Consensus-Based Distributed Filtering with Fusion Step Analysis

Jiachen Qian\textsuperscript{a}, Peihu Duan\textsuperscript{c}, Zhisheng Duan\textsuperscript{a}, Guanrong Chen\textsuperscript{b}, Ling Shi\textsuperscript{c}

\textsuperscript{a}State Key Laboratory for Turbulence and Complex Systems, Department of Mechanics and Engineering Science, College of Engineering, Peking University, Beijing 100871, China
\textsuperscript{b}Department of Electrical Engineering, City University of Hong Kong, Hong Kong SAR, China
\textsuperscript{c}Department of Electronic and Computer Engineering, the Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China

Abstract

For consensus on measurement-based distributed filtering (CMDF), through infinite consensus fusion operations during each sampling interval, each node in the sensor network can achieve optimal filtering performance with centralized filtering. However, due to the limited communication resources in physical systems, the number of fusion steps cannot be infinite. To deal with this issue, the present paper analyzes the performance of CMDF with finite consensus fusion operations. First, by introducing a modified discrete-time algebraic Riccati equation and several novel techniques, the convergence of the estimation error covariance matrix of each sensor is guaranteed under a collective observability condition. In particular, the steady-state covariance matrix can be simplified as the solution to a discrete-time Lyapunov equation. Moreover, the performance degradation induced by reduced fusion frequency is obtained in closed form, which establishes an analytical relation between the performance of the CMDF with finite fusion steps and that of centralized filtering. Meanwhile, it provides a trade-off between the filtering performance and the communication cost. Furthermore, it is shown that the steady-state estimation error covariance matrix exponentially converges to the centralized optimal steady-state matrix with fusion operations tending to infinity during each sampling interval. Finally, the theoretical results are verified with illustrative numerical experiments.

Key words: Distributed filtering, Consensus, Information fusion, Algebraic Riccati Equation

1 Introduction

In the past three decades, the development of wireless sensor network (WSN) facilitates a wide range of engineering applications, e.g., environmental monitoring \cite{1} and spacecraft navigation \cite{2}. In these scenarios, sensors measure partial states of the target system and reconstruct the state of the system, which can be used to execute various kinds of tasks.

For WSNs consisting of multiple sensors, frameworks of state estimation are divided into two main categories: centralized state estimation and distributed state estimation. For centralized state estimation, a global fusion center receives and processes the measurements from sensors in the WSN. For distributed state estimation, each sensor acts as a local fusion center that collects observation information from its neighbors to reconstruct the state of the system. Compared with centralized state estimation, distributed state estimation is easier to implement and more robust, while its filtering mechanism is more challenging to design.

With the development of multi-agent consensus theory, some typical consensus-based distributed filtering algorithms were proposed and extensively studied \cite{3–13}. When equipped with the average consensus technique, each sensor first attempts to collect the global information and then altogether achieve the centralized optimal filtering performance. To name a few approaches, Olfati-Saber \cite{3} first formulated a framework of measurement-
based distributed filtering (CMDF), where each sensor
obtains the average measurement at every sampling in-
stant by exchanging information with neighbors. Later,
Kamal et al. [6] and Battistelli et al. [7,8] proposed some
distributed information fusion techniques such as con-
sensus on information matrix, hybrid consensus on mea-
surement and information matrix. Cattivelli et al. [5]
and Talebi et al. [13] developed several distributed filter-
ing structures based on average fusion with neighboring
nodes, with the convergence of the estimation error co-
variance guaranteed.

In the above references, the number of fusion step \( L \) be-
tween two successive sampling instants \( k \) and \( k+1 \) has a
noticeable effect on the performance of the proposed al-
gorithm. Kamal et al. [6], Kamagpour and Tomlin [14]
and Battistelli et al. [15] pointed out that each sensor
node in a distributed sensor network can achieve perfect
consensus on the observation information as the number
of fusion step \( L \) tends to infinity. Consequently, the per-
formance of the distributed filtering algorithm converges
to the centralized optimal steady-state performance [16].
However, the infinite fusion step \( L \) is impossible to re-
alize in practice, which renders the limited applicability
of the consensus-based distributed filtering algorithm.
To deal with this issue, Battistelli et al. [8] and Li et al.
[12,17] analyzed the CMDF algorithm with finite fu-
sion step \( L \). Specifically, Battistelli et al. [8] pointed out
that the consensus weight takes an integral part in the
performance of the CMDF algorithm. Li et al. [12,17]
carried out stability analysis of the CMDF algorithm
and performed optimization to obtain proper consensus
weights. Battilotti et al. [15] proposed a consensus-based
distributed algorithm, and show the convergence of the
steady-state performance to the centralized one with the
increase of fusion step \( L \). However, in the field of dis-
tributed filtering, the performance analysis was mainly
on the boundedness and convergence of the matrix iter-
ation, while more significant performance indexes, such
as convergence properties of steady-state performance
gap with the centralized filtering as \( L \) tends to infinity,
still require further exploration.

In this paper, an in-depth investigation on the perform-
ance of the CMDF algorithm is carried out, along with
an analysis on the number of fusion frequency \( L \). The
main contributions of this paper are summarized as fol-
lows:

1. The convergence condition of the CMDF algorithm
   with a finite fusion step \( L \) is derived (Theorem 1).
   It is shown that as long as the pair of state transi-
tion matrix and observation matrix is collectively
observable, the steady-state performance of the es-
timation error covariance matrix can be simplified
as the solution to a discrete-time Lyapunov equa-
tion (DLE). This result quantitatively describes the
steady-state performance degradation induced by

2. The effect of fusion step \( L \) on the steady-state per-
formance of CMDF is discussed. The infinite series
expression of the difference between two discrete-
time algebraic Riccati equation (DARE) solutions
is formulated. Based on this, it is shown that the
convergence rate of the performance gap between
the centralized optimal case and the distributed
case is not slower than exponential convergence,
with \( L \) tending to infinity (Theorem 2,3). This
result is a theoretical completion of performance
analysis of CMDF, which reflects the trade-off be-
tween estimation accuracy and communication cost
in the distributed setting.

3. Some new properties of the DARE are derived, in-
cluding the expression of the difference between
solutions to two different DARE and the uniform
property of solutions to a group of DARE. These
novel properties play an essential role in facilitating
the applicability of the CMDF algorithm.

The remainder of this paper is organized as follows. Some
preliminaries, including useful lemmas and the problem
formulation, are presented in Section 2. The main re-
results, including the derivation and analysis of the steady-
state performance of the CM-based algorithm, are pre-
sented in Section 3. Some illustrative numerical simula-
tions are presented in Section 4. Conclusions are drawn
in Section 5.

Notation: For two symmetric matrices \( X_1 \) and \( X_2 \), \( X_1 \geq X_2 \) (\( X_1 \geq X_2 \)) means \( X_1 - X_2 \) is positive definite (posi-
tive semi-definite). \( \exp (\cdot) \) denotes the exponential function. \( |a| \) denotes the absolute value of real number \( a \) or
the norm of complex number. \( \mathcal{L} \geq 0 \) \( (\mathcal{L} \geq 0) \) means that all the elements of matrix \( \mathcal{L} \) are positive (non-negative).
\( \mathbb{E} \{ x \} \) denotes the expectation of a random variable \( x \). \( \lambda (A) \) denotes the eigenvalue of matrix \( A \). \( \rho (\mathcal{A}) \) denotes
the spectral radius of \( A \). \( \lVert A \rVert_2 \) denotes the 2-norm (the largest singular value) of matrix \( A \).

