A Distributed Kalman-like Observer with Dynamic Inversion-Based Correction for Multi-Agent Estimation

Nicola De Carli, Dimos V. Dimarogonas

***Abstract*— We present a novel distributed Kalman-like observer for cooperative state estimation in multi-agent systems. Our approach builds on a class of existing Kalman-like observers that replace the process covariance matrix with a forgetting factor. We show that this replacement enables the propagation of the information matrix dynamics in a fully distributed manner, while preserving key stability properties. We compute the observer's correction term by solving a linear equation dynamically in a distributed manner, circumventing the need for direct centralized matrix inversion. Unlike existing methods that partially discard cross-information to allow distributed computations, our approach preserves inter-agent coupling. Rigorous stability guarantees are provided, and numerical simulations in a cooperative localization scenario demonstrate the effectiveness of the approach in estimating agent states.**

***Index Terms*— Multi-agent systems, cooperative localization, observer design**

I. Introduction

COOPERATIVE state estimation plays a crucial role in multi-agent systems, enabling robots, sensor networks, and distributed control systems to reconstruct their states using local information and edge measurements (i.e. measurements involving the state of two neighboring agents) [1], [2], with applications in cooperative localization and distributed monitoring in power networks. Key challenges include achieving scalability, low communication overhead, and stability guarantees. While scalable solutions exist when estimating an external process observed by a sensor network [3], [4], they are generally unsuitable when the goal is to estimate the internal state of the multi-agent system itself—particularly because the state dimension scales with the number of agents.

A widely adopted approach in robotics is the distributed Kalman filter (DKF) proposed in [1], which approximates the steps of a centralized extended Kalman filter by leveraging sparsity in the output coupling and applying approximations that reduce dependence on non-involved estimates. Recent work [5] links some of these approximations to maximum determinant matrix completions. However, the DKF lacks formal stability guarantees and discards useful information for the sake of distributed implementation.

This work was supported in part by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg (KAW) Foundation, the Horizon Europe EIC project SymAware (101070802), the ERC LEAFHOUND Project, the Swedish Research Council (VR), and Digital Futures.

Nicola De Carli and Dimos V. Dimarogonas are with the Division of Decision and Control Systems, KTH Royal Institute of Technology, Stockholm, Sweden. E-mail: {ndc,dimos}@kth.se

Parallel to the developments in robotics, distributed Kalman filtering techniques have also been explored in the context of power network monitoring, seemingly independently from the robotics community. In [6], a DKF variant was proposed that discards cross-terms in the Riccati equation, ensuring convergence only for acyclic graphs. Similarly, [2] introduced another DKF formulation, again limited to acyclic network structures. These methods highlight an ongoing need for more generalizable distributed estimation frameworks that can operate under broader network topologies.

Motivated by the above limitations, we explore the use of a class of Kalman-like observers [7]–[9] in a distributed setting. These observers replace the process noise covariance in the Riccati equation with a forgetting factor. Originally introduced for linear time-varying systems as an optimal solution to a deterministic optimization problem [8], [9], these observers were later extended to nonlinear triangular systems [10] and, more recently, to more general nonlinear systems [11].

We show that, this formulation enables the distributed propagation of the information matrix dynamics, enabling each agent to compute its correction term by dynamically solving a linear equation in a distributed manner which only requires communication between neighboring agents. Unlike conventional DKF approaches, the proposed observer retains inter-agent information, preserving key dependencies between neighboring estimates. Furthermore, it requires only joint observability, i.e. that the system is observable when aggregating the measurements of all agents, as in the centralized case.

We establish rigorous stability guarantees by leveraging arguments from singular perturbation theory, showing that the proposed observer ensures uniform global exponential convergence for a proper choice of a time-scale separation parameter. We demonstrate the effectiveness of the observer through numerical simulations in a cooperative localization scenario, where double integrator agents estimate their states using relative position measurements and a limited number of anchor robots, which are able to measure their own position.

The rest of this paper is structured as follows. Section II presents the system modeling and problem formulation. Section III introduces the proposed distributed observer and its implementation. Section IV provides a rigorous stability analysis. Section V presents simulation results, and Section VI concludes with future research directions.

II. System Modeling

Consider a multi-agent system composed of N agents with decoupled linear dynamics:

$$\dot{\boldsymbol{x}}_i = \boldsymbol{A}_i(t)\boldsymbol{x}_i + \boldsymbol{B}_i(t)\boldsymbol{u}_i \quad \forall i \in \{1, ..., N\} \tag{1}$$

with state $\boldsymbol{x}_i \in \mathbb{R}^{d_i}$ and input $\boldsymbol{u}_i \in \mathbb{R}^{m_i}$, where matrices $\boldsymbol{A}_i \in \mathbb{R}^{d_i \times d_i}$ and $\boldsymbol{B}_i \in \mathbb{R}^{d_i \times m_i}$ may be time-varying. For brevity, time dependence is occasionally omitted. While heterogeneous dimensions ($d_i \neq d_j$) are allowed, we assume homogeneous states and inputs with $d_i = d$, $m_i = m$, $\forall i \in \{1, \dots, N\}$ to simplify notation.

Agent interactions are modeled via a *static directed sensing* graph $\mathcal{G}_s := (\mathcal{V}, \mathcal{E}_s)$ and an *undirected communication* graph $\mathcal{G}_c := (\mathcal{V}, \mathcal{E}_c)$, where $\mathcal{V} = \{1, 2, ..., N\}$ is the vertex set and $\mathcal{E}_* \subseteq \mathcal{V} \times \mathcal{V}$ is the edge set with cardinality $M_* := |\mathcal{E}_*|$. A sensing edge $(i, j) \in \mathcal{E}_s$ indicates that agent i measures agent j, which need not be reciprocal (e.g., due to field-of-view limitations). We assume the communication graph is the undirected counterpart of the sensing graph, i.e., $(i, j) \in \mathcal{E}_c$ implies that either $(i, j) \in \mathcal{E}_s$ or $(j, i) \in \mathcal{E}_s$, enabling bidirectional information exchange. Let $\mathcal{N}_i := \{j \mid (i, j) \in \mathcal{E}_c\}$ denote the (communication) neighbors of agent i. In the following, we say that the sparsity pattern of a matrix is *consistent* with the graph, if its (i, j)-th entry (or block) is nonzero if and only if $(i, j) \in \mathcal{E}_c$.

