrices, respectively. The set of complex numbers is denoted by $\mathbb{C}$. The inner product of matrices is defined by $\langle X_1, X_2\rangle = \mathrm{tr}(X_1^{\mathrm{T}}X_2)$, where $\mathrm{tr}(\cdot)$ denotes the trace of a matrix. Let $\|\cdot\|$ and $\|\cdot\|_F$ be the Euclidean norm of a vector and the Frobenius norm of a matrix, respectively. Let $\mathrm{col}(X_1,\ldots,X_n) \in \mathbb{R}^{n^2\times n}$ be the stacked column matrix of $X_i \in \mathbb{R}^{n\times n}$. Let $\mathrm{diag}(x)$ where $x \in \mathbb{R}^n$ denote a n-by-n diagonal matrix with diagonal elements x.

For a convex set $\mathcal{X} \subset \mathbb{R}^n$, the projection of $x \in \mathbb{R}^n$ onto $\mathcal{X}$ is defined by

$$P_{\mathcal{X}}(x) = \arg\min_{y\in\mathcal{X}} \|x-y\| .$$

While for a symmetric matrix $X \in \mathbb{R}^{n\times n}$ which satisfies the eigenvalue decomposition $X = U^{\mathrm{T}}\mathrm{diag}[\lambda_1,\ldots,\lambda_n]U$, the projection of X onto $\mathbb{S}_+^n$ can be written as

$$P_{\mathbb{S}_+^n}(X) = U\mathrm{diag}[(\lambda_1)_+,\ldots,(\lambda_n)_+]U^{\mathrm{T}},$$

where $(\lambda_i)_+ = \max\{\lambda_i, 0\}$.

According to Li, Zeng et al. (2021), the basic property of the projection operator is

$$\langle X - P_\Omega(X), Y - P_\Omega(X)\rangle \le 0, \tag{1}$$

where $\Omega \subset \mathbb{R}^{n\times n}$, $Y \in \Omega$ and $X \in \mathbb{R}^{n\times n}$.

The first-order convex condition holds for a convex function $F(\cdot) : \mathbb{R}^{m\times n} \to \mathbb{R}$, i.e.,

$$F(Y) - F(X) \ge \langle \nabla F(X), Y - X\rangle , \tag{2}$$

where $X \in \mathbb{R}^{m\times n}$, $Y \in \mathbb{R}^{m\times n}$, and $\nabla F(X) \in \mathbb{R}^{m\times n}$.

2.2. A unified description for measurement-induced linear equations

The distance measurement between sensors $i \in \mathcal{V}$ and $j \in \mathcal{V}$ can be described by

$$d_{ij} = \|p_i - p_j\| \in \mathbb{R}^+. \tag{3}$$

The bearing measurement from sensor i to sensor j is a unit vector, which can be written as

$$b_{ij} = \frac{p_j - p_i}{\|p_j - p_i\|} \in \mathbb{R}^2. \tag{4}$$

The interior angle $\alpha_{kij} \in [0, 2\pi)$ rotating from ray $\overrightarrow{ij}$ to ray $\overrightarrow{ik}$ under the counterclockwise direction can be described by

$$\alpha_{kij} = \begin{cases} \arccos(b_{ij}^{\mathrm{T}}b_{ik}), & \text{if } b_{ij}^{\mathrm{T}}R(\frac{\pi}{2})b_{ik} \ge 0,\\ 2\pi - \arccos(b_{ij}^{\mathrm{T}}b_{ik}), & \text{otherwise,}\end{cases} \tag{5}$$

which can be obtained from local bearing measurements b_{ij}^i and b_{ik}^i. By saying local bearings (resp. bearings), we mean that the measurements are obtained in the sensor's local (resp. aligned) coordinate frame.

Since (3)–(5) are nonlinear with respect to p, finding solution p_f directly from p_a and some inter-sensor distance/bearing/angle measurements is a nonlinear problem (Bishop et al., 2009). Instead, to solve the network localization problem in a linear manner, Diao et al. (2014) associates four sensors' inter-sensor distance measurements together to establish a distance-induced linear equation, i.e.,

$$I_2p_i - a_{ij}(d)I_2p_j - a_{ik}(d)I_2p_k - a_{il}(d)I_2p_l = 0, \tag{6}$$

where $a_{ij}(d) \in \mathbb{R}, a_{ik}(d) \in \mathbb{R}, a_{il}(d) \in \mathbb{R}$ are the barycentric coordinates of p_i with respect to p_j, p_k and p_l, which are only related to distance measurements among the four sensors i, j, k, l (see Diao et al., 2014 for the detailed calculation). Subsequently, we will give the formulation of the bearing-induced linear equation. There are two cases for bearing measurements, namely bearing measurements when sensors have aligned and unaligned coordinate frames, respectively. For the case where sensors share an aligned coordinate frame (Zhao & Zelazo, 2016), the basic unit to establish a bearing-induced linear equation is an edge (i,j), i.e., $P_{b_{ij}}(p_j - p_i) = 0$, where $P_{b_{ij}} = (I_2 - b_{ij}b_{ij}^{\mathrm{T}}) \in \mathbb{R}^{2\times 2}$. While for the case where sensors have unaligned coordinate frames (Lin et al., 2016), the basic unit to establish a bearing-induced linear equation is a triple (i,j,k), i.e., $b'_{ij}(p'_j - p'_i) + b'_{ik}(p'_k - p'_i) = 0$, where $x' = x_1 + ix_2 \in \mathbb{C}$ is defined as the complex expression of $x = [x_1, x_2]^{\mathrm{T}} \in \mathbb{R}^2$. In this paper, we consider the latter case and the former case can be seen as a special form of the latter case. According to the fact that the field of matrices of the special form $\left[\begin{smallmatrix} x_1 & -x_2\\ x_2 & x_1\end{smallmatrix}\right]$ combined by matrix addition and matrix multiplication, is isomorphic to the field of complex numbers x' combined by complex addition and complex multiplication (Ahlfors, 1953, Section 1.3), the bearing-induced linear equation can be reformulated as $B_{ij}(P_j - P_i) + B_{ik}(P_k - P_i) = \mathbf{0}_{2\times 2}$, where $B_{ij} \in \mathbb{R}^{2\times 2}$, $B_{ik} \in \mathbb{R}^{2\times 2}$, $P_i \in \mathbb{R}^{2\times 2}$, $P_j \in \mathbb{R}^{2\times 2}$, $P_k \in \mathbb{R}^{2\times 2}$ denote the matrix representation of b'_{ij}, b'_{ik}, p'_i, p'_j and p'_k, respectively. Multiplying $[1, 0]^{\mathrm{T}}$ on both sides of this equation, one can obtain the bearing-induced linear equation

$$B_{ij}(p_j - p_i) + B_{ik}(p_k - p_i) = 0. \tag{7}$$

Also, an angle-induced linear equation is established in Chen (2022) as

$$\begin{aligned}&\left(\sin\alpha_{jki}I_2 - \sin\alpha_{ijk}R^{\mathrm{T}}(\alpha_{kij})\right)p_i\\ &+ \left(\sin\alpha_{ijk}R^{\mathrm{T}}(\alpha_{kij})\right)p_j - \sin\alpha_{jki}I_2p_k = 0,\end{aligned} \tag{8}$$

whose coefficient matrices are only related to angle measurements α_{jki}, α_{kij} and α_{ijk}.

Before showing the important role of (6)–(8) in the network linear localization, we first describe them in a unified framework. Since (6)–(8) associate with more than two sensors, instead of using graphs consisting of inter-sensor edges to describe the network topology, we define a multigraph $\mathcal{M} = \{\mathcal{U}_1,\ldots,\mathcal{U}_m | m \in \mathbb{N}^+\} \subset \prod_{i=1}^m \mathcal{V}$ to describe the communication and sensing topology of the sensor network, where $\mathcal{U}_l = \{(i_1,\ldots,i_s) | i_1 \ne i_2 \ne \cdots \ne i_s, \{i_1,\ldots,i_s\} \subset \mathcal{V}, s \in \mathbb{N}^+\}$ denotes an associated basic unit of the network topology, $l \in \{1,2,\ldots,m\}$ denotes the index of each unit, and $s \in \mathbb{N}^+$ is the number of vertices in each unit. Then, the sensor network with every type of measurements can be described by a multi-point framework $\mathbb{F}(\mathcal{V}, \mathcal{M}, p)$, where $\mathcal{V}$ is the vertex set with $|\mathcal{V}| = n \in \mathbb{N}^+$, $\mathcal{M}$ is the multigraph with $|\mathcal{M}| = m \in \mathbb{N}^+$, and $p \in \mathbb{R}^{2n}$. Clearly, for distance measurements, the basic unit is a quadruple, while for bearing measurements the basic unit is an edge, and for angle measurements, the basic unit is a triplet. Since each unit $\mathcal{U}_l$ of $\mathcal{M}$ gives one linear equation, one can construct m linear equations from $\mathcal{M}$. The compact form of these m linear equations can be written as $Mp = 0$ where $M \in \mathbb{R}^{2m\times 2n}$ is called as the measurement matrix with the following structure

$$\begin{array}{c} \\ \text{unit } \mathcal{U}_1\\ \cdots\\ \text{unit } \mathcal{U}_l\\ \cdots\\ \text{unit } \mathcal{U}_m\end{array}\begin{array}{c}\begin{array}{ccccccc}\cdots & \text{sensor } i & \cdots & \text{sensor } j & \cdots & \text{sensor } k & \cdots\end{array}\\ \begin{bmatrix} \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ \cdots & A_i^{\mathcal{U}_l} & \cdots & A_j^{\mathcal{U}_l} & \cdots & A_k^{\mathcal{U}_l} & \cdots\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\end{bmatrix}\end{array}, \tag{9}$$

where $\{i, j, k\} \subset \mathcal{V}$, $A_i^{\mathcal{U}_l} \in \mathbb{R}^{2\times 2}$, $A_j^{\mathcal{U}_l} \in \mathbb{R}^{2\times 2}$, $A_k^{\mathcal{U}_l} \in \mathbb{R}^{2\times 2}$ correspond to those *coefficient matrices* established from (6)–(8), which are only related to those available sensor measurements. To clearly show the communication relationship within $\mathbb{F}(\mathcal{V}, \mathcal{M}, p)$, we define an undirected graph $\mathcal{G}_{\mathcal{M}}(\mathcal{V}, \mathcal{E})$ where $(i, j) \in \mathcal{E}$ if $\exists(i_1, \ldots, i_s) \in \mathcal{M}$ and $\{i, j\} \subseteq \{i_1, \ldots, i_s\}$.

2.3. A unified description for linear network localization algorithms

Based on the established linear equations, we now describe the existing linear localization algorithms in a unified form. Following the ingenious formulation of bearing-based linear network localization (Zhao & Zelazo, 2016), one can formulate the general linear localization problem as a least-square optimization problem with the following cost function

$$\begin{aligned} J(\hat{p}_f) = &\sum_{\mathcal{U}_l \in \mathcal{M}} \|A_{i_1}^{\mathcal{U}_l}\hat{p}_{i_1} + A_{i_2}^{\mathcal{U}_l}\hat{p}_{i_2} + \cdots + A_{i_s}^{\mathcal{U}_l}\hat{p}_{i_s}\|^2 \\ =&\hat{p}^{\mathrm{T}} M^{\mathrm{T}} M \hat{p}, \end{aligned} \tag{10}$$

where $\hat{p}_{i_1}$ represents the estimate of sensor i_1's position, $\hat{p}_i = p_i, \forall i \in \mathcal{V}_a$, and the definitions of matrices $A_{i_1}^{\mathcal{U}_l}, \cdots, A_{i_s}^{\mathcal{U}_l}$ can be found in (9). Let $L = M^{\mathrm{T}}M \in \mathbb{R}^{2n\times 2n}$. By partitioning matrix $M = [M_a \ \ M_f]$ into anchor sensors' part $M_a \in \mathbb{R}^{2m\times 2n_a}$ and free sensors' part $M_f \in \mathbb{R}^{2m\times 2n_f}$, the matrix L can be written in the form of

$$L = M^{\mathrm{T}}M = \begin{bmatrix} L_{aa} & L_{af} \\ L_{fa} & L_{ff} \end{bmatrix} \in \mathbb{R}^{2n\times 2n}, \tag{11}$$

where $L_{aa} = M_a^{\mathrm{T}}M_a \in \mathbb{R}^{2n_a\times 2n_a}$, $L_{af} = M_a^{\mathrm{T}}M_f \in \mathbb{R}^{2n_a\times 2n_f}$, $L_{fa} = M_f^{\mathrm{T}}M_a \in \mathbb{R}^{2n_f\times 2n_a}$, and $L_{ff} = M_f^{\mathrm{T}}M_f \in \mathbb{R}^{2n_f\times 2n_f}$. We call L_{ff} the network's localization matrix since it plays a key role in network linear localization. Taking the gradient of (10) along $\hat{p}_f = [\hat{p}_{n_a+1}^{\mathrm{T}}, \ldots, \hat{p}_n^{\mathrm{T}}]^{\mathrm{T}}$ yields the unified localization algorithm

$$\begin{aligned} \dot{\hat{p}}_f(t) =& -\nabla_{\hat{p}_f} J(\hat{p}_f) = -L_{ff}\hat{p}_f(t) - L_{fa}p_a \\ =& -L_{ff}(\hat{p}_f(t) - p_f), \end{aligned} \tag{12}$$

where $p_f = -L_{ff}^{-1}L_{fa}p_a$ since $L_{ff}p_f + L_{fa}p_a = 0$, and L_{ff} is nonsingular if and only if the corresponding network is localizable (Chen, 2022; Diao et al., 2014; Lin et al., 2016). Writing (12) into component forms under different measurement cases easily verifies that (12) is distributed (Chen, 2022; Diao et al., 2014; Zhao & Zelazo, 2016).

When considering the existence of measurement noises, we denote the error matrices due to measurement noises by ΔL_{ff} and ΔL_{fa}, and define $\hat{L}_{ff} = L_{ff} + \Delta L_{ff}$ and $\hat{L}_{fa} = L_{fa} + \Delta L_{fa}$. Then, the distributed localization algorithm (12) becomes

$$\begin{aligned} \dot{\hat{p}}_f(t) =& -\hat{L}_{ff}\hat{p}_f(t) - \hat{L}_{fa}p_a \\ =& -\hat{L}_{ff}\left(\hat{p}_f(t) + \hat{L}_{ff}^{-1}\hat{L}_{fa}p_a\right), \end{aligned} \tag{13}$$

which holds if $\hat{L}_{ff}$ is nonsingular.

Now, we give the condition for the nonsingularity of $\hat{L}_{ff}$.

Lemma 1 (*Chen, Cao et al., 2022; Fang et al., 2020; Lin et al., 2016; Zhao & Zelazo, 2016*). *Given a localizable network $\mathbb{F}(\mathcal{V}, \mathcal{M}, p)$ with nonsingular L_{ff}, then $\hat{L}_{ff}$ is nonsingular if the error matrix ΔL_{ff} satisfies*

$$\left\|\Delta L_{ff}\right\| < \lambda_{\min}\left(L_{ff}\right). \tag{14}$$

Proposition 2 (*Chen, 2022; Diao et al., 2014; Zhao & Zelazo, 2016*). *For a sensor network $\mathbb{F}(\mathcal{V}, \mathcal{M}, p)$ with continuous localization algorithm (12),*

(a) the system (12) is asymptotically stable iff the network is localizable which holds iff L_{ff} is nonsingular;
(b) the convergence speed of localization error $\tilde{p}_f(t) = \hat{p}_f(t) - p_f$ is higher if $\lambda_{\min}(L_{ff})$ is larger;
(c) the robustness of (13) against measurement noises is higher if $\lambda_{\min}(L_{ff})$ is larger.

Proposition 2 indicates that a larger $\lambda_{\min}(L_{ff})$ will make positive effects on the performance of linear localization algorithms. Proposition 2 serves as a general result for linear localization approaches based on different types of measurements (Chen, 2022; Diao et al., 2014; Zhao & Zelazo, 2016). This is because all these localization approaches in the literature are developed based on the corresponding measurement-induced linear equations (6), (7), (8) and the gradient descent of their least square errors (12).

