Fast Attack-Resilient Distributed State Estimator for Cyber-Physical Systems

Feng Yu, Raj Gautam Dutta, *Member, IEEE*, Teng Zhang, Yaodan Hu, *Student Member, IEEE*,
Yier Jin, *Senior Member, IEEE*

***Abstract*—The performance of resilient state estimators developed for cyber-physical systems (CPS) decreases as the number of compromised sensors of the system increases. Furthermore, some of these algorithms leverage computationally expensive optimization techniques to incorporate resiliency. As such, we propose Fast Resilient Distributed State Estimator (FRDSE), which is a novel resilient distributed algorithm that produces bounded state estimation errors regardless of the magnitude of the attack and the number of compromised sensors. Our algorithm converges to the true state in an attack-free and noise-free scenario and it produces bounded estimation errors during an attack. Compared to existing algorithms, FRDSE is more computationally efficient. We provide theoretical guarantees on the convergence of FRDSE in attack free scenario and prove its resiliency during an attack. We demonstrate the performance of our algorithm against False Data Injection (FDI) attack in a platoon of vehicles and compare its run time against existing algorithms. We observe that on a platoon of eight vehicles, run time of our algorithm is 0.102 seconds, much lower than the state-of-the-art solutions.**

***Index Terms*—Cyber-Physical System, Vehicle Platoon, LTI system, Security, Distributed Estimation, Kalman Filter**

I. Introduction

Decentralized Kalman filter uses the information of all components of the distributed system for state estimation [1], [2], resulting in a computational complexity of $\mathcal{O}(n^2)$. Due to the computational inefficiency of decentralized Kalman filter, efforts were made to design the Distributed Kalman Filter (DKF), where each node of the network communicates with only its neighbors [3]–[8]. DKF has an estimation step and an average-consensus step that fuses sensor data and covariance data [3]. Under certain conditions, these filters are optimal for linear stochastic distributed systems. However, they are not resilient to adversarial attacks. Consequently, attempts have been made to develop attack-resilient distributed state estimators [9]–[13].

Dadras et al. [9] proposed a detection scheme against gain modification attack and destabilizing attack on vehicle platoons. Their method combined the system identification approach with a thresholding/clustering method. In their scheme, the system matrix of each vehicle was identified by considering the input-output data; but it did not require any knowledge of normal and adversarial parameters and the number and the locations of attackers. Sajjad et al. [10] designed a sliding mode controller (SMC) with an attack detection scheme to reduce the damage caused by a collision in a platoon. Their attack detector was decentralized and relied on local sensor information. Khan and Stanković [11] proposed attack detection and single message exchange state estimation methods for compromised communication and sensing scenarios. Their method relied on local consistency and nodal consistency of data sets. Matei et al. [12] designed a multi-agent filtering scheme in conjunction with a trust-based mechanism to secure the state estimates of power grids under a false data injection attack. In their approach, an agent of the grid computed local state estimates based on their own measurement and of their trusted neighbors. However, both [6], [12] did not provide any theoretical guarantees of their methods.

Mitra and Sundaram [14] developed a secure distributed observer for the Byzantine adversary model, where some nodes of the network were compromised by adversaries. Prior to the state estimation, they decomposed the linear system model using Kalman's observability decomposition method. Then, Luenberger observers at each node estimated the states corresponding to detectable eigenvalues [15]. The undetectable portions of the states at each node were estimated using a secure consensus algorithm using measurements of well-behaving neighboring nodes. However, their method required the network to be highly connected and they assumed that only a small number of nodes are corrupted, which was known by their algorithm. In addition, they assumed that the system matrix only had simple and real eigenvalues, which might not hold in practice.

Dutta et al. [16] developed Resilient Distributed Kalman Filter (RDKF) based on distributed Kalman filter, which was resilient to sensor attacks on a distributed system. Compared to previous methods, their method showed an asymptotic convergence of estimation error to zero when there was no attack and that during an attack, the disturbance on the state estimates of RDKF was bounded. In their method, they used a convex optimization library to solve the minimization problem during each iteration, which was computationally inefficient. The RDKF was based on the idea of minimizing the total innovation of the Kalman filter at each time-step and was motivated by the distributed multi-agent optimization of [17]. However, it was more computationally expensive as each step of the algorithm required solving an optimization problem

Manuscript received April 17, 2020; revised June 17, 2020; accepted July 6, 2020. This article was presented in the International Conference on Embedded Software 2020 and appears as part of the ESWEEK-TCAD special issue.

Feng Yu and Teng Zhang are with Department of Mathematics, University of Central Florida, Orlando, FL 32816 USA (email: {feng.yu, teng.zhang}@ucf.edu).

Raj Gautam Dutta, Yaodan Hu and Yier Jin are with Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA ({r.dutta, cindy.hu}@ufl.edu, yier.jin@ece.ufl.edu).

Corresponding Author: Yier Jin

using an external optimization library.

To improve the computational efficiency of the method in [16], we look into the literature on distributed multi-agent optimization. The objective of the distributed multi-agent optimization is to cooperatively minimize the cost function $\sum_{i=1}^{n} f_i(x)$, where f_i is the cost of i-th vehicle and x is the state of the system. Rabbat and Nowak [18] solved the optimization problem by using an incremental subgradient approach for a ring-type network. For other types of networks, Nedić and Ozdaglar [19] proposed an algorithm called decentralized gradient descent (DGD) that assumed each cost function to be convex and had a bounded (sub)gradient. Other existing decentralized algorithms for solving the distributed multi-agent optimization problem were [20]–[22]. With a fixed step size, these algorithms converged to a point which lied on a neighborhood of the true solution, irrespective of the differentiability or non-differentiability of the cost function [23]. To guarantee convergence to the true solution, an approach was to use a diminishing step size that generally leaded to a lower convergence rate [20]–[22]. Shi et al. [24] proposed an algorithm called EXTRA which converged to the true solution, but the authors considered a fixed step size to obtain a higher convergence rate. The EXTRA algorithm required the cost function to be differentiable, and PG-EXTRA [25] was the non-differentiable version of the algorithm.

We make the following contributions:

- We design Fast Resilient Distributed State Estimator (FRDSE) that can detect data integrity attacks on sensor measurements of a distributed system such as vehicle platoon.
- FRDSE has two unique characteristics: i) the number of neighbors of an agent that can be compromised is not restricted and ii) the estimator's performance does not degrade (beyond an upper bound) with the magnitude of the attack.
- Improving from the Resilient Distributed Kalman Filter (RDKF) [16], FRDSE significantly enhances the computational efficiency because of its closed-form solution.
- We provide theoretical guarantees on FRDSE's convergence when there is no attack and resilience when attack presents.

The rest of the paper is organized as follows: In Section II, we present the notations and the communication graph of the system, describe the dynamic system (of the platoon) and attack models, formulate the resilient distributed estimation problem, and show some convergence results of DKF. In Section III, we present our fast resilient distributed state estimator (FRDSE) and give its performance analysis. The effectiveness of FRDSE is demonstrated on numerical examples in Section IV. Final conclusions are drawn in Section V. Proofs of theorems and lemmas are given in Appendix.

II. Preliminaries and Problem Description

A. Notations

We consider a platoon of n vehicles and the communication between these vehicles is described as an undirected graph $G = (\mathcal{V}, \mathcal{E})$. In the graph, the vertex set $\mathcal{V} = \{1, 2, \cdots, n\}$ stands for the vehicles/nodes/agents and the edges $\mathcal{E} = \mathcal{V} \times \mathcal{V}$ represent the communication links between them. A bi-directional edge, $(i, j) \in \mathcal{E}$, between the i-th and the j-th vehicles, enables them to send and receive messages between each other, but not simultaneously. We also assume that each vehicle has its own information, i.e., $(i, i) \in \mathcal{E}$ for all $i = 1, 2, \cdots, n$. The set $N^{(i)} = \{i\} \cup \{j \in \mathcal{V} : (i, j) \in \mathcal{E}\}$ denotes the set of neighbors of the i-th vehicle, wherein the node i itself is accounted to avoid the special case that the i-th vehicle does not have neighbors (e.g. the leading vehicle). Furthermore, we assume that each agent has an observer composed of q distinct sensors. The sensor measurements are leveraged for estimating the states of the system.

Throughout the paper, we use $\mathbf{P}^{(i)-1}$ and $\mathbf{P}^{(i)T}$ to denote the inverse matrix and the transpose matrix of $\mathbf{P}^{(i)}$, respectively.

B. System and Measurement Models

We model the dynamics of the distributed system (platoon) of n agents/vehicles as a discrete linear time-invariant (LTI) model,

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k, \tag{1}$$

$$\mathbf{y}_k^{(i),a} = \mathbf{C}^{(i)}\mathbf{x}_k + \mathbf{a}_k^{(i)}, 1 \le i \le n \tag{2}$$

where $\mathbf{x}_k = [\mathbf{x}_{k,1}; \mathbf{x}_{k,2}; \cdots; \mathbf{x}_{k,n}] \in \mathbb{R}^{ns}$ is the state vector of the system at time $k \in \mathbb{N}$ with $\mathbf{x}_{k,i} \in \mathbb{R}^s$ being the state of the i-th agent, $\mathbf{A} \in \mathbb{R}^{(ns)\times(ns)}$ is the system matrix, $\mathbf{y}_k^{(i)} \in \mathbb{R}^q$ is the measurement vector of the q sensors at time k of the i-th agent, $\mathbf{C}^{(i)} \in \mathbb{R}^{q\times(ns)}$ is the observation matrix of the i-th agent, and $\mathbf{a}_k^{(i)}$ is the attack vector. Here, $\mathbf{y}_k^{(i),a} \in \mathbb{R}^q$ is corrupted when $\mathbf{a}_k^{(i)} \neq 0$ for any k. The vector $\mathbf{a}_k^{(i)}$ denotes the attack vector at time k and its value depends on the attacker. In this paper, we assume that each vehicle estimates the state vector based on the measurements from its neighbors and its own measurements. Also, an agent (good or malicious) is assumed to transmit the same information to all its neighbors, which appears in many practical scenarios such as in vehicular ad-hoc networks.