Each agent may be able to obtain private measurements $\boldsymbol{y}_i^p(\boldsymbol{x}_i) \in \mathbb{R}^{q_{pi}}$ and/or relative measurements $\boldsymbol{y}_{ij}^r(\boldsymbol{x}_i, \boldsymbol{x}_j) \in \mathbb{R}^{q_{rij}}$, with $(i, j) \in \mathcal{E}_s$, assumed to be linear:

$$\boldsymbol{y}_i^p = \delta_i^p \boldsymbol{H}_i^p(t)\boldsymbol{x}_i \qquad \boldsymbol{y}_{ij}^r = \boldsymbol{H}_{iji}^r(t)\boldsymbol{x}_i + \boldsymbol{H}_{ijj}^r(t)\boldsymbol{x}_j \tag{2}$$

where $\delta_i^p = 1$ if a private measurement is available to agent i and $\delta_i^p = 0$ otherwise, and similarly we introduce, for later use, the indicator $\delta_{ij}^r = 1$ if $(i, j) \in \mathcal{E}_s$, $\delta_{ij}^r = 0$ otherwise. A common special case is when the relative measurement $\boldsymbol{y}_{ij}^r$ only depends on the relative state among the agents:

$$\boldsymbol{y}_{ij}^r = \boldsymbol{H}_{ij}^r(t)(\boldsymbol{x}_j - \boldsymbol{x}_i). \tag{3}$$

While measurement models may vary across edges (e.g., position, velocity), we assume uniform dimensions for simplicity: $q_{pi} = q_p$ for all i, and $q_{rij} = q_r$ for all $(i, j) \in \mathcal{E}_s$.

We indicate the collective state as $\boldsymbol{x} := \left[\boldsymbol{x}_1^\top ..., \boldsymbol{x}_N^\top\right]^\top \in \mathbb{R}^{Nd}$ and input $\boldsymbol{u} := \left[\boldsymbol{u}_1^\top ..., \boldsymbol{u}_N^\top\right]^\top \in \mathbb{R}^{Nm}$, the stack of private output measurements as $\boldsymbol{y}^p := \left[\boldsymbol{y}_1^{p\top}, ..., \boldsymbol{y}_N^{p\top}\right]^\top \in \mathbb{R}^{Nq_p}$ and the stack of relative measurements as $\boldsymbol{y}^r := \left[\boldsymbol{y}_1^{r\top}, ..., \boldsymbol{y}_{M_s}^{r\top}\right]^\top \in \mathbb{R}^{M_s q_r}$, where, with a slight abuse of notation, we identify $\boldsymbol{y}_k^r$ with the relative measurement $\boldsymbol{y}_{ij}^r$ corresponding to the sensing edge $e_{sk} = (i, j)$. Furthermore, we define the block diagonal concatenation of local matrices $\boldsymbol{A}_i$ as $\boldsymbol{A} := \text{blkdiag}(\boldsymbol{A}_1, ..., \boldsymbol{A}_N) \in \mathbb{R}^{Nd}$ and, similarly, $\boldsymbol{B} = \text{blkdiag}(\boldsymbol{B}_1, ..., \boldsymbol{B}_N)$. The block diagonal concatenation of private measurement matrices is denoted as $\boldsymbol{H}^p := \text{blkdiag}(\boldsymbol{H}_1^p, ..., \boldsymbol{H}_N^p) \in \mathbb{R}^{Nq_p \times Nd}$, and we also define the sparse matrix $\boldsymbol{H}^r \in \mathbb{R}^{M_s q_r \times Nd}$, which has block components

$$\left[\boldsymbol{H}^r(t)\right]_{k\ell} := \begin{cases} \boldsymbol{H}_{iji}^r(t) & \text{if } e_{sk} = (i, j) \text{ and } \ell = i \\ \boldsymbol{H}_{ijj}^r(t) & \text{if } e_{sk} = (i, j) \text{ and } \ell = j \\ \boldsymbol{0}_{q_r \times d} & \text{otherwise,} \end{cases} \tag{4}$$

which allows us to compactly write the relative measurements output as $\boldsymbol{y}^r = \boldsymbol{H}^r(t)\boldsymbol{x}$. Notice that $\boldsymbol{H}^r$ has the sparsity pattern of an incidence matrix. We also introduce a selection matrix $\boldsymbol{\Delta}^p := \text{blkdiag}(\{\delta_i^p \boldsymbol{I}_{q_p}\}_{i=1}^N) \in \mathbb{R}^{Nd \times Nd}$, with $\boldsymbol{I}_a \in \mathbb{R}^{a \times a}$ denoting the identity matrix, to select the available private measurements, so that, we can compactly write the private output equation as $\boldsymbol{y}^p = \boldsymbol{\Delta}^p \boldsymbol{H}^p(t)\boldsymbol{x}$. Moreover, in the special case where all relative measurements follow the form of (3), the output equation can be expressed as:

$$\boldsymbol{y}^r = \text{blkdiag}(\{\boldsymbol{H}_k^r(t)\}_{k=1}^{M_s})\boldsymbol{E}_d^\top \boldsymbol{x}, \tag{5}$$

where $\boldsymbol{E}_d := \boldsymbol{E} \otimes \boldsymbol{I}_d$, with '$\otimes$' being the Kronecker product and $\boldsymbol{E}$ being the incidence matrix of the sensing graph, defined as

$$[\boldsymbol{E}]_{ik} := \begin{cases} -1 & \text{if } i \text{ is the tail of } e_k \\ 1 & \text{if } i \text{ is the head of } e_k \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

To conclude, the system model can be compactly written as:

$$\begin{aligned} \dot{\boldsymbol{x}} &= \boldsymbol{A}(t)\boldsymbol{x} + \boldsymbol{B}(t)\boldsymbol{u} \\ \boldsymbol{y} &= \begin{bmatrix} \boldsymbol{y}^p \\ \boldsymbol{y}^r \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Delta}^p \boldsymbol{H}^p(t) \\ \boldsymbol{H}^r(t) \end{bmatrix} \boldsymbol{x} = \boldsymbol{H}(t)\boldsymbol{x}. \end{aligned} \tag{7}$$

We make the following assumption regarding the system matrices being uniformly bounded and sufficiently smooth:

Assumption 1. *The system matrices $\boldsymbol{A}(t)$ and $\boldsymbol{H}(t)$ are uniformly bounded, i.e. $\|\boldsymbol{A}(t)\| \leq \bar{a}$, $\|\boldsymbol{H}(t)\| \leq \bar{h}$.*

Assumption 2. *The time derivative of $\boldsymbol{A}(t)$ and $\boldsymbol{H}(t)$ is uniformly bounded, i.e. $\left\|\dot{\boldsymbol{A}}(t)\right\| \leq \bar{a}_d$, $\left\|\dot{\boldsymbol{H}}(t)\right\| \leq \bar{h}_d$.*

III. Distributed Kalman-Like Observer

Probably the most popular observer for centralized systems, the Kalman-Bucy filter has the following dynamics:

$$\begin{aligned} \dot{\hat{\boldsymbol{x}}} &= \boldsymbol{A}\hat{\boldsymbol{x}} + \boldsymbol{B}\boldsymbol{u} + \boldsymbol{P}\boldsymbol{H}^\top \boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{H}\hat{\boldsymbol{x}}) \\ \dot{\boldsymbol{P}} &= \boldsymbol{A}\boldsymbol{P} + \boldsymbol{P}\boldsymbol{A}^\top + \boldsymbol{Q} - \boldsymbol{P}\boldsymbol{H}^\top \boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{P} \end{aligned} \tag{8}$$

where $\boldsymbol{R} \succ \boldsymbol{0}$ and $\boldsymbol{Q} \succ \boldsymbol{0}$ are positive definite matrices, typically representing the measurement and process noise covariance matrices, respectively. They are both assumed block-diagonal with $\boldsymbol{R} = \text{blkdiag}(\{\boldsymbol{R}_k\}_{k=1}^M)$ and $\boldsymbol{Q} = \text{blkdiag}(\{\boldsymbol{Q}_i\}_{i=1}^N)$, where each $\boldsymbol{R}_k$ matches the corresponding output size and $\boldsymbol{Q}_i \in \mathbb{R}^{d \times d}$. We will also refer the blocks of $\boldsymbol{R}$ as $\boldsymbol{R}_{ij}$ for relative measurements associated with the edge (i, j) and $\boldsymbol{R}_{ii}$ for private measurements of the agent i and $\boldsymbol{R}^p := \text{blkdiag}(\{\boldsymbol{R}_{ii}\}_{i=1}^N)$.