3. Centralized eigenvalue optimization algorithm

3.1. Selection of the decision variable

It is expected to maximize $\lambda_{\min}(L_{ff})$ since a larger $\lambda_{\min}(L_{ff})$ implies better performance of linear localization algorithms according to Proposition 2. Note that if $A_{i_1}^{\mathcal{U}_l}p_{i_1} + A_{i_2}^{\mathcal{U}_l}p_{i_2} + \cdots + A_{i_s}^{\mathcal{U}_l}p_{i_s} = 0$, $\mathcal{U}_l \in \mathcal{M}$ is a measurement-induced linear equation, $\sqrt{\beta_l}A_{i_1}^{\mathcal{U}_l}p_{i_1} + \sqrt{\beta_l}A_{i_2}^{\mathcal{U}_l}p_{i_2} + \cdots + \sqrt{\beta_l}A_{i_s}^{\mathcal{U}_l}p_{i_s} = 0, \forall \beta_l \in \mathbb{R}^+$ is also a measurement-induced linear equation of the network, which provides us a freedom to adjust the values of $\beta_l, l = 1, \ldots, m$ such that $\lambda_{\min}(L_{ff})$ can be improved. Specifically, since $Mp = 0$ are linear equations of $\mathbb{F}$,

$$(\mathrm{diag}[\sqrt{\beta_1}, \ldots, \sqrt{\beta_m}] \otimes I_2)Mp = 0 \tag{15}$$

are also linear equations of $\mathbb{F}$. From (15), the new L_{ff} is modified to

$$\begin{aligned} L_{ff}(\beta) =& \left(\left(\mathrm{diag}[\sqrt{\beta_1}, \ldots, \sqrt{\beta_m}] \otimes I_2\right)M_f\right)^{\mathrm{T}} \\ &\cdot \left(\mathrm{diag}[\sqrt{\beta_1}, \ldots, \sqrt{\beta_m}] \otimes I_2\right)M_f \\ =& \beta_1 e_1^{\mathrm{T}} e_1 + \beta_2 e_2^{\mathrm{T}} e_2 + \cdots + \beta_m e_m^{\mathrm{T}} e_m \\ =& \beta_1 E_1 + \beta_2 E_2 + \cdots + \beta_m E_m, \end{aligned} \tag{16}$$

where $M_f = [e_1^{\mathrm{T}}, \ldots, e_m^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{2m\times 2n_f}$, and $e_i \in \mathbb{R}^{2\times 2n_f}$. Note that a normalization is needed for $\beta = [\beta_1, \ldots, \beta_m]^{\mathrm{T}} \in \mathbb{R}^m$ to guarantee the feasibility of the problem, namely $\sum_{i=1}^m \beta_i = 1$. Then the problem to be solved is formulated as follows:

$$\begin{aligned} &\max_{\beta} \ \lambda_{\min}(L_{ff}(\beta)) = \max_{\beta} \ \lambda_{\min}\Big(\sum_{i=1}^m \beta_i E_i\Big) \\ &\ \text{s.t. } \beta_i > 0, \sum_{i=1}^m \beta_i = 1. \end{aligned} \tag{17}$$

In the above, β is the weight of units in the network $\mathbb{F}(\mathcal{V}, \mathcal{M}, p)$. To optimize the weight of units for localization, we introduce a normalized constraint on the sum of β_i. Notably, the optimal value of β_i maintains proportionality when the sum of β_i takes different values. In other words, if the constraint is changed to $\sum_{i=1}^m \beta_i = c > 0$, the optimal weight becomes $\beta_i = c\beta_i^*$, and the optimal eigenvalue $\lambda_i = c\lambda_i^*, \forall i = 1, \ldots, m$, where β_i^*, λ_i^* are the optimal weight and eigenvalue of (17), respectively. This indicates the trivial extension from the current results to the scenario where $\sum_{i=1}^m \beta_i$ has different values.

Remark 3. Directly increasing the gain k_c of the localization law (13), denoted as $\dot{\hat{p}}_f(t) = -k_c\hat{L}_{ff}\left(\hat{p}_f(t) + \hat{L}_{ff}^{-1}\hat{L}_{fa}p_a\right)$ is not a suitable option for improving system performance due to two main reasons, despite its ability to improve the convergence rate.

Firstly, raising k_c leads to a larger variance of the localization error, compromising robustness. This conflicts with our goal of simultaneously improving both convergence rate and robustness. Secondly, introducing k_c is a special case of (17), where $\beta_1 = \cdots = \beta_m = k_c$. However, under the constraint $\sum_{i=1}^{m}\beta_i = mk_c$, this is not the optimal solution, which implies suboptimal energy utilization when considering $\sum_{i=1}^{m}\beta_i = mk_c$ as an energy constraint. Additionally, finite-time localization algorithms, commonly used, show two drawbacks. Firstly, they involve a large magnitude of gain (Galicki, 2015) in the localization law, whose side effects have been mentioned above. Secondly, their impact on estimation dynamics under measurement noise is unknown due to nonlinearity from measurements to overall dynamics. Specifically, while a nonlinear finite-time localization law may perform well without noise, it may diverge with the presence of noise (Aldana-López, Seeber, Haimovich, & Gómez-Gutiérrez, 2023; Orlov, Kairuz, & Aguilar, 2021).

3.2. Multiplicity analysis for $\lambda(L_{ff}(\beta))$

In this subsection, we analyze the algebraic multiplicity of $\lambda_{\min}(L_{ff}(\beta))$. We first present the following lemma.

Lemma 4. *For $\beta_i > 0, i = 1, \ldots, m$, $L_{ff}(\beta)$ can be described by*

$$L_{ff}(\beta) = \begin{bmatrix} a_1 I_2 & b_{12}R(\theta_{12}) & \cdots & b_{1n_f}R(\theta_{1n_f}) \\ b_{12}R^T(\theta_{12}) & a_2 I_2 & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ b_{1n_f}R^T(\theta_{1n_f}) & \cdots & \cdots & a_{n_f} I_2 \end{bmatrix}, \tag{18}$$

where $a_i = \sum_{\mathcal{U}_l \in \mathcal{M}} \beta_l d_i^{\mathcal{U}_l}, \forall i = 1, \ldots, n_f$, $d_i^{\mathcal{U}_l} \in \mathbb{R}^+$, $b_{ij} \in \mathbb{R}^+$ and $\theta_{ij} \in \mathbb{R}$ denote the corresponding coefficient and rotation angle after normalizing $\sum_{\mathcal{U}_l \in \mathcal{M}} \beta_l d_{ij}^{\mathcal{U}_l} R(\theta_{ij}^{\mathcal{U}_l})$, $d_{ij}^{\mathcal{U}_l} \in \mathbb{R}^+$, $\theta_{ij}^{\mathcal{U}_l} \in \mathbb{R}$, $i \neq j$, respectively.

The proof of Lemma 4 can be found in Appendix A. Lemma 4 provides a unified description for $L_{ff}(\beta)$ of different linear localization algorithms. Based on Lemma 4, we present the following results.

Theorem 5 (*Liang et al., 2023, Theorem 2*). *Suppose that $\beta_i > 0, \forall i = 1, \ldots, m$. The algebraic multiplicity of $\lambda_j(L_{ff}(\beta))$ is equal to the geometric multiplicity of $\lambda_j(L_{ff}(\beta))$, $\forall j = 1, \ldots, 2n_f$, which is always an even number.* □

From Theorem 5, one can derive that the algebraic multiplicity of $\lambda_{\min}\left(L_{ff}(\beta)\right)$ is always at least two, which indicates that $\lambda_{\min}\left(L_{ff}(\beta)\right)$ is non-differentiable according to Overton (1988, 1992). We refer readers to our CDC paper (Liang et al., 2023) for the proof of the theorem.

3.3. Eigenvalue optimization approach

To overcome the difficulty due to the non-differentiability of (17), we transform the problem into a SDP (Helmberg, Rendl, Vanderbei, & Wolkowicz, 1996; Jarre, 1993), i.e.,

$$\begin{aligned} &\min_{\lambda,\beta} \lambda \\ &\text{s.t } \lambda I + \sum_{i=1}^{m}\beta_i E_i \succeq \mathbf{0}, \\ &\quad \beta_i > 0, i = 1, \ldots, m, \\ &\quad \sum_{i=1}^{m}\beta_i = 1, \end{aligned} \tag{19}$$

where $\beta = [\beta_1, \ldots, \beta_m]^{\mathrm{T}} \in \mathcal{B}$, which denotes the feasible set of β. In more detail, let $\mathcal{B} = \prod_{i=1}^{m}\mathcal{B}_i$, where $\mathcal{B}_i = \{0 < \beta_i < 1\}$. Here we use a relaxed convex set $\mathcal{B}$ motivated by Van Tran et al. (2022) due to the intractability of the exact set. We have the following result.

Lemma 6. *The set of solutions of the problem (19) is nonempty and compact.*

The proof of Lemma 6 can be found in Appendix B. Lemma 6 indicates that there always exists an optimal solution of the problem (19), which shows that the problem is meaningful to solve.

We use the interior-point algorithm (Boyd & Vandenberghe, 2004; Helmberg et al., 1996; Jarre, 1993) to solve the problem (19), since this algorithm is effective to solve eigenvalue optimization problems with inequality constraints. By utilizing the logarithmic barrier functions $\log(\cdot)$ and $\log\det(\cdot)$, we approximately reformulate the equality and inequality constrained optimization problem (19) as an equality constrained problem

$$\begin{aligned} &\min_{x} wf_c(x) + \phi(x) \\ &\text{s.t. } b_c^{\mathrm{T}}x = 1, \end{aligned} \tag{20}$$

where

$$f_c(x) = a_c^{\mathrm{T}}x,$$

$$\phi(x) = -\sum_{i=1}^{m}\log x_i - \log\det E(x),$$

$$E(x) = \lambda I + \sum_{i=1}^{m}\beta_i E_i.$$

Here, $x = \left[\beta^{\mathrm{T}}, \lambda\right]^{\mathrm{T}} \in \mathbb{R}^{m+1}$, $a_c = [0, \ldots, 0, 1]^{\mathrm{T}} \in \mathbb{R}^{m+1}$, $b_c = [1, \ldots, 1, 0]^{\mathrm{T}} \in \mathbb{R}^{m+1}$, and $w \in \mathbb{R}^+$. Denote the domain of $\phi(x)$ by $\operatorname{dom}\phi = \{x \in \mathbb{R}^{m+1} | x_i > 0, i = 1, \ldots, m, x_{m+1} < 0, E(x) \succeq \mathbf{0}\}$. Note that if x is a strictly feasible solution of (20), namely, $x \in \operatorname{dom}\phi \cap \left\{x | x \in \mathbb{R}^{m+1}, b_c^{\mathrm{T}}x = 1\right\}$, it also satisfies all the constraints in (19). The cost function $wf_c(x) + \phi(x)$ is convex since the sum of convex functions $wf_c(x)$ and $\phi(x)$ is also convex, which indicates that the problem (20) has a global optimal solution. Besides, since the cost function in (20) is second-order differentiable due to the second-order differentiability of $wf_c(x)$ and $\phi(x)$, we can use Newton's methods (Boyd & Vandenberghe, 2004, Section 9.5) to solve (20) when w is fixed. The scalar w denotes a weighted parameter, which is updated according to

$$w_{k+1} = \mu w_k, \tag{21}$$

where $\mu > 1$ and $k = 1, 2, \ldots$ denotes the iteration sequence number. The choice of the parameter μ involves a trade-off in the numbers of inner and outer steps required in Algorithm 1. According to Boyd and Vandenberghe (2004, Section 11.3), for μ in a range from around 3 to 100 or so, the total number of steps, namely the result of the number of the inner steps multiplies the number of the outer steps, remains approximately constant, which indicates that choice of μ is not particularly critical.

The complete process of the interior-point method is shown in Algorithm 1. At each iteration k, we compute the central point $x^*(w_k)$ starting from the previously computed central point by utilizing the Newton's method to solve (20) when $w = w_k$. Then we calculate w_{k+1} following (21). The scalar $\epsilon \in \mathbb{R}^+$ denotes the accurate threshold and the stop criterion $\frac{m_c}{w_k} < \epsilon$, where m_c denotes the number of inequality constraints in (19).

Denote the optimal value of $f_c(x)$ in (19) by f^*. Note that x^* is the optimal solution of (20), which is an approximation problem of (19). The gap between f^* and $f(x^*)$ can be described by the following proposition. Let $m_c = m+1$ denote the number of terms $\log(\cdot)$ in $\phi(x)$ of (20).

Algorithm 1 Interior-point method to solve (20)

Require: Strictly feasible x, initial $w_0 > 0$, factor μ, and tolerance ϵ.

for $k = 1, 2, \cdots$ **do**
  Compute the central point $x^*(w_k)$ by minimizing (20), starting at x.
  Update x by $x = x^*(w_k)$.
  if $\frac{m_c}{w_k} < \epsilon$ **then**
    Quit the loop.
  end if
  Calculate w_{k+1} following (21).
end for
return The optimization solution x to the problem (19).

Proposition 7 (Boyd & Vandenberghe, 2004, Section 11.2). *Consider the optimization problem (20). Assume $w > 0$, $\mu > 1$ and $\epsilon > 0$. For every strictly feasible initial $x(w_0)$, under Algorithm 1, the gap associated with $x^*(w)$ is described as*

$$f_c\left(x^*(w)\right) - f_c^* \le \frac{m_c}{w}, \tag{22}$$

i.e., $x^*(w)$ is no more than $\frac{m_c}{w}$-suboptimal.

Proposition 7 shows that the gap $\frac{m_c}{w}$ converges to zero when w goes to infinity, which indicates that the optimal solution $x^*(w)$ of (20) is close enough to that of (19) when the gap $\frac{m_c}{w}$ is less than the specified accurate threshold ϵ.

Remark 8. We give a complexity analysis for Algorithm 1. We observe that the expected threshold ϵ is achieved after $\left\lceil \frac{\log(m_c/\epsilon w_0)}{\log \mu} \right\rceil$ centering steps. This determination arises from solving the inequality $\frac{m_c}{w} \le \epsilon$, where w is defined in (21) and $\frac{m_c}{w}$ represents the actual accuracy bound given in Proposition 7. According to Boyd and Vandenberghe (2004, Section 11.5.2), the required number of Newton steps at every centering step is $\frac{m_c(\mu-1-\log\mu)}{\gamma_1} + \gamma_2$, where $\gamma_1 \in \mathbb{R}, \gamma_2 \in \mathbb{R}$ are some constants. Then the total number of Newton steps is $\left\lceil \frac{\log((m+1)/\epsilon w_0)}{\log \mu} \right\rceil \left(\frac{(m+1)(\mu-1-\log\mu)}{\gamma_1} + \gamma_2 \right)$, which indicates that the total number of Newton steps increases with m, i.e., the number of units in the network, increases.

4. Distributed eigenvalue optimization algorithm

In this section, we present a distributed eigenvalue optimization approach to obtain the optimal β^*. To develop a distributed algorithm, we need to decouple (19) while maintaining equivalence with the original problem. Firstly, we define the decision variable of unit i as $x_i = [\lambda_i, \beta_i]^{\mathrm{T}}$ with the feasible set $\mathcal{X}_i = \left\{x_i \in \mathbb{R}^2 | \lambda_i \in \mathbb{R}, \beta_i \in \mathcal{B}_i\right\}$, $\forall i \in \mathcal{M}$. Letting $\lambda = \sum_{i=1}^m \lambda_i$ results in the decomposition of the cost function into m separable local functions, namely $\sum_{i=1}^m a^{\mathrm{T}} x_i$, where $a = [1, 0]^{\mathrm{T}} \in \mathbb{R}^2$. With the introduction of x_i, the LMI in (19) becomes $\sum_{i=1}^m (\lambda_i I + \beta_i E_i) \succeq \mathbf{0}$, successfully decoupling this constraint. Lastly, the equality constraint $\sum_{i=1}^m \beta_i = 1$ is equivalently expressed as $\sum_{i=1}^m c^{\mathrm{T}} x_i - 1_{i=1}(i)$, where $c = [0, 1]^{\mathrm{T}} \in \mathbb{R}^2$, $1_{i=1}(i)$ denotes the indicator function, i.e., $1_{i=1}(i) = 1$ if $i = 1$ and $1_{i=1}(i) = 0$ otherwise. After the decomposition, the optimization problem is transformed into

$$\begin{aligned} \min_{\boldsymbol{x} \in \mathcal{X}} \quad & \sum_{i=1}^m a^{\mathrm{T}} x_i \\ \text{s.t.} \quad & \sum_{i=1}^m (\lambda_i I + \beta_i E_i) \succeq \mathbf{0}, \\ & \sum_{i=1}^m \left(c^{\mathrm{T}} x_i - 1_{i=1}(i)\right) = 0, \end{aligned} \tag{23}$$

where $\boldsymbol{x} = [x_1^{\mathrm{T}}, \ldots, x_m^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{2m}$, and $\mathcal{X} = \prod_{i=1}^m \mathcal{X}_i \in \mathbb{R}^{2m}$. The optimization problem (23) is effectively decoupled while preserving its equivalence with the original problem (19). It is noteworthy that no unit needs to be aware of its index i. We just need to formulate the equality constraint of arbitrary a unit i as $h_i(x_i) = c^{\mathrm{T}} x_i - 1$, while the equality constraints of other units are denoted as $h_j(x_j) = c^{\mathrm{T}} x_j, j \in \mathcal{M} \setminus \{i\}$, before deploying the sensor network to a practice scenario.