2 Preliminaries

2.1 Graph Theory

The topology of a sensor network is denoted by a graph
\( \mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{L}) \), where \( \mathcal{V} = \{1,2, \ldots, N\} \) is the node set, \( \mathcal{E} \subseteq \mathcal{V} \times \mathcal{V} \) is the edge set, and \( \mathcal{L} = \{l_{ij}\} \) is the adjacency
matrix of the network. The adjacency matrix reflects the
communication among the nodes, \( l_{ij} > 0 \iff (i,j) \in \mathcal{E} \),
which means that sensor \( i \) can receive the information
from sensor \( j \), thus sensor \( j \) is denoted as the in-neighbor
of sensor \( i \). Sensor \( i \) is denoted as the out-neighbor
of sensor \( j \). \( N_i \) denotes the in-neighbor set of sensor \( i \),
and \( l_i \) denotes the \( i \)-th row of \( \mathcal{L} \), which contains the informa-
tion of \( N_i \). In addition, the matrix \( \mathcal{L} \) is doubly stochas-
tic, i.e., \( \sum_{j=1}^{N} l_{ij} = \sum_{j=1}^{N} l_{ji} = 1, \forall i \in \mathcal{V} \), which ensures
the average weighting of the information from the neighboring nodes. The diameter $d$ of graph $\mathcal{G}$ is the length of the longest path between two nodes in the graph.

### 2.2 Target System and Sensor Model

Consider a network of $N$ sensors, which measure and estimate the states of a target system, described as

$$
\begin{align*}
x_{k+1} &= Ax_k + \omega_k, \quad k = 0, 1, 2, \ldots \\
y_{i,k} &= C_i x_k + v_{i,k}, \quad i = 1, 2, \ldots, N,
\end{align*}
$$

where $x_k \in \mathbb{R}^n$ is the state vector of the system and $y_{i,k} \in \mathbb{R}^{n_i}$ is the measurement vector of sensor $i$, $\omega_k \in \mathbb{R}^n$ is the process noise with covariance matrix $Q \in \mathbb{R}^{n \times n}$ and $v_{i,k} \in \mathbb{R}^{m_i}$ is the observation noise with covariance matrix $R_i \in \mathbb{R}^{n_i \times n_i}$. The sequences $\{\omega_k\}_{k=0}^{\infty}$ and $\{v_{i,k}\}_{k=0,i=1}^{\infty,N}$ are mutually uncorrelated white Gaussian noise. $Q$ and $R_i$ are positive definite. Besides, $A$ is the state-transition matrix and $C_i$ is the observation matrix of sensor $i$. For simplicity, use $i \in \mathcal{V}$ to represent the $i$-th sensor of the network. $C = [C_1^T, C_2^T, \ldots, C_N^T]^T$ is the observation matrix and $R = \text{diag}(R_1, \ldots, R_N)$ is the covariance matrix of observation noise of the whole network. $l_{ij}^L$ is the $(i,j)$-th element of matrix $L^k$, which is the network adjacency matrix at step $L$.

#### 2.3 Some Useful Lemmas

**Lemma 1** If the communication topology corresponding to $\mathcal{L}$ is strongly connected and $\mathcal{L}$ is doubly stochastic with positive diagonal elements, then the matrix $L^k$ converges to $\frac{1}{N} 11^T$ exponentially as $k \to \infty$.

**Proof.** The strongly connected property guarantees that the matrix $\mathcal{L}$ is irreducible. With Theorem 8.5.2 and Lemma 8.5.5 in [18], the matrix $\mathcal{L}$ is also primitive with positive diagonal elements. Thus, there is only one eigenvalue of $\mathcal{L}$ equal to the spectral radius 1, and the norms of all other eigenvalues are strictly less than one. Consider any eigenvector corresponding to an eigenvalue $|\lambda| < 1$, denoted as $x$. Then, one has

$$1^T \mathcal{L} x = 1^T x = \lambda 1^T x,$$

where the first equality follows from the fact that $\mathcal{L}$ is doubly stochastic. Therefore one has $1^T x = 0$. Thus, all the eigenvalues of $\mathcal{L}$ except 1 are also eigenvalues of $\mathcal{L} - \frac{1}{N} 11^T$. With Theorem 1 in [19], the matrix $L^k$ converges to $\frac{1}{N} 11^T$ exponentially with the increase of $k$, and the convergence rate is not slower than the norm of the second largest eigenvalue of $\mathcal{L}$. 

**Lemma 2** (Matrix Inversion Lemma) For any matrix $P, Q, C$ of proper dimensions, if $P^{-1}$ and $Q^{-1}$ exist, then the following equality holds:

$$(P^{-1} + C^T Q^{-1} C)^{-1} = P - PC^T (CPC^T + Q)^{-1} CP.$$

**Lemma 3** For any matrix $A \in \mathbb{R}^{n \times n}$, the following inequality holds:

$$\|A^k\|_2 \leq \sqrt{n} \sum_{j=0}^{n-1} \binom{n-1}{j} \|A\|_2^j \rho(A)^{k-j},$$

where $(\begin{pmatrix} m \end{pmatrix}_n)$ is the combinatorial number, and $\rho(A)$ and $\|A\|_2$ denote the spectral radius and maximum singular value of $A$, respectively.

The proof is given in Appendix A.

#### 2.4 Problem Formulation

In this paper, the fundamental problem of CM-based distributed filtering algorithm with finite fusion step $L$ is formulated as follows:

1. Evaluate the effect induced by insufficient fusion on the steady-state performance of the CM-based distributed filtering algorithm.
2. Based on the above analysis, find out the relationship between fusion step $L$ and performance degradation compared with the centralized optimal performance.

### 3 Main Results

In this section, first, the convergence property of the proposed CMDF algorithm is analyzed, including the iteration of matrix and error covariance matrix. Then, the steady-state performance between the CM-based filter and the centralized filter are compared. To do so, the following two assumptions are needed.

**Assumption 1** The communication topology is strongly connected.

**Assumption 2** The system pair $(A, C)$ is collectively observable.

As proposed in [6, 7, 10, 20], the above two assumptions are both mild for distributed filtering problems. Generally speaking, Assumption 1 is to ensure globally average weighting of the information from the nodes of the sensor network so as to achieve the performance of the centralized optimal case, and Assumption 2 is essential.
for the stability of the filtering algorithm. It is worth mentioning that the doubly stochastic matrix \( \mathcal{L} \) can also be obtained in a distributed way as long as the communication graph is strongly connected [21].

### 3.1 Algorithm Convergence Analysis

In this subsection, consider the following CM-based distributed Kalman filtering algorithm proposed in [3,8,14], as shown in Algorithm 1, where \( N \) is the a priori parameter shared by all the nodes in the sensor network.