We assume adversaries can manipulate any number of sensors of compromised nodes, $\mathcal{V}_a \subset \mathcal{V}$, of the network and has the knowledge of observation matrices of the corrupted nodes, the system matrix $\mathbf{A}$, and the communication topology G. The malicious measurements affect the state estimation of the corrupted vehicles and also their neighbors. In this way, the attack influences the state estimation of the distributed system. Formally, we define a compromised agent (vehicle) as:

Definition 1: Compromised Agent: An agent i is compromised at time $k \in \mathbb{N}$ if its attack vector $\mathbf{a}_k^{(i)} \neq 0$.

C. Resilient Distributed Estimation Problem

Given an LTI distributed system of n agents with a linear measurement model and a communication graph G, we would like to use a distributed algorithm in the form of

$$\hat{\mathbf{x}}_k^{(i)} = f(\{\hat{\mathbf{x}}_{k-1}^{(j)}\}_{j\in N(i)}, \mathbf{y}_k^{(i),a}, \mathbf{A}, \{\mathbf{C}^{(i)}\}_{1\le i\le n})$$

to estimate $\mathbf{x}_k^{(i)}$. The goal is to control the overall estimation errors, $\|\mathbf{e}_k^{(i)}\| = \|\hat{\mathbf{x}}_k^{(i)} - \mathbf{x}_k\|$, i.e., to make the overall errors converge to zero in the attack-free case and the errors are bounded when the attack is launched. Throughout this paper, we make the following assumptions:

- The pair $(\mathbf{A}, \mathcal{C}) := (\mathbf{A}, [\mathbf{C}^{(i)}; \cdots ; \mathbf{C}^{(n)}])$ is detectable of the system.
- Agents/Vehicles transfer estimated state information to its neighbors through a secure communication channel. Thus, we assume that there is no attack on the network.
- We assume that the agents/vehicles are not capable of detecting sensor attacks on its neighbors and thus, accept both malicious and non-malicious state estimates from its neighbors.

D. Distributed Kalman Filter

First of all, let us briefly revisit Kalman filter from a Bayesian interpretation. Assume the vectors $w_k^{(i)} \overset{i.i.d}{\sim} \mathcal{N}(0, \Sigma_w^{(i)})$ and $v_k^{(i)} \overset{i.i.d}{\sim} \mathcal{N}(0, \Sigma_v^{(i)})$ are additive white Gaussian noise. We follow the convention that $\mathcal{N}(\mu, \Sigma)$ represents the Gaussian distribution with mean μ and covariance Σ. A Kalman filter consists of two stages–the *prediction* stage and the *correction* stage. For succinctness, we abuse the notation $\mathbf{x}_k$ to denote $\mathbf{x}_k^{(i)}$ of agent i. In the prediction stage, we assume that the distribution of $\mathbf{x}_k$ follows the Gaussian distribution $\mathcal{N}(\hat{\mathbf{x}}_k, \mathbf{P}_k)$ to obtain *a prior* distribution, $z_k \sim \mathcal{N}(\hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1})$, where

$$\begin{aligned}\hat{\mathbf{x}}_{k|k-1} &= \mathbf{A}\hat{\mathbf{x}}_{k-1},\\ \mathbf{P}_{k|k-1} &= \operatorname{Cov}(\hat{\mathbf{x}}_{k|k-1} - \mathbf{x}_k)\\ &= \mathbf{A}\operatorname{Cov}(\hat{\mathbf{x}}_{k-1} - \mathbf{x}_{k-1})\mathbf{A}^T + \operatorname{Cov}(w_{k-1})\\ &= \mathbf{A}\mathbf{P}_{k-1}\mathbf{A}^T + \Sigma_w\end{aligned}$$

In the correction step, the predicted estimate $\hat{\mathbf{x}}_k$ and the error covariance $\mathbf{P}_k$ are updated using MLE (maximum likelihood estimation) based on the current measurements containing the measurement noise. Combining the prior distribution $\mathbf{z}_k$ with $\mathbf{y}_k \sim \mathcal{N}(\mathbf{C}x_k, \Sigma_v)$, we apply Bayes' rule to have the posterior distribution of $\mathbf{x}_k$ proportional to the product of probability density functions of $\mathbf{z}_k$ and $\mathbf{y}_k$, i.e.

$$\begin{aligned}\mathbf{x}_k \propto \exp\Big(-\frac{1}{2}\Big[&(\mathbf{y}_k - \mathbf{C}x_k)^T\Sigma_v^{-1}(\mathbf{y}_k - \mathbf{C}x_k)\\ &+ (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T\mathbf{P}_{k|k-1}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})\Big]\Big)\end{aligned}$$

As a result, we have $\mathbf{x}_{k+1} \sim \mathcal{N}(\hat{\mathbf{x}}_{k+1}, \mathbf{P}_{k+1})$ with

$$\begin{aligned}\mathbf{P}_k &= \left(\mathbf{C}^T\Sigma_v^{-1}\mathbf{C} + \mathbf{P}_{k|k-1}^{-1}\right)^{-1}\\ \hat{\mathbf{x}}_k &= \mathbf{P}_k\left(\mathbf{C}^T\Sigma_v^{-1}\mathbf{y}_k + \mathbf{P}_{k|k-1}^{-1}\hat{\mathbf{x}}_{k|k-1}\right)\end{aligned}$$

The distributed Kalman filter consists of *local prediction* and *distributed correction*. In local prediction stage, each agent i has the predictions

$$\begin{aligned}\hat{\mathbf{x}}_{k|k-1}^{(i)} &= \mathbf{A}\hat{\mathbf{x}}_{k-1}^{(i)},\\ \mathbf{P}_{k|k-1}^{(i)} &= \mathbf{A}\mathbf{P}_{k-1}^{(i)}\mathbf{A}^T + \Sigma_w^{(i)},\end{aligned}$$

The difference between the centralized Kalman filter and the distributed Kalman filter is that the impacts of i's neighbors on i's predictions are considered in a distributed manner. Related works include [3], [5], [6], [26]. The Distributed Kalman filter (DKF) has the following prediction rules,

$$\mathbf{P}_{k|k-1}^{(i)} = \mathbf{A}\mathbf{P}_{k-1}^{(i)}\mathbf{A}^T + \Sigma_w^{(i)} \tag{3}$$

$$\hat{\mathbf{x}}_k^{(i)} = \mathbf{P}_k^{(i)}\left(\frac{1}{d_i}\sum_{j\in N^{(i)}}\mathbf{P}_{|}^{(j)\,-1}\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(j)} + \mathbf{C}^{(i)\,T}\Sigma_v^{(i)\,-1}\mathbf{y}_k^{(i),a}\right) \tag{4}$$

where $\hat{\mathbf{x}}_k^{(i)}$ is the state estimate at time k, $\mathbf{P}_{|}^{(i)}$ is *a priori* estimation error covariance of agent i, and the estimation error covariance matrix, $\mathbf{P}^{(i)}$, is chosen according to the following equation,

$$\mathbf{P}^{(i)} = \left(\frac{1}{d_i}\sum_{j\in N^{(i)}}(\mathbf{A}\mathbf{P}^{(j)}\mathbf{A}^T + \Sigma_w^{(j)})^{-1} + \mathbf{C}^{(i)\,T}\Sigma_v^{(i)\,-1}\mathbf{C}^{(i)}\right)^{-1} \tag{5}$$

where $N^{(i)} = \{i\} \cup \{j \in \mathcal{V} : (i, j) \in \mathcal{E}\}$ is the set of i's neighbors and $d_i = |N^{(i)}|$ is the degree of node i.

Usually $\Sigma_v^{(i)}$ and $\Sigma_w^{(i)}$ are used to denote the covariance matrices of the noise in Kalman filter. In this paper, we treat them as parameters for developing our algorithm in the noise-free setting (Kalman filter application in the noise-free setting is discussed in [27]). In practice, $\Sigma_v^{(i)}$ and $\Sigma_w^{(i)}$ can be chosen to be any positive definite matrices.

By assuming $(\mathbf{A}, \mathcal{C})$ is observable in our model, we obtain a steady-state DKF with error covariance matrix $\mathbf{P}^{(i)} = \lim_{k\to\infty}\mathbf{P}_k^{(i)}$. There are several ways to prove the convergence result of $\mathbf{P}_k^{(i)}$. In [16], Theorem III.1 proves convergence of the covariance matrices of the estimator by showing that when $\{\mathbf{P}_k^{(i)}\}_{k\geq 0}$ is increasing and is bounded above, then the limit exists. [28] proves the convergence using probability theory and [6] performs convergence analysis on a modified DKF which has one prediction/update step at each time point.

The following theorem states the main result of distributed estimation without attacks and noises:

Theorem 1 (Convergence of DKF, Theorem III.2 [16]): If the graph G is connected, $(\mathbf{A}, \mathbf{C})$ is observable, and $\Sigma_v^{(i)}$ is full rank for all $1 \leq i \leq n$, the estimation of equation (4) converges to the real states, i.e. $\lim_{k\to\infty}\|\hat{\mathbf{x}}_k^{(i)} - \mathbf{x}_k\| \to 0$ for all $1 \leq i \leq n$ and the convergence rate is linear.

The result described here is called "omniscience property" in [5], [13], which is proved under the same system setting as Theorem 1, but for different estimation algorithms. We remark that while the condition "$(\mathbf{A}, \mathbf{C})$ is observable" is slightly more restrictive than the condition "$(\mathbf{A}, \mathbf{C})$ is detectable" in [5]. A system is detectable if all the unobservable states are stable [29].