4.1. Distributed optimization algorithm design

For notation simplicity, let $f(\boldsymbol{x}) = \sum_{i=1}^m f_i(x_i) = \sum_{i=1}^m a^{\mathrm{T}} x_i$, $G(\boldsymbol{x}) = \sum_{i=1}^m G_i(x_i) = \sum_{i=1}^m -B_i^{\mathrm{T}}(x_i \otimes I_{2n_f})$, where $B_i = \left[I_{2n_f}, E_i^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{R}^{4n_f \times 2n_f}$, and $h(\boldsymbol{x}) = \sum_{i=1}^m h_i(x_i) = \sum_{i=1}^m \left(c^{\mathrm{T}} x_i - 1_{i=1}(i)\right)$. The Lagrange function $L_1(\boldsymbol{x}, R, \nu)$ for (23) is given by

$$L_1(\boldsymbol{x}, R, \nu) = \sum_{i=1}^m f_i(x_i) + \sum_{i=1}^m \langle R, G_i(x) \rangle + \nu \sum_{i=1}^m h_i(x), \tag{24}$$

where $R \in \mathbb{S}_+^{2n_f}$ and $\nu \in \mathbb{R}$ are the Lagrange multipliers.

Lemma 9. *There exists a dual solution (R^*, ν^*) of the problem (24).*

Proof. Obviously, (23) is convex and satisfies Slater's condition, ensuring strong duality (Boyd & Vandenberghe, 2004, Section 5.2.3). This indicates the existence of a dual solution (R^*, ν^*) such that $f(\boldsymbol{x}^*) = L_1(\boldsymbol{x}^*, R^*, \nu^*)$, where $x^* = \arg\min_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x})$. $\square$

For such a convex optimization problem (23) with zero duality gap, based on Lemma 9, finding its optimal solution is equivalent to seeking the saddle point $(\boldsymbol{x}^*, R^*, \nu^*)$ of (24).

Lagrange multipliers R and ν in (24) serve as global dual variables, associated with decision variables of all units. However, each unit is unable to directly access the values of R and ν as there is no central agent. To solve it, we construct a modified Lagrange function, i.e.,

$$\begin{aligned} & L_2(\boldsymbol{x}, \boldsymbol{R}, \boldsymbol{\nu}) = \sum_{i=1}^m \left(f_i(x_i) + \langle R_i, G_i(x_i) \rangle + \nu_i h_i(x_i)\right) \\ & \text{s.t.} \quad R_i = R_j, \nu_i = \nu_j, \forall i, j \in \mathcal{M}, \end{aligned} \tag{25}$$

where $\boldsymbol{R} = [R_1^{\mathrm{T}}, \ldots, R_m^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{2mn_f \times 2n_f}$, $\boldsymbol{\nu} = [\nu_1, \ldots, \nu_m]^{\mathrm{T}} \in \mathbb{R}^m$. The dual variables R_i and ν_i of unit i are local estimated Lagrangian multipliers for R and ν, respectively. Consensus constraints for R_i and ν_i are introduced and once satisfied, we can say that R_i and ν_i converge to R and ν, respectively. Thus, the optimal solution of (25) equals to that of (24).

The distributed optimization algorithm for solving (25) is designed as

$$\begin{aligned} \dot{\boldsymbol{x}}(t) &= 2k_1(\hat{\boldsymbol{x}} - \boldsymbol{x}), \\ \dot{\boldsymbol{R}}(t) &= k_1(\hat{\boldsymbol{R}} - \boldsymbol{R}), \\ \dot{\boldsymbol{\nu}}(t) &= k_1(\hat{\boldsymbol{\nu}} - \boldsymbol{\nu}), \\ \dot{\boldsymbol{U}}(t) &= k_2 L_{\boldsymbol{R}} \boldsymbol{R}, \\ \dot{\boldsymbol{\eta}}(t) &= k_2 L \boldsymbol{\nu}, \end{aligned} \tag{26}$$

where $\hat{\boldsymbol{x}} = P_{\mathcal{X}}\left(\boldsymbol{x} - \nabla_{\boldsymbol{x}} L_2(\boldsymbol{x}, \boldsymbol{R} + \dot{\boldsymbol{R}}, \boldsymbol{\nu} + \dot{\boldsymbol{\nu}})\right)$, $\hat{\boldsymbol{R}} = P_{\left(\mathbb{S}_+^{2n_f}\right)^m}\left(\boldsymbol{R} + \nabla_{\boldsymbol{R}} L_2(\boldsymbol{x}, \boldsymbol{R}, \boldsymbol{\nu}) - L_{\boldsymbol{R}} \boldsymbol{R} - L_{\boldsymbol{R}} \boldsymbol{U}\right)$, $\hat{\boldsymbol{\nu}} = P_{\mathbb{R}^m}\left(\boldsymbol{\nu} + \nabla_{\boldsymbol{\nu}} L_2(\boldsymbol{x}, \boldsymbol{R}, \boldsymbol{\nu}) - L\boldsymbol{\nu} - L\boldsymbol{\eta}\right)$, $L_{\boldsymbol{R}} = L \otimes I_{2n_f}$, L denotes the corresponding Laplace matrix of the graph $\mathcal{G}_{\mathcal{M}}$, $\boldsymbol{U} = [U_1^{\mathrm{T}}, \ldots, U_m^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{2mn_f \times 2n_f}$, $\boldsymbol{\eta} = [\eta_1, \ldots, \eta_m]^{\mathrm{T}} \in \mathbb{R}^m$, $\left(\mathbb{S}_+^{2n_f}\right)^m = \prod_{i=1}^m \mathbb{S}_+^{2n_f}$, and $k_1, k_2 \in \mathbb{R}$ are constant gains to be specified. The component form of (26) for each unit $i \in \mathcal{M}$ is

$$\begin{aligned} \dot{x}_i(t) &= 2k_1\left(\hat{x}_i - x_i\right), \\ \dot{R}_i(t) &= k_1\big(\hat{R}_i - R_i\big), \\ \dot{\nu}_i(t) &= k_1\left(\hat{\nu}_i - \nu_i\right), \\ \dot{U}_i(t) &= k_2 \sum_{j\in\mathcal{N}_i}\left(R_i - R_j\right), \\ \dot{\eta}_i(t) &= k_2 \sum_{j\in\mathcal{N}_i}\left(\nu_i - \nu_j\right), \end{aligned} \tag{27}$$

where $\hat{x}_i = 2k_1 P_{\mathcal{X}_i}\big(x_i - a + \left[\,\langle\, I_{2n_f}, R_i + \dot{R}_i \,\rangle\,, \langle\, E_i, R_i + \dot{R}_i \,\rangle\,\right]^{\mathrm{T}} - (\nu_i + \dot{\nu}_i)\, c\big)$, $\hat{R}_i = P_{\mathbb{S}_+^{2n_f}}\big(R_i - B_i^{\mathrm{T}}(x_i \otimes I_{2n_f}) - \sum_{j\in\mathcal{N}_i}(R_i - R_j) - \sum_{j\in\mathcal{N}_i}(U_i - U_j)\big)$, $\hat{\nu}_i = \big(\nu_i + c^{\mathrm{T}}x_i - 1_{i=1}(i) - \sum_{j\in\mathcal{N}_i}\left(\nu_i - \nu_j\right) - \sum_{j\in\mathcal{N}_i}\left(\eta_i - \eta_j\right)\big)$, and $\mathcal{N}_i$ denotes the neighbor set of unit i.

Indeed, algorithm (26) can be viewed as a projected primal–dual algorithm. To handle the feasible set constraint, we introduce the projection operators $P_{\mathcal{X}}$ and $P_{(\mathbb{S}_+^{2n_f})^m}$ for $\boldsymbol{x}$ and $\boldsymbol{R}$, respectively. Referring to Antipin (1994, Theorem 1), damping terms $\dot{\boldsymbol{R}}$ and $\dot{\boldsymbol{\nu}}$ are included in the update law of $\boldsymbol{x}$ to ensure the convergence of dynamics (26) from a control perspective. The auxiliary variables $\boldsymbol{U}$ and $\boldsymbol{\eta}$ represent integral terms of $\boldsymbol{R}$ and $\boldsymbol{\nu}$, respectively, which are necessary. That's because substituting $\boldsymbol{U}$ into the update law of $\boldsymbol{R}$ reveals two components: one is $\nabla_{\boldsymbol{R}} L_2(\boldsymbol{x}, \boldsymbol{R}, \boldsymbol{\nu})$ used to maximize the Lagrange function (25), and the other is the proportional-integral (PI) controller ensuring each local estimation R_i dynamically tracks the true dual variable R in (24) and achieves consensus. The role of $\boldsymbol{\eta}$ is analogous.

Remark 10. It is worth mentioning that our algorithm has two different aspects comparing to the existing works. Firstly, the algorithm proposed in this paper can deal with the optimization problem with three different types of constraints, simultaneously, which can be uncoupled and coupled, vector and matrix constraints, demonstrating its flexibility and applicability. Secondly, existing works such as Li, Deng et al. (2021) and Li, Zeng et al. (2021) decouple LMI constraints using high-dimensional slack variables. The introduction of slack variables pose a significant burden on storage and computation for sensors. In contrast, our approach avoids it, thus easing burdens on sensors.

Remark 11. Here we discuss the identification, interaction and scalability of this algorithm. For the identification of basic units, each sensor only needs to establish communication links and conduct measurements with respect to its neighbors that form a unit, which can be done in a distributed manner (Chen, Lin, & Xie, 2024) or in the network design stage to ensure the network localizability (Chen, 2022). Since the interaction is among units and each unit consists of multiple sensors, the communication between units is similar to that among routers. Thus, those low-level interacting mechanisms from communication engineering can be an alternative here for implementing the inter-unit communication. Based on the definition of scalability in Rana and Stout (2000), our algorithm is scalable since the proposed algorithm for each unit only depends on the communication with its neighbor units, no matter how many number of units in the network. It is worth mentioning that the overall localization system's performance is a property of the whole network, rather than a local property and that is why the algorithm for performance optimization requires to compute eigenvalues of the localization matrix with its dimension scaling up with the number of free sensors.

4.2. Optimality and convergence analysis

According to Le, Chen, Li, Yan, and Xi (2019) and Li, Zeng et al. (2021), the graph $\mathcal{G}_{\mathcal{M}}$ needs to be connected if we expect the algorithm (26) works. Then we give the following assumption.

Assumption 12. The graph $\mathcal{G}_{\mathcal{M}}$ is connected and undirected.

The relationship between the optimal solution of problem (25) and the equilibrium point of dynamics (26) is addressed in the following theorem.

Theorem 13. *Under Assumption 12, $\boldsymbol{x}^*$ is an optimal solution of (23) iff there exists $(\boldsymbol{R}^*, \boldsymbol{\nu}^*, \boldsymbol{U}^*, \boldsymbol{\eta}^*)$ such that $(\boldsymbol{x}^*, \boldsymbol{R}^*, \boldsymbol{\nu}^*, \boldsymbol{U}^*, \boldsymbol{\eta}^*)$ is an equilibrium point of (26).*

The proof of this theorem is provided in Appendix C. As stated in Theorem 13 we know that finding the equilibrium point of (26) is equivalent to finding the optimal solution of (25). Here, we establish the following theorem on the convergence of dynamics (26).

Theorem 14. *Suppose Assumption 12 holds. If the constant gains k_1 and k_2 satisfy*

$$\frac{k_1}{k_2} \geq \frac{\lambda_{\max}(L)}{2}, \tag{28}$$

where $\lambda_{\max}(L)$ denotes the largest eigenvalue of the Laplacian matrix L, the trajectory $(\boldsymbol{x}(t), \boldsymbol{R}(t), \boldsymbol{\nu}(t), \boldsymbol{U}(t), \boldsymbol{\eta}(t))$ converges to an equilibrium point $(\boldsymbol{x}^*, \boldsymbol{R}^*, \boldsymbol{\nu}^*, \boldsymbol{U}^*, \boldsymbol{\eta}^*)$ under (26).

The proof of Theorem 14 can be found in Appendix D. Theorems 13 and 14 indicate that under the proposed algorithm (26), one can find the optimal solution of (25).

Next, the convergence rate of the algorithm (26) is discussed. It is difficult to analyze the convergence rate of (26) directly, since the existence of various constraints and non-smooth terms. Referring to Li, Xie and Hong (2019), we provide an indirect convergence rate.

Theorem 15. *Assume the conditions in Theorem 13 hold, then*

$$\begin{aligned} &\lim_{t\to\infty} \inf\left(\left\|\dot{\boldsymbol{x}}\right\| + \left\|\dot{\boldsymbol{R}}\right\|_F + \left\|\dot{\boldsymbol{U}}\right\|_F + \left\|\dot{\boldsymbol{\nu}}\right\| + \left\|\dot{\boldsymbol{\eta}}\right\|\right) \\ =&O\Big(\frac{1}{\sqrt{t}}\Big). \end{aligned} \tag{29}$$

The proof of Theorem 15 can be found in Appendix E.

5. Simulation results

This section presents simulation examples to validate the effectiveness of the proposed algorithms.

The sensor network is an angle-based network shown in Fig. 1, which consists of $m = 13$ triangles. The form of the triangles is (3, 6, 11), (4, 6, 7), (6, 7, 9), (4, 5, 7), (5, 7, 8), (7, 8, 10), (1, 9, 10), (2, 4, 5), (1, 8, 10), (3, 6, 9), (2, 5, 8), (6, 11, 4) and (4, 6, 12). The positions of three anchor sensors are $p_1 = [2.5, -21.7]^{\mathrm{T}}$, $p_2 = [13.7, 10.6]^{\mathrm{T}}$ and $p_3 = [-19.3, 0]^{\mathrm{T}}$. The positions of free sensors are $p_4 = [-2.0, 16.8]^{\mathrm{T}}$, $p_5 = [4.5, 10.5]^{\mathrm{T}}$, $p_6 = [-13.4, 1.0]^{\mathrm{T}}$, $p_7 = [0.6, -0.3]^{\mathrm{T}}$, $p_8 = [9.2, 0.5]^{\mathrm{T}}$, $p_9 = [-8.7, -15.4]^{\mathrm{T}}$, $p_{10} = [0, -8.3]^{\mathrm{T}}$, $p_{11} = [-7.7, 12.7]^{\mathrm{T}}$ and $p_{12} = [-8.3, 9.9]^{\mathrm{T}}$. The initial decision variables are $x_i(0) = [\lambda_i, \beta_i]^{\mathrm{T}} = [0, \frac{1}{m}]^{\mathrm{T}}$, and the dual variables are $R_i(0) = I_{2n_f}$ and $\nu_i(0) = 1$, $\forall i \in \mathcal{M}$. The gain coefficients are specified as $k_1 = 1.2$, $k_2 = 0.03$.