**Algorithm 1** Distributed Information Fusion Algorithm

**Input:**
\[
\hat{x}_{i,0}, P_{i,0} \quad i = 1, 2, \ldots, N
\]

**Prediction:**
\[
\hat{x}_{i,k|k-1} = A\hat{x}_{i,k-1} \\
P_{i,k|k-1} = AP_{i,k-1|k-1}A^T + Q
\]

**Fusion:**
\[
S_{i,k}^{(0)} = NC_i^T R_i^{-1} C_i, \quad I_{i,k}^{(0)} = NC_i^T R_i^{-1} y_{i,k}
\]

For \( i = 1, 2, \ldots, L \)
\[
S_{i,k}^{(i)} = \sum_{j=1}^{N} l_{ij} S_{i,k}^{(i-1)}, \quad I_{i,k}^{(i)} = \sum_{j=1}^{N} l_{ij} I_{i,k}^{(i-1)}
\]

**Correction:**
\[
P_{i,k|k} = \left( P_{i,k|k-1}^{-1} + S_{i,k}^{(L)} \right)^{-1} \\
\hat{x}_{i,k|k} = P_{i,k|k} \left( P_{i,k|k-1}^{-1} \hat{x}_{i,k|k-1} + I_{i,k}^{(L)} \right)
\]

With this algorithm, as the number of fusion step \( L \) tends to infinity, each sensor node can precisely obtain the information \( \sum_{i=1}^{N} C_i^T R_i^{-1} C_i \) and \( \sum_{i=1}^{N} C_i^T R_i^{-1} y_{i,k} \) in a distributed way. Based on this premise, Kamgar-pour and Tomlin [14] proved that, with \( k \) tending to infinity, the matrix \( P_{i,k|k-1} \) converges to the centralized optimal steady-state performance, i.e., the solution \( P \) to the DARE

\[
P = APA^T + Q - APC^T (CPC^T + R)^{-1} CPA^T.
\]

However, due to the limitation of the communication rate, it is impossible to communicate with the neighboring sensor node for infinitely many times between two sampling instants. Thus, the objective of this subsection is to derive the convergence property of the algorithm based on a finite fusion step \( L \).

Let the modified observation matrix for each sensor be
\[
\tilde{C}_i^{(L)} = \left[ \text{sign}(l_{i1}^{(L)}) C_1^T, \ldots, \text{sign}(l_{iN}^{(L)}) C_N^T \right]^T,
\]

and the modified noise covariance matrix be
\[
\tilde{R}_i^{(L)} = \text{diag} \left( \frac{1}{N l_{i1}^{(L)}} R_1, \ldots, \frac{1}{N l_{iN}^{(L)}} R_N \right),
\]

where
\[
\text{sign} (x) = \begin{cases} 
-1, & x < 0 \\
0, & x = 0 \\
1, & x > 0
\end{cases}
\]

and \( \frac{1}{N l_{ij}^{(L)}} \) is set to 0 if the denominator \( N l_{ij}^{(L)} = 0 \). As \( \mathcal{L} \) is a non-negative matrix, i.e., \( \mathcal{L}^T \geq 0 \), the term \( \text{sign}(l_{ij}^{(L)}) \) only takes the value of 0 or 1. The following Lemma describes the convergence property of the iteration \( P_{i,k+1|k} \).

**Lemma 4** If \( (A, \tilde{C}_i^{(L)}) \) is observable for all \( i \in \mathcal{V} \), then \( P_{i,k+1|k} \) converges to the solution of the DARE

\[
P_i^{(L)} = AP_i^{(L)} A^T + Q - AP_i^{(L)} (\tilde{C}_i^{(L)})^T \\
\times \left( \tilde{C}_i^{(L)} P_i^{(L)} (\tilde{C}_i^{(L)})^T + \tilde{R}_i^{(L)} \right)^{-1} \tilde{C}_i^{(L)} P_i^{(L)} A^T,
\]

i.e.,
\[
\lim_{k \to \infty} P_{i,k+1|k} = P_i^{(L)}, \quad \forall i \in \mathcal{V}.
\]

**Proof.** The iteration of \( P_{i,k+1|k} \) can be reformulated as

\[
P_{i,k+1|k} = A \left( P_{i,k|k-1}^{-1} + (\tilde{C}_i^{(L)})^T (\tilde{R}_i^{(L)})^{-1} \tilde{C}_i^{(L)} \right)^{-1} A^T + Q.
\]

With Lemma 2, the following equality is obtained:
Denote the estimation error covariance matrix as

\[
(P_{i,k|k-1}^{-1} + (\tilde{C}_i^{(L)})^T (\tilde{R}_i^{(L)})^{-1} \tilde{C}_i^{(L)})^{-1}
\]

\[
= P_{i,k|k-1}^{-1} - P_{i,k|k-1}(\tilde{C}_i^{(L)})^T \times (\tilde{C}_i^{(L)} P_{i,k|k-1}(\tilde{C}_i^{(L)})^T + \tilde{R}_i^{(L)})^{-1} \tilde{C}_i^{(L)} P_{i,k|k-1}^{-1}.
\]

The equation (5) can be reformulated as

\[
P_{i,k+1|k} = A P_{i,k|k-1} A^T + Q - A P_{i,k|k-1}(\tilde{C}_i^{(L)})^T \times (\tilde{C}_i^{(L)} P_{i,k|k-1}(\tilde{C}_i^{(L)})^T + \tilde{R}_i^{(L)})^{-1} \tilde{C}_i^{(L)} P_{i,k|k-1} A^T.
\]

Thus, equation (4) and the steady-state form of equation (5) are equivalent. With the result in [16], if \((A, \tilde{C}_i^{(L)})\) is observable, then \(P_{i,k+1|k}\) converges to the solution of the DARE (4).

**Remark 1** If the number of fusion step is less than the diameter of the communication topology, i.e., \(L < d\), then the matrix \(\tilde{R}_i^{(L)}\) is not invertible since some \(l_{ij}^{(L)}\) may be 0. However, one can eliminate the corresponding zero blocks in \(\tilde{C}_i^{(L)}\) and \(\tilde{R}_i^{(L)}\) to make the modified \(\tilde{R}_i^{(L)}\) invertible, or replace the inverse sign with generalized inverse, i.e., \((\tilde{R}_i^{(L)})^{-1}\). These two kinds of modification do not affect the observability of the pair \((A, \tilde{C}_i^{(L)})\) and the equivalent relationship

\[
\tilde{C}_i^{(L)} (\tilde{R}_i^{(L)})^{-1} \tilde{C}_i^{(L)} = N \sum_{j=1}^N l_{ij}^{(L)} C_j^T R_j^{-1} C_j.
\]

In order to simplify the notation, \(\tilde{C}_i^{(L)}, \tilde{R}_i^{(L)}, (\tilde{R}_i^{(L)})^{-1}\) and \(\tilde{R}_i^{(L)}\) will be kept to denote the modified observation and noise matrices.

With the DARE (4), it is clear that the steady-state performance of the CMDF algorithm is closely related to the number of fusion step \(L\) between two sampling instants. Furthermore, the matrix \(P_i^{(L)}\) is determined by \(A, \tilde{C}_i^{(L)}, \tilde{R}_i^{(L)}\), which is apparently not the real state estimation error covariance matrix for sensor \(i\). Therefore, in addition to the above analysis on \(P_{i,k+1|k}\), it also needs to analyze the steady-state performance of the real error covariance matrix of each sensor \(i\).

Denote the estimation error of sensor node \(i\) as \(\hat{x}_{i,k|k-1} = x_k - \hat{x}_{i,k|k-1}, \hat{x}_{i,k|k} = x_k - \hat{x}_{i,k|k}\), and denote the estimation error covariance as 

\[
P_{i,k|k-1} = \mathbb{E}\{(\hat{x}_{i,k|k-1} - \hat{x}_{i,k|k-1})^T (\hat{x}_{i,k|k-1} - \hat{x}_{i,k|k-1})\}, \quad P_{i,k|k} = \mathbb{E}\{(\hat{x}_{i,k|k} - \hat{x}_{i,k|k})^T (\hat{x}_{i,k|k} - \hat{x}_{i,k|k})\}.
\]

**Theorem 1** If \((A, \tilde{C}_i^{(L)})\) is observable for \(i \in V\), then the iteration of \(P_{i,k|k-1}\) will converge to the solution to the discrete-time Lyapunov equation (DLE)

\[
\tilde{P}_i^{(L)} = \tilde{A}_i^{(L)} \tilde{P}_i^{(L)} \tilde{A}_i^{(L)}^T + Q + K_{P_i^{(L)}} \tilde{R}_i^{(L)} K_{P_i^{(L)}}^T,
\]

where

\[
\tilde{A}_i^{(L)} = A - A P_i^{(L)} (\tilde{C}_i^{(L)})^T \times (\tilde{C}_i^{(L)} P_i^{(L)} (\tilde{C}_i^{(L)})^T + \tilde{R}_i^{(L)})^{-1} \tilde{C}_i^{(L)}
\]

\[
K_{P_i^{(L)}} = A P_i^{(L)} (\tilde{C}_i^{(L)})^T \times (\tilde{C}_i^{(L)} P_i^{(L)} (\tilde{C}_i^{(L)})^T + \tilde{R}_i^{(L)})^{-1}.
\]

The proof is given in Appendix B.