The Distributed Kalman filter (4) is obtained by the following optimization problem,

$$\begin{aligned}\hat{\mathbf{x}}_k^{(i)} = \arg\min_{\mathbf{x}}&(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{x})^T\Sigma_v^{(i)-1}(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{x})\\ &+ \frac{1}{d_i}\sum_{j\in N^{(i)}}(\mathbf{x} - \mathbf{A}\hat{\mathbf{x}}_{k-1}^{(j)})^T\mathbf{P}_{|}^{(j)\,-1}(\mathbf{x} - \mathbf{A}\hat{\mathbf{x}}_{k-1}^{(j)}).\end{aligned} \tag{6}$$

Motivated by this optimization-based estimator, the Resilient Distributed Kalman Filter (RDKF) proposed in [16]

$$\begin{aligned}\hat{\mathbf{x}}_k^{(i)} &= \arg\min_{\mathbf{x}} \lambda \left\| \Sigma_v^{(i)\,-\frac{1}{2}}(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{x}) \right\| \\ &+ \frac{1}{d_i}\sum_{j\in N^{(i)}} (\mathbf{x}-\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(j)})^T \mathbf{P}_{|}^{(j)\,-1}(\mathbf{x}-\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(j)}), \qquad (7)\end{aligned}$$

where $\mathbf{y}_k^{(i),a}$ is the corrupted measurement at time k of the i-th data, the matrix $\mathbf{P}_{|}^{(j)\,-1}$ is defined by (3), and $\lambda > 0$ is a parameter balancing the local information and the data collected from its neighbors. Instead of using the term $(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{x})^T \Sigma_v^{(i)-1}(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{x})$ in (6) directly, the first term in (7) considers the use of square root. The term $\left\| \Sigma_v^{(i)\,-\frac{1}{2}}(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{x}) \right\|$ with l_1 norm and the introduction of the penalty parameter, λ, make the algorithm more resilient to attacks. The second term considers the effects of its neighbors as it also presents in DKF (6), which makes RDKF a distributed scheme. The centralized estimator requires the simultaneous knowledge of parameters and measurements from all agents to carry out the estimation, while a distributed estimator only needs information from its neighbors. The optimization (7) is convex, which is broadly studied by many researchers [30]. Though the solution of (7) can be obtained by some general packages (e.g. "cvx" [31] in Matlab), each update $\hat{\mathbf{x}}_k^{(i)}$ takes several iterates to be solved. Moreover, because of the limitation of the machine and the solver, the accuracy only reaches up to 10^{-4} (see Figure 2). To make the estimator implement faster and have a better accuracy, we propose a new optimization-based estimator, Fast Resilient Distributed Estimator.

III. Fast Secured Distributed Estimation Methods

In this section, we introduce a novel estimator which is motivated by the Resilient Distributed State Estimator (RDKF) presented in (7).

In order to make an optimization-based estimator more robust to attacks, a commonly used strategy is to use optimization with l_1 norm on the terms affected by the attack [32]. As the Distributed Kalman filter (DKF) (6) contains the square term $(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{x})^T \Sigma_v^{(i)-1}(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{x})$, RDKF uses the term $\|\Sigma_v^{(i)\,-\frac{1}{2}}(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{x})\|$ to lower the impact of attacks, which makes the estimator more robust. However, RDKF does not have a good structure as the DKF does, though it provides the resilience against the attacks. To tackle this issue, we propose the Fast Resilient Distributed State Estimator (FRDSE) based on the following optimization:

$$\begin{aligned}\hat{\mathbf{x}}_k^{(i)} &= \arg\min_{\mathbf{x}} \lambda \frac{\left\| \Sigma_v^{(i)\,-\frac{1}{2}}(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{x}) \right\|^2}{2\left\| \Sigma_v^{(i)\,-\frac{1}{2}}(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(i)}) \right\|} \\ &+ \frac{1}{d_i}\sum_{j\in N^{(i)}} (\mathbf{x}-\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(j)})^T \mathbf{P}_{|}^{(j)\,-1}(\mathbf{x}-\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(j)}) \qquad (8)\end{aligned}$$

Unlike RDKF in (7), we consider the square term $\|\Sigma_v^{(i)\,-\frac{1}{2}}(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{x})\|^2$ divided by $\|\Sigma_v^{(i)\,-\frac{1}{2}}(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(i)})\|$ in (8). Notice that the term affected by the attack, $\|\Sigma_v^{(i)\,-\frac{1}{2}}(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{x})\|$, can be approximated by the denominator from (1). The fraction term in (8) has a similar effect as the term $\|\Sigma_v^{(i)\,-\frac{1}{2}}(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{x})\|$ in (7). Such a form of the first term in (8) dramatically reduces the influence of the attack like RDKF, which also provides the robustness to the estimator. Furthermore, the convex and quadratic structure of FRDSE in (8) gives computational advantages over RDKF.

The parameter λ gives a balance between the terms,

$$\|\Sigma_v^{(i)\,-\frac{1}{2}}(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{x}_k)\|^2 / \|\Sigma_v^{(i)\,-\frac{1}{2}}(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(i)})\|$$

and

$$\sum_{j\in N^{(i)}} (\mathbf{x}_k-\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(j)})^T \mathbf{P}_{|}^{(j)\,-1}(\mathbf{x}_k-\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(j)}).$$

A large λ implies more weight is placed on $\mathbf{y}_k^{(i),a}$, which has true and corrupted sensor measurements. Although choosing a large λ will make the error convergence faster, it makes the system more unstable in the presence of attack. On the contrary, when λ is small, it takes more steps for the estimation errors to converge to zero, but it makes the algorithm more resilient to attack. The parameter $\Sigma_v^{(i)}$ of equation (8) has an opposite impact on the system: when $\Sigma_v^{(i)}$ is large, it takes longer for the estimation errors to converge to zero, but it makes FRDSE more resilient to attacks. The parameter $\Sigma_w^{(i)}$ has an influence in $\mathbf{P}_{|}^{(i)}$ and it makes an impact similar to $\Sigma_v^{(i)}$ on the system. The numerical results in Section IV verify the impacts of the three parameters on the estimation errors.

To discuss convergence and resiliency of FRDSE, we consider two scenarios: 1) All agents/vehicles are benign and the system operates normally; 2) Some agents/vehicles are compromised in the distributed system. The next theorem (Proof of Theorem 2 is available in the Appendix) provides a theoretical guarantee for the convergence of FRDSE in the first scenario. The algorithm obeys the "omniscience property" and the estimation error converges to zero if the initial estimation errors $\mathbf{e}_0^{(i)}$ are not too large.

Theorem 2 (Convergence of FRDSE): Under the same assumptions of Theorem 1, if the initial estimation errors $\{\mathbf{e}_0^{(i)}\}_{1\leq i\leq n}$ satisfy the following condition: for any $\mathbf{x}$ that satisfies $\|\Sigma_v^{(i)-\frac{1}{2}}\mathbf{C}^{(i)}\mathbf{A}\mathbf{x}\| = \frac{\lambda}{2}$, it has the property that $\mathbf{x}^T\mathbf{P}^{(i)-1}\mathbf{x} \geq \mathbf{e}_0^{(i)T}\mathbf{P}^{(i)-1}\mathbf{e}_0^{(i)}$. When there is no attack, i.e., $\mathbf{a}_k^{(i)} = 0$ for all $k \geq 1$ and $1 \leq i \leq n$, the sequence $\{\hat{x}_k^{(i)}\}_{1\leq i\leq n}$ produced by equation (8) converges to the real state $\mathbf{x}_k$ i.e., the estimation errors $\|\mathbf{e}_k^{(i)}\| = \|\hat{\mathbf{x}}_k^{(i)} - \mathbf{x}_k\|$ converges to zero.

As for the second scenario, Theorem 3 (Proof is available in the Appendix) states that no matter how large the magnitude of the attacks, the deviation of the state estimate of FRDSE is upper bounded. This result suggests that the estimation errors are bounded during an attack, while the traditional DKF may have an unbounded estimation error caused by an unbounded attack. Furthermore, compared to RDKF, our method has huge computational advantages by its closed-form of the solution.

The following lemma shows that the optimization problem in equation (8) has a closed-form solution, which makes implementation time is much shorter than RDKF.

Lemma 1: Consider the optimization problem of equation (8). The solution $\hat{\mathbf{x}}_k^{(i)}$ is based on $\mathbf{y}_k^{(i),a}$ and $\{\hat{x}_{k-1}^{(j)}\}_{j\in N^{(i)}}$ and has a closed-form expression,

$$\hat{\mathbf{x}}_k^{(i)} = \left(\frac{\lambda \mathbf{C}^{(i)T}\Sigma_v^{(i)-1}\mathbf{C}^{(i)}}{\left\| \Sigma_v^{(i)\,-\frac{1}{2}}(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(i)}) \right\|} + \frac{2}{d_i}\sum_{j\in N^{(i)}} \mathbf{P}_{|}^{(j)-1} \right)^{-1} \cdot \left(\frac{\lambda \mathbf{C}^{(i)T}\Sigma_v^{(i)-1}\mathbf{y}_k^{(i),a}}{\left\| \Sigma_v^{(i)\,-\frac{1}{2}}(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(i)}) \right\|} + \frac{2}{d_i}\sum_{j\in N^{(i)}} \mathbf{P}_{|}^{(j)-1}\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(j)} \right) \tag{9}$$

Theorem 3 (Resiliency of FRDSE): The estimation $\hat{x}_k^{(i)}$ of equation (8) given by equation (9) is resilient to attacks on sensor measurements, $\mathbf{y}_k^{(i),a}$, in the sense that $\hat{\mathbf{x}}_k^{(i)}$ is bounded for each k and i.