Through Algorithm 1, we obtain the optimal solution

$$\begin{aligned} \beta^* = [\,&0.053323\ 0.060983\ 0.001535\ 0.090534\ 0.044165 \\ &0.010316\ 0.047960\ 0.056141\ 0.057161\ 0.109001 \\ &0.092911\ 0.078914\ 0.297038\,]^{\mathrm{T}}. \end{aligned} \tag{30}$$

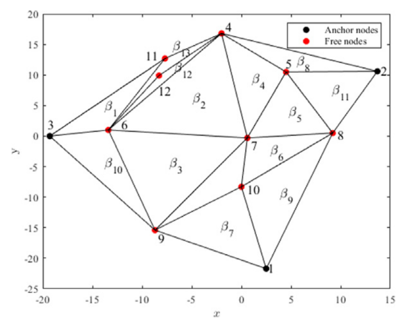


Fig. 1. Network topology with 12 nodes and 13 triangles.

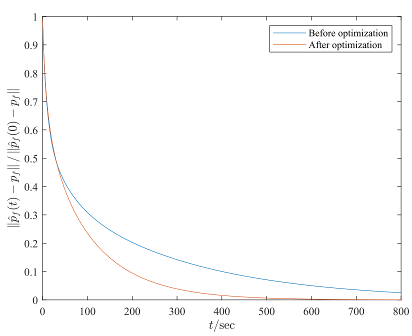


Fig. 2. Position estimation errors without measurement noises.

Note that $\lambda_{\min}\left(L_{ff}(\beta^*)\right) = 0.008876 \approx 2.5 \cdot \lambda_{\min}(L_{ff}(\beta(0)))$, which is improved effectively.

It is worth noting that the weight β_{13}^* corresponding to triangle $(4, 6, 12)$ obviously increases a lot with respect to $\beta_{13}(0)$, while the weight β_3^* corresponding to triangle $(6, 7, 9)$ has a significant decrease from its initial value $\beta_3(0)$. One physical interpretation for this is that the eigenvalues of E_{13} defined in (16) are relatively small, which is one of the reasons that makes $\lambda_{\min}(L_{ff}(\beta))$ small. If large coefficient β_{13} is given in the front of E_{13}, $\lambda_{\min}(L_{ff}(\beta))$ can be increased. While the eigenvalues of E_3 is relatively large, and due to the normalization constraint of β, giving a small coefficient β_3 in the front of E_3 will not cause a decrease in $\lambda_{\min}(L_{ff}(\beta))$.

Fig. 2 shows the evolution of the position estimation error $\left\|\hat{p}_f(t) - p_f\right\|$ without the existence of measurement noises under the localization algorithm (12). The convergence rate of the position estimation error to zero is faster after using the proposed performance optimization method, which validates that our performance optimization method can improve the convergence rate of the localization algorithm (12).

To simulate noisy sensor environments, each element of the error matrices ΔL_{ff} and ΔL_{fa} is generated by white noises. Fig. 3 shows the evolution of the position estimation error $\left\|\hat{p}_f(t) - p_f\right\|$ when considering the existence of measurement noises under the localization algorithm (13). When $\left\|\Delta L_{ff}\right\| = 0.002710 < \lambda_{\min}\left(L_{ff}(\beta(0))\right) < \lambda_{\min}\left(L_{ff}(\beta^*)\right)$, the left panel of Fig. 3 shows that the stable position estimation error is less after using the proposed performance optimization method. When $\lambda_{\min}\left(L_{ff}(\beta(0))\right) < \left\|\Delta L_{ff}\right\| = 0.006166 < \lambda_{\min}\left(L_{ff}(\beta^*)\right)$, The right panel of Fig. 3 shows that the localization algorithm (13) has better robustness against measurement noises after using the performance optimization method. These two simulation examples validate that our performance optimization method can improve the stability margin and robustness of the localization algorithm.

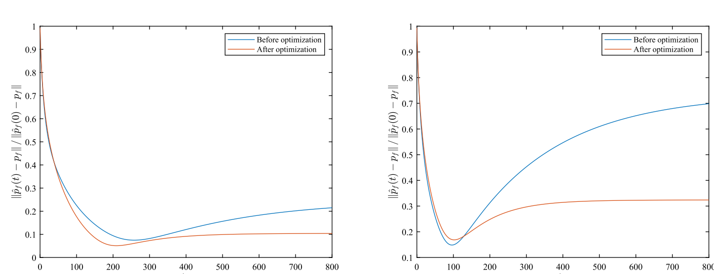
To validate the effectiveness of the distributed optimization algorithm (26), we primarily demonstrate its capability to yield the optimal solution, which coincides with the optimal solution (30) obtained through the centralized Algorithm 1. Let $\lambda_{\min}\left(L_{ff}(\beta^*)\right)$ be abbreviated as $\lambda_{\min}^*$ for simplicity. The minimum eigenvalue of L_{ff} at time instant t is represented as $\lambda_{\min}(t) = \lambda_{\min}(\sum_{i=1}^{m} \beta_i(t)E_i)$.

Two panels of Fig. 4 show the evolution of convergence errors $\left|\lambda_{\min}(t) - \lambda_{\min}^*\right|$ and $\|\beta(t) - \beta^*\|$, respectively. Both of them asymptotically converge to zero under the proposed distributed optimization algorithm. These two simulation results verify the results of Theorems 13 and 14, i.e., the algorithm (26) will converge to the equilibrium point, which is also the optimal solution of the problem (25).

Fig. 5 shows the trajectories of consensus errors $\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\left\|R_i(t) - R_j(t)\right\|_F$ and $\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\left\|\nu_i(t) - \nu_j(t)\right\|$. The results show that R_i and ν_i, $i \in \mathcal{M}$ could converge to the same asymptotically.

6. Conclusions

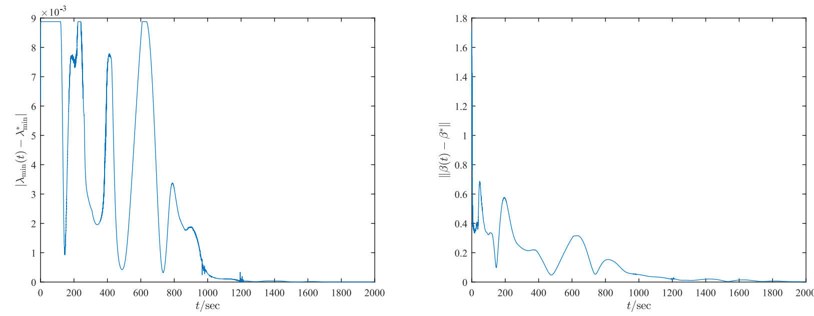
This paper aims to improve the performance of linear localization algorithms. Firstly, we have proposed a unified description for linear localization algorithms, and shown that the stability margin, convergence rate and robustness of linear localization algorithms are commonly determined by the minimum eigenvalue of the network's localization matrix. Secondly, by carefully choosing the decision variable, we have formulated the performance optimization problem as an eigenvalue optimization problem. Thirdly, we have proposed centralized and distributed optimization algorithms to obtain the optimal solution of the



(a) Position estimation errors when $\|\Delta L_{ff}\| = 0.002710$

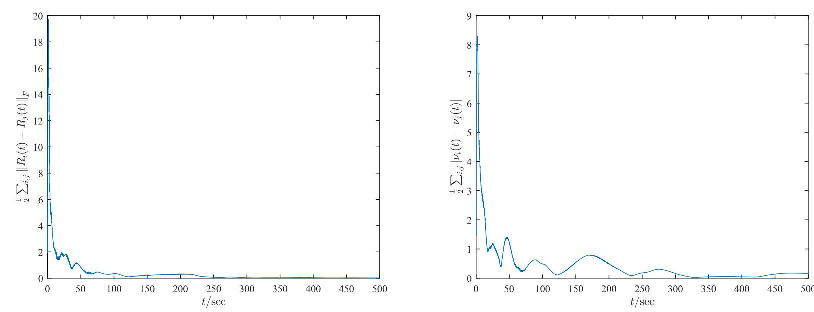
(b) Position estimation errors when $\|\Delta L_{ff}\| = 0.006166$

Fig. 3. Position estimation errors with the existence of measurement noises.



(a) Convergence error of $\lambda_{\min}(t)$ (b) Convergence error of $\beta(t)$

Fig. 4. Convergence errors of $\lambda_{\min}(t)$ and $\beta(t)$.



(a) Consensus error of $\boldsymbol{R}(t)$ (b) Consensus error of $\boldsymbol{\nu}(t)$

Fig. 5. Consensus errors of $\boldsymbol{R}(t)$ and $\boldsymbol{\nu}(t)$.

eigenvalue optimization problem. Finally, simulation examples have validated the algorithms' effectiveness.

Acknowledgments

The work of L. Chen is supported by Shenzhen Key Laboratory of Control Theory and Intelligent Systems, China under grant No. ZDSYS20220330161800001, Guangdong Provincial Natural Science Foundation, China under grant No. 2024A1515011523 and National Natural Science Foundation of China under Grant No. 62473190. The work of J. Mei and C. Liang is supported in part by the National Natural Science Foundation of China under Grant No. 62073098, in part by the Guangdong Basic and Applied Basic Research Foundation, China under Grant No. 2023B1515120018, and in part by the Preeminent Youth Team Project of Guangdong Basic and Applied Basic Research Foundation, China with Project No. 2024B1515040008.

Appendix A. Proof of Lemma 4

Firstly, we conduct basic trigonometric calculations for the angle-induced coefficient matrices. According to (8), the coefficient matrices are specialized to $A_i^{\mathcal{U}_l} = \sin\alpha_{jki} I_2 - \sin\alpha_{ijk} R^{\mathrm{T}}(\alpha_{kij}) \in \mathbb{R}^{2\times 2}$, $A_j^{\mathcal{U}_l} = \sin\alpha_{ijk} R^{\mathrm{T}}(\alpha_{kij}) \in \mathbb{R}^{2\times 2}$, and $A_k^{\mathcal{U}_l} = -\sin\alpha_{jki} I_2 \in \mathbb{R}^{2\times 2}$. One has

$$
\begin{aligned}
&(A_i^{\mathcal{U}_l})^{\mathrm{T}} A_i^{\mathcal{U}_l} = \varepsilon_i^{\mathcal{U}_l} I_2, \\
&(A_i^{\mathcal{U}_l})^{\mathrm{T}} A_j^{\mathcal{U}_l} = \begin{bmatrix} \varepsilon_{ij,1}^{\mathcal{U}_l} & \varepsilon_{ij,2}^{\mathcal{U}_l} \\ -\varepsilon_{ij,2}^{\mathcal{U}_l} & \varepsilon_{ij,1}^{\mathcal{U}_l} \end{bmatrix} = \varepsilon_{ij}^{\mathcal{U}_l} R(\theta_{ij}^{\mathcal{U}_l}),
\end{aligned} \tag{A.1}
$$

where $\varepsilon_i^{\mathcal{U}_l} \in \mathbb{R}^+$, $\varepsilon_{ij,1}^{\mathcal{U}_l} \in \mathbb{R}$ and $\varepsilon_{ij,2}^{\mathcal{U}_l} \in \mathbb{R}$ are constants related to those interior angles within $\mathcal{U}_l$, and $\theta_{ij}^{\mathcal{U}_l} \in \mathbb{R}$ denotes the corresponding rotation angle. Then, according to the definition of $L_{ff}(\beta)$ given in (16) and (A.1), one can have

$$
\begin{aligned}
L_{ff}(\beta) &= \beta_1 e_1^{\mathrm{T}} e_1 + \cdots + \beta_m e_m^{\mathrm{T}} e_m \\
&= \begin{bmatrix} a_1 I_2 & \cdots & b_{1n_f} R(\theta_{1n_f}) \\ \vdots & \vdots & \vdots \\ b_{1n_f} R^{\mathrm{T}}(\theta_{1n_f}) & \cdots & a_{n_f} I_2 \end{bmatrix},
\end{aligned} \tag{A.2}
$$

where $a_i = \sum_{\mathcal{U}_l \in \mathcal{M}} \beta_l \varepsilon_i^{\mathcal{U}_l}$, $b_{ij} \in \mathbb{R}$ and $\theta_{ij} \in \mathbb{R}$ are the corresponding coefficient and rotation angle after normalizing $\sum_{\mathcal{U}_l \in \mathcal{M}} \beta_l \varepsilon_{ij}^{\mathcal{U}_l} R(\theta_{ij}^{\mathcal{U}_l})$.

Secondly, from the generalized barycentric coordinate defined in Diao et al. (2014) we know that the physical meaning of the unit $\mathcal{U}_l$ is the lth distance-induced linear equation. Let $a_{li}^{\mathcal{U}_l} \in \mathbb{R}$, $i \in \{1, \ldots, n\}$ denotes the lth sensor's generalized barycentric coordinate with respect to sensor i. Then for the distance-induced linear Eq. (6), if we have a barycentric coordinate for every sensor's position, the localization matrix L_{ff} can be specialized to

$$
\begin{aligned}
L_{ff}(\beta) &= \left(\beta_1 e_1^{\mathrm{T}} e_1 + \cdots + \beta_m e_m^{\mathrm{T}} e_m\right) \otimes I_2 \\
&= \begin{bmatrix} a_1 I_2 & \cdots & b_{1n_f} R(0) \\ \vdots & \vdots & \vdots \\ b_{1n_f} R^{\mathrm{T}}(0) & \cdots & a_{n_f} I_2 \end{bmatrix}.
\end{aligned}
$$

The diagonal element is defined as $a_i I_2 = \sum_{\mathcal{U}_l \in \mathcal{M}} \beta_l (a_{li}^{\mathcal{U}_l})^2 I_2$, and the off-diagonal element is defined as $b_{ij} R(0) = \sum_{\mathcal{U}_l \in \mathcal{M}} \beta_l a_{li}^{\mathcal{U}_l} a_{lj}^{\mathcal{U}_l} \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$, $i \neq j$.

Finally we consider the bearing-based localization network. Let $B_{ij}^{\mathcal{U}_l} = \begin{bmatrix} a_{ij}^{\mathcal{U}_l} & -b_{ij}^{\mathcal{U}_l} \\ b_{ij}^{\mathcal{U}_l} & a_{ij}^{\mathcal{U}_l} \end{bmatrix} \in \mathbb{R}^{2\times 2}$, which denotes the weight matrix for sensor i with respect to sensor j in unit $\mathcal{U}_l$. We have

$$(B_{ij}^{\mathcal{U}_l})^{\mathrm{T}} B_{ij}^{\mathcal{U}_l} = \begin{bmatrix} a_{ij}^{\mathcal{U}_l} & -b_{ij}^{\mathcal{U}_l} \\ b_{ij}^{\mathcal{U}_l} & a_{ij}^{\mathcal{U}_l} \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} a_{ij}^{\mathcal{U}_l} & -b_{ij}^{\mathcal{U}_l} \\ b_{ij}^{\mathcal{U}_l} & a_{ij}^{\mathcal{U}_l} \end{bmatrix} = (r_{ij}^{\mathcal{U}_l})^2 I_2,$$

$$(B_{ij}^{\mathcal{U}_l})^{\mathrm{T}} B_{ik}^{\mathcal{U}_l} = \begin{bmatrix} a_{ij}^{\mathcal{U}_l} & -b_{ij}^{\mathcal{U}_l} \\ b_{ij}^{\mathcal{U}_l} & a_{ij}^{\mathcal{U}_l} \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} a_{ik}^{\mathcal{U}_l} & -b_{ik}^{\mathcal{U}_l} \\ b_{ik}^{\mathcal{U}_l} & a_{ik}^{\mathcal{U}_l} \end{bmatrix} = r_{ij}^{\mathcal{U}_l} r_{ik}^{\mathcal{U}_l} R(\varphi_{ik}^{\mathcal{U}_l} - \varphi_{ij}^{\mathcal{U}_l}),$$

where $i \neq j \neq k$, $r_{ij}^{\mathcal{U}_l} = \sqrt{(a_{ij}^{\mathcal{U}_l})^2 + (b_{ij}^{\mathcal{U}_l})^2}$, $\cos\varphi_{ij}^{\mathcal{U}_l} = a_{ij}^{\mathcal{U}_l}/r_{ij}^{\mathcal{U}_l}$, and $\sin\varphi_{ij}^{\mathcal{U}_l} = b_{ij}^{\mathcal{U}_l}/r_{ij}^{\mathcal{U}_l}$. Based on the bearing-induced linear Eq. (7), $L_{ff}(\beta)$ can be described by

$$\begin{aligned} L_{ff}(\beta) &= \beta_1 e_1^{\mathrm{T}} e_1 + \cdots + \beta_m e_m^{\mathrm{T}} e_m \\ &= \begin{bmatrix} a_1 I_2 & \cdots & b_{1n_f} R(\theta_{1n_f}) \\ \vdots & \vdots & \vdots \\ b_{1n_f} R^{\mathrm{T}}(\theta_{1n_f}) & \cdots & a_{n_f} I_2 \end{bmatrix}, \end{aligned}$$

where $a_i = \sum_{\mathcal{U}_l \in \mathcal{M}} \beta_l (r_{l1}^{\mathcal{U}_l})^2$, $b_{ij} \in \mathbb{R}$, $\theta_{ij} \in \mathbb{R}$ are the corresponding coefficient and rotation angle respectively after normalizing $\sum_{\mathcal{U}_l \in \mathcal{M}} \beta_l r_{li}^{\mathcal{U}_l} r_{lj}^{\mathcal{U}_l} R(\varphi_{lj}^{\mathcal{U}_l} - \varphi_{li}^{\mathcal{U}_l})$. Consequently, we can conclude the result.