Challenges in Distributed Computation. In a distributed setting, our goal is for each agent to estimate its own state and compute the relevant blocks of the Riccati equation using only local information and communication with neighboring agents. To understand the challenges of implementing the Riccati equation in a distributed setting, assume the initial covariance matrix $\boldsymbol{P}(0)$ has a sparse structure consistent with the communication graph. Under these conditions:

- Each agent can compute its corresponding blocks in $\boldsymbol{A}(0)\boldsymbol{P}(0)$ (similarly in $\boldsymbol{P}(0)\boldsymbol{A}(0)^\top$) using local information and communication with neighbors, i.e.,

$$[\boldsymbol{A}(0)\boldsymbol{P}(0)]_{ij} = \begin{cases} \boldsymbol{A}_i(0)\boldsymbol{P}_i(0), & \text{if } i = j \\ \boldsymbol{A}_i(0)\boldsymbol{P}_j(0), & \text{if } j \in \mathcal{N}_i \\ \boldsymbol{0}_{d \times d}, & \text{otherwise} \end{cases} \tag{9}$$

- The block-diagonal nature of $\boldsymbol{Q}$ preserves sparsity over time.
- The term $\boldsymbol{H}^\top \boldsymbol{R}^{-1}\boldsymbol{H}$, with block-diagonal $\boldsymbol{R}$, has a sparsity pattern consistent with the graph, i.e.

$$[\boldsymbol{H}^\top \boldsymbol{R}^{-1}\boldsymbol{H}]_{ij} = \begin{cases} \sum_{j\in\mathcal{N}_i} \left(\delta^r_{ij}\boldsymbol{H}^\top_{iji}\boldsymbol{R}^{-1}_{ij}\boldsymbol{H}_{iji} + \delta^r_{ji}\boldsymbol{H}^\top_{jii}\boldsymbol{R}^{-1}_{ji}\boldsymbol{H}_{jii}\right) \\ \quad + \delta^p_i \boldsymbol{H}^\top_i \boldsymbol{R}^{-1}_{ii}\boldsymbol{H}_i, & \text{if } i=j \\ \delta^r_{ij}\boldsymbol{H}^\top_{iji}\boldsymbol{R}^{-1}_{ij}\boldsymbol{H}_{ijj} + \delta^r_{ji}\boldsymbol{H}^\top_{jii}\boldsymbol{R}^{-1}_{ji}\boldsymbol{H}_{jij}, & \text{if } j\in\mathcal{N}_i \\ \boldsymbol{0}_{d\times d}, & \text{otherwise} \end{cases} \tag{10}$$

If $\boldsymbol{P}(0)$ has a sparsity pattern consistent with the graph, then the matrix multiplication $\boldsymbol{P}(0)\boldsymbol{H}^\top(0)\boldsymbol{R}^{-1}(0)\boldsymbol{H}(0)\boldsymbol{P}(0)$ requires information from three-hop neighbors, as it involves the product of three matrices with a sparsity pattern consistent with the graph [12, Sect. 4.2]. If the graph is connected, the matrix will become a full matrix. Similarly, if the initial $\boldsymbol{P}(0)$ is block diagonal, the last term retains a sparsity pattern consistent with the graph, causing $\boldsymbol{P}(t)$ to quickly adopt the same structure. As in the previous case, this structure is not preserved over time, and $\boldsymbol{P}(t)$ eventually becomes fully dense.

Reformulating the Observer with the Information Matrix. Instead of using the covariance matrix $\boldsymbol{P}$, we consider the observer in terms of the information matrix $\boldsymbol{S} = \boldsymbol{P}^{-1}$, for which $\dot{\boldsymbol{S}} = -\boldsymbol{S}\dot{\boldsymbol{P}}\boldsymbol{S}$. This yields:

$$\begin{aligned} \dot{\hat{\boldsymbol{x}}} &= \boldsymbol{A}\hat{\boldsymbol{x}} + \boldsymbol{B}\boldsymbol{u} + \boldsymbol{S}^{-1}\boldsymbol{H}^\top\boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{H}\hat{\boldsymbol{x}}) \\ \dot{\boldsymbol{S}} &= -\boldsymbol{A}^\top\boldsymbol{S} - \boldsymbol{S}\boldsymbol{A} - \boldsymbol{S}\boldsymbol{Q}\boldsymbol{S} + \boldsymbol{H}^\top\boldsymbol{R}^{-1}\boldsymbol{H}. \end{aligned} \tag{11}$$

In this case, the last term in the Riccati equation, $\boldsymbol{H}^\top\boldsymbol{R}^{-1}\boldsymbol{H}$, can be computed in a distributed way (see (10)), but the term $\boldsymbol{S}\boldsymbol{Q}\boldsymbol{S}$ cannot, as it results in the product of two matrices consistent with the graph. Furthermore, even if $\boldsymbol{S}$ maintained a sparsity pattern consistent with the graph, computing the correction term in the observer equation requires inverting $\boldsymbol{S}$—which is not feasible in a distributed manner.

Kalman-Like Observers and the Forgetting Factor. An alternative class of Kalman-like observers [11], avoids explicit process noise modeling by setting $\boldsymbol{Q} = \boldsymbol{0}_{d\times d}$ and instead introducing a forgetting factor γ to ensure stability. As discussed in [13, Sect. 6.2], the forgetting factor is equivalent to process noise augmentation, but its exact value depends on the evolution of $\boldsymbol{S}$. This leads to a linear matrix differential equation:

$$\dot{\boldsymbol{S}} = -\left(\boldsymbol{A} + \frac{\gamma}{2}\boldsymbol{I}_{Nd}\right)^\top \boldsymbol{S} - \boldsymbol{S}\left(\boldsymbol{A} + \frac{\gamma}{2}\boldsymbol{I}_{Nd}\right) + \boldsymbol{H}^\top\boldsymbol{R}^{-1}\boldsymbol{H} \tag{12}$$

where $\gamma > 0$ ensures that past measurements are gradually discounted. Assuming the initial condition $\boldsymbol{S}(0)$ follows a sparsity pattern consistent with the graph, this structure is preserved over time under (12). Specifically, for relative measurements (5), this yields

$$\begin{aligned} \dot{\boldsymbol{S}} = &-\left(\boldsymbol{A} + \frac{\gamma}{2}\boldsymbol{I}_{Nd}\right)^\top \boldsymbol{S} - \boldsymbol{S}\left(\boldsymbol{A} + \frac{\gamma}{2}\boldsymbol{I}_{Nd}\right) \\ &+ \boldsymbol{H}^{p\top}\boldsymbol{\Delta}^p\boldsymbol{R}^{p-1}\boldsymbol{\Delta}^p\boldsymbol{H}^p + \boldsymbol{E}_d\mathrm{blkdiag}(\{\boldsymbol{H}^{r\top}_k\boldsymbol{R}^{-1}_k\boldsymbol{H}^r_k\}_{k=1}^M)^\top\boldsymbol{E}_d^\top \end{aligned} \tag{13}$$

Here, the last term represents a matrix-weighted Laplacian matrix [14], where each edge is weighted by the positive semidefinite matrix $\boldsymbol{H}^{r\top}_k\boldsymbol{R}^{-1}_k\boldsymbol{H}^r_k$.