**Remark 2** For CMDF, the explicit values of \(\sum_{i=1}^N C_i^T \cdot R_i^{-1} y_i\) and \(\sum_{i=1}^N C_i^T \cdot R_i^{-1} C_i\) in each sampling instant cannot be obtained by each node with finite fusion step \(L\). Compared with the centralized optimal setting, the effect of imprecise information on the state estimation performance will accumulate, and the corresponding degraded estimation performance will become steady. A more in-depth discussion on the performance degradation induced by finite fusion step \(L\) is proposed in Theorem 2 and 3.

In addition to these, Theorem 1 also provides the least requirement on the stability of the CM-based distributed filtering algorithm. Denote the \(L\)-step neighbor of node \(i\) as the nodes that can reach node \(i\) in \(L\) steps. Lemma 4 and Theorem 1 indicate that the lowest bound of \(L\) for the boundedness of \(P_{i,k|k-1}, \forall i \in V\), is that each pair \((A, \tilde{C}_i^{(L)})\) is observable for all \(i \in V\), namely, for each node \(i \in V\), its \(L\)-step neighbors need to be collectively observable. Note that if the fusion is not sufficient enough, i.e., the pairs \((A, \tilde{C}_i^{(L)})\) are not observable for some \(i \in V\), then the real covariance matrix \(\tilde{P}_i^{(L)}\) may diverge for an unstable \(A\). However, the CM-based distributed filtering algorithm can achieve the centralized optimal performance with any degree of precision for sufficiently large \(L\), as shown below.

Now, one is ready to complete the analysis about the convergence of \(P_{i,k|k-1}\) and \(P_{i,k|k-1}\) to the centralized performance, and to complete the formulation of the steady-state performance of iterations as the solutions to DARE (4) and DLE (6), respectively.

When \(L\) is larger than the diameter of the communication graph, i.e., \(L \geq d\), one has \(l_{ij}^{(L)} > 0\) for all \(i, j\) [7],
Thus, one can rewrite the DARE (4) for each sensor node as
\[
P_i^{(L)} = AP_i^{(L)}A^T + Q - AP_i^{(L)}C^T (CP_i^{(L)}C^T + \tilde{R}_i^{(L)})^{-1} CP_i^{(L)}A^T.
\]

It is obvious that, compared to the centralized optimal setting, the solution of the matrix \(P_i^{(L)}\) is based on a mismatched \(\tilde{R}_i^{(L)}\) instead of \(R\). Therefore, the effect of non-sufficient fusion is formulated as the mismatch of noise covariance matrix \(R\). In the next subsection, the effect of mismatched \(R\) on the properties of \(P_i^{(L)}\) and \(\tilde{P}_i^{(L)}\) with \(L \geq d\) will be further discussed.

### 3.2 Performance Analysis

Before discussing the comparison of filtering performance, the following two Lemmas are established to illustrate the properties of solutions to DARE. For simplicity, use the term \(\text{dare}(A, C, Q, R)\) to denote the solution \(P\) to DARE (2).

**Lemma 5** The solution \(P\) to the DARE (2) is monotonically increasing with \(R\), i.e., if \(R_1 \geq R_2 > 0\), \(P_1 = \text{dare}(A, C, Q, R_1)\) and \(P_2 = \text{dare}(A, C, Q, R_2)\), then \(P_1 \geq P_2\).

The proof to this Lemma is similar to the proof of Theorem 1 in [22], thus is omitted.

Similarly, denote \(\tilde{A}_P = A - APC^T (CPC^T + R)^{-1} C\), where \(P\) is the solution to DARE (2).

**Lemma 6** The matrix \(\tilde{A}_P\) is Schur stable and
\[
\rho(\tilde{A}_P) \leq \sqrt{1 - \frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)}} , \quad \|\tilde{A}_P\|_2 \leq \sqrt{\frac{\lambda_{\max}(P)}{\lambda_{\min}(Q)}}.
\]

The proof is given in Appendix C.

**Remark 3** Lemma 6 shows the property of the feedback matrix \(\tilde{A}_P\), in terms of spectral radius and singular values. This Lemma is of vital importance to analyze the uniform property of the solutions to a group of Riccati Equations, as can be seen in the proof of Theorem 2.

In order to discuss the performance degradation due to the effect of insufficient fusion, it needs to compare the difference between two DARE first. Introduce two Riccati Equations with different \(R\):
\[
P_1 = AP_1 A^T + Q - AP_1 C^T (CP_1 C^T + R_1)^{-1} CP_1 A^T
\]
\[
P_2 = AP_2 A^T + Q - AP_2 C^T (CP_2 C^T + R_2)^{-1} CP_2 A^T.
\]

**Lemma 7** Suppose \(P_1\) and \(P_2\) satisfy the DARE (7) respectively. Then, one has
\[
P_1 - P_2 = \sum_{k=0}^{\infty} \tilde{A}_P^k AP_1 C^T (R_2^{-1} - R_1^{-1}) C \tilde{P}_2 A^T (\tilde{A}_P^T)^k,
\]
where
\[
\tilde{P}_1 = P_1 - A^{-1} K_P_1 C P_1, \quad \tilde{P}_2 = P_2 - A^{-1} K_P_2 C P_2
\]
\[
K_P_1 = AP_1 C^T (CP_1 C^T + R_1)^{-1}
\]
\[
K_P_2 = AP_2 C^T (CP_2 C^T + R_2)^{-1}.
\]

The proof is given in Appendix D.

With Lemma 7, one can estimate the difference of two solutions to DARE with different \(R\) by a sum of infinite series:
\[
\|P_1 - P_2\|_2 \leq \|\tilde{P}_1\|_2 \|\tilde{P}_2\|_2 \|C\|_2 \|R_2^{-1} - R_1^{-1}\|_2
\]
\[
\times \|\tilde{A}_P\|_2^2 \sum_{k=0}^{\infty} \|\tilde{A}_P^k\|_2 \|\tilde{A}_P^k\|_2.
\]

Let \(R_1 = R\), and \(R_2 = \tilde{R}_i^{(L)} (L \geq d)\). Then, an estimation of the performance degradation compared with the centralized setting is given by
\[
\|P - P_i^{(L)}\|_2 \leq \|\tilde{P}_1\|_2 \|\tilde{P}_i^{(L)}\|_2 \|C\|_2 \|\tilde{R}_i^{(L)}\|^{-1} - R_1^{-1}\|_2
\]
\[
\times \|\tilde{A}_P\|_2^2 \sum_{k=0}^{\infty} \|\tilde{A}_P^k\|_2 \|\tilde{A}_P^k\|_2.
\]

Next, it remains to analyze the property of the infinite sum \(\sum_{k=0}^{\infty} \|\tilde{A}_P^k\|_2 \|\tilde{A}_P^k\|_2\) to obtain the relationship between \(\|P - P_i^{(L)}\|_2\) and \(\|\tilde{R}_i^{(L)}\|^{-1} - R_1^{-1}\|_2\), where Lemma 3 and Lemma 6 will be utilized.