Appendix B. Proof of Lemma 6

Firstly, we will prove that $\mathcal{B}$ is compact and convex. We can suppose on the contrary that $\beta_i = 0, \exists i \in \mathcal{M}$ is possible. It is obvious that $E_i, \forall i \in \mathcal{M}$ is positive definite, so if the initial $\beta(0)$ satisfies the constraints, then one has straightforward $\lambda_{\min}(L_{ff}(\beta(0))) > 0$. If $\beta_i^* = 0$ after optimization, it is easy to prove that $\lambda_{\min}\left(L_{ff}(\beta^*)\right) = 0 < \lambda_{\min}(L_{ff}(\beta(0)))$, which implies a contradiction since the purpose of the optimization is to maximize the minimum eigenvalue. Therefore, we conclude that β_i will not converge to zero and must be greater than a lower bound when we optimize β, i.e., $\beta_i \geq \epsilon_1$, where $1 > \epsilon_1 > 0$. Similarly, we can derive that β_i will not converge to one and must be less than an upper bound when we optimize β, i.e., $\beta_i \leq 1 - \epsilon_2$, where $1 > \epsilon_2 > 0$, since $\beta_i = 1$ indicates $\beta_j, j \in \{1, \ldots, m\} \backslash i = 0$ due to the constraint $\sum_{i=1}^{m} \beta_i = 1$ and $\beta_i > 0, i = 1, \ldots, m$. Using the above mentioned analysis, we conclude that $\mathcal{B}$ is compact and convex.

Secondly, we prove that the set of solutions of the problem (19) is nonempty and compact. Since the set $\mathcal{B}$ is convex and bounded, there exists no nonzero direction of recession according to Bertsekas (2009, Proposition 1.4.2), which indicates that the domain of the cost function $f(\lambda)$ and $\mathcal{B}$ has no common nonzero direction of recession. Since $f(\lambda) = \lambda$ is convex, and the convergence direction of $f(\lambda)$ is $-\infty$, we can always find a scalar $w \in \mathbb{R}$ that satisfies $w \leq \lambda_{\min}(L_{ff}(\beta(0)))$, which indicates that the epigraph of the cost function $\mathrm{epi}\,(f(\lambda)) = \{(\lambda, w) | f(\lambda) \leq w\}$ is closed. Under the above analysis, the set of solutions of $f(\lambda)$ is nonempty and compact according to Bertsekas (2009, Proposition 3.3.2).

Appendix C. Proof of Theorem 13

Assume that $(\boldsymbol{x}^*, R^*, \nu^*)$ is the saddle point of (24), which satisfies the KKT condition according to Scherer and Weiland (2000, Theorem 1.16) and Antipin (1994), i.e.,

$$\boldsymbol{x}^* = P_{\mathcal{X}}\left(\boldsymbol{x}^* - \nabla_{\boldsymbol{x}} L_2(\boldsymbol{x}^*, R^*, \nu^*)\right), \tag{C.1a}$$

$$G(\boldsymbol{x}^*) \preceq \boldsymbol{0}, \left\langle R^*, G(\boldsymbol{x}^*)\right\rangle = 0, \tag{C.1b}$$

$$h(\boldsymbol{x}^*) = 0. \tag{C.1c}$$

Define $\boldsymbol{U}^* = \left[(U_1^*)^{\mathrm{T}}, \ldots, (U_m^*)^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{R}^{2mn_f \times 2n_f}$ and $\boldsymbol{\eta}^* = \left[\eta_1^*, \ldots, \eta_m^*\right]^{\mathrm{T}} \in \mathbb{R}^m$. Assume $(\boldsymbol{x}^*, \boldsymbol{R}^*, \boldsymbol{\nu}^*, \boldsymbol{U}^*, \boldsymbol{\eta}^*)$ is an equilibrium point of (26), which satisfies

$$\boldsymbol{0} = P_{\mathcal{X}}\left(\boldsymbol{x}^* - \nabla\bar{f}(\boldsymbol{x}^*) - \nabla\bar{G}(\boldsymbol{x}^*)\boldsymbol{R}^* - \nabla\bar{h}(\boldsymbol{x}^*)\boldsymbol{\nu}^*\right) - \boldsymbol{x}^*, \tag{C.2a}$$

$$\boldsymbol{0} = P_{\mathbb{S}_+^{(2n_f)^m}}\left(\boldsymbol{R}^* + \bar{G}(\boldsymbol{x}^*) - L_{\boldsymbol{R}}\boldsymbol{R}^* - L_{\boldsymbol{R}}\boldsymbol{U}^*\right) - \boldsymbol{R}^*, \tag{C.2b}$$

$$\boldsymbol{0} = \bar{h}(\boldsymbol{x}^*) - L\boldsymbol{\nu}^* - L\boldsymbol{\eta}^*, \tag{C.2c}$$

$$\boldsymbol{0} = L_{\boldsymbol{R}}\boldsymbol{R}^*, \tag{C.2d}$$

$$\boldsymbol{0} = L\boldsymbol{\nu}^*, \tag{C.2e}$$

where $\nabla\bar{f}(\boldsymbol{x}) = [\nabla^{\mathrm{T}} f_1(x_1), \ldots, \nabla^{\mathrm{T}} f_m(x_m)]^{\mathrm{T}} \in \mathbb{R}^{2m}$, $\bar{G}(\boldsymbol{x}) = [G_1^{\mathrm{T}}(x_1), \ldots, G_m^{\mathrm{T}}(x_m)]^{\mathrm{T}} \in \mathbb{R}^{4mn_f \times 2n_f}$, $\nabla\bar{G}(\boldsymbol{x}^*)\boldsymbol{R}^* = [\langle I_{2n_f}, R_1^*\rangle, \langle E_1, R_1^*\rangle, \ldots, \langle I_{2n_f}, R_m^*\rangle, \langle E_m, R_m^*\rangle]^{\mathrm{T}} \in \mathbb{R}^{2m}$, $\bar{h}(\boldsymbol{x}) = [h_1(x_1), \ldots, h_m(x_m)]^{\mathrm{T}} \in \mathbb{R}^m$, $\nabla\bar{h}(\boldsymbol{x}^*)\boldsymbol{\nu}^* = [\nabla^{\mathrm{T}} h_1(x_1)\nu_1^*, \ldots, \nabla^{\mathrm{T}} h_m(x_m)\nu_m^*]^{\mathrm{T}} \in \mathbb{R}^{2m}$, and $\nabla h_i(x_i) = c \in \mathbb{R}^2$. Strictly speaking, $\nabla\bar{G}(\boldsymbol{x}^*)\boldsymbol{R}^*$ and $\nabla\bar{h}(\boldsymbol{x}^*)\boldsymbol{\nu}^*$ are imprecise mathematical expressions relative to their definitions, and however, we use this expressions for easy reading.

We first prove (C.1) $\Rightarrow$ (C.2). Define $\boldsymbol{R}^* = \boldsymbol{1}_m \otimes R^*$ and $\boldsymbol{\nu}^* = \boldsymbol{1}_m \nu^*$, and then $\boldsymbol{R}^*$ and $\boldsymbol{\nu}^*$ hold for (C.2d) and (C.2e), respectively.

Take any $s \in \mathbb{R}$, and let $\boldsymbol{s} = \boldsymbol{1}_m \otimes s$, and then $\boldsymbol{s}$ holds for $L\boldsymbol{s} = \boldsymbol{0}$ and $\boldsymbol{s} \in \ker(L)$, where $\ker(L) = \{x | x \in \mathbb{R}^m, Lx = 0\}$ denotes the null space of L. Since (C.1c) holds, $\boldsymbol{s}^{\mathrm{T}}\bar{h}(\boldsymbol{x}^*) = msh(\boldsymbol{x}^*) = 0$, which indicates that $\bar{h}(\boldsymbol{x}^*) \in \mathrm{range}(L)$, where $\mathrm{range}(L) = \{y \in \mathbb{R}^m | y = Lx, \exists x \in \mathbb{R}^m\}$ according to Horn and Johnson (2012, Section 0.6.6). Therefore there exists $\boldsymbol{\eta}^*$ satisfying (C.2c).

Let $\boldsymbol{R}^* = \boldsymbol{1}_m \otimes R^*$, which indicates that $\boldsymbol{R}^* \in \ker(L_{\boldsymbol{R}})$. Based on (C.1b), one can derive that $\left\langle \boldsymbol{R}^*, \bar{G}(\boldsymbol{x}^*) - L_{\boldsymbol{R}}(\boldsymbol{R}^* + \boldsymbol{U}^*)\right\rangle = m\left\langle R^*, G(x^*)\right\rangle - \left\langle \boldsymbol{R}^*, L_{\boldsymbol{R}}(\boldsymbol{R}^* + \boldsymbol{U}^*)\right\rangle = \boldsymbol{0}$, which indicates that $\bar{G}(\boldsymbol{x}^*) - L_{\boldsymbol{R}}(\boldsymbol{R}^* + \boldsymbol{U}^*) \in N_{\mathbb{S}_+^{(2n_f)^m}}(\boldsymbol{R}^*)$, where $N_{\mathbb{S}_+^{(2n_f)^m}}(\boldsymbol{R}^*) = \{\boldsymbol{X} \in \mathbb{S}_+^{(2n_f)^m} | \left\langle \boldsymbol{X}, \boldsymbol{R}^*\right\rangle \preceq 0\}$ denotes the normal cone of $\boldsymbol{R}^*$. Then it is easy to verify that $P_{\mathbb{S}_+^{(2n_f)^m}}(\boldsymbol{R}^* + \bar{G}(\boldsymbol{x}^*) - L_{\boldsymbol{R}}\boldsymbol{R}^* - L_{\boldsymbol{R}}\boldsymbol{U}^*) = \boldsymbol{R}^*$, which indicates that (C.2b) holds.

Besides, from (C.1a), $P_{\mathcal{X}}\left(\boldsymbol{x} - \nabla_{\boldsymbol{x}} L_2(\boldsymbol{x}^*, R^*, \nu^*)\right)$ is equal to $P_{\mathcal{X}}(\nabla\bar{G}(\boldsymbol{x}^*)\boldsymbol{R}^* - \nabla\bar{h}(\boldsymbol{x}^*)\boldsymbol{\nu}^*)$, which indicates that (C.2a) holds.

Secondly, we prove (C.1) $\Leftarrow$ (C.2). If $(\boldsymbol{x}^*, \boldsymbol{R}^*, \boldsymbol{\nu}^*, \boldsymbol{U}^*, \boldsymbol{\eta}^*)$ is an equilibrium point of (C.2), we obtain

$$\dot{\boldsymbol{U}} = \boldsymbol{0}, \dot{\boldsymbol{\eta}} = \boldsymbol{0},$$

which indicates that the block-matrix elements of $\boldsymbol{R}^*$ and $\boldsymbol{\nu}^*$ converge to the same, respectively.

Assume $\boldsymbol{R}^* = \boldsymbol{1}_m \otimes R^*$ and $\boldsymbol{\nu}^* = \boldsymbol{1}_m \otimes \nu^*$. The Eq. (C.2b) is equivalent to

$$\boldsymbol{R}^* \succeq \boldsymbol{0}, \tag{C.3a}$$

$$\left\langle \boldsymbol{R}^*, \bar{G}(\boldsymbol{x}^*) - L_{\boldsymbol{R}}\boldsymbol{R}^* - L_{\boldsymbol{R}}\boldsymbol{U}^*\right\rangle \preceq 0, \tag{C.3b}$$

where (C.3b) holds since only when (C.3b) holds, $P_{\mathbb{S}_+^{(2n_f)^m}}(\boldsymbol{R}^* + \bar{G}(\boldsymbol{x}^*) - L_{\boldsymbol{R}}\boldsymbol{R}^* - L_{\boldsymbol{R}}\boldsymbol{U}^*) = \boldsymbol{R}^*$ holds. Since $\boldsymbol{R}^*$ is semi-definite positive, one can further infer that (C.3b) holds only when

$$\bar{G}(\boldsymbol{x}^*) - L_{\boldsymbol{R}}\boldsymbol{R}^* - L_{\boldsymbol{R}}\boldsymbol{U}^* \preceq \boldsymbol{0}. \tag{C.4}$$

Left multiplying $\boldsymbol{1}_m^{\mathrm{T}} \otimes I_{2n_f}$ on both sides of (C.4), one has

$$\begin{aligned} &\left(\boldsymbol{1}_m^{\mathrm{T}} \otimes I_{2n_f}\right)\left(\bar{G}(\boldsymbol{x}^*) - L_{\boldsymbol{R}}\boldsymbol{R}^* - L_{\boldsymbol{R}}\boldsymbol{U}^*\right) \\ &= \left(\boldsymbol{1}_m^{\mathrm{T}} \otimes I_{2n_f}\right)\bar{G}(\boldsymbol{x}^*) = G(\boldsymbol{x}^*) \preceq \boldsymbol{0}, \end{aligned} \tag{C.5}$$

where the first equation holds under the fact $(\boldsymbol{1}_m^{\mathrm{T}} \otimes I_{2n_f})L_{\boldsymbol{R}} = \boldsymbol{0}$. Since $P_{\mathbb{S}_+^{(2n_f)^m}}(\boldsymbol{R}^* + \bar{G}(\boldsymbol{x}^*) - L_{\boldsymbol{R}}\boldsymbol{R}^* - L_{\boldsymbol{R}}\boldsymbol{U}^*) = \boldsymbol{R}^*$ and $L_{\boldsymbol{R}}\boldsymbol{R}^* = \boldsymbol{0}$, we have

$$\left\langle R^*, G(\boldsymbol{x}^*)\right\rangle = \boldsymbol{0}. \tag{C.6}$$

In light of (C.3), (C.5) and (C.6), (C.1b) holds.

Then left multiplying $\mathbf{1}_m^{\mathrm{T}}$ on both sides of (C.2c), one has

$$\mathbf{1}_m^{\mathrm{T}}\left(\bar{h}(\boldsymbol{x}^*) - L\boldsymbol{\nu}^* - L\boldsymbol{\eta}^*\right) = \mathbf{1}_m^{\mathrm{T}}\bar{h}(\boldsymbol{x}^*) = h(\boldsymbol{x}^*) = 0,$$

which shows that (C.1c) holds.

From (C.2a) we have

$$\begin{aligned}
&P_{\mathcal{X}}\left(\boldsymbol{x}^* - \nabla\bar{f}(\boldsymbol{x}^*) - \nabla\bar{G}(\boldsymbol{x}^*)\boldsymbol{R}^* - \nabla\bar{h}(\boldsymbol{x}^*)\boldsymbol{\nu}^*\right) - \boldsymbol{x}^*\\
=&P_{\mathcal{X}}\Big(\boldsymbol{x}^* - \nabla\bar{f}(\boldsymbol{x}^*) - \nabla\bar{G}(\boldsymbol{x}^*)(\mathbf{1}_m\otimes R^*)\\
&- \nabla\bar{h}(\boldsymbol{x}^*)\left(\mathbf{1}_m\otimes\nu^*\right)\Big) - \boldsymbol{x}^*\\
=&P_{\mathcal{X}}\left(\boldsymbol{x}^* - \nabla_{\boldsymbol{x}}L_2(\boldsymbol{x}^*, R^*, \nu^*)\right) - \boldsymbol{x}^*\\
=&\mathbf{0},
\end{aligned}$$

which indicates that (C.1a) holds. The result follows.