Theorem 3 implies that the disturbance on the state estimate caused by an arbitrary attack on $\mathbf{y}_k^{(a)}$ is bounded. The bound on estimation error is independent of time. Moreover, the bound of $\hat{\mathbf{x}}_k^{(i)}$ partially explains the observation made in the beginning of Section III that large $\Sigma_v^{(i)}$ corresponds to more stable performance of the estimator during an attack, as from equation (9) we observe that large $\Sigma_v^{(i)}$ gives a smaller upper bound on the estimation error.

We remark that Theorem 3 only captures the impact of sporadic attack (an attack which does not occur continuously for a long duration of time) on the estimation of $\hat{\mathbf{x}}_k^{(i)}$. If the estimation error is small enough to satisfy the condition of Theorem 2 after an attack, then we can consider such an estimation error as the "initial estimation error" and use it to show that despite the attack, the estimation errors of FRDSE still converge to zero, provided we have attack-free measurements after the sporadic attack.

IV. Experimental Results

In this section, we compare the performances of our proposed algorithm against the distributed Kalman filter (DKF) and the resilient distributed Kalman filter (RDKF). Simulation results are presented to justify our algorithm's high efficiency and attack-resiliency.

A. Experiment Setup

We simulate the dynamics of a 5-vehicle platoon, communication graph, and the sensor attack on vehicles with CPU AMD 2600X and MATLAB R2018b for $t = 200$ simulation time units (X-axis of figures). Each vehicle is equipped with 4 sensors such as radars and LIDARs, and the state vector of each vehicle is $\mathbf{x}^{(i)} = [d^{(i)}, v^{(i)}, a^{(i)}, u^{(i)}] \in \mathbb{R}^4$, where $d^{(i)}, v^{(i)}, a^{(i)}$ are the distance, the velocity, and the acceleration of the i-th vehicle, and $u^{(i)}$ is the control input of the plant. The dimension of $\mathbf{x}$ is quadruple the number of vehicles, i.e. $\mathbf{x}_k \in \mathbb{R}^{20}$. The choice of matrices $\mathbf{A}, \{\mathbf{C}^{(i)}\}_{i=1}^n$ see [26]. We choose Predecessor-leader following topology (PLF) as our communication graph as shown in Fig. 1.

Remark: While implementing our proposed algorithm, equation (7) and equation (9) contain small denominators as the estimation of states $\hat{\mathbf{x}}_k^{(i)}$ are close to the true states. Dividing by a very small number will result in numerical instability. To address this issue, we instead use $\max(\|\Sigma_v^{(i)\,-\frac{1}{2}}(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(i)})\|, \epsilon)$, where $\epsilon = 0.001$.

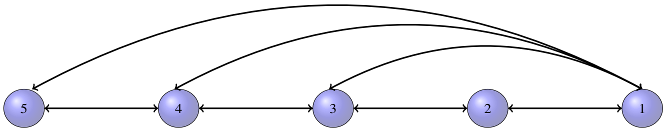


Fig. 1: Undirected graph of homogeneous predecessor-leader following topology of a 5-vehicle platoon. (1)-(5) represents numbering of vehicles (nodes) as (1) is the leading vehicle of the platoon. The edges represent sensor and V2V communications among vehicles.

B. Experiment Results and Discussions

We design the numerical simulations under both attack-free and false data injection attack scenarios. In the attack-free scenario, the convergence speed and the estimation error amplitude of the three algorithms, DKF, RDKF, and FRDSE (the proposed algorithm), are compared in Fig. 2. In the FDI scenario, three experiments are implemented in respect of resilience, efficiency and parameter tuning. The first experiment is to show the attack-resiliency of the proposed algorithm. We simulate the 5-vehicle platoon under the circumstances where different vehicles (single vehicle or multiple vehicles) are corrupted or the attacks occur at different time periods (single period or multiple periods). Though our proposed algorithm has similar resilience as RDKF, the implementation takes much less time than RDKF. We compare the execution time of DKF, RDKF, and FRDSE as the number of vehicles increases to illustrate computational advantages of our proposed algorithm. Ultimately, different parameter settings (λ, Σ_v, Σ_w) are examined on the 5-vehicle platoon to exhibit their impacts on the performance of the proposed algorithm. The details of the simulation results are presented in the following.

Case 1: Attack free scenario

Consider an attack-free scenario. The performances of our proposed algorithm (FRDSE), Resilient Distributed Kalman Filter (RDKF), and the Distributed Kalman Filter (DKF) are shown in Fig. 2 (a)-(c), respectively. The parameters are set as $(\lambda, \Sigma_v, \Sigma_w) = (100, 10\mathbf{I}, \mathbf{I})$. Instead of choosing the estimation error of $\mathbf{x}_k^{(i)}$, we focus on the error of estimated distance $\hat{d}_k^{(i)}$ as in practice we need to control the distances between vehicles to prevent a collision. We observe that the estimation error of distance of each vehicle produced by our proposed algorithm is up to 10^{-8}, which is much smaller than RDKF as expected. Moreover, the error of our proposed algorithm converges to a small value (less than 1 m) faster than DKF.

Case 2: False Data Injection (FDI) attack

We consider the case where sensor measurements of some vehicles are compromised in certain times. In FDI attack, we add random data to some sensor outputs during certain periods of time.

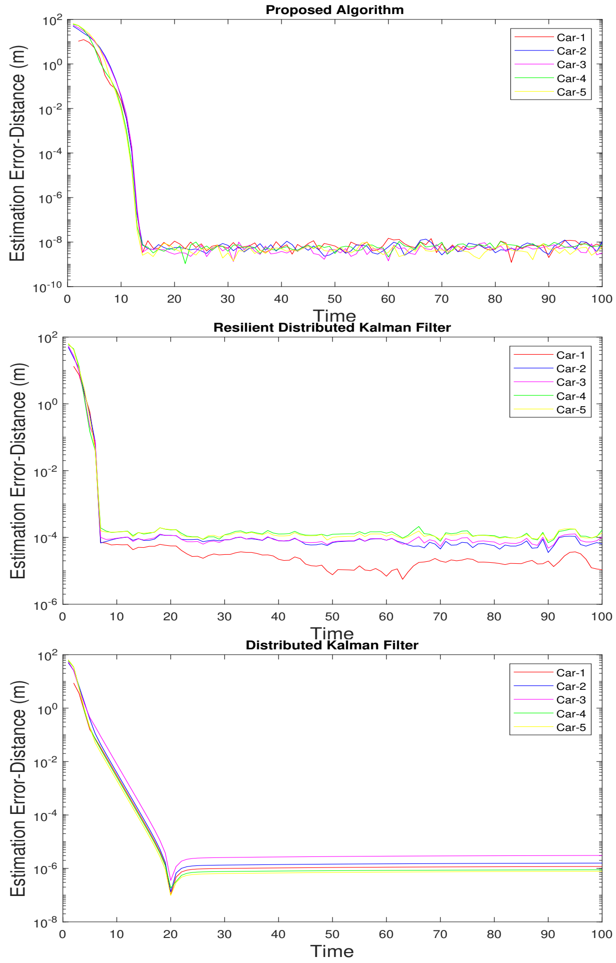
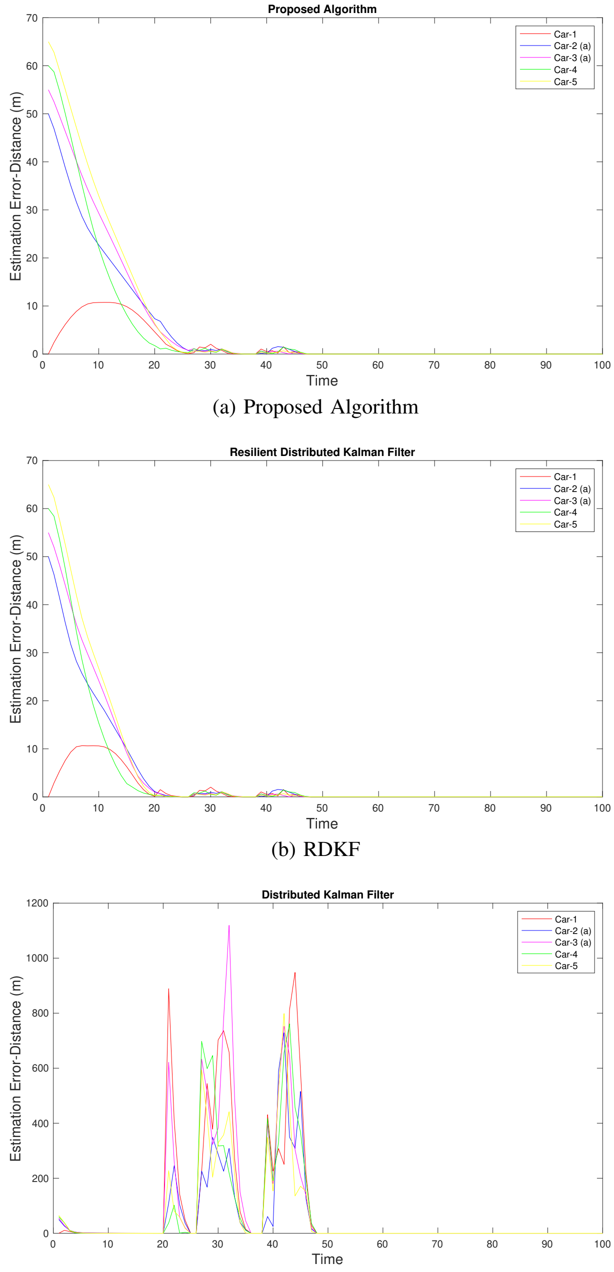


Fig. 2: Distance estimation errors of all vehicles produced by our proposed algorithm (FRDSE), Resilient Distributed Kalman Filter (RDKF), and Distributed Kalman Filter (DKF) in attack free scenario. The parameters are set as $(\lambda, \Sigma_v, \Sigma_w) = (100, 10\mathbf{I}, \mathbf{I})$. Time in X-axis is simulation time.