As the matrices \(\tilde{A}_P\) and \(\tilde{A}_P^{(L)}\) are Schur stable, one can find a bound of the infinite sum \(\sum_{k=0}^{\infty} \|\tilde{A}_P^k\|_2 \|\tilde{A}_P^k\|_2\) and obtain an approximate result as
\[
\|P - P_i^{(L)}\|_2 \leq M \|\tilde{R}_i^{(L)}\|^{-1} - R_1^{-1}\|_2,
\]
where $M$ is a constant. However, for different $\tilde{R}_i^{(L)}$, the constant $M$ may also be different. This issue makes the approximately linear relationship between $\|P - P_i^{(L)}\|_2$ and $\|\tilde{R}_i^{(L)} - R\|_2$, not so trivial to obtain, meanwhile it brings about difficulties to the analysis of the uniform property of $\|P - P_i^{(L)}\|_2$ for all $i \in \mathcal{V}$ and $L \geq d$. Therefore, it needs to discuss the property of the infinite sum $\sum_{k=0}^{\infty} \|\tilde{A}_i^k P\|_2 \|\tilde{A}_i^k P_i^{(L)}\|_2$ for a group of $\tilde{R}_i^{(L)}$ to reveal the trade-off between the number of fusion step $L$ and the estimation performance.

**Theorem 2** For any given $(A, C)$ and $Q, R$, there exist two constants $M_1 > 0, 0 < q_1 < 1$, such that

$$\|P - P_i^{(L)}\|_2 \leq M_1 q_i^L, \quad \forall i \in \mathcal{V}, \forall L \geq d.$$  

The proof is given in Appendix E.

**Remark 4** The main difficulty in proving Theorem 2 is to formulate a uniform upper bound in terms of $\sum_{k=0}^{\infty} \|\tilde{A}_i^k P\|_2 \|\tilde{A}_i^k P_i^{(L)}\|_2$ for all $L \geq d$ and $i \in \mathcal{V}$. As one can use an exponentially convergent bound for $\|P_i^{(L)} - P\|_2$, the convergence speed of $\|P_i^{(L)} - P\|_2$ with increasing $L$ is not slower than exponential convergence. Theorem 2 also illustrates that, with the number of fusion step tending to infinity, the steady-state performance matrix $P_i^{(L)}$ will exponentially converge to the centralized optimal steady-state performance, which reflects the trade-off between estimation performance and communication cost among sensor nodes in distributed setting.

**Remark 5** It is worth mentioning that the results proposed in [3, 6, 14, 15] only discussed the asymptotic optimal property of the corresponding algorithm, i.e., the case when $L \rightarrow \infty$. However, the case with finite iteration step $L$ was rarely discussed in the literature. Meanwhile, the references [7, 23–25] discussed the performance of distributed information fusion algorithms with finite fusion steps but did not compare the filtering performance with the centralized optimal setting. Thus, Theorem 2 not only characterizes the property of the filtering algorithm with insufficient information fusion but also compares the performance degradation with the centralized setting, which is a more in-depth result in the area of distributed filtering.

It is apparent that the matrix $P_i^{(L)}$ is not the real error covariance matrix, i.e.,

$$P_i^{(L)} \neq \lim_{k \rightarrow \infty} \mathbb{E}\{\tilde{x}_{i,k|k-1} \tilde{x}_{i,k|k-1}^T\}.$$  

With the aforementioned analysis, the iteration of $\mathbb{E}\{\tilde{x}_{i,k|k-1} \tilde{x}_{i,k|k-1}^T\}$ will finally converge to a solution to DLE. Thus, with the technique developed above, one can further discuss the real performance degradation of $\mathbb{E}\{\tilde{x}_{i,k|k-1} \tilde{x}_{i,k|k-1}^T\}$ due to insufficient fusion $L$.

First, the aim is to analyze the distance between the matrix $P_i^{(L)}$ and the matrix $\bar{P}_i^{(L)}$.

**Theorem 3** For any given $(A, C)$ and $Q, R$, there exist two constants $M_2 > 0, 0 < q_2 < 1$, such that

$$\|\bar{P}_i^{(L)} - P_i^{(L)}\|_2 \leq M_2 q_i^L, \quad \forall i \in \mathcal{V}, \forall L \geq d.$$  

The proof is given in Appendix F.

The following Corollary finally illustrates the performance degradation of $\bar{P}_i^{(L)}$ compared with the centralized steady-state performance $P$.

**Corollary 1** For any given $(A, C)$ and $Q, R$, there exist two constants $M_3 > 0, 0 < q_3 < 1$, such that

$$\|\bar{P}_i^{(L)} - P\|_2 \leq M_3 q_i^L, \quad \forall i \in \mathcal{V}, \forall L \geq d.$$  

**Proof.** The corollary can be proved with Theorem 2, Theorem 3 and the triangular inequality of 2-norm:

$$\|\bar{P}_i^{(L)} - P\|_2 \leq \|\bar{P}_i^{(L)} - \bar{P}_i^{(L)}\|_2 + \|\bar{P}_i^{(L)} - P\|_2.$$  

□

**Corollary 2** The parameters $q_1, q_2, q_3$ mentioned in Theorem 2, Theorem 3 and Corollary 1 are not larger than the norm of the second largest eigenvalue of the stochastic matrix $L$.

**Proof.** From the proof of Theorem 2 and Theorem 3, the decay rates $q_1, q_2, q_3$ are actually bounded by that of $N_{ii}^{(L)} - 1$. Thus, the result follows. □

3.3 Discussion

In Theorem 3, the deviation between $\bar{P}_i^{(L)}$ and $P_i^{(L)}$ are reformulated in the series form (F.1), which indicates that the deviation is in essence determined by the difference of the noise terms $R$ and $\tilde{R}_i^{(L)}$. With the Kalman equality $K_{P^{(L)}} = A \bar{P}_i^{(L)} C^T (\bar{P}_i^{(L)})^{-1}$ and the uniform matrix.
boundedness of $\sum_{k=0}^{\infty} \| \hat{A}_{(\ell)}^{k} \|^2_{2}$, a more inherent result is that the decay rates of $\| P_i^{(L)} - \hat{P}_i^{(L)} \|_2$ are uniformly bounded by that of $\tilde{L}_{ij}^{(L)} = (N_{ij}^{(L)} - 1)^2 + (N_{ij}^{(L)} - 1)$, in which the linear term $(N_{ij}^{(L)} - 1)$ dominates the convergence speed for sufficiently large $L$. Based on these properties, the exponential convergence of $\| P_i^{(L)} - \hat{P}_i^{(L)} \|_2$ is finally proved.

Corollary 2 also reflects the superiority of the exponential convergence. An intuitive idea from Corollary 2 is that the parameter $l_{ij}$ could be designed to minimize the second largest eigenvalue of $L$ or minimize the $L$-step consensus error to reach better performance. Some distributed parameter tuning techniques such as subgradient method [19] or graph filtering [26] can also be utilized, which is an interesting direction for future exploration.

To summarize, the contents and proof of Theorem 3 are the basis for the derivation of Corollary 1 and 2, and Corollary 2 reveals the connection between the convergence rate $q$ and the spectral property of $L$.