Appendix D. Proof of Theorem 14

Define an auxiliary function as

$$M(\boldsymbol{x}, \boldsymbol{R}, \boldsymbol{U}, \boldsymbol{\nu}, \boldsymbol{\eta}) = f(\boldsymbol{x}) + \frac{1}{2}\left\|\hat{\boldsymbol{R}}\right\|_F^2 + \frac{1}{2}\left\|\hat{\boldsymbol{\nu}}\right\|^2.$$

In the following, let M and M^* denote $M(\boldsymbol{x}, \boldsymbol{R}, \boldsymbol{U}, \boldsymbol{\nu}, \boldsymbol{\eta})$ and $M(\boldsymbol{x}^*, \boldsymbol{R}^*, \boldsymbol{U}^*, \boldsymbol{\nu}^*, \boldsymbol{\eta}^*)$ for simplicity, respectively.

Define the Lyapunov function as

$$V = V_1 + V_2 + V_3 + V_4 + V_5 + V_6, \tag{D.1}$$

where

$$\begin{aligned}
V_1 =& M - M^* - \left(\boldsymbol{x} - \boldsymbol{x}^*\right)^{\mathrm{T}}\nabla_{\boldsymbol{x}}M^* - \left\langle\boldsymbol{R} - \boldsymbol{R}^*, \nabla_{\boldsymbol{R}}M^*\right\rangle\\
&- \left\langle\boldsymbol{U} - \boldsymbol{U}^*, \nabla_{\boldsymbol{U}}M^*\right\rangle - \left(\boldsymbol{\nu} - \boldsymbol{\nu}^*\right)^{\mathrm{T}}\nabla_{\boldsymbol{\nu}}M^*\\
&- \left(\boldsymbol{\eta} - \boldsymbol{\eta}^*\right)^{\mathrm{T}}\nabla_{\boldsymbol{\eta}}M^*,\\
V_2 =& \frac{1}{2}\left\|\boldsymbol{x} - \boldsymbol{x}^*\right\|^2,\\
V_3 =& \frac{1}{2}\left\|\boldsymbol{R} - \boldsymbol{R}^*\right\|_F^2,\\
V_4 =& \Big\langle\mathrm{col}(\boldsymbol{R} - \boldsymbol{R}^*, \boldsymbol{U} - \boldsymbol{U}^*), \begin{bmatrix}\frac{3}{2}L_{\boldsymbol{R}} + \frac{k_2}{2k_1}L_{\boldsymbol{R}}^2 & L_{\boldsymbol{R}}\\ L_{\boldsymbol{R}} & \frac{k_1}{k_2}I\end{bmatrix}\\
&\mathrm{col}(\boldsymbol{R} - \boldsymbol{R}^*, \boldsymbol{U} - \boldsymbol{U}^*)\Big\rangle,\\
V_5 =& \frac{1}{2}\left\|\boldsymbol{\nu} - \boldsymbol{\nu}^*\right\|^2,\\
V_6 =& \mathrm{col}^{\mathrm{T}}(\boldsymbol{\nu} - \boldsymbol{\nu}^*, \boldsymbol{\eta} - \boldsymbol{\eta}^*)\begin{bmatrix}\frac{3}{2}L + \frac{k_2}{2k_1}L^2 & L\\ L & \frac{k_1}{k_2}I\end{bmatrix}\\
&\mathrm{col}(\boldsymbol{\nu} - \boldsymbol{\nu}^*, \boldsymbol{\eta} - \boldsymbol{\eta}^*).
\end{aligned}$$

Because of the convexity of the function M, it holds $V_1 \geq 0$. To see the positive semi-definiteness of $V_3 + V_4$, it needs to verify the matrix $\begin{bmatrix}\frac{1}{2}I + \frac{3}{2}L_{\boldsymbol{R}} + \frac{k_2}{2k_1}L_{\boldsymbol{R}}^2 & L_{\boldsymbol{R}}\\ L_{\boldsymbol{R}} & \frac{k_1}{k_2}I\end{bmatrix}$ is positive semi-definite, which is satisfied if $\frac{k_1}{k_2} \geq \frac{\lambda_{\max}^2(L)}{1+3\lambda_{\max}(L)}$ based on Schur complements (Horn & Johnson, 2012, 0.8.5). This condition is satisfied if $\frac{k_1}{k_2} \geq \frac{\lambda_{\max}(L)}{2}$, since $\frac{\lambda_{\max}^2(L)}{1+3\lambda_{\max}(L)} \leq \frac{\lambda_{\max}(L)}{2}$. Similarly, $V_5 + V_6$ is also positive semi-definite. Consequently, the Lyapunov function $V(t) \geq 0$ for all $t \geq 0$.

Firstly, we derive the analytical form of $\dot{V}_1$, i.e.,

$$\begin{aligned}
\dot{V}_1 =& \left(\nabla_{\boldsymbol{x}}M - \nabla_{\boldsymbol{x}}M^*\right)^{\mathrm{T}}\dot{\boldsymbol{x}} + \left\langle\hat{\boldsymbol{R}} - \boldsymbol{R}^*, (I - L_{\boldsymbol{R}})\dot{\hat{\boldsymbol{R}}}\right\rangle\\
&- \left\langle\hat{\boldsymbol{R}} - \boldsymbol{R}^*, L_{\boldsymbol{R}}\dot{\boldsymbol{U}}\right\rangle + \left(\hat{\boldsymbol{\nu}} - \boldsymbol{\nu}^*\right)^{\mathrm{T}}(I - L)\dot{\boldsymbol{\nu}}\\
&- \left(\hat{\boldsymbol{\nu}} - \boldsymbol{\nu}^*\right)^{\mathrm{T}}L\dot{\boldsymbol{\eta}}\\
=& \left(\nabla_{\boldsymbol{x}}M - \nabla_{\boldsymbol{x}}M^*\right)^{\mathrm{T}}\dot{\boldsymbol{x}} + \left\langle\hat{\boldsymbol{R}} - \boldsymbol{R}^*, \dot{\boldsymbol{R}}\right\rangle\\
&- \left\langle\hat{\boldsymbol{R}} - \boldsymbol{R}^*, L_{\boldsymbol{R}}\left(\dot{\boldsymbol{U}} + \dot{\boldsymbol{R}}\right)\right\rangle + \left(\hat{\boldsymbol{\nu}} - \boldsymbol{\nu}^*\right)^{\mathrm{T}}\dot{\boldsymbol{\nu}}\\
&- \left(\hat{\boldsymbol{\nu}} - \boldsymbol{\nu}^*\right)^{\mathrm{T}}L\left(\dot{\boldsymbol{\eta}} + \dot{\boldsymbol{\nu}}\right).
\end{aligned} \tag{D.2}$$

Combining the first term in (D.2) and the definition of V_2, we have

$$\begin{aligned}
&\left(\nabla_{\boldsymbol{x}}M - \nabla_{\boldsymbol{x}}M^*\right)^{\mathrm{T}}\dot{\boldsymbol{x}} + \dot{V}_2\\
=&2k_1\left(\nabla_{\boldsymbol{x}}M - \boldsymbol{x} + \hat{\boldsymbol{x}}\right)^{\mathrm{T}}\left(\hat{\boldsymbol{x}} - \boldsymbol{x}^*\right) + 2k_1\left(\boldsymbol{x} - \hat{\boldsymbol{x}} - \nabla_{\boldsymbol{x}}M^*\right)^{\mathrm{T}}\\
&\left(\hat{\boldsymbol{x}} - \boldsymbol{x}^*\right) + 2k_1\left(\nabla_{\boldsymbol{x}}M - \nabla_{\boldsymbol{x}}M^*\right)^{\mathrm{T}}\left(\boldsymbol{x}^* - \boldsymbol{x}\right)\\
&+ 2k_1\left(\boldsymbol{x} - \boldsymbol{x}^*\right)^{\mathrm{T}}\left(\hat{\boldsymbol{x}} - \boldsymbol{x}\right)\\
\leq&2k_1\left(\nabla_{\boldsymbol{x}}M - \nabla_{\boldsymbol{x}}M^*\right)^{\mathrm{T}}\left(\boldsymbol{x}^* - \boldsymbol{x}\right) - 2k_1\left\|\hat{\boldsymbol{x}} - \boldsymbol{x}\right\|^2\\
=& - 2k_1\left(\nabla\bar{f}(\boldsymbol{x}) - \nabla\bar{f}(\boldsymbol{x}^*)\right)^{\mathrm{T}}\left(\boldsymbol{x} - \boldsymbol{x}^*\right) - 2k_1\big(\nabla\bar{G}(\boldsymbol{x})\hat{\boldsymbol{R}}\\
&- \nabla\bar{G}(\boldsymbol{x}^*)\boldsymbol{R}^*\big)^{\mathrm{T}}\left(\boldsymbol{x} - \boldsymbol{x}^*\right) - 2k_1\left(\nabla\bar{h}(\boldsymbol{x})\hat{\boldsymbol{\nu}} - \nabla\bar{h}(\boldsymbol{x}^*)\boldsymbol{\nu}^*\right)^{\mathrm{T}}\\
&\left(\boldsymbol{x} - \boldsymbol{x}^*\right) - 2k_1\left\|\hat{\boldsymbol{x}} - \boldsymbol{x}\right\|^2\\
\leq& - 2k_1\left(\nabla\bar{G}(\boldsymbol{x})\hat{\boldsymbol{R}} - \nabla\bar{G}(\boldsymbol{x}^*)\boldsymbol{R}^*\right)^{\mathrm{T}}\left(\boldsymbol{x} - \boldsymbol{x}^*\right)\\
&- 2k_1\left(\nabla\bar{h}(\boldsymbol{x})\hat{\boldsymbol{\nu}} - \nabla\bar{h}(\boldsymbol{x}^*)\boldsymbol{\nu}^*\right)^{\mathrm{T}}\left(\boldsymbol{x} - \boldsymbol{x}^*\right)\\
&- 2k_1\left\|\hat{\boldsymbol{x}} - \boldsymbol{x}\right\|^2,
\end{aligned} \tag{D.3}$$

where $\nabla\bar{G}(\boldsymbol{x})\hat{\boldsymbol{R}}$ is defined as $[\langle I_{2n_f}, \hat{R}_1\rangle, \langle E_1, \hat{R}_1\rangle, \ldots, \langle I_{2n_f}, \hat{R}_m\rangle, \langle E_m, \hat{R}_m\rangle]^{\mathrm{T}} \in \mathbb{R}^{2m}$, $\nabla\bar{h}(\boldsymbol{x})\hat{\boldsymbol{\nu}}$ is defined as $\nabla\bar{h}(\boldsymbol{x})\hat{\boldsymbol{\nu}} = [\nabla^{\mathrm{T}}h_1(x_1)\hat{\nu}_1, \ldots, \nabla^{\mathrm{T}}h_m(x_m)\hat{\nu}_m]^{\mathrm{T}} \in \mathbb{R}^{2m}$, $\nabla h_i(x_i) = c \in \mathbb{R}^2$, and the definitions of $\nabla\bar{G}(\boldsymbol{x}^*)\boldsymbol{R}^*$ and $\nabla\bar{h}(\boldsymbol{x}^*)\boldsymbol{\nu}^*$ are similar with $\nabla\bar{G}(\boldsymbol{x})\hat{\boldsymbol{R}}$ and $\nabla\bar{h}(\boldsymbol{x})\hat{\boldsymbol{\nu}}$, respectively. The first inequality holds since $2k_1(\nabla_{\boldsymbol{x}}M - \boldsymbol{x} + \hat{\boldsymbol{x}})^{\mathrm{T}}\left(\hat{\boldsymbol{x}} - \boldsymbol{x}^*\right) \leq 0$ and $-2k_1\nabla_{\boldsymbol{x}}^{\mathrm{T}}M^*\left(\hat{\boldsymbol{x}} - \boldsymbol{x}^*\right) \leq 0$ according to (1), and the second inequality holds since $-2k_1(\nabla\bar{f}(\boldsymbol{x}) - \nabla\bar{f}(\boldsymbol{x}^*))^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}^*) \leq 0$ due to the convexity of $f(\boldsymbol{x})$.

From the second term in (D.2) and the definition of V_3, we have

$$\begin{aligned}
&\left\langle\hat{\boldsymbol{R}} - \boldsymbol{R}^*, \dot{\boldsymbol{R}}\right\rangle + \dot{V}_3\\
=& - k_1\left\|\hat{\boldsymbol{R}} - \boldsymbol{R}\right\|_F^2 + 2k_1\left\langle\hat{\boldsymbol{R}} - \boldsymbol{R}^*, \bar{G}(\boldsymbol{x}) - L_{\boldsymbol{R}}(\boldsymbol{R} + \boldsymbol{U})\right\rangle\\
&+ 2k_1\left\langle\hat{\boldsymbol{R}} - \boldsymbol{R}^*, \hat{\boldsymbol{R}} - \boldsymbol{R} - \bar{G}(\boldsymbol{x}) + L_{\boldsymbol{R}}(\boldsymbol{R} + \boldsymbol{U})\right\rangle\\
\leq& - k_1\left\|\hat{\boldsymbol{R}} - \boldsymbol{R}\right\|_F^2 + 2k_1\left\langle\hat{\boldsymbol{R}} - \boldsymbol{R}^*, \bar{G}(\boldsymbol{x}) - L_{\boldsymbol{R}}(\boldsymbol{R} + \boldsymbol{U})\right\rangle\\
=& - k_1\left\|\hat{\boldsymbol{R}} - \boldsymbol{R}\right\|_F^2 + 2k_1\left\langle\hat{\boldsymbol{R}} - \boldsymbol{R}^*, \bar{G}(\boldsymbol{x}) - \bar{G}(\boldsymbol{x}^*)\right\rangle\\
&+ 2k_1\left\langle\hat{\boldsymbol{R}} - \boldsymbol{R}^*, \bar{G}(\boldsymbol{x}^*) - L_{\boldsymbol{R}}\boldsymbol{U}^*\right\rangle + 2k_1\left\langle\hat{\boldsymbol{R}} - \boldsymbol{R}^*, L_{\boldsymbol{R}}\boldsymbol{U}^*\right\rangle\\
&- 2k_1\left\langle\hat{\boldsymbol{R}} - \boldsymbol{R}^*, L_{\boldsymbol{R}}(\boldsymbol{R} + \boldsymbol{U})\right\rangle\\
\leq& - k_1\left\|\hat{\boldsymbol{R}} - \boldsymbol{R}\right\|_F^2 + 2k_1\left\langle\hat{\boldsymbol{R}} - \boldsymbol{R}^*, \bar{G}(\boldsymbol{x}) - \bar{G}(\boldsymbol{x}^*)\right\rangle\\
&+ 2k_1\left\langle\hat{\boldsymbol{R}} - \boldsymbol{R}^*, L_{\boldsymbol{R}}\boldsymbol{U}^*\right\rangle - 2k_1\left\langle\hat{\boldsymbol{R}} - \boldsymbol{R}^*, L_{\boldsymbol{R}}\left(\boldsymbol{R} + \boldsymbol{U}\right)\right\rangle,
\end{aligned} \tag{D.4}$$

where the first inequality holds since $\langle\hat{\boldsymbol{R}} - \boldsymbol{R}^*, (\hat{\boldsymbol{R}} - \boldsymbol{R} - \bar{G}(\boldsymbol{x}) + L_{\boldsymbol{R}}(\boldsymbol{R} + \boldsymbol{U}))\rangle \leq 0$ according to (1), and the second inequality holds since $\langle\hat{\boldsymbol{R}} - \boldsymbol{R}^*, \bar{G}(\boldsymbol{x}^*) - L_{\boldsymbol{R}}\boldsymbol{U}^*\rangle = \langle\hat{\boldsymbol{R}} - \boldsymbol{R}^*, \boldsymbol{R}^* + \bar{G}(\boldsymbol{x}^*) - L_{\boldsymbol{R}}(\boldsymbol{R}^* + \boldsymbol{U}^*) - \boldsymbol{R}^*\rangle \leq 0$. Besides, the second term in (D.4) can be derived as