Dynamic Distributed Approximate Inversion of $\boldsymbol{S}$. A remaining challenge is computing $\boldsymbol{S}^{-1}$. As we will discuss in more detail in the next section, under suitable observability assumptions, $\boldsymbol{S}$ remains uniformly positive definite. Moreover, since $\boldsymbol{S}$ has a sparsity pattern consistent with the network topology, the corresponding linear system can be dynamically solved in a distributed manner to obtain the observer correction term using a continuous-time Richardson iteration scheme [15], [16].

Let us define the centralized observer correction term

$$\boldsymbol{\xi} := \boldsymbol{S}^{-1}\boldsymbol{H}^\top\boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{H}\hat{\boldsymbol{x}}). \tag{14}$$

We propose a distributed observer defined as

$$\begin{aligned} \dot{\hat{\boldsymbol{x}}} &= \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{u} + \hat{\boldsymbol{\xi}} \\ \mu\dot{\hat{\boldsymbol{\xi}}} &= -\left(\boldsymbol{S}\hat{\boldsymbol{\xi}} - \boldsymbol{H}^T\boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{H}\hat{\boldsymbol{x}})\right) \\ \dot{\boldsymbol{S}} &= -\left(\boldsymbol{A} + \frac{\gamma}{2}\boldsymbol{I}_{Nd}\right)^\top \boldsymbol{S} - \boldsymbol{S}\left(\boldsymbol{A} + \frac{\gamma}{2}\boldsymbol{I}_{Nd}\right) + \boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{H} \end{aligned} \tag{15}$$

with $\mu > 0$ representing a time-scale separation parameter. We point out that since $\boldsymbol{S}$ is uniformly positive definite, the only equilibrium for the dynamics of $\hat{\boldsymbol{\xi}}$ is given by $\hat{\boldsymbol{\xi}} = \boldsymbol{\xi}$ and, at the equilibrium, we recover the Kalman-like observer from [11].

The dynamics of the observer (15) can be computed in a distributed way. Each agent executes the following:

$$\begin{aligned} \dot{\hat{\boldsymbol{x}}}_i &= \boldsymbol{A}_i\hat{\boldsymbol{x}}_i + \boldsymbol{B}_i\boldsymbol{u}_i + \hat{\boldsymbol{\xi}}_i \\ \mu\dot{\hat{\boldsymbol{\xi}}}_i &= -\sum_{j\in\mathcal{N}_i}\Big(\boldsymbol{S}_{ij}\hat{\boldsymbol{\xi}}_j - \delta^r_{ij}\boldsymbol{H}^{r\top}_{iji}\boldsymbol{R}^{-1}_{ij}(\boldsymbol{y}^r_{ij} - \boldsymbol{H}^r_{iji}\hat{\boldsymbol{x}}_i - \boldsymbol{H}^r_{ijj}\hat{\boldsymbol{x}}_j) \\ &\quad -\delta^r_{ji}\boldsymbol{H}^{r\top}_{jii}\boldsymbol{R}^{-1}_{ji}(\boldsymbol{y}^r_{ji} - \boldsymbol{H}^r_{jii}\hat{\boldsymbol{x}}_i - \boldsymbol{H}^r_{jij}\hat{\boldsymbol{x}}_j)\Big) \\ &\quad - \left(\boldsymbol{S}_{ii}\hat{\boldsymbol{\xi}}_i - \delta^p_i\boldsymbol{H}^{p\top}_i\boldsymbol{R}^{-1}_{ii}(\boldsymbol{y}^p_i - \boldsymbol{H}^p_i\hat{\boldsymbol{x}}_i)\right) \\ \dot{\boldsymbol{S}}_{ii} &= -\left(\boldsymbol{A}_i + \frac{\gamma}{2}\boldsymbol{I}_d\right)^\top \boldsymbol{S}_{ii} - \boldsymbol{S}_{ii}\left(\boldsymbol{A}_i + \frac{\gamma}{2}\boldsymbol{I}_d\right) + \delta^p_i\boldsymbol{H}^{p\top}_i\boldsymbol{R}^{-1}_{ii}\boldsymbol{H}^p_i \\ &\quad + \sum_{j\in\mathcal{N}_i}\left(\delta^r_{ij}\boldsymbol{H}^{r\top}_{iji}\boldsymbol{R}^{-1}_{ij}\boldsymbol{H}^r_{iji} + \delta^r_{ji}\boldsymbol{H}^{r\top}_{jii}\boldsymbol{R}^{-1}_{ji}\boldsymbol{H}^r_{jii}\right) \\ \dot{\boldsymbol{S}}_{ij} &= -\left(\boldsymbol{A}_i + \frac{\gamma}{2}\boldsymbol{I}_d\right)^\top \boldsymbol{S}_{ij} - \boldsymbol{S}_{ij}\left(\boldsymbol{A}_j + \frac{\gamma}{2}\boldsymbol{I}_d\right) \\ &\quad + \delta^r_{ij}\boldsymbol{H}^{rT}_{iji}\boldsymbol{R}^{-1}_{ij}\boldsymbol{H}^r_{ijj} + \delta^r_{ji}\boldsymbol{H}^{r\top}_{jii}\boldsymbol{R}^{-1}_{ji}\boldsymbol{H}^r_{jij} \qquad \text{if } j\in\mathcal{N}_i \end{aligned} \tag{16}$$

To implement this observer, each agent i needs to receive from each neighbor j the following quantities: $\hat{\boldsymbol{\xi}}_j$, $\hat{\boldsymbol{x}}_j$, $\boldsymbol{y}^r_{ji}$; at the same time, we assume the neighbors system matrices to be known, if not, they should be communicated, possibly only once.

IV. Stability Analysis

In this section, we analyse the stability of the proposed observer. First, we discuss the connection between the observability of the system and some uniform properties of $\boldsymbol{S}$ and then use the aforementioned properties in the stability analysis.

A. The observability Gramian

The *observability Gramian* (OG) has been classically used to study the observability of linear time-varying systems.