4 Simulation

In this section, a target tracking numerical experiment is provided to illustrate the effectiveness of the proposed algorithm. The state transition matrix has the expression

$$a_k = \begin{pmatrix} 1 & T \\ 0 & 1 \end{pmatrix} \quad A_k = \begin{pmatrix} a_k & 0_{2 \times 2} \\ 0_{2 \times 2} & a_k \end{pmatrix}.$$ 

The error covariance matrix $Q$ takes the form of

$$G = \begin{pmatrix} T^3 & T^2 \\ T^2 & T \end{pmatrix} \quad Q = \begin{pmatrix} G & 0.5G \\ 0.5G & G \end{pmatrix},$$

where the sample interval is set to be $T = 1$. There are three kinds of sensors with the observation matrix:

$$C^{(1)} = [1, 0, 0, 0]$$

$$C^{(2)} = [0, 0, 1, 0]$$

$$C^{(3)} = [0, 0, 0, 0].$$

The whole network consists of 20 sensor nodes, including 3 sensors of kind $C^{(1)}$, 3 sensors of kind $C^{(2)}$, 14 sensors of kind $C^{(3)}$, and $R_i = 1, \forall i \in V$. The locations of the sensors are randomly set in a $300 \times 300$ region and each sensor is with a communication radius of 130. The communication topology of sensor network is randomly generated in the numerical experiments, as presented in Fig.1.

![Fig. 1. Illustration figure for communication topology of the sensor network.](image)

The expression of three kinds of observation matrices indicates that each kind of sensor is able to measure some partial states of the system but none of the pair $(A, C_i)$ is fully observable. Thus, the numerical experiments can fully illustrate the effectiveness of the CMDF algorithm, as well as the relationship between the filtering performance and fusion step $L$. Furthermore, the existence of $C^{(3)}$ implies that some sensor nodes in the network do not have the ability to observe the target system. This kind of node is called a naive node [20].

In the experiment, the steady-state performance of the CMDF algorithm is evaluated by the mean square error (MSE). Using Monte Carlo method, the filtering process is run 200 steps for each simulation and 1000 times in total. With the result obtained in Section 3.1, in each simulation, the performance of the CMDF algorithm converges to the steady-state performance. Thus, the estimated value in the steady-state zone is used to calculate the mean square error, i.e.,

$$MSE_{i,k} = \frac{1}{1000} \sum_{l=1}^{1000} \| \hat{x}_{i,\infty}^{(l)} - x_{i,\infty}^{(l)} \|^2_2,$$  \hspace{1cm} (8)

where $\hat{x}_{i,\infty}^{(l)}$ and $x_{i,\infty}^{(l)}$ denote the estimated state and real state in the $l$-th simulation, respectively, and $\infty$ means sufficiently large sampling instant $k$.

Through changing the value of fusion step $L$, one can calculate the steady-state performance of each sensor with different values of fusion step $L$. The performance of CM-based distributed filtering algorithm with different fusion step $L$ is presented in Fig.2 and Fig.3, where Fig.2 illustrates the steady-state performance of mean square error for partial sensors to fully illustrate the exponential convergence and Fig.3 illustrates the steady-state performance with different $L$ for the whole sensor network. Fig.2 and Fig.3 verify the results of Theorem 2 and Theorem 3.

Remark 6 It is worth mentioning that the filtering per-
Compared with the literature, the results established in this paper theoretically supplement to the problem of consensus-based distributed filtering. In addition to this, the results on the DARE can be utilized in other fields of state estimation. Future work will include a preciser analysis on the exponential bound of $\|P_i^{(L)} - P\|_2$, including parameter selection for controlling the steadystate performance of the algorithm, and parameter tuning technique for a better convergence rate.

A Proof of Lemma 3

Consider the Schur Decomposition of matrix $A$ as $A = U^HTU$, where $U$ is a unitary matrix and $T$ is an upper-triangular matrix with the diagonal elements being the eigenvalues of $A$. As the singular value is unitary invariant, it is easy to obtain

$$\lambda_i (T) = \lambda_i (A), \quad \|T\|_2 = \|A\|_2.$$ 

By the definition of 2-norm, one has

$$\|A\|_2^2 = \max (\lambda (AA^T)) = \max (\lambda (TT^T)) .$$

As the maximum eigenvalue of the positive definite matrix $AA^T$ is larger than all the diagonal elements of $AA^T$, one has

$$\|A\|_2^2 \geq \max_i \sum_{j=1}^{n} t_{ij}^2, \quad \max |t_{ij}| \leq \|A\|_2,$$

where $t_{ij}$ is the $(i, j)$-th element of $T$. Denote $M_T = \max_{i,j} |t_{ij}|$. By Corollary 3.15 in [27], one has

$$\|A^k\|_2 \leq \sqrt{n} \sum_{j=0}^{n-1} \binom{n-1}{j} M_T^j \max |A|^{k-j}.$$ 

B Proof of Theorem 1

With Algorithm 1, one obtains

$$\tilde{x}_{i,k|k} = P_{i,k|k} \tilde{P}_{i,k|k}^{-1} x_k - P_{i,k|k} \times \left( P_{i,k|k-1}^{-1} \tilde{x}_{i,k|k-1} + \sum_{j \in \mathcal{N}_i} t_{ij} C_{ij}^T R_{j}^{-1} y_j \right)$$

$$= P_{i,k|k} \left( P_{i,k|k-1}^{-1} \tilde{x}_{i,k|k-1} - (\tilde{C}_i^{(L)})^T (\tilde{R}_i^{(L)})^{-1} \tilde{v}_i \right)$$

$$\tilde{P}_{i,k|k} = P_{i,k|k} \left[ \tilde{P}_{i,k|k-1}^{-1} \tilde{P}_{i,k|k-1} P_{i,k|k-1} \tilde{P}_{i,k|k} \right]$$

$$+ P_{i,k|k} (\tilde{C}_i^{(L)})^T (\tilde{R}_i^{(L)})^{-1} \tilde{R}_i^{(L)} (\tilde{R}_i^{(L)})^{-1} \tilde{C}_i^{(L)} P_{i,k|k}$$
and
\[ \tilde{x}_{i,k+1|k} = A\tilde{x}_{i,k|k} + \omega_{k+1} \]
\[ \tilde{P}_{i,k+1|k} = A\tilde{P}_{i,k|k}A^T + Q. \]

It follows that
\[
AP_{i,k|k}P_{i,k|k-1}^{-1} = A - AP_{i,k|k-1}(\tilde{C}_i^{(L)})^T \\
\times \left( \tilde{C}_i^{(L)} P_{i,k|k-1}(\tilde{C}_i^{(L)})^T + \tilde{R}_i^{(L)} \right)^{-1} \tilde{C}_i^{(L)},
\]
and
\[
P_{i,k|k}(\tilde{C}_i^{(L)})^T (\tilde{R}_i^{(L)})^{-1} = P_{i,k|k-1}(\tilde{C}_i^{(L)})^T \\
\times \left( \tilde{C}_i^{(L)} P_{i,k|k-1}(\tilde{C}_i^{(L)})^T + \tilde{R}_i^{(L)} \right)^{-1}.
\]

As \( k \) tends to infinity, these two matrices will converge to the steady-state forms, i.e.,
\[
\lim_{k \to \infty} AP_{i,k|k}P_{i,k|k-1}^{-1} = \hat{A}_P^{(L)} \]
\[
\lim_{k \to \infty} AP_{i,k|k}\tilde{C}_i^{(L)} (\tilde{R}_i^{(L)})^{-1} = K_P^{(L)}.
\]

Being the solution to DARE, the above matrix \( \hat{A}_P^{(L)} \) is Schur stable, i.e., the spectral radius of \( \hat{A}_P^{(L)} \) is less than 1 [28]. Thus, with Theorem 1 in [5], the steady-state performance of \( \tilde{P}_{i,k|k-1} \) will also converge, i.e.,
\[
\lim_{k \to \infty} \tilde{P}_{i,k|k-1} = \tilde{P}_i^{(L)} \]
and the convergent solution satisfies the DLE (6).

\[ \square \]