$$\begin{aligned}
&2k_1\left\langle\hat{\boldsymbol{R}} - \boldsymbol{R}^*, \bar{G}(\boldsymbol{x}) - \bar{G}(\boldsymbol{x}^*)\right\rangle\\
=&2k_1\left\langle\hat{\boldsymbol{R}}, \bar{G}(\boldsymbol{x}) - \bar{G}(\boldsymbol{x}^*) - \nabla\bar{G}(\boldsymbol{x})(\boldsymbol{x} - \boldsymbol{x}^*)\right\rangle\\
&+ 2k_1\left\langle\hat{\boldsymbol{R}}, \nabla\bar{G}(\boldsymbol{x})(\boldsymbol{x} - \boldsymbol{x}^*)\right\rangle\\
&- 2k_1\left\langle\boldsymbol{R}^*, \bar{G}(\boldsymbol{x}) - \bar{G}(\boldsymbol{x}^*) - \nabla G(\boldsymbol{x}^*)(\boldsymbol{x} - \boldsymbol{x}^*)\right\rangle\\
&- 2k_1\left\langle\boldsymbol{R}^*, \nabla G(\boldsymbol{x}^*)(\boldsymbol{x} - \boldsymbol{x}^*)\right\rangle\\
\leq&2k_1\left(\nabla\bar{G}(\boldsymbol{x})\hat{\boldsymbol{R}} - \nabla\bar{G}(\boldsymbol{x}^*)\boldsymbol{R}^*\right)^{\mathrm{T}}\left(\boldsymbol{x} - \boldsymbol{x}^*\right),
\end{aligned} \tag{D.5}$$

where $\nabla\bar{G}(\boldsymbol{x}^*)(\boldsymbol{x} - \boldsymbol{x}^*)$ is defined as $[\nabla^{\mathrm{T}}G_1(x_1)(x_1 - x^*)\otimes I_{2n_f}, \ldots, G_m^{\mathrm{T}}(x_m)(x_m - x^*)\otimes I_{2n_f}] \in \mathbb{R}^{2mn_f\times 2n_f}$, $\nabla G_i(x_i) = \left[I_{2n_f}, E_i^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{R}^{4n_f\times 2n_f}$, and the equalities $\langle\hat{\boldsymbol{R}}, \bar{G}(\boldsymbol{x}) - \bar{G}(\boldsymbol{x}^*) - \nabla\bar{G}(\boldsymbol{x})(\boldsymbol{x} - \boldsymbol{x}^*)\rangle = 0$

and $-\langle \boldsymbol{R}^*, \bar{G}(\boldsymbol{x}) - \bar{G}(\boldsymbol{x}^*) - \nabla\bar{G}(\boldsymbol{x}^*)(\boldsymbol{x}-\boldsymbol{x}^*)\rangle = 0$ hold since $G(\boldsymbol{x})$ is an affine function, i.e., $\bar{G}(\boldsymbol{x}) - \bar{G}(\boldsymbol{x}^*) = \nabla\bar{G}(\boldsymbol{x})(\boldsymbol{x}-\boldsymbol{x}^*) = \nabla\bar{G}(\boldsymbol{x}^*)(\boldsymbol{x}-\boldsymbol{x}^*)$. Furthermore, the third and fourth terms in (D.4) can be formulated as

$$\begin{aligned}
&2k_1\langle \hat{\boldsymbol{R}} - \boldsymbol{R}^*, L_{\boldsymbol{R}}\boldsymbol{U}^*\rangle - 2k_1\langle \hat{\boldsymbol{R}} - \boldsymbol{R}^*, L_{\boldsymbol{R}}(\boldsymbol{R}+\boldsymbol{U})\rangle \\
=&2k_1\langle \hat{\boldsymbol{R}} - \boldsymbol{R}, L_{\boldsymbol{R}}\boldsymbol{U}^*\rangle + 2k_1\langle \boldsymbol{R} - \boldsymbol{R}^*, L_{\boldsymbol{R}}\boldsymbol{U}^*\rangle \\
&- 2k_1\langle \hat{\boldsymbol{R}} - \boldsymbol{R}, L_{\boldsymbol{R}}\boldsymbol{R}\rangle - 2k_1\langle \hat{\boldsymbol{R}} - \boldsymbol{R}, L_{\boldsymbol{R}}\boldsymbol{U}\rangle \\
&- 2k_1\langle \boldsymbol{R} - \boldsymbol{R}^*, L_{\boldsymbol{R}}\boldsymbol{R}\rangle - 2k_1\langle \boldsymbol{R} - \boldsymbol{R}^*, L_{\boldsymbol{R}}\boldsymbol{U}\rangle \\
=&2\langle \boldsymbol{U}^*, L_{\boldsymbol{R}}\dot{\boldsymbol{R}}\rangle + \frac{2k_1}{k_2}\langle \boldsymbol{U}^*, \dot{\boldsymbol{U}}\rangle - 2\langle \boldsymbol{R} - \boldsymbol{R}^*, L_{\boldsymbol{R}}\dot{\boldsymbol{R}}\rangle \\
&- 2\langle \boldsymbol{U}, L_{\boldsymbol{R}}\dot{\boldsymbol{R}}\rangle - 2k_1\langle \boldsymbol{R}, L_{\boldsymbol{R}}\boldsymbol{R}\rangle - \frac{2k_1}{k_2}\langle \boldsymbol{U}, \dot{\boldsymbol{U}}\rangle \\
=&-2\langle \boldsymbol{U} - \boldsymbol{U}^*, L_{\boldsymbol{R}}\dot{\boldsymbol{R}}\rangle - \frac{2k_1}{k_2}\langle \boldsymbol{U} - \boldsymbol{U}^*, \dot{\boldsymbol{U}}\rangle \\
&- 2\langle \boldsymbol{R} - \boldsymbol{R}^*, L_{\boldsymbol{R}}\dot{\boldsymbol{R}}\rangle - 2k_1\langle \boldsymbol{R}, L_{\boldsymbol{R}}\boldsymbol{R}\rangle,
\end{aligned} \tag{D.6}$$

in which the derivation is based on the fact $L_{\boldsymbol{R}}\boldsymbol{R}^* = \boldsymbol{0}$ and the dynamics (26). Substituting (D.5) and (D.6) into (D.4), one can have

$$\begin{aligned}
&\langle \hat{\boldsymbol{R}} - \boldsymbol{R}^*, \dot{\boldsymbol{R}}\rangle + \dot{V}_3 \\
\le& -k_1\|\hat{\boldsymbol{R}} - \boldsymbol{R}\|_F^2 + 2k_1\left(\nabla\bar{G}(\boldsymbol{x})\hat{\boldsymbol{R}} - \nabla\bar{G}(\boldsymbol{x}^*)\boldsymbol{R}^*\right)^{\mathrm{T}} \\
&(\boldsymbol{x}-\boldsymbol{x}^*) - 2\langle \boldsymbol{U} - \boldsymbol{U}^*, L_{\boldsymbol{R}}\dot{\boldsymbol{R}}\rangle - \frac{2k_1}{k_2}\langle \boldsymbol{U} - \boldsymbol{U}^*, \dot{\boldsymbol{U}}\rangle \\
&- 2\langle \boldsymbol{R} - \boldsymbol{R}^*, L_{\boldsymbol{R}}\dot{\boldsymbol{R}}\rangle - 2k_1\langle \boldsymbol{R}, L_{\boldsymbol{R}}\boldsymbol{R}\rangle.
\end{aligned} \tag{D.7}$$

From the third term in (D.2) and the definition of V_4, one has

$$\begin{aligned}
&-\langle \hat{\boldsymbol{R}} - \boldsymbol{R}^*, L_{\boldsymbol{R}}(\dot{\boldsymbol{U}} + \dot{\boldsymbol{R}})\rangle + \dot{V}_4 \\
=&-\langle \hat{\boldsymbol{R}} - \boldsymbol{R}, L_{\boldsymbol{R}}\dot{\boldsymbol{R}}\rangle - \langle \boldsymbol{R} - \boldsymbol{R}^*, L_{\boldsymbol{R}}\dot{\boldsymbol{R}}\rangle - \langle \hat{\boldsymbol{R}} - \boldsymbol{R}, L_{\boldsymbol{R}}\dot{\boldsymbol{U}}\rangle \\
&- \langle \boldsymbol{R} - \boldsymbol{R}^*, L_{\boldsymbol{R}}\dot{\boldsymbol{U}}\rangle + \langle \boldsymbol{R} - \boldsymbol{R}^*, (3L_{\boldsymbol{R}} + \frac{k_2}{k_1}L_{\boldsymbol{R}}^2)\dot{\boldsymbol{R}}\rangle \\
&+ \frac{2k_1}{k_2}\langle \boldsymbol{U} - \boldsymbol{U}^*, \dot{\boldsymbol{U}}\rangle + 2\langle \boldsymbol{R} - \boldsymbol{R}^*, L_{\boldsymbol{R}}\dot{\boldsymbol{U}}\rangle \\
&+ 2\langle \boldsymbol{U} - \boldsymbol{U}^*, L_{\boldsymbol{R}}\dot{\boldsymbol{R}}\rangle \\
=&-\frac{1}{k_1}\langle \dot{\boldsymbol{R}}, L_{\boldsymbol{R}}\dot{\boldsymbol{R}}\rangle + 2\langle \boldsymbol{R} - \boldsymbol{R}^*, L_{\boldsymbol{R}}\dot{\boldsymbol{R}}\rangle + k_2\langle \boldsymbol{R}, L_{\boldsymbol{R}}^2\boldsymbol{R}\rangle \\
&+ \frac{2k_1}{k_2}\langle \boldsymbol{U} - \boldsymbol{U}^*, \dot{\boldsymbol{U}}\rangle + 2\langle \boldsymbol{U} - \boldsymbol{U}^*, L_{\boldsymbol{R}}\dot{\boldsymbol{R}}\rangle \\
\le&2\langle \boldsymbol{R} - \boldsymbol{R}^*, L_{\boldsymbol{R}}\dot{\boldsymbol{R}}\rangle + k_2\langle \boldsymbol{R}, L_{\boldsymbol{R}}^2\boldsymbol{R}\rangle + \frac{2k_1}{k_2}\langle \boldsymbol{U} - \boldsymbol{U}^*, \dot{\boldsymbol{U}}\rangle \\
&+ 2\langle \boldsymbol{U} - \boldsymbol{U}^*, L_{\boldsymbol{R}}\dot{\boldsymbol{R}}\rangle,
\end{aligned} \tag{D.8}$$

where the inequality holds since $-\frac{1}{k_1}\langle \dot{\boldsymbol{R}}, L_{\boldsymbol{R}}\dot{\boldsymbol{R}}\rangle \le 0$ and $-\frac{1}{k_2}\langle \dot{\boldsymbol{U}}, \dot{\boldsymbol{U}}\rangle \le 0$.

Based on (D.7) and (D.8), one can derive that

$$\begin{aligned}
&\langle \hat{\boldsymbol{R}} - \boldsymbol{R}^*, \dot{\boldsymbol{R}}\rangle - \langle \hat{\boldsymbol{R}} - \boldsymbol{R}^*, L_{\boldsymbol{R}}(\dot{\boldsymbol{U}} + \dot{\boldsymbol{R}})\rangle + \dot{V}_3 + \dot{V}_4 \\
\le& -k_1\|\hat{\boldsymbol{R}} - \boldsymbol{R}\|_F^2 + 2k_1\left(\nabla\bar{G}(\boldsymbol{x})\hat{\boldsymbol{R}} - \nabla\bar{G}(\boldsymbol{x}^*)\boldsymbol{R}^*\right)^{\mathrm{T}} \\
&(\boldsymbol{x}-\boldsymbol{x}^*) - \langle \boldsymbol{R}, (2k_1L_{\boldsymbol{R}} - k_2L_{\boldsymbol{R}}^2)\boldsymbol{R}\rangle.
\end{aligned} \tag{D.9}$$

Similar with the derivation of (D.9), one has

$$\begin{aligned}
&(\hat{\boldsymbol{v}} - \boldsymbol{v}^*)^{\mathrm{T}}\dot{\boldsymbol{v}} - (\hat{\boldsymbol{v}} - \boldsymbol{v}^*)^{\mathrm{T}}L(\dot{\boldsymbol{\eta}} + \dot{\boldsymbol{v}}) + \dot{V}_5 + \dot{V}_6 \\
\le& -k_1\|\hat{\boldsymbol{v}} - \boldsymbol{v}\|^2 + 2k_1\left(\nabla\bar{h}(\boldsymbol{x})\hat{\boldsymbol{v}} - \nabla\bar{h}(\boldsymbol{x}^*)\boldsymbol{v}^*\right)^{\mathrm{T}} \\
&(\boldsymbol{x}-\boldsymbol{x}^*) - \boldsymbol{v}^{\mathrm{T}}(2k_1L - k_2L^2)\boldsymbol{v}.
\end{aligned} \tag{D.10}$$

Based on (D.1), (D.2), (D.9) and (D.10), the derivative of the Lyapunov function can be calculated as

$$\begin{aligned}
\dot{V} =&\dot{V}_1 + \dot{V}_2 + \dot{V}_3 + \dot{V}_4 + \dot{V}_5 + \dot{V}_6 \\
\le& -2k_1\|\hat{\boldsymbol{x}} - \boldsymbol{x}\|^2 - k_1\|\hat{\boldsymbol{R}} - \boldsymbol{R}\|_F^2 - k_1\|\hat{\boldsymbol{v}} - \boldsymbol{v}\|^2 \\
&- \boldsymbol{v}^{\mathrm{T}}(2k_1L - k_2L^2)\boldsymbol{v} - \langle \boldsymbol{R}, (2k_1L_{\boldsymbol{R}} - k_2L_{\boldsymbol{R}}^2)\boldsymbol{R}\rangle \\
\le& 0,
\end{aligned} \tag{D.11}$$

where the last inequality holds since the matrices $2k_1L - k_2L^2$ and $2k_1L_{\boldsymbol{R}} - k_2L_{\boldsymbol{R}}^2$ are both positive semi-definite under the condition in Theorem 14, and then we can easily conclude that all terms in (D.11) are negative semi-definite.

By the definition of V and the fact $V(t) \le V(0)$, one can indicate that the variables $\boldsymbol{x}, \boldsymbol{R}, \boldsymbol{U}, \boldsymbol{v}, \boldsymbol{\eta}$ are all bounded. Note that according to the definition of V, if any variable in $(\boldsymbol{x}, \boldsymbol{R}, \boldsymbol{U}, \boldsymbol{v}, \boldsymbol{\eta})$ goes to infinity, the Lyapunov function V will go to infinity, i.e., $V \to \infty$ as $\|\boldsymbol{x}\| + \|\boldsymbol{v}\| + \|\boldsymbol{\eta}\| + \|\boldsymbol{R}\|_F + \|\boldsymbol{U}\|_F \to \infty$, which shows that the set $\mathcal{D} = \{(\boldsymbol{x}, \boldsymbol{R}, \boldsymbol{U}, \boldsymbol{v}, \boldsymbol{\eta})|\dot{V}(\boldsymbol{x}, \boldsymbol{R}, \boldsymbol{U}, \boldsymbol{v}, \boldsymbol{\eta}) \le 0\}$ is unbounded according to Khalil (2002). Let $\mathcal{E}$ be the set of all points in $\mathcal{D}$ where $\dot{V} = 0$, and let $\mathcal{L}$ be the largest invariant set in $\mathcal{E}$. According to LaSalle's theorem (Khalil, 2002, Theorem 4.4), the trajectory $(\boldsymbol{x}(t), \boldsymbol{R}(t), \boldsymbol{U}(t), \boldsymbol{v}(t), \boldsymbol{\eta}(t))$ will converge to the set $\mathcal{L}$ as $t \to \infty$. By analyzing the result in (D.11), one can indicate that the set $\mathcal{L} = \{(\boldsymbol{x}, \boldsymbol{R}, \boldsymbol{U}, \boldsymbol{v}, \boldsymbol{\eta})|\hat{\boldsymbol{x}} = \boldsymbol{x}, \hat{\boldsymbol{R}} = \boldsymbol{R}, \hat{\boldsymbol{v}} = \boldsymbol{v}, L_{\boldsymbol{R}}\boldsymbol{R} = \boldsymbol{0}, L\boldsymbol{v} = 0\}$. Obviously, $\mathcal{L}$ is the set of equilibrium points of (26). Consequently, the trajectory $(\boldsymbol{x}(t), \boldsymbol{R}(t), \boldsymbol{U}(t), \boldsymbol{v}(t), \boldsymbol{\eta}(t))$ will converge to the equilibrium point of (26). This ends the proof.