Definition IV.1. *Given a linear time-varying system defined by the pair $(\boldsymbol{A}(t), \boldsymbol{H}(t))$ with $\boldsymbol{A}(t)\in\mathbb{R}^{d\times d}$ and $\boldsymbol{H}(t)\in\mathbb{R}^{q\times d}$, the associated observability Gramian weighted by a positive definite matrix $\boldsymbol{R}^{-1}$ on an interval $[t_0, t_1]\subset[0,\infty)$ is the positive semidefinite matrix defined by*

$$\boldsymbol{\mathcal{G}}(t_0, t_1) := \int_{t_0}^{t_1}\boldsymbol{\Phi}(\tau, t_0)^\top\boldsymbol{H}(\tau)^\top\boldsymbol{R}^{-1}\boldsymbol{H}(\tau)\boldsymbol{\Phi}(\tau, t_0)\,d\tau \tag{17}$$

where $\mathbf{\Phi}(\tau, t_0) \in$ is the state transition matrix associated to the system, which is the unique solution to

$$\dot{\mathbf{\Phi}}(\tau, t_0) = \boldsymbol{A}(\tau)\mathbf{\Phi}(\tau, t_0) \quad \mathbf{\Phi}(t_0, t_0) = \boldsymbol{I}_d \tag{18}$$

It is well-known [17] that, the pair $(\boldsymbol{A}(t), \boldsymbol{H}(t))$ is observable at time t_0 if and only if there exists a finite $t_1 > t_0$ such that the OG in (17) is nonsingular.

Before proceeding, we make the following set of assumptions.

Assumption 3. *The following conditions are satisfied.*

- $\boldsymbol{S}(0) \succ \mathbf{0}$
- *The system satisfies the following observability condition*

$$\boldsymbol{\mathcal{G}}(t - \bar{t}, t) \succeq \alpha \boldsymbol{I}_{Nd} \quad \forall t \geq t_0 \geq \bar{t} \quad \alpha > 0 \tag{19}$$

- $\gamma > 2\bar{a}$

Notice that the first and third assumptionss directly depend on design parameters. Together, these assumptions ensure that the matrix $\boldsymbol{S}$ is uniformly bounded from below and from above, as discussed in the following proposition directly adapted from [11, Lemma 1].

Proposition 1. *Any solution to* (12)*, initialized at* $\boldsymbol{S}(0) \succ \mathbf{0}$ *is positive definite for all* t*, and if* (19) *is satisfied, then,*

$$\underline{s}_\gamma \boldsymbol{I}_{Nd} \preceq \boldsymbol{S}(t) \quad \forall t \geq t_0 \text{ with } \underline{s}_\gamma := \alpha e^{-\gamma \bar{t}}. \tag{20}$$

Furthermore, under assumption 1 and assuming $\gamma > 2\bar{a}$, every solution to (12) *is bounded*

$$\boldsymbol{S}(t) \preceq \bar{s}_\gamma \boldsymbol{I}_{Nd} \quad \forall t \geq 0 \tag{21}$$

with $\bar{s}_\gamma := \max\{\bar{\lambda}(\boldsymbol{S}(0)), \frac{\bar{h}^2}{(\gamma - 2\bar{a})\underline{\lambda}(\boldsymbol{R})}\}$, where we used the symbols $\bar{\lambda}(\cdot)$ and $\underline{\lambda}(\cdot)$ to denote the maximum and minimum eigenvalue functions.

Remark 1. *In certain cases where simplified dynamics are used and the state transition matrix is known in closed form, checking observability via* (19) *becomes more tractable by leveraging the connection between the observability Gramian (OG) and matrix-weighted Laplacians. For instance, in robotics, the single integrator model is a common reduced-order approximation, where $\boldsymbol{A} = \mathbf{0}_{d\times d}$ and the state transition matrix is the identity. Assuming relative measurements as in* (5) *and block-diagonal $\boldsymbol{R}$, the OG in* (17) *simplifies to:*

$$\begin{aligned} \boldsymbol{\mathcal{G}}(t_0, t_1) &= \int_{t_0}^{t_1} \boldsymbol{E}_d \text{blkdiag}(\{\boldsymbol{H}_k^{r\top} \boldsymbol{R}_k^{-1} \boldsymbol{H}_k^r\}_{k=1}^M)^\top \boldsymbol{E}_d^\top \, d\tau \\ &= \boldsymbol{E}_d \text{blkdiag}\left(\left\{\int_{t_0}^{t_1} \boldsymbol{H}_k^{r\top} \boldsymbol{R}_k^{-1} \boldsymbol{H}_k^r \, d\tau\right\}_{k=1}^M\right)^\top \boldsymbol{E}_d^\top. \end{aligned} \tag{22}$$

This shows that the OG reduces to a matrix-weighted Laplacian matrix [14]. Including private measurements yields a grounded Laplacian [18], [19], for which several results facilitate verification of positive definiteness [12]. Similarly, for double integrator dynamics, observability reduces to checking positive definiteness of a grounded matrix-weighted Laplacian, as shown in [20, Theorem 1].

B. Lyapunov analysis

We analyze the stability of the proposed observer using singular perturbation theory and a composite Lyapunov function. We show that for sufficiently small μ, the error system exhibits uniform global exponential stability (UGES) under Assumptions 1-3.

Define the errors:

$$\tilde{\boldsymbol{x}} = \boldsymbol{x} - \hat{\boldsymbol{x}} \qquad \tilde{\boldsymbol{\xi}} = \boldsymbol{\xi} - \hat{\boldsymbol{\xi}}. \tag{23}$$

Using (7), (14) and (15), the dynamics of the error $\tilde{\boldsymbol{x}}$ can be expressed as:

$$\begin{aligned} \dot{\tilde{\boldsymbol{x}}} &= \boldsymbol{A}\tilde{\boldsymbol{x}} - \hat{\boldsymbol{\xi}} + \boldsymbol{\xi} - \boldsymbol{\xi} \\ &= (\boldsymbol{A} - \boldsymbol{S}^{-1}\boldsymbol{H}^\top \boldsymbol{R}^{-1}\boldsymbol{H})\tilde{\boldsymbol{x}} + \tilde{\boldsymbol{\xi}}. \end{aligned} \tag{24}$$

Using (14) and (15), the correction term error dynamics are instead given as:

$$\begin{aligned} \dot{\tilde{\boldsymbol{\xi}}} &= \dot{\boldsymbol{\xi}} + \frac{1}{\mu}\left(\boldsymbol{S}\hat{\boldsymbol{\xi}} - \boldsymbol{H}^\top \boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{H}\hat{\boldsymbol{x}})\right) \\ &= \dot{\boldsymbol{\xi}} + \frac{1}{\mu}\boldsymbol{S}\left(\hat{\boldsymbol{\xi}} - \boldsymbol{S}^{-1}\boldsymbol{H}^\top \boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{H}\hat{\boldsymbol{x}})\right) \\ &= -\frac{1}{\mu}\boldsymbol{S}\tilde{\boldsymbol{\xi}} + \dot{\boldsymbol{\xi}}. \end{aligned} \tag{25}$$