Case 2-a: Resiliency of 5-vehicle platoon

In the first simulation, we set attacks to occur with probability $p_a = 0.99$ (i.e. the probability of attack being successful is high during the attack duration) on $t = (21, 50)$ and we consider sensors of Car-2 and Car-3 being compromised. The attack vector is chosen as Gaussian vector with mean 10,000. We set the parameters $(\lambda, \Sigma_v, \Sigma_w) = (10, 10\mathbf{I}, \mathbf{I})$, where $\mathbf{I}$ is the identity matrix. Fig. 3 compares the performances of our proposed algorithm with RDKF and DKF. We observe that the estimated distance errors of our proposed algorithm and RDKF caused by the attack on $t = (21, 50)$ are small and the system is stable, while the error of DKF goes up to 1800 m. As perturbation in estimation error caused by the malware is small in our algorithm and RDKF, the likelihood of preventing a collision is high as they keep the distance error below the minimum stopping distance, $\tau = 10$m.

One additional simulation with different attacks is completed in Fig. 4. The large random malicious data with high probability $p_a = 0.99$ are injected in the two periods of time with different attack strategy. On $t = (31, 50)$ Car-2 and Car-3 are compromised and Car-4 and Car-5 are attacked on $t = (71, 90)$. We set parameters as $(\lambda, \Sigma_v, \Sigma_w) = (10, 10\mathbf{I}, \mathbf{I})$. Fig. 4-(a) shows the evolution of the distance estimation errors of our proposed algorithm. During the attack, the estimation errors are small enough to prevent vehicle collision for both FRDKF and RDKF, while the DKF (Fig. 4-(b)) has large estimation errors that can lead to collision.

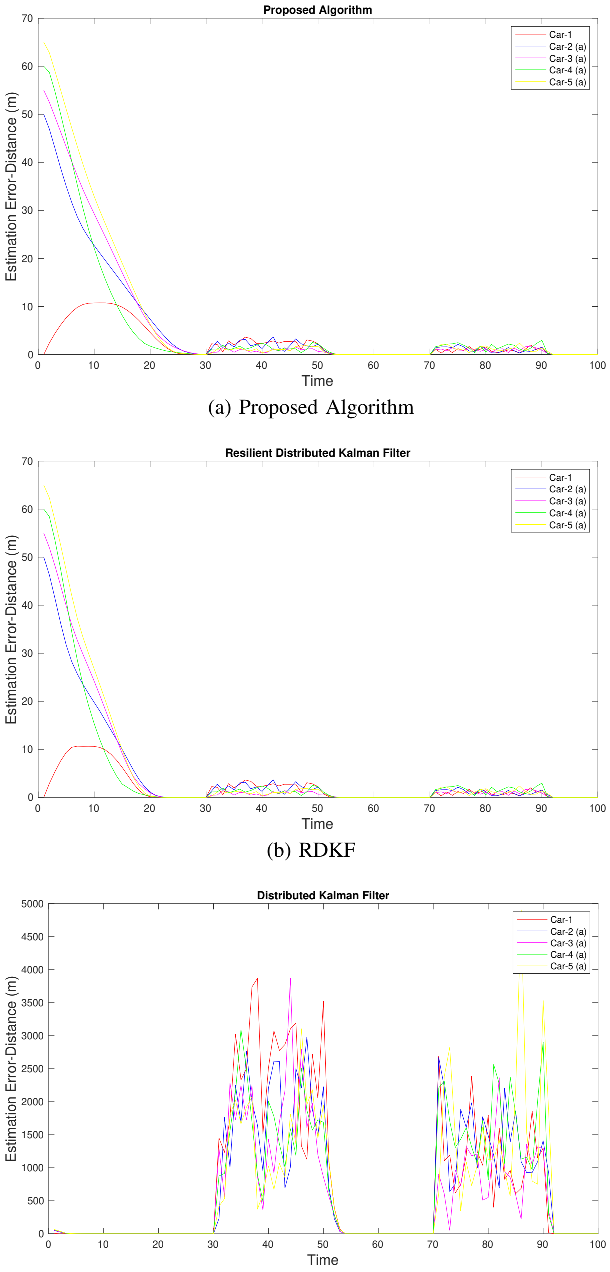


(a) Proposed Algorithm

(b) RDKF

(c) DKF

Fig. 3: Performances of our proposed algorithm, RDKF and DKF against FDI attack. The attack occurs with probability $p_a = 0.99$ from $t = (21, 50)$ and Car-2 and Car-3 are compromised. We consider $(\lambda, \Sigma_v, \Sigma_w) = (10, 10\mathbf{I}, \mathbf{I})$ and we mark corrupted vehicles as (a). Time in X-axis is simulation time.



(a) Proposed Algorithm

(b) RDKF

(c) DKF

Fig. 4: Performances of the proposed algorithm (FRDSE), RDKF and DKF against FDI attack. The attacks are launched with probability $p_a = 0.99$. Car-2 and Car-3 are compromised on $t = (31, 50)$ and Car-4 and Car-5 are attacked on $t = (71, 90)$.
We consider $(\lambda, \Sigma_v, \Sigma_w) = (10, 10\mathbf{I}, \mathbf{I})$. The corrupted vehicles are marked with (a). Time in X-axis is simulation time.

TABLE I: Execution time (Seconds) of DKF, RDKF, and FRDSE with different number of vehicles. Simulation time is set as $t = 50$.

No. Vehicles	4	5	6	7	8
FRDSE	0.0221	0.0489	0.0771	0.1098	0.1019
DKF	0.009	0.0142	0.0148	0.0272	0.0277
RDKF	34.4239	37.4263	45.1442	52.4194	60.8529

Case 2-b: Execution time with more vehicles

The execution time of our algorithm is much less than RDKF by the convex and quadratic structure of FRDSE in equation (8). To illustrate FRDSE's computational advantages over RDKF, we implement three algorithms on the platoon with more vehicles. The experiments are implemented based on the CPU AMD 2600X and MATLAB R2018b. We compare the execution time of DKF, RDKF, and FRDSE (our algorithm) with different lengths of vehicle platoons (4, 5, 6, 7, 8 vehicles) for simulation time, $t = (1, 50)$. The results are shown in Table I.

For a platoon of eight vehicles, the execution time of RDKF is 60.8529 seconds, while the execution time of our proposed algorithm is only 0.1019 seconds. Overall, our proposed algorithm is around 500 times faster than RDKF. Though DKF takes less time to our proposed algorithm, its vulnerability against attack makes DKF less practicable than FRDSE.

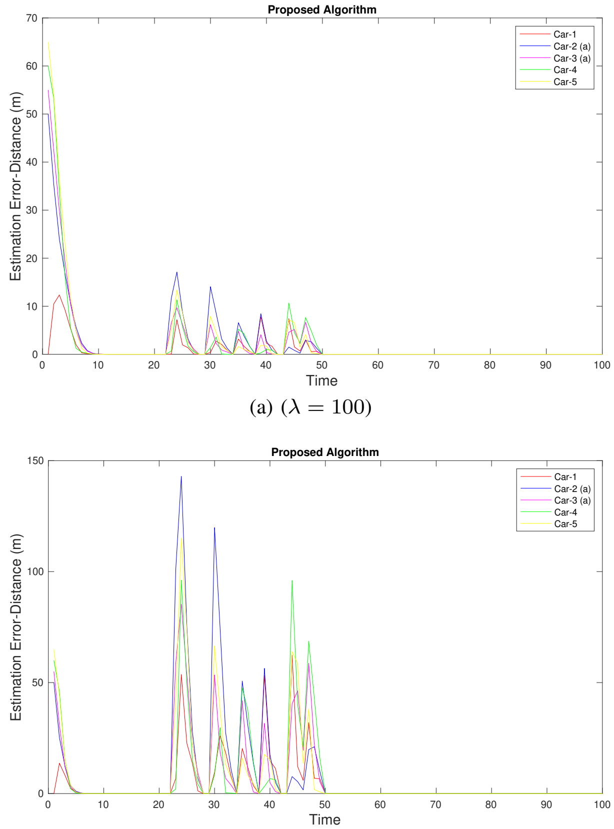
Case 2-c: Parameter tuning

We also verify the discussions about the influence of the parameters in Section III. Fig. 5 shows experiments of our proposed algorithm FRDSE with two sets of parameters, $(\Sigma_v, \Sigma_w, \lambda) = (10\mathbf{I}, \mathbf{I}, 100)$ and $(\Sigma_v, \Sigma_w, \lambda) = (10\mathbf{I}, \mathbf{I}, 1000)$. The attacks are of FDI type with high probability on $t = (21, 50)$ and the sensors of Car-2, Car-3 are corrupted. The results confirm that smaller λ (Fig. 5-(a)) gives slower convergence at the beginning, but makes the algorithm more resilient to attacks. On the contrary, larger λ (Fig. 5-(b)) gives faster convergence at the beginning, but makes it less resilient to attacks.

Under the same settings of the attack, Fig. 6 shows the results of two experiments with $(\lambda, \Sigma_v, \Sigma_w) = (100, \mathbf{I}, \mathbf{I})$ and $(\lambda, \Sigma_v, \Sigma_w) = (1000, \mathbf{I}, 10\mathbf{I})$. We can see that larger Σ_w (Fig. 6-(b)) gives faster convergence at the beginning, but make our algorithm less resilient to attacks. However, smaller Σ_w (Fig. 6-a(a)) makes the algorithm more resilient at the expense of longer convergence time at the beginning.