\section{Proof of Lemma 6}

Consider the DARE
\[ P = APA^T + Q - APCT \left( CPC^T + R \right)^{-1} CPA. \]

Following the similar idea in [28], the above equation is rewritten as
\[ P = \hat{A}_P P \hat{A}_P^T + Q + K_P RK_P^T, \quad (C.1) \]
where \( K_P = APCT \left( CPC^T + R \right)^{-1} \). Consider the left eigenvector \( x \) of \( \hat{A}_P \) with corresponding eigenvalue \( \lambda \), i.e., \( x\hat{A}_P = \lambda x \). Pre- and post-multiplying the DARE (C.1) with \( x \) and \( x^H \), respectively, one has
\[
xPxx^H = |\lambda|^2 xPxx^H + xQxx^H + xKPRK_P^TxPxx^H \\
\geq |\lambda|^2 xPxx^H + xQxx^H.
\]

The inequality holds due to the fact that the matrix \( K_P RK_P^T \) is positive semi-definite. Dividing both side of the inequality with \( xPxx^H \), one has
\[ |\lambda|^2 \leq 1 - \frac{xQxx^H}{xPxx^H} \leq 1 - \frac{\min \{ xQxx^H \} }{xPxx^H} \leq 1 - \frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)}, \]
thus the inequality of the spectral radius is proved.

For the inequality on the 2-norm, consider the singular value decomposition \( \hat{A}_P = USV^T \), where \( \Sigma = \text{diag} \{ \sigma_1, \ldots, \sigma_n \} \), \( \sigma_1 \geq \cdots \geq \sigma_n \geq 0 \). Let \( x = (1, 0, \ldots, 0)^T \). Pre- and post-multiplying the DARE (C.1) with \( x \) and \( x^T \), respectively, one has
\[ xPxx^T = \sigma_1^2 (V^TPV)_{1,1} + xQxx^T + xKPRK^Tx^T, \]
where \((V^TPV)_{1,1}\) is the (1, 1)-th element of matrix \( V^TPV \). Note that \( (V^TPV)_{1,1} \geq \lambda_{\min} (V^TPV) = \lambda_{\min} (P) \). Thus, one has
\[ \sigma_1^2 \leq \frac{\lambda_{\max}(P)}{\lambda_{\min}(Q)} \leq \frac{\lambda_{\max}(P)}{\lambda_{\min}(Q)}, \]
where the last inequality holds because \( P \geq Q \).

\[ \square \]

\section{Proof of Lemma 7}

As \( P_1 \) and \( P_2 \) satisfy the DARE (7), one has
\[ P_1 - P_2 = A(P_1 - P_2)A^T + K_{P_1} C(CP_1 C^T + R) K_{P_2} + K_{P_2} C^T P_1 A^T + A P_1 C^T K_{P_2}^T. \]

Note that
\[
AP_1 C^T K_{P_2}^T - K_{P_1} CP_2 A^T \\
= K_{P_1}(CP_1 C^T + R_1) K_{P_2} - K_{P_1}(CP_2 C^T + R_2) K_{P_2} \\
= K_{P_1} C(CP_1 - P_2) C^T K_{P_2}^T + K_{P_1} (R_1 - R_2) K_{P_2}^T.
\]

Thus, the original equation (D.1) can be rewritten as
\[
P_1 - P_2 = (A - K_{P_1} C)(P_1 - P_2)(A - K_{P_2} C)^T \\
+ K_{P_1} (R_1 - R_2) K_{P_2}^T \\
= \hat{A}_{P_1}(P_1 - P_2) \hat{A}_{P_2}^T + K_{P_1} (R_1 - R_2) K_{P_2}^T.
\]

Performing the above iteration for \( k \) times, one has
\[
P_1 - P_2 = \hat{A}_{P_1}^k (P_1 - P_2) (\hat{A}_{P_2}^T)^k \\
+ \sum_{i=0}^{k} \hat{A}_{P_1}^i K_{P_1} (R_1 - R_2) K_{P_2}^T (\hat{A}_{P_2}^T)^i.
\]

Due to the fact that \( \hat{A}_{P_1} = A - K_{P_1} C \) and \( \hat{A}_{P_2} = A - K_{P_2} C \) are Schur stable, by performing the iteration for infinitely many times, one has
\[
\lim_{k \to \infty} \hat{A}_{P_1}^k (P_1 - P_2) (\hat{A}_{P_2}^T)^k = 0.
\]
and

\[ P_1 - P_2 = \sum_{i=0}^{\infty} \hat{A}_{P_i}^i K_{P_i} (R_1 - R_2) K_{P_i}^T (\hat{A}_{P_i}^T)^i. \]  

(D.2)

With the Kalman identical equation, one has

\[ K_{P_1} = A\hat{P}_1 C^T R_1^{-1} \quad K_{P_2} = A\hat{P}_2 C^T R_2^{-1}. \]

Therefore, the equation (D.2) can be rewritten as

\[ P_1 - P_2 = \sum_{i=0}^{\infty} \hat{A}_{P_i}^i A\hat{P}_1 C^T (R_2^{-1} - R_1^{-1}) C\hat{P}_2 A^T (\hat{A}_{P_i}^T)^i. \]

\[ \square \]

E Proof of Theorem 2

First, analyze the property of \( \| (\hat{R}_i^{(L)})^{-1} - R^{-1} \|_2 \).

Using the designed algorithm, one can obtain the following relationship:

\[ (\hat{R}_i^{(L)})^{-1} = N\text{diag}((l_i^{(L)})_1 I_{n_1}, \cdots, (l_i^{(L)})_{n_i} I_{n_i}) R^{-1}. \]

By Lemma 1, the matrix \( L^L \) converges to \( \frac{1}{N} 1_{N} L^L \) exponentially with increasing \( L \), which indicates that each term \( N l_i^{(L)} \) converges to 1 exponentially. Thus, there exist two constants \( M_0 > 0 \) and \( 0 < q < 1 \) such that, for any \( L \geq d \) and \( i \in \mathcal{V} \),

\[ \| (\hat{R}_i^{(L)})^{-1} - R^{-1} \|_2 < M_0 q^L. \]

The main idea of the following proof is to find a uniform bound on \( \sum_{i=0}^{\infty} \| \hat{A}_{P_i}^k \|_2 \| \hat{A}_{P_i}^k L^L \|_2 \) for a group of \( \hat{R}_i^{(L)} \), i.e., \( \forall i \in \mathcal{V}, L \geq d \). With Lemma 3, one can obtain

\[ \| \hat{A}_{P_i}^k \|_2 \leq \sqrt{n} \sum_{j=0}^{n-1} \binom{n-1}{j} k^j \| \hat{A}_{P_i}^j \|_2 \| \rho(\hat{A}_{P_i}^j) \|_2^{k-j}, \]

where \( n \) is the dimension of the system. Choose a sufficiently large \( R \), such that \( \hat{R}_i^{(L)} \leq R \) for all \( i \in \mathcal{V} \) and \( L \geq d \). Denote \( P = \text{dare} (A, C, Q, R) \), With Lemma 5 and Lemma 6, one can obtain

\[ \| A_{P_i}^j \| \leq \sqrt{\frac{\lambda_{\max}(P)}{\lambda_{\min}(Q)}} \leq \sqrt{\frac{\lambda_{\max}(A)}{\lambda_{\min}(Q)}} \leq \sigma_1 \]

\[ \rho(A_{P_i}^j) \leq \frac{1 - \lambda_{\min}(Q)}{\lambda_{\max}(P)} \leq 1 - \frac{1 - \lambda_{\min}(Q)}{\lambda_{\max}(P)} \leq \rho_p. \]