Summing up, the full error system dynamics can be written as

$$\begin{aligned} \dot{\tilde{\boldsymbol{x}}} &= (\boldsymbol{A} - \boldsymbol{S}^{-1}\boldsymbol{H}^\top \boldsymbol{R}^{-1}\boldsymbol{H})\tilde{\boldsymbol{x}} + \tilde{\boldsymbol{\xi}} \\ \dot{\tilde{\boldsymbol{\xi}}} &= -\frac{1}{\mu}\boldsymbol{S}\tilde{\boldsymbol{\xi}} + \dot{\boldsymbol{\xi}} \\ \dot{\boldsymbol{S}} &= -\left(\boldsymbol{A} + \frac{\gamma}{2}\boldsymbol{I}_{Nd}\right)^\top \boldsymbol{S} - \boldsymbol{S}\left(\boldsymbol{A} + \frac{\gamma}{2}\boldsymbol{I}_{Nd}\right) + \boldsymbol{H}^\top \boldsymbol{R}^{-1}\boldsymbol{H}. \end{aligned} \tag{26}$$

We first analyze the fast subsystem, then the slow subsystem, and finally derive uniform global exponential stability using a composite Lyapunov function. Due to the small parameter μ, the system exhibits time-scale separation: $\tilde{\boldsymbol{\xi}}$ evolves on a fast time scale $\mathcal{O}(1/\mu)$, while $\tilde{\boldsymbol{x}}$ evolves more slowly. For $\mu \to 0$, $\tilde{\boldsymbol{x}}$ appears quasi-static to the fast dynamics, allowing us to treat it as a constant when analyzing $\tilde{\boldsymbol{\xi}}$. Changing the time variable from t to $\tau = (t - t_0)/\mu$:

$$\frac{d\boldsymbol{\xi}}{d\tau} = \frac{\partial \boldsymbol{\xi}(\tilde{\boldsymbol{x}})}{\partial \tilde{\boldsymbol{x}}}\frac{d\tilde{\boldsymbol{x}}}{d\tau} \to \mathbf{0}. \tag{27}$$

Lemma 1 (Exponential Stability of the Fast Subsystem)**.** *For $\mu \to 0$, under Assumption 3, the fast subsystem*

$$\frac{d\tilde{\boldsymbol{\xi}}}{d\tau} = -\boldsymbol{S}(\tau)\tilde{\boldsymbol{\xi}} \tag{28}$$

is UGES with rate $\underline{s}_\gamma$.

Proof. Consider the candidate Lyapunov function for the fast subsystem:

$$V_\xi = \frac{1}{2}\tilde{\boldsymbol{\xi}}^\top \tilde{\boldsymbol{\xi}} \tag{29}$$

Its derivative along the trajectories of (28) satisfies:

$$\frac{dV_\xi}{d\tau} = -\tilde{\boldsymbol{\xi}}^\top \boldsymbol{S}(\tau)\tilde{\boldsymbol{\xi}} \quad \leq -\underline{s}_\gamma \left\|\tilde{\boldsymbol{\xi}}\right\|^2 \tag{30}$$

where we used the lower bound in (20). □

Lemma 2 (Exponential Stability of the Slow Subsystem)**.** *Under the Assumption 3, when $\hat{\boldsymbol{\xi}} = \boldsymbol{\xi}$, the slow subsystem*

$$\dot{\tilde{\boldsymbol{x}}} = (\boldsymbol{A} - \boldsymbol{S}^{-1}\boldsymbol{H}^\top \boldsymbol{R}^{-1}\boldsymbol{H})\tilde{\boldsymbol{x}} \tag{31}$$

is UGES with a rate of at least γ.

Proof. Define the candidate Lyapunov function:

$$V_x = \tilde{\boldsymbol{x}}^\top \boldsymbol{S}(t)\tilde{\boldsymbol{x}}, \tag{32}$$

which satisfies $\underline{s}_\gamma \|\tilde{\boldsymbol{x}}\|^2 \le V_x \le \bar{s}_\gamma \|\tilde{\boldsymbol{x}}\|^2$. Its time derivative satisfies:

$$\begin{aligned}\dot{V}_x &= \tilde{\boldsymbol{x}}^\top \dot{\boldsymbol{S}}\tilde{\boldsymbol{x}} + \tilde{\boldsymbol{x}}^\top(\boldsymbol{A}^\top\boldsymbol{S} + \boldsymbol{S}\boldsymbol{A})\tilde{\boldsymbol{x}} - 2\tilde{\boldsymbol{x}}^\top\boldsymbol{H}^\top\boldsymbol{R}^{-1}\boldsymbol{H}\tilde{\boldsymbol{x}} \\ &= -\tilde{\boldsymbol{x}}^\top\left(\gamma\boldsymbol{S} + \boldsymbol{H}^\top\boldsymbol{R}^{-1}\boldsymbol{H}\right)\tilde{\boldsymbol{x}} \le -\gamma\underline{s}_\gamma\|\tilde{\boldsymbol{x}}\|^2,\end{aligned} \tag{33}$$

where $\underline{s}_\gamma$ is strictly positive, confirming the exponential stability for the slow subsystem. □

Lemma 3. *Under Assumptions 1- 3, the following holds:*

$$\left\|\dot{\boldsymbol{\xi}}\right\| \le k_\xi \left\|\tilde{\boldsymbol{\xi}}\right\| + k_x \|\tilde{\boldsymbol{x}}\| \tag{34}$$

Proof. Expanding $\dot{\boldsymbol{\xi}}$ and performing standard bookkeeping, we obtain:

$$\begin{aligned}\dot{\boldsymbol{\xi}} = &\left[\frac{d}{dt}(\boldsymbol{S}^{-1}\boldsymbol{H}^\top\boldsymbol{R}^{-1})\boldsymbol{H}\right. \\ &\left.+(\boldsymbol{S}^{-1}\boldsymbol{H}^\top\boldsymbol{R}^{-1}(\dot{\boldsymbol{H}} + \boldsymbol{H}(\boldsymbol{A} + \boldsymbol{S}^{-1}\boldsymbol{H}^\top\boldsymbol{R}^{-1}\boldsymbol{H})))\right]\tilde{\boldsymbol{x}} \\ &- \boldsymbol{S}^{-1}\boldsymbol{H}^\top\boldsymbol{R}^{-1}\boldsymbol{H}\tilde{\boldsymbol{\xi}},\end{aligned} \tag{35}$$

all matrices involved are uniformly bounded, allowing us to upper bound each term. This yields (34), where k_ξ and k_x are determined by the product of the respective upper bounds in (35). □

We are now ready to state the main theorem.