Appendix E. Proof of Theorem 15

According to (D.11), one can obtain that

$$\begin{aligned}
\dot{V} \le& -2k_1\|\hat{\boldsymbol{x}} - \boldsymbol{x}\|^2 - k_1\|\hat{\boldsymbol{v}} - \boldsymbol{v}\|^2 - k_1\|\hat{\boldsymbol{R}} - \boldsymbol{R}\|_F^2 \\
&- \langle \boldsymbol{R}, (2k_1L_{\boldsymbol{R}} - k_2L_{\boldsymbol{R}}^2)\boldsymbol{R}\rangle - \boldsymbol{v}^{\mathrm{T}}(2k_1L - k_2L^2)\boldsymbol{v} \\
\le& -\frac{1}{2k_1}\|\dot{\boldsymbol{x}}\|^2 - \frac{1}{k_1}\|\dot{\boldsymbol{R}}\|_F^2 - \frac{\epsilon}{k_2^2}\|\dot{\boldsymbol{U}}\|_F^2 \\
&- \frac{1}{k_1}\|\dot{\boldsymbol{v}}\|^2 - \frac{\epsilon}{k_2^2}\|\dot{\boldsymbol{\eta}}\|^2 \\
\le& -d_1\left(\|\dot{\boldsymbol{x}}\|^2 + \|\dot{\boldsymbol{R}}\|_F^2 + \|\dot{\boldsymbol{U}}\|_F^2 + \|\dot{\boldsymbol{v}}\|^2 + \|\dot{\boldsymbol{\eta}}\|^2\right) \\
\le& -\frac{d_1}{4}\left(\|\dot{\boldsymbol{x}}\| + \|\dot{\boldsymbol{R}}\|_F + \|\dot{\boldsymbol{U}}\|_F + \|\dot{\boldsymbol{v}}\| + \|\dot{\boldsymbol{\eta}}\|\right)^2,
\end{aligned}$$

where $\epsilon \in \mathbb{R}^+$ satisfies $2k_1L_{\boldsymbol{R}} - k_2L_{\boldsymbol{R}}^2 \succeq \epsilon L_{\boldsymbol{R}}^2$ and $2k_1L - k_2L^2 \succeq \epsilon L^2$, and $d_1 = \min\{\frac{1}{2k_1}, \frac{\epsilon}{k_2^2}\}$.

By contradiction, if $\lim_{t\to\infty}\inf\left(\|\dot{\boldsymbol{x}}\| + \|\dot{\boldsymbol{R}}\|_F + \|\dot{\boldsymbol{U}}\|_F + \|\dot{\boldsymbol{v}}\| + \|\dot{\boldsymbol{\eta}}\|\right) \neq O(\frac{1}{\sqrt{t}})$, we assume that $\|\dot{\boldsymbol{x}}\| + \|\dot{\boldsymbol{R}}\|_F + \|\dot{\boldsymbol{U}}\|_F + \|\dot{\boldsymbol{v}}\| + \|\dot{\boldsymbol{\eta}}\| \ge \frac{d_2}{\sqrt{t}}$ when $t \ge t_{th}$, where $d_2 \in \mathbb{R}^+$, and t_{th} denotes some threshold time. Then we have

$$\begin{aligned}
V(t) \le&V(t_{th}) + \int_{t_{th}}^{t} -\frac{d_1}{4}\cdot\frac{d_2^2}{s^2}ds \\
=&V(t_{th}) + \frac{d_1d_2^2}{4}\ln t_{th} - \frac{d_1d_2^2}{4}\ln t,
\end{aligned}$$

from which we can indicate that $V(t) < 0$ when t is large enough, which is a contradiction with the fact that $V(t) \ge 0$ for all $t \ge 0$. This ends the proof.

References

Ahlfors, Lars V. (1953). *vol. 177, Complex Analysis: an introduction to the theory of analytic functions of one complex variable*. New York, London.

Aldana-López, Rodrigo, Seeber, Richard, Haimovich, Hernan, & Gómez-Gutiérrez, David (2023). On inherent limitations in robustness and performance for a class of prescribed-time algorithms. *Automatica*, *158*, Article 111284.

Antipin, Anatoly S. (1994). Feedback-controlled saddle gradient processes. *Automation and Remote Control*, *55*(3), 311–320.

Bertsekas, Dimitri (2009). *Convex optimization theory*: *vol. 1*, Athena Scientific.

Bishop, Adrian N., Anderson, Brian D. O., Fidan, Baris, Pathirana, Pubudu N., & Mao, Guoqiang (2009). Bearing-only localization using geometrically constrained optimization. *IEEE Transactions on Aerospace and Electronic Systems*, *45*(1), 308–320.

Boyd, Stephen, & Vandenberghe, Lieven (2004). *Convex optimization*. Cambridge University Press.

Cao, Kun, Han, Zhimin, Lin, Zhiyun, & Xie, Lihua (2021). Bearing-only distributed localization: A unified barycentric approach. *Automatica*, *133*, Article 109834.

Chen, Liangming (2022). Triangular angle rigidity for distributed localization in 2D. *Automatica*, *143*, Article 110414.

Chen, Liangming, Cao, Kun, Xie, Lihua, Li, Xiaolei, & Feroskhan, Mir (2022). 3-d network localization using angle measurements and reduced communication. *IEEE Transactions on Signal Processing*, *70*, 2402–2415.

Chen, Liangming, Lin, Zhiyun, & Xie, Lihua (2024). Angle-based distributed node localizability and localization. *IEEE Transactions on Automatic Control*, *69*(3), 1890–1897.

Chen, Liangming, Xie, Lihua, Li, Xiaolei, Fang, Xu, & Feroskhan, Mir (2022). Simultaneous localization and formation using angle-only measurements in 2D. *Automatica*, *146*, Article 110605.

Diao, Yingfei, Lin, Zhiyun, & Fu, Minyue (2014). A barycentric coordinate based distributed localization algorithm for sensor networks. *IEEE Transactions on Signal Processing*, *62*(18), 4760–4771.

Fang, Xu, Li, Xiaolei, & Xie, Lihua (2020). Angle-displacement rigidity theory with application to distributed network localization. *IEEE Transactions on Automatic Control*, *66*(6), 2574–2587.

Galicki, Mirosław (2015). Finite-time control of robotic manipulators. *Automatica*, *51*, 49–54.

Helmberg, Christoph, Rendl, Franz, Vanderbei, Robert J., & Wolkowicz, Henry (1996). An interior-point method for semidefinite programming. *SIAM Journal on Optimization*, *6*(2), 342–361.

Horn, Roger A., & Johnson, Charles R. (2012). *Matrix analysis*. Cambridge University Press.

Jarre, Florian (1993). An interior-point method for minimizing the maximum eigenvalue of a linear combination of matrices. *SIAM Journal on Control and Optimization*, *31*(5), 1360–1377.

Jing, Gangshan, Wan, Changhuang, & Dai, Ran (2021). Angle-based sensor network localization. *IEEE Transactions on Automatic Control*, *67*(2), 840–855.

Khalil, Hassan K. (2002). *Nonlinear systems*. Prentice Hall.

Kleiner, Alexander, & Dornhege, Christian (2007). Real-time localization and elevation mapping within urban search and rescue scenarios. *Journal of Field Robotics*, *24*(8–9), 723–745.

Le, Xinyi, Chen, Sijie, Li, Fei, Yan, Zheng, & Xi, Juntong (2019). Distributed neurodynamic optimization for energy internet management. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *49*(8), 1624–1633.

Li, Weijian, Deng, Wen, Zeng, Xianlin, & Hong, Yiguang (2021). Distributed solver for linear matrix inequalities: an optimization perspective. *Control Theory and Technology*, *19*, 507–515.

Li, Xiaolei, Luo, Xiaoyuan, & Zhao, Shiyu (2019). Globally convergent distributed network localization using locally measured bearings. *IEEE Transactions on Control of Network Systems*, *7*(1), 245–253.

Li, Xiuxian, Xie, Lihua, & Hong, Yiguang (2019). Distributed continuous-time nonsmooth convex optimization with coupled inequality constraints. *IEEE Transactions on Control of Network Systems*, *7*(1), 74–84.

Li, Weijian, Zeng, Xianlin, Hong, Yiguang, & Ji, Haibo (2021). Distributed design for nuclear norm minimization of linear matrix equations with constraints. *IEEE Transactions on Automatic Control*, *66*(2), 745–752.

Liang, Chenyang, Chen, Liangming, Li, Yibei, Mei, Jie, & Xie, Lihua (2023). Performance optimization of angle-based network localization. In *2023 62nd IEEE Conference on Decision and Control* (pp. 5159–5164). IEEE.

Lin, Zhiyun, Fu, Minyue, & Diao, Yingfei (2015). Distributed self localization for relative position sensing networks in 2D space. *IEEE Transactions on Signal Processing*, *63*(14), 3751–3761.

Lin, Zhiyun, Han, Tingrui, Zheng, Ronghao, & Fu, Minyue (2016). Distributed localization for 2-D sensor networks with bearing-only measurements under switching topologies. *IEEE Transactions on Signal Processing*, *64*(23), 6345–6359.

Nedic, Angelia, Ozdaglar, Asuman, & Parrilo, Pablo A. (2010). Constrained consensus and optimization in multi-agent networks. *IEEE Transactions on Automatic Control*, *55*(4), 922–938.

Orlov, Yury, Kairuz, Ramón I. Verdés, & Aguilar, Luis T. (2021). Prescribed-time robust differentiator design using finite varying gains. *IEEE Control Systems Letters*, *6*, 620–625.

Overton, Michael L. (1988). On minimizing the maximum eigenvalue of a symmetric matrix. *SIAM Journal on Matrix Analysis and Applications*, *9*(2), 256–268.

Overton, Michael L. (1992). Large-scale optimization of eigenvalues. *SIAM Journal on Optimization*, *2*(1), 88–120.

Rana, Omer F., & Stout, Kate (2000). What is scalability in multi-agent systems? In *Proceedings of the Fourth International Conference on Autonomous Agents* (pp. 56–63).

Scherer, Carsten, & Weiland, Siep (2000). *Lecture Notes*: *vol. 3*, *Linear matrix inequalities in control*. Delft, the Netherlands: Dutch Institute for Systems and Control, 2.

Schoof, Eric, Chapman, Airlie, & Mesbahi, Mehran (2017). Weighted bearing-compass dynamics: edge and leader selection. *IEEE Transactions on Network Science and Engineering*, *5*(3), 247–260.

Sun, Zhiyong, Yu, Changbin, & Anderson, Brian D. O. (2015). Distributed optimization on proximity network rigidity via robotic movements. In *2015 34th Chinese Control Conference* (pp. 6954–6960). IEEE.

Trinh, Minh Hoang, Van Tran, Quoc, & Ahn, Hyo-Sung (2019). Minimal and redundant bearing rigidity: Conditions and applications. *IEEE Transactions on Automatic Control*, *65*(10), 4186–4200.

Van Tran, Quoc, Sun, Zhiyong, Anderson, Brian D. O., & Ahn, Hyo-Sung (2022). Distributed optimization for graph matching. *IEEE Transactions on Cybernetics*.

Wang, Xiaoxuan, Yang, Shaofu, Guo, Zhenyuan, Wen, Shiping, & Huang, Tingwen (2022). A distributed network system for nonsmooth coupled-constrained optimization. *IEEE Transactions on Network Science and Engineering*, *9*(5), 3691–3700.

Xia, Yinqiu, Yu, Chengpu, & He, Chengyang (2022). An exploratory distributed localization algorithm based on 3D barycentric coordinates. *IEEE Transactions on Signal and Information Processing over Networks*, *8*, 702–712.

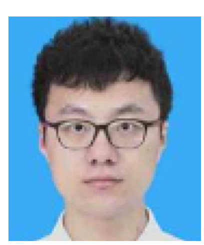
Yang, Tao, Yi, Xinlei, Wu, Junfeng, Yuan, Ye, Wu, Di, Meng, Ziyang, et al. (2019). A survey of distributed optimization. *Annual Reviews in Control*, *47*, 278–305.

Zhao, Shiyu (2018). Affine formation maneuver control of multiagent systems. *IEEE Transactions on Automatic Control*, *63*(12), 4140–4155.

Zhao, Shiyu, & Zelazo, Daniel (2016). Localizability and distributed protocols for bearing-based network localization in arbitrary dimensions. *Automatica*, *69*, 334–341.

Zhu, Guangwei, & Hu, Jianghai (2009). Stiffness matrix and quantitative measure of formation rigidity. In *Proceedings of the 48h IEEE conference on decision and control (CDC) held jointly with 2009 28th Chinese control conference* (pp. 3057–3062). IEEE.

Liangming Chen is an Associate Professor with the Southern University of Science and Technology, China. He received the B.E. degree in Automation from Southwest Jiaotong University, China, in 2015. From 2015 to 2021, he was enrolled jointly in the Ph.D. programs of Harbin Institute of Technology, China and University of Groningen, the Netherlands. From 2021 to 2022, he was a research fellow in Nanyang Technological University, Singapore. He serves in the conference editorial board of EUCA. His research interests are in multi-agent formations and distributed localization.



Chenyang Liang received the B.S. degree in Automation and the M.E. degree in Control Engineering from Harbin Institute of Technology, Harbin, China, in 2017 and 2019, respectively. He is currently working toward the Ph.D. degree in Electronics Information at Harbin Institute of Technology, Shenzhen, Guangdong, China. His current interests include multi-agent distributed localization and distributed optimization.

Yibei Li received her B.E. degree in Automation from Harbin Institute of Technology, China, in 2015. She received her Licentiate degree and Ph.D. degree in Applied and Computational Mathematics from KTH Royal Institute of Technology, Sweden, in 2019 and 2022 respectively. She currently is a Wallenberg-NTU Postdoctoral Fellow at Nanyang Technological University, Singapore. Her research interests include inverse optimization, nonlinear control, game theory and multi-agent systems.

Jie Mei (Member, IEEE) received the B.S. degree in information and computing science from Jilin University, Changchun, China, in 2007, and the Ph.D. degree in control science and engineering from the Harbin Institute of Technology, Harbin, China, in 2011. He was an exchange Ph.D. student supported by the China Scholarship Council with the Department of Electrical and Computer Engineering, Utah State University, Logan, UT, USA, from 2009 to 2011. He held postdoctoral research positions with the Harbin Institute of Technology Shenzhen Graduate School, Guangdong, China, the City University of Hong Kong, Hong Kong, China, and the University of California at Riverside, Riverside, CA, USA, from 2012 to 2015. Since June 2015, he has been with the School of Mechanical Engineering and Automation, Harbin Institute of Technology, Shenzhen, Guangdong, China, where he is currently a Professor. His current research interests include coordination of distributed multi-agent systems, perception and control of intelligent unmanned vehicles, and their applications.

Lihua Xie (Fellow, IEEE) received 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 Chair Professor and the Director of the Center for Advanced Robotics Technology Innovation. He served as the Head of Division of Control and Instrumentation, and the Director of Delta-NTU Corporate Lab for Cyber-Physical Systems. His research interests include robust control and estimation, networked control systems, multiagent networks, localization, and unmanned systems. He is an Editor-in-Chief for Unmanned Systems. He was the Editor of IET Book Series in Control and an Associate Editor for a number of journals including the IEEE Transactions on Automatic Control, Automatica, the IEEE Transactions on Control Systems Technology, and the IEEE Transactions on Control of Network Systems. He was an IEEE Distinguished Lecturer and an Elected Member of Board of Governors, the IEEE Control System Society. He is a Fellow of IFAC, Fellow of Academy of Engineering Singapore, and vice president of IEEE Control Systems Society.