V. Conclusion

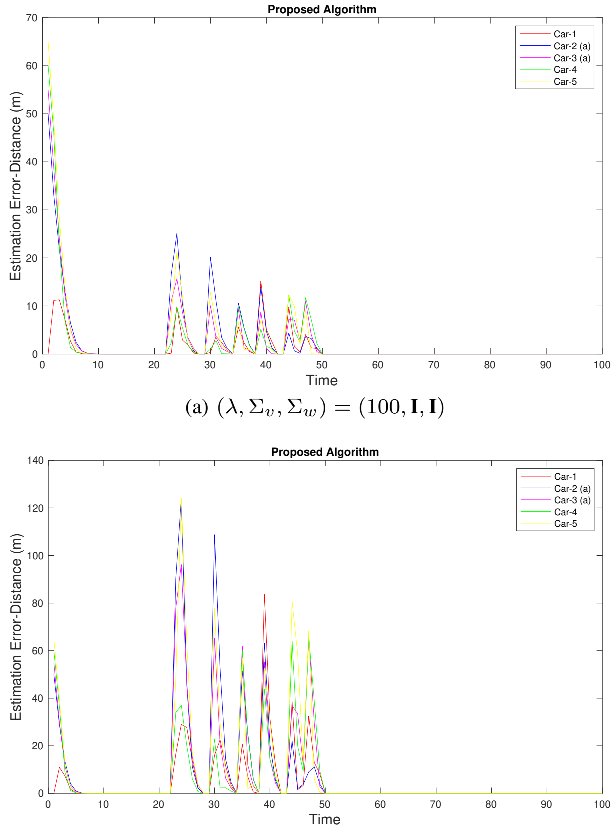
In this paper, we have proposed Fast Resilient Distributed State Estimator (FRDSE), a novel computationally efficient attack resilient distributed state estimation scheme that can recursively estimates states and bounds the disturbance on the state estimate caused by an attack. We prove that the estimation error of our method asymptotically converges to zero when there is no attack and noise and has an upper bound during an attack. By tuning the parameters of our estimator, we obtain smaller disturbance in estimation errors even in the presence of an attack and it is more resilient than the DKF. Our proposed algorithm has more computationally advantages over the Resilient Distributed Kalman filter (RDKF) and it is more practicable in real-time computing. In the future, we plan to extend our algorithm for a time-varying graph.



(a) ($\lambda = 100$)

(b) ($\lambda = 1000$)

Fig. 5: Performance of our proposed algorithm against FDI attack. The attack occurs with probability $p_a = 0.99$ from $t = (21, 50)$ and Car-2 and Car-3 are compromised. We consider $(\lambda, \Sigma_v, \Sigma_w) = (100, 10\mathbf{I}, \mathbf{I})$ and $(\lambda, \Sigma_v, \Sigma_w) = (1000, 10\mathbf{I}, \mathbf{I})$ respectively. The corrupted vehicles are marked with (a). Time in X-axis is simulation time.



(a) $(\lambda, \Sigma_v, \Sigma_w) = (100, \mathbf{I}, \mathbf{I})$

(b) $(\lambda, \Sigma_v, \Sigma_w) = (100, \mathbf{I}, 10\mathbf{I})$

Fig. 6: Performance of our proposed algorithm against FDI attack. The attack occurs with probability $p_a = 0.99$ from $t = (21, 50)$ and Car-2 and Car-3 are compromised. We consider $(\lambda, \Sigma_v, \Sigma_w) = (100, \mathbf{I}, \mathbf{I})$ and $(\lambda, \Sigma_v, \Sigma_w) = (1000, \mathbf{I}, 10\mathbf{I})$ respectively. The corrupted vehicles are marked with (a). Time in X-axis is simulation time.

Acknowledgements

This work is partially supported by National Science Foundation (CNS-1818500) and CyberFlorida Collaborative Seed Award Program.

References

[1] J. Speyer, "Computation and transmission requirements for a decentralized linear-quadratic-gaussian control problem," *IEEE Transactions on Automatic Control*, vol. 24, no. 2, pp. 266–269, 1979.

[2] B. Rao, H. F. Durrant-Whyte, and J. Sheen, "A fully decentralized multi-sensor system for tracking and surveillance," *The International Journal of Robotics Research*, vol. 12, no. 1, pp. 20–44, 1993.

[3] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[4] R. Olfati-Saber, "Distributed kalman filtering for sensor networks," in *2007 46th IEEE Conference on Decision and Control*, Dec 2007, pp. 5492–5498.

[5] S. Park and N. C. Martins, "Design of distributed lti observers for state omniscience," *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 561–576, Feb 2017.

[6] D. Marelli, M. Zamani, and M. Fu, "Distributed kalman filter in a network of linear dynamical systems," *arXiv preprint arXiv:1711.07625*, 2017.

[7] A. Abdelgawad, "Distributed kalman filter with fast consensus for wireless sensor networks," *International Journal of Wireless Information Networks*, vol. 23, no. 1, pp. 82–88, Mar 2016.

[8] R. Carli, A. Chiuso, L. Schenato, and S. Zampieri, "Distributed kalman filtering based on consensus strategies," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 4, pp. 622–633, May 2008.

[9] S. Dadras, S. Dadras, and C. Winstead, "Identification of the attacker in cyber-physical systems with an application to vehicular platooning in adversarial environment," in *2018 Annual American Control Conference (ACC)*. IEEE, 2018, pp. 5560–5567.

[10] I. Sajjad, D. D. Dunn, R. Sharma, and R. Gerdes, "Attack mitigation in adversarial platooning using detection-based sliding mode control," in *Proceedings of the First ACM Workshop on Cyber-Physical Systems-Security and/or PrivaCy*. ACM, 2015, pp. 43–53.

[11] U. A. Khan and A. M. Stanković, "Secure distributed estimation in cyber-physical systems," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 5209–5213.

[12] I. Matei, J. S. Baras, and V. Srinivasan, "Trust-based multi-agent filtering for increased smart grid security," in *Control & Automation (MED), 2012 20th Mediterranean Conference on*. IEEE, 2012, pp. 716–721.

[13] A. Mitra and S. Sundaram, "Secure distributed observers for a class of linear time invariant systems in the presence of byzantine adversaries," in *2016 IEEE 55th Conference on Decision and Control (CDC)*, Dec 2016, pp. 2709–2714.

[14] A. Pelc and D. Peleg, "Broadcasting with locally bounded byzantine faults," *Information Processing Letters*, vol. 93, no. 3, pp. 109–115, 2005.

[15] G. Ellis, *Observers in control systems: a practical guide*. Elsevier, 2002.

[16] R. G. Dutta, T. Zhang, and Y. Jin, "Resilient distributed filter for state estimation of cyber-physical systems under attack," in *2019 American Control Conference (ACC)*. IEEE, 2019, pp. 5141–5147.

[17] Y. Cao, W. Yu, W. Ren, and G. Chen, "An overview of recent progress in the study of distributed multi-agent coordination," *IEEE Transactions on Industrial informatics*, vol. 9, no. 1, pp. 427–438, 2013.

[18] M. Rabbat and R. Nowak, "Distributed optimization in sensor networks," in *Proceedings of the 3rd international symposium on Information processing in sensor networks*. ACM, 2004, pp. 20–27.

[19] A. Nedić and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 48–61, 1 2009.

[20] I. Matei and J. S. Baras, "Performance evaluation of the consensus-based distributed subgradient method under random communication topologies," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 4, pp. 754–771, 2011.

[21] D. Jakovetić, J. Xavier, and J. M. Moura, "Fast distributed gradient methods," *IEEE Transactions on Automatic Control*, vol. 59, no. 5, pp. 1131–1146, 2014.

[22] J. C. Duchi, A. Agarwal, and M. J. Wainwright, "Dual averaging for distributed optimization: Convergence analysis and network scaling," *IEEE Transactions on Automatic control*, vol. 57, no. 3, pp. 592–606, 2012.

[23] K. Yuan, Q. Ling, and W. Yin, "On the convergence of decentralized gradient descent," *SIAM Journal on Optimization*, vol. 26, no. 3, pp. 1835–1854, 2016.

[24] W. Shi, Q. Ling, G. Wu, and W. Yin, "Extra: An exact first-order algorithm for decentralized consensus optimization," *SIAM Journal on Optimization*, vol. 25, no. 2, pp. 944–966, 2015.

[25] ——, "A proximal gradient algorithm for decentralized composite optimization," *IEEE Transactions on Signal Processing*, vol. 63, no. 22, pp. 6013–6023, 2015.

[26] R. G. Dutta, F. Yu, T. Zhang, Y. Hu, and Y. Jin, "Security for safety: a path toward building trusted autonomous vehicles," in *Proceedings of the International Conference on Computer-Aided Design*. ACM, 2018, p. 92.

[27] K. Rapp and P.-O. Nyman, "Stability properties of the discrete-time extended kalman filter," *IFAC Proceedings Volumes*, vol. 37, no. 13, pp. 1377–1382, 2004.

[28] S. Kar, S. Cui, H. V. Poor, and J. M. F. Moura, "Convergence results in distributed kalman filtering," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2011, pp. 2500–2503.

[29] J. Hespanha, *Linear Systems Theory: Second Edition*. Princeton University Press, 2018.

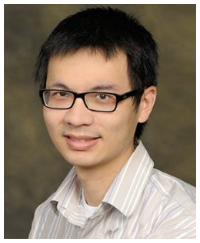
[30] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[31] M. Grant and S. Boyd, "Cvx: Matlab software for disciplined convex programming."