Theorem 1 (Exponential Stability of the Interconnected System)**.** *Under Assumptions 1-3, for sufficiently small μ, the error system* (26) *is UGES.*

Proof. Consider the composite candidate Lyapunov function:

$$V = bV_x + V_\xi \tag{36}$$

with $b > 0$ being a constant to be chosen. The time derivative satisfies:

$$\begin{aligned}\dot{V} &\le -b\gamma\underline{s}_\gamma\|\tilde{\boldsymbol{x}}\|^2 + (2b\bar{s}_\gamma + k_x)\|\boldsymbol{\xi}\|\|\boldsymbol{x}\| - \left(\frac{\underline{s}_\gamma}{\mu} - k_\xi\right)\left\|\tilde{\boldsymbol{\xi}}\right\|^2 \\ &= \begin{bmatrix}\|\tilde{\boldsymbol{x}}\| \\ \|\tilde{\boldsymbol{\xi}}\|\end{bmatrix}^\top \begin{bmatrix}-b\gamma\underline{s}_\gamma & \frac{1}{2}(2b\bar{s}_\gamma + k_x) \\ \frac{1}{2}(2b\bar{s}_\gamma + k_x) & -\left(\frac{\underline{s}_\gamma}{\mu} - k_\xi\right)\end{bmatrix} \begin{bmatrix}\|\tilde{\boldsymbol{x}}\| \\ \|\tilde{\boldsymbol{\xi}}\|\end{bmatrix}.\end{aligned} \tag{37}$$

Applying Sylvester's criterion, $\dot{V}$ is negative definite when:

$$b\gamma\underline{s}_\gamma\left(\frac{\underline{s}_\gamma}{\mu} - k_\xi\right) > \frac{1}{4}(2b\bar{s}_\gamma + k_x)^2. \tag{38}$$

This condition holds for:

$$\mu < \frac{4b\gamma}{(2b\bar{s}_\gamma/\underline{s}_\gamma + k_x/\underline{s}_\gamma)^2 + 4b\gamma k_\xi/\underline{s}_\gamma}. \tag{39}$$

It follows that, the system is UGES for μ satisfying (39). □

The bound (39) highlights key design principles:

- Excitation level $\underline{s}_\gamma$: higher $\underline{s}_\gamma$ improves robustness and reduces the effect of time-varying terms in $\dot{\boldsymbol{\xi}}$.
- Condition number $\frac{\bar{s}_\gamma}{\underline{s}_\gamma}$: well-conditioned $\boldsymbol{S}$ improves stability, aligning with the intuition that solving ill-conditioned equations leads to poor stability properties.
- Forgetting factor γ:
  - higher γ accelerates slow dynamics convergence.
  - excessive γ reduces $\boldsymbol{S}$ faster, inducing a high gain which amplifies noise effects.

Thus, γ and the conditioning of $\boldsymbol{S}$ must be carefully tuned. While for simplicity we considered a scalar γ, a diagonal matrix $\boldsymbol{\Gamma}$ could be used to weight different state components differently. Furthermore, if the time dependence of $\boldsymbol{A}(t)$ and $\boldsymbol{H}(t)$ is input-driven, active sensing strategies [9], [21], [22] could be employed to optimize $\underline{s}_\gamma$ or the conditioning of $\boldsymbol{S}$.

Remark 2. *A potential modification to improve performance, inspired by the forgetting factor approach in parameter estimation [23, Chapter 8], is to enforce a lower bound on $\boldsymbol{S}$. This can be achieved by modifying* (12) *as follows:*

$$\dot{\boldsymbol{S}} = -\gamma(\boldsymbol{S} - \boldsymbol{S}_{\min}) - \boldsymbol{A}^\top\boldsymbol{S} - \boldsymbol{S}\boldsymbol{A} + \boldsymbol{H}^\top\boldsymbol{R}^{-1}\boldsymbol{H}, \tag{40}$$

where $\boldsymbol{S}_{\min} \succ \boldsymbol{0}$ ensures that $\boldsymbol{S}$ remains uniformly above a certain minimum threshold. This modification prevents $\boldsymbol{S}$ from decreasing too much, mitigating excessive sensitivity to newly acquired measurements.

Remark 3. *The correction term dynamics introduce a low-pass filtering effect on newly acquired measurements, thereby reducing sensitivity to high-frequency measurement noise.*

V. Simulation Results

We evaluate the proposed distributed Kalman-like observer in a cooperative localization scenario where $N = 15$ mobile robots estimate their positions using relative position measurements, with $A = 4$ anchors having access to absolute positions. The robots follow double integrator dynamics:

$$\dot{\boldsymbol{p}}_i = \boldsymbol{v}_i \quad \dot{\boldsymbol{v}}_i = \boldsymbol{u}_i \tag{41}$$

where $\boldsymbol{p}_i \in \mathbb{R}^2$ represents the position of the i-th robot, $\boldsymbol{v}_i \in \mathbb{R}^2$ its velocity and $\boldsymbol{u}_i \in \mathbb{R}^2$ the control input. The measurement model is given by:

$$\boldsymbol{y}_{ij}^r = \boldsymbol{p}_j - \boldsymbol{p}_i \qquad \boldsymbol{y}_i^p = \delta_i\boldsymbol{p}_i. \tag{42}$$

A random initial configuration and graph topology are generated (Fig. 1), and robots move at constant speed $v = 1\,\mathrm{m\,s^{-1}}$ with random direction. The initial estimate is obtained by adding gaussian noise $\nu \sim \mathcal{N}(0, \sigma^2)$ to each component of the actual robot positions, with $\sigma = 1.3\,\mathrm{m}$. The modification

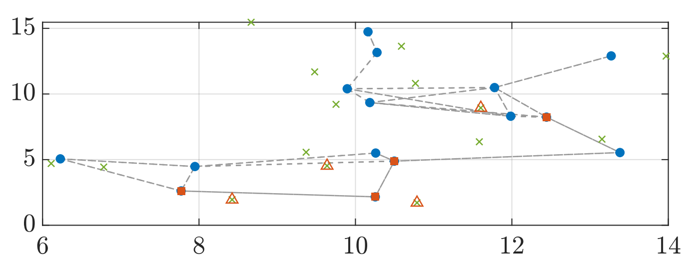

Fig. 1: Initial configuration of the real robots (circle markers) and estimated robots (cross markers). The red circles represent the anchors and their initial estimate is shown with a triangle marker around. The network communication graph is also shown.

(40) is adopted as well as a diagonal forgetting factor $\boldsymbol{\Gamma}$, which applies different decay rates to position and velocity, with components γ_p and γ_v, respectively. The initial condition $\boldsymbol{S}(0)$ is obtained from the inverse of the initial covariance matrix. The used parameters are $\boldsymbol{S}_{\min} = 0.5\boldsymbol{S}(0)$, $\gamma_p = 6.0$, $\gamma_v = 40.0$

and output gain matrices $\boldsymbol{R}_i^p = \boldsymbol{I}_2$, $\forall i \in \mathcal{V}$ and $\boldsymbol{R}_{ij}^r = \boldsymbol{I}_2$, $\forall (i,j) \in \mathcal{E}_s$ are tuned so as to ensure a good conditioning of $\boldsymbol{S}$. The parameter μ is selected as $\mu = 0.1$, which leads to satisfactory convergence of the fast dynamics (Fig. 2 and Fig. 3). As a result the estimation error converges to zero 4, as expected.