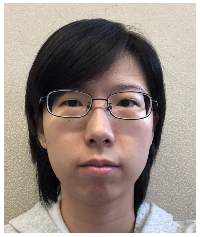
[32] Y. Mo and E. Garone, "Secure dynamic state estimation via local estimators," in *2016 IEEE 55th Conference on Decision and Control (CDC)*, Dec 2016, pp. 5073–5078.

Feng Yu is a PhD student in Department of Mathematics at the University of Central Florida (UCF). He received his M.S. degree in mathematics in 2018 from the University of Central Florida after he got B.S. in mathematics and B.E. in economics from Southwestern University of Finance and Economics, China, in 2016. His research focuses on PDE, optimization. He is currently on the areas of Cyber-Physical System and optimization. He received the ORC fellowship from the University of Central Florida in 2016.

Raj Gautam Dutta is postdoctoral associate at the Department of Electrical and Computer Engineering (ECE) at the University of Florida. He received his Ph.D degree in Computer Engineering in 2018 from University of Central Florida. He obtained his B.Tech in Electronics and Communication from Visvesvaraya Technological University, India in 2007 and M.S degree in Electrical Engineering with an emphasis on control systems from the University of Central Florida, USA in 2011. His current research interest includes design of attack detection and resilient algorithms for autonomous CPS. In the past, he developed security solutions for semiconductor soft IP cores by using formal verification



Teng Zhang received the B.S. degree in Mathematics from Fudan University, China, in 2006, and the Ph.D. degree in Mathematics from University of Minnesota, Minneapolis, MN, USA, in 2011. From September 2011 to August 2015, he was a Postdoctoral Fellow at Institute for Mathematics and its Applications (IMA), Minneapolis, MN, and the Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ, USA. He is currently a faculty member at the department of Mathematics at the University of Central Florida, Orlando, FL. His research interests include robust statistics, high-dimensional data analysis, matrix analysis/random matrix analysis, convex programming, and their applications to computer vision and data mining.



Yaodan Hu is a PhD student in Department of Electrical and Computer Engineering at the University of Florida (UF). She received her Bachelor of Science degree in Applied Physics from the University of Science and Technology of China (USTC), China in 2015. Her research interests include attack and defense mechanism design in Cyber-Physical Systems (CPS).

Yier Jin (M'12-SM'19) is an Associate Professor and IoT Term Professor in the Department of Electrical and Computer Engineering (ECE) in the University of Florida (UF). He received his PhD degree in Electrical Engineering in 2012 from Yale University after he got the B.S. and M.S. degrees in Electrical Engineering from Zhejiang University, China, in 2005 and 2007, respectively. His research focuses on the areas of hardware security, embedded systems design and security, trusted hardware intellectual property (IP) cores and hardware-software co-design for modern computing systems. He is also interested in the security analysis on Internet of Things (IoT) and wearable devices with particular emphasis on information integrity and privacy protection in the IoT era. Dr. Jin is a recipient of the DoE Early CAREER Award in 2016 and ONR Young Investigator Award in 2019. He received Best Paper Award at DAC'15, ASP-DAC'16, HOST'17, ACM TODAES'18, GLSVLSI'18, and DATE'19. He is also the IEEE Council on Electronic Design Automation (CEDA) Distinguished Lecturer.

APPENDIX A
TECHNICAL PROOFS

Proof of Theorem 2: We denote the estimation error as $\mathbf{e}_k^{(i)} = \hat{\mathbf{x}}_k^{(i)} - \mathbf{x}_k$. From equation (8), one has

$$\mathbf{e}_k^{(i)} = \arg\min_{\mathbf{e}} \lambda \frac{\left\|\Sigma_v^{(i)\,-\frac{1}{2}}\left(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}(\mathbf{e}+\mathbf{x}_k)\right)\right\|^2}{2\left\|\Sigma_v^{(i)\,-\frac{1}{2}}\left(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{A}(\mathbf{e}_{k-1}^{(i)}+\mathbf{x}_{k-1})\right)\right\|} + \frac{1}{d_i}\sum_{j\in N^{(i)}}(\mathbf{e}+\mathbf{x}_k-\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(j)})^T\mathbf{P}_{|}^{(j)\,-1}(\mathbf{e}+\mathbf{x}_k-\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(j)}) \tag{10}$$

For conciseness, we denote the objective function in equation (10) as $\mathbf{e}_k^{(i)} := \arg\min_{\mathbf{e}} g(\mathbf{e})$. When no attacks present, it follows that $\mathbf{y}_k^{(i),a} = \mathbf{C}^{(i)}\mathbf{x}_k = \mathbf{C}^{(i)}\mathbf{A}\mathbf{x}_{k-1}$ and $\mathbf{A}\mathbf{e}_{k-1}^{(i)} = \mathbf{A}\mathbf{x}_{k-1}^{(i)} - \mathbf{A}\mathbf{x}_{k-1} = \mathbf{A}\mathbf{x}_{k-1}^{(i)} - \mathbf{x}_k$. The objective function $g(\mathbf{e})$ can be simplified as

$$g(\mathbf{e}) = \frac{\lambda}{2m_k}\left(\mathbf{C}^{(i)}\mathbf{e}\right)^T\Sigma_v^{-1}\left(\mathbf{C}^{(i)}\mathbf{e}\right) + \frac{1}{d_i}\sum_{j\in N^{(i)}}(\mathbf{e}-\mathbf{A}\mathbf{e}_{k-1}^{(j)})^T\mathbf{P}_{|}^{(j)\,-1}(\mathbf{e}-\mathbf{A}\mathbf{e}_{k-1}^{(j)}),$$

where $m_k := \|\Sigma_v^{-\frac{1}{2}}\mathbf{C}^{(i)}\mathbf{A}\mathbf{e}_{k-1}^{(i)}\|$. Differentiating the objective function $g(\mathbf{e})$, we get

$$\nabla g(\mathbf{e}) = \frac{\lambda}{m_k}\mathbf{C}^{(i)T}\Sigma_v^{-1}\mathbf{C}^{(i)}\mathbf{e} + \frac{2}{d_i}\sum_{j\in N^{(i)}}\mathbf{P}_{|}^{(j)-1}(\mathbf{e}-\mathbf{A}\mathbf{e}_{k-1}^{(j)}) \tag{11}$$

Using the fact that $\nabla g(\mathbf{e}_k^{(i)}) = \mathbf{0}$, we have that

$$\left(\nabla g(\mathbf{e}_k^{(i)})\right)^T\mathbf{e}_k^{(i)} = \frac{\lambda}{m_k}\mathbf{e}_k^{(i)T}\mathbf{C}^{(i)T}\Sigma_v^{-1}\mathbf{C}^{(i)}\mathbf{e}_k^{(i)} + \frac{2}{d_i}\sum_{j\in N^{(i)}}(\mathbf{e}_k^{(i)}-\mathbf{A}\mathbf{e}_{k-1}^{(j)})^T\mathbf{P}_{|}^{(j)\,-1}\mathbf{e}_k^{(i)} = 0.$$

Note that the covariance matrix $\mathbf{P}_{|}^{(i)-1} \succeq 0$, then

$$\frac{1}{d_i}\sum_{j\in N^{(i)}}(\mathbf{e}_k^{(i)}-\mathbf{A}\mathbf{e}_{k-1}^{(j)})^T\mathbf{P}_{|}^{(j)\,-1}(\mathbf{e}_k^{(i)}-\mathbf{A}\mathbf{e}_{k-1}^{(j)}) \geq 0 \tag{12}$$

Since $g(\mathbf{e}_k^{(i)}) \geq 0$ and using equation (12), it follows that

$$\begin{aligned}\frac{1}{d_i}\sum_{j\in N^{(i)}}&(\mathbf{A}\mathbf{e}_{k-1}^{(j)})^T\mathbf{P}_{|}^{(j)\,-1}\mathbf{A}\mathbf{e}_{k-1}^{(j)}\\ \geq& -\frac{\lambda}{2m_k}\left(\mathbf{C}^{(i)}\mathbf{e}_k^{(i)}\right)^T\Sigma_v^{-1}\left(\mathbf{C}^{(i)}\mathbf{e}_k^{(i)}\right)\\ &+\frac{2}{d_i}\sum_{j\in N^{(i)}}(\mathbf{A}\mathbf{e}_{k-1}^{(j)})^T\mathbf{P}_{|}^{(j)\,-1}\mathbf{e}_k^{(i)}\\ &-\frac{1}{d_i}\sum_{j\in N^{(i)}}\mathbf{e}_k^{(i)T}\mathbf{P}_{|}^{(j)\,-1}\mathbf{e}_k^{(i)}\\ =&\frac{\lambda}{2m_k}\left(\mathbf{C}^{(i)}\mathbf{e}_k^{(i)}\right)^T\Sigma_v^{-1}\left(\mathbf{C}^{(i)}\mathbf{e}_k^{(i)}\right)\\ &+\frac{1}{d_i}\sum_{j\in N^{(i)}}\mathbf{e}_k^{(i)T}\mathbf{P}_{|}^{(j)\,-1}\mathbf{e}_k^{(i)}\end{aligned} \tag{13}$$