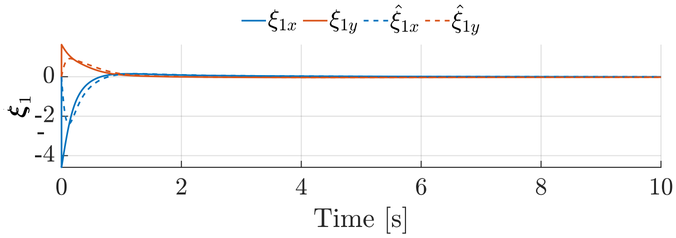

Fig. 2: Ideal correction term (solid lines) and estimated correction term (dashed lines) for robot 1, i.e. $\boldsymbol{\xi}_1$ and $\hat{\boldsymbol{\xi}}_1$.

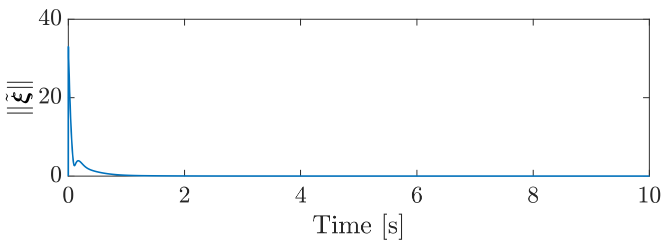

Fig. 3: Norm of the error on the correction term, i.e. $||\tilde{\boldsymbol{\xi}}||$.

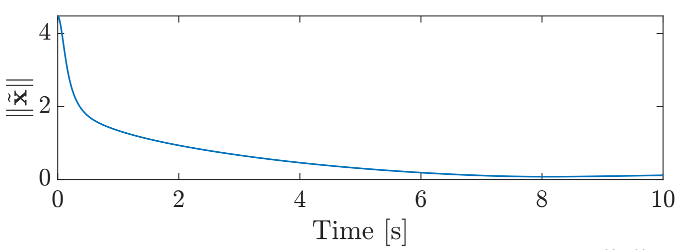

Fig. 4: Norm of the estimation error on the full state, i.e. $||\tilde{\boldsymbol{x}}||$.

VI. Conclusions

This paper proposed a distributed Kalman-like observer for cooperative state estimation in multi-agent systems. By replacing the explicit process noise covariance with a forgetting factor, the observer ensures distributed computation of the information matrix while preserving stability properties. A dynamically computed correction term enables each agent to update its estimate without requiring matrix inversion.

The observer requires only joint observability, making it applicable to systems with sparse or heterogeneous sensing. Simulations in a cooperative localization scenario confirm its effectiveness, demonstrating improved estimation accuracy with minimal communication overhead.

References

[1] L. Luft, T. Schubert, S. I. Roumeliotis, and W. Burgard, "Recursive decentralized localization for multi-robot systems with asynchronous pairwise communication," *The International Journal of Robotics Research*, vol. 37, no. 10, pp. 1152–1167, 2018.

[2] J. Yang, W.-A. Zhang, and F. Guo, "Distributed kalman-like filtering and bad data detection in the large-scale power system," *IEEE transactions on industrial informatics*, vol. 18, no. 8, pp. 5096–5104, 2021.

[3] G. Battistelli, L. Chisci, G. Mugnai, A. Farina, and A. Graziano, "Consensus-based linear and nonlinear filtering," *IEEE Transactions on Automatic Control*, vol. 60, no. 5, pp. 1410–1415, 2014.

[4] A. T. Kamal, J. A. Farrell, and A. K. Roy-Chowdhury, "Information weighted consensus filters and their application in distributed camera networks," *IEEE Transactions on Automatic Control*, vol. 58, no. 12, pp. 3112–3125, 2013.

[5] R. Jung, L. Luft, and S. Weiss, "Isolated kalman filtering: theory and decoupled estimator design," *Autonomous Robots*, vol. 49, no. 1, pp. 1–25, 2025.

[6] Y. Sun, M. Fu, B. Wang, H. Zhang, and D. Marelli, "Dynamic state estimation for power networks using distributed map technique," *Automatica*, vol. 73, pp. 27–37, 2016.

[7] G. Bornard, N. Couenne, and F. Celle, "Regularly persistent observers for bilinear systems," in *New Trends in Nonlinear Control Theory: Proceedings of an International Conference on Nonlinear Systems, Nantes, France, June 13–17, 1988*. Springer, 1989, pp. 130–140.

[8] H. Hammouri and J. de Leon Morales, "Observer synthesis for state-affine systems," in *29th IEEE Conference on Decision and Control*. IEEE, 1990, pp. 784–785.

[9] P. Bernard, V. Andrieu, and D. Astolfi, "Observer design for continuous-time dynamical systems," *Annual Reviews in Control*, vol. 53, pp. 224–248, 2022.

[10] G. Besançon, G. Bornard, and H. Hammouri, "Observer synthesis for a class of nonlinear control systems," *European Journal of control*, vol. 2, no. 3, pp. 176–192, 1996.

[11] P. Bernard, N. Mimmo, and L. Marconi, "On the semi-global stability of an ek-like filter," *IEEE Control Systems Letters*, vol. 5, no. 5, pp. 1771–1776, 2020.

[12] F. Bullo, J. Cortes, F. Dörfler, and S. Martinez, *Lectures on network systems*. CreateSpace, 2018, vol. 1.

[13] B. D. Anderson and J. B. Moore, *Optimal filtering*. Courier Corporation, 2005.

[14] S. E. Tuna, "Synchronization under matrix-weighted laplacian," *Automatica*, vol. 73, pp. 76–81, 2016.

[15] D. Bertsekas and J. Tsitsiklis, *Parallel and distributed computation: numerical methods*. Athena Scientific, 2015.

[16] N. De Carli, P. Salaris, and P. R. Giordano, "Distributed control barrier functions for global connectivity maintenance," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 12 048–12 054.

[17] C.-T. Chen, *Linear system theory and design*. Saunders college publishing, 1984.

[18] F. Dorfler and F. Bullo, "Kron reduction of graphs with applications to electrical networks," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 1, pp. 150–163, 2012.

[19] M. Pirani and S. Sundaram, "On the smallest eigenvalue of grounded laplacian matrices," *IEEE Transactions on Automatic Control*, vol. 61, no. 2, pp. 509–514, 2015.

[20] Z. Tang, B. Fidan, K. H. Johansson, J. Martensson, and T. Hamel, "Observer-based control of second-order multi-vehicle systems in bearing-persistently exciting formations," in *2024 IEEE 63rd Conference on Decision and Control (CDC)*. IEEE, 2024, pp. 7522–7527.

[21] P. Salaris, M. Cognetti, R. Spica, and P. R. Giordano, "Online optimal perception-aware trajectory generation," *IEEE Transactions on Robotics*, vol. 35, no. 6, pp. 1307–1322, 2019.

[22] N. De Carli, P. Salaris, and P. R. Giordano, "Multi-robot active sensing for bearing formations," in *2023 International Symposium on Multi-Robot and Multi-Agent Systems (MRS)*. IEEE, 2023, pp. 184–190.

[23] J.-J. E. Slotine, W. Li *et al.*, *Applied nonlinear control*. Prentice hall Englewood Cliffs, NJ, 1991, vol. 199, no. 1.