If $m_k = \|\Sigma_v^{(i)\,-\frac{1}{2}}\mathbf{C}^{(i)}\mathbf{A}\mathbf{e}_{k-1}^{(i)}\| \leq \frac{\lambda}{2}$, from equation (13) one has

$$\begin{aligned}\frac{1}{d_i}\sum_{j\in N^{(i)}}&(\mathbf{A}\mathbf{e}_{k-1}^{(j)})^T\mathbf{P}_{|}^{(j)\,-1}\mathbf{A}\mathbf{e}_{k-1}^{(j)}\\ \geq& \left(\mathbf{C}^{(i)}\mathbf{e}_k^{(i)}\right)^T\Sigma_v^{-1}\left(\mathbf{C}^{(i)}\mathbf{e}_k^{(i)}\right) + \frac{1}{d_i}\sum_{j\in N^{(i)}}\mathbf{e}_k^{(i)T}\mathbf{P}_{|}^{(j)\,-1}\mathbf{e}_k^{(i)}\\ =& \mathbf{e}_k^{(i)T}\mathbf{P}^{(i)-1}\mathbf{e}_k^{(i)},\end{aligned} \tag{14}$$

which follows from equation (3) and equation (5). Since $\mathbf{P}^{(i)-1}-\mathbf{A}^T\mathbf{P}_{|}^{(i)-1}\mathbf{A}=\mathbf{P}^{(i)-1}-\mathbf{A}^T\left(\mathbf{A}\mathbf{P}^{(i)}\mathbf{A}^T+\Sigma_w^{(i)}\right)^{-1}\mathbf{A}=\mathbf{P}^{(i)-1}-\left(\mathbf{P}^{(i)}+\mathbf{A}^{-1}\Sigma_w^{(i)}(\mathbf{A}^{-1})^T\right)^{-1}$ is positive semi-definite for any i, we have

$$\begin{aligned}\mathbf{e}_k^{(i)T}\mathbf{P}^{(i)-1}\mathbf{e}_k^{(i)} &\leq \frac{1}{d_i}\sum_{j\in N^{(i)}}(\mathbf{A}\mathbf{e}_{k-1}^{(j)})^T\mathbf{P}_{|}^{(j)\,-1}\mathbf{A}\mathbf{e}_{k-1}^{(j)}\\ &\leq \frac{1}{d_i}\sum_{j\in N^{(i)}}\mathbf{e}_{k-1}^{(j)T}\mathbf{P}^{(j)\,-1}\mathbf{e}_{k-1}^{(j)} \leq \max_{1\leq i\leq n}\mathbf{e}_{k-1}^{(i)T}\mathbf{P}^{(i)-1}\mathbf{e}_{k-1}^{(i)},\end{aligned}$$

which implies that $\max_{1\leq i\leq n}\mathbf{e}_k^{(i)T}\mathbf{P}^{(i)-1}\mathbf{e}_k^{(i)}$ does not increase as k increase and thus, it converges. Next, we show that $\max_{1\leq i\leq n}\mathbf{e}_k^{(i)T}\mathbf{P}^{(i)-1}\mathbf{e}_k^{(i)}$ converges to zero. If this is not the case, we assume $\eta \geq 0$ is the limit of $\max_{1\leq i\leq n}\mathbf{e}_k^{(i)T}\mathbf{P}^{(i)-1}\mathbf{e}_k^{(i)}$. For any $\epsilon > 0$, there exists a large $K(\epsilon)$ such that $\eta < \mathbf{e}_k^{(l)T}\mathbf{P}^{(l)-1}\mathbf{e}_k^{(l)} < \eta+\epsilon, \forall k > K(\epsilon)$ where l satisfies $\mathbf{e}_k^{(l)T}\mathbf{P}^{(l)-1}\mathbf{e}_k^{(l)} = \max_{1\leq i\leq n}\mathbf{e}_k^{(i)T}\mathbf{P}^{(i)-1}\mathbf{e}_k^{(i)}$. Then it implies that

$$\left|\frac{1}{d_l}\sum_{j\in N^{(l)}}\mathbf{e}_{k-1}^{(j)T}(\mathbf{A}^T\mathbf{P}_{|}^{(j)-1}\mathbf{A}-\mathbf{P}^{(j)\,-1})\mathbf{e}_{k-1}^{(j)}\right| < \epsilon.$$

Therefore,

$$\|\mathbf{e}_{k-1}^{(j)}\|^2 < \left\|\left(\mathbf{A}^T\mathbf{P}_{|}^{(j)-1}\mathbf{A}-\mathbf{P}^{(j)\,-1}\right)^{-1}\right\|d_l\epsilon, \quad \forall j\in N^{(l)}$$

It reaches that

$$\begin{aligned}\eta < \mathbf{e}_k^{(l)T}\mathbf{P}^{(l)-1}\mathbf{e}_k^{(l)} &\leq \|\mathbf{e}_{k-1}^{(l)}\|^2\|\mathbf{P}^{(l)-1}\|\\ &< \left\|\left(\mathbf{A}^T\mathbf{P}_{|}^{(j)-1}\mathbf{A}-\mathbf{P}^{(j)\,-1}\right)^{-1}\right\|\cdot\|\mathbf{P}^{(l)-1}\|d_l\epsilon,\end{aligned}$$

which implies that $\eta = 0$ as $\epsilon > 0$ is arbitrary small. Hence, the estimation error $\|e_k^{(i)}\|$ converges to zero as $\max_{1\leq i\leq n}\mathbf{e}_k^{(i)T}\mathbf{P}^{(i)-1}\mathbf{e}_k^{(i)}$ converges to zero. ■

Proof of Lemma 1: Note that the objective function of equation (8) is quadratic and the minimizer can be found by taking derivative of equation (8). It is followed that

$$0 = \frac{2}{d_i}\sum_{j\in N^{(i)}}\left(\mathbf{P}_{|}^{(j)-1}\hat{\mathbf{x}}_k - \mathbf{P}_{|}^{(j)\,-1}\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(j)}\right) + \frac{\lambda(\mathbf{C}^{(i)T}\Sigma_v^{(i)-1}\mathbf{C}^{(i)}\hat{\mathbf{x}}_k - \mathbf{C}^{(i)T}\Sigma_v^{(i)-1}\mathbf{y}_k^{(i),a})}{\left\|\Sigma_v^{(i)\,-\frac{1}{2}}(\mathbf{y}_k^{(i),a}-\mathbf{C}^{(i)}\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(i)})\right\|}$$

Organize the equation above and we obtain that

$$\hat{\mathbf{x}}_k^{(i)} = \left(\frac{\lambda \mathbf{C}^{(i)T}\Sigma_v^{(i)-1}\mathbf{C}^{(i)}}{\left\|\Sigma_v^{(i)\,-\frac{1}{2}}(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(i)})\right\|} + \frac{2}{d_i}\sum_{j\in N^{(i)}} \mathbf{P}_{|}^{(j)-1}\right)^{-1}$$
$$\cdot \left(\frac{\lambda \mathbf{C}^{(i)T}\Sigma_v^{(i)-1}\mathbf{y}_k^{(i),a}}{\left\|\Sigma_v^{(i)\,-\frac{1}{2}}(\mathbf{y}_k^{(i),a} - \mathbf{C}^{(i)}\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(i)})\right\|} + \frac{2}{d_i}\sum_{j\in N^{(i)}} \mathbf{P}_{|}^{(j)-1}\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(j)}\right)$$

■

Proof of Theorem 3: Let $\mathbf{y}_k^{(i),a} = \mathbf{C}^{(i)}\mathbf{A}\hat{\mathbf{x}}_{k-1}^{(i)} + \eta^{(i)}$ for some $\eta^{(i)} \in \mathbb{R}^q$ and for simplicity, we denote $\mathbf{h}^i = \mathbf{A}\hat{\mathbf{x}}_{k-1}^{(i)}$. Then, we rewrite equation (9) as

$$\hat{\mathbf{x}}_k^{(i)} = \left(\frac{\lambda}{\left\|\Sigma_v^{-\frac{1}{2}}\eta^{(i)}\right\|}\mathbf{C}^{(i)T}\Sigma_v^{-1}\mathbf{C}^{(i)} + \frac{2}{d_i}\sum_{j\in N^{(i)}} \mathbf{P}_{|}^{(j)-1}\right)^{-1}$$
$$\cdot \left(\frac{\lambda}{\left\|\Sigma_v^{-\frac{1}{2}}\eta^{(i)}\right\|}\mathbf{C}^{(i)T}\Sigma_v^{-1}(\mathbf{C}^{(i)}\mathbf{h}^i + \eta^{(i)}) + \frac{2}{d_i}\sum_{j\in N^{(i)}} \mathbf{P}_{|}^{(j)-1}\mathbf{h}^j\right)$$
$$= \mathbf{h}^i + \left(\frac{\lambda}{\left\|\Sigma_v^{-\frac{1}{2}}\eta^{(i)}\right\|}\mathbf{C}^{(i)T}\Sigma_v^{-1}\mathbf{C}^{(i)} + \frac{2}{d_i}\sum_{j\in N^{(i)}} \mathbf{P}_{|}^{(j)-1}\right)^{-1}$$
$$\cdot \left(\frac{\lambda}{\left\|\Sigma_v^{-\frac{1}{2}}\eta^{(i)}\right\|}\mathbf{C}^{(i)T}\Sigma_v^{-1}\eta^{(i)} + \frac{2}{d_i}\sum_{j\in N^{(i)}} \mathbf{P}_{|}^{(j)-1}(\mathbf{h}^j - \mathbf{h}^i)\right)$$

Since, for each i and k, $\mathbf{C}^{(i)}, \hat{\mathbf{x}}_{k-1}^{(i)}, \mathbf{P}_{|}^{(i)-1}$ are fixed and bounded, then for any $\mathbf{y}_k^{(i),a}$, we have that

$$\|\hat{\mathbf{x}}_k^{(i)}\| \leq \|\mathbf{h}^i\| + \frac{2}{d_i}\left\|\left(\sum_{j\in N^{(i)}} \mathbf{P}_{|}^{(j)-1}\right)^{-1}\right\| \cdot$$
$$\left(\lambda\|\mathbf{C}^{(i)}\Sigma_v^{-\frac{1}{2}}\| + \frac{2}{d_i}\sum_{j\in N^{(i)}} \|\mathbf{P}_{|}^{(j)-1}\| \cdot \|\mathbf{h}^j - \mathbf{h}^i\|\right), \tag{15}$$

which implies that $\|\hat{\mathbf{x}}_k^{(i)}\|$ is bounded. ■