Thus, one has

\[ \sum_{k=0}^{\infty} \| \hat{A}_{P_i}^k \|_2 \| \hat{A}_{P_i}^k L^L \|_2 \leq \sum_{k=0}^{\infty} \sqrt{n} \sum_{j=0}^{n-1} \binom{n-1}{j} k^j \| \rho(\hat{A}_{P_i}^j) \|_2^{k-j} \]

\[ \leq \sum_{k=0}^{\infty} \sqrt{n} \sum_{j=0}^{n} \lambda_{\max}(P) \binom{n-1}{j} k^{n} \| \rho(\hat{A}_{P_i}^j) \|_2^{k-j} \]

\[ \text{for all } i \text{ and } L \geq d. \]

It is easy to verify the convergence of the infinite sum \( \sum_{k=0}^{\infty} \| \hat{A}_{P_i}^k \|_2 \) with \( \rho_p < 1 \). Based on the fact that the 2-norm \( \| \hat{A}_{P_i}^k \|_2 \) is uniformly bounded for all \( k \), one can find a real number \( M_0 ' \) such that

\[ \sum_{k=0}^{\infty} \| \hat{A}_{P_i}^k \|_2 \| \hat{A}_{P_i}^k L^L \|_2 \leq M_0 ' \quad \forall i \in \mathcal{V}, \forall L \geq d. \]

With the above analysis, the theorem is proved. \( \square \)

F Proof of Theorem 3

The matrices \( P_i^{(L)} \) and \( \tilde{P}_i^{(L)} \) satisfy the following equations:

\[ P_i^{(L)} = A_{P_i}^{(L)} \tilde{P}_i^{(L)} A_i^{(L)^T} + Q + K_{P_i}^{(L)} R_i^{(L)} K_{P_i}^{(L)^T} \]

\[ \tilde{P}_i^{(L)} = A_{P_i}^{(L)} \tilde{P}_i^{(L)} A_i^{(L)^T} + Q + K_{P_i}^{(L)} R_i^{(L)} K_{P_i}^{(L)^T}. \]

Then, one can obtain

\[ \tilde{P}_i^{(L)} - P_i^{(L)} = A_{P_i}^{(L)} (\tilde{P}_i^{(L)} - P_i^{(L)}) A_i^{(L)^T} + K_{P_i}^{(L)} (R_i^{(L)} - R_i^{(L)}) K_{P_i}^{(L)^T} \]

\[ \text{and} \]

\[ \tilde{P}_i^{(L)} - P_i^{(L)} = \sum_{k=0}^{\infty} \hat{A}_{P_i}^k L^L \left( R_i^{(L)} - R_i^{(L)} \right) K_{P_i}^{(L)^T} (\hat{A}_{P_i}^k L^L)^k. \]

(F.1)

Using \( K_{P_i}^{(L)} = A_{P_i}^{(L)} C_i^{(L)^T} R_i^{(L)^{-1}} \), one can rewrite (F.1) as

\[ \tilde{P}_i^{(L)} - P_i^{(L)} = \sum_{k=0}^{\infty} \hat{A}_{P_i}^k A_{P_i}^{(L)} C_i^{(L)^T} (R_i^{(L)^{-1}})^{-1} \times (R_i^{(L)} - R_i^{(L)}) (R_i^{(L)^{-1}})^{-1} C_i^{(L)^T} (\hat{A}_{P_i}^k L^L)^k. \]
Thus, it follows that
\[
\left\| \hat{P}_i^{(L)} - P_i^{(L)} \right\|_2 \leq \left\| (R_i^{(L)})^{-1} (R - R_i^{(L)}) (R_i^{(L)})^{-1} \right\|_2 \\
\times \left\| \hat{P}_i^{(L)} \right\|_2 \left\| C \right\|_2 \left\| A \right\|_2^2 \sum_{k=0}^{\infty} \left\| \hat{A}_{k}^{(L)} \right\|_2^2.
\]

Similarly to the proof in Theorem 2, one can find a uniform bound as \( \sum_{k=0}^{\infty} \left\| \hat{A}_{k}^{(L)} \right\|_2^2 \leq M_2, \ \forall i \in \mathcal{V}, \forall L \geq d. \) Thus, it suffices to discuss the term \((R_i^{(L)})^{-1} (R - R_i^{(L)}) (R_i^{(L)})^{-1}\) to show the convergence with increasing \(L\).

Both matrices \(R\) and \(R_i^{(L)}\) are block diagonal. Thus, one can obtain
\[
(R_i^{(L)})^{-1} (R - R_i^{(L)}) (R_i^{(L)})^{-1} = \text{diag}(\tilde{l}_{i_1}^{(L)}, \ldots, \tilde{l}_{i_N}^{(L)}) R^{-1},
\]
where \(\tilde{l}_{ij}^{(L)} = N^2 (l_{ij}^{(L)})^2 - N l_{ij}^{(L)}\). One can rewrite the expression of \(\tilde{l}_{ij}^{(L)}\) as
\[
\tilde{l}_{ij}^{(L)} = (N l_{ij}^{(L)} - 1)^2 + (N l_{ij}^{(L)} - 1).
\]

With Lemma 1, the term \(N l_{ij}^{(L)} - 1\) converges to 0 exponentially with increasing \(L\), i.e., there exists \(M_3 > 0\), \(0 < q_3 < 1\), such that \(\forall i, j\) one has \(\left| N l_{ij}^{(L)} - 1 \right| \leq M_3 q_3^L\).

For \(L \geq \log_{1 + M_3} M_2\), one has \(\left| N l_{ij}^{(L)} - 1 \right| < \left| N l_{ij}^{(L)} - 1 \right|\), therefore there exist two parameters \(M_2'' > 0\) and \(0 < q_2 < 1\), such that
\[
\left\| (R_i^{(L)})^{-1} (R - R_i^{(L)}) (R_i^{(L)})^{-1} \right\|_2 \leq M_2'' q_2^L
\]
for all \(i \in \mathcal{V}\) and \(L \geq d\). The theorem is proved. \(\square\)

References

[1] B. J. Silva, R. M. Fisher, A. Kumar, and G. P. Hancke, “Experimental link quality characterization of wireless sensor networks for underground monitoring,” IEEE Transactions on Industrial Informatics, vol. 11, no. 5, pp. 1099–1110, 2015.

[2] T. T. Vu and A. R. Rahmani, “Distributed consensus-based Kalman filter estimation and control of formation flying spacecraft: Simulation and validation,” AIAA Guidance, Navigation, and Control Conference 2015, MGC 2015 - Held at the AIAA SciTech Forum 2015, 2015.

[3] R. Olfati-Saber, “Distributed Kalman filtering for sensor networks,” in Proceedings of the IEEE Conference on Decision and Control, 2007, pp. 5492–5498.

[4] ——, “Kalman-consensus filter: Optimality, stability, and performance,” in Proceedings of the IEEE Conference on Decision and Control held jointly with 2009 28th Chinese Control Conference, 2009, pp. 7036–7042.

[5] F. S. Cattivelli and A. H. Sayed, “Diffusion strategies for distributed Kalman filtering and smoothing,” IEEE Transactions on Automatic Control, vol. 55, no. 9, pp. 2069–2084, 2010.

[6] 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.

[7] G. Battistelli and L. Chisci, “Kulback-leibler average consensus on probability density functions, and distributed state estimation with guaranteed stability,” Automatica, vol. 50, no. 3, pp. 707–718, 2014.

[8] 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, 2015.

[9] S. Wang and W. Ren, “On the convergence conditions of distributed dynamic state estimation using sensor networks: A unified framework,” IEEE Transactions on Control Systems Technology, vol. 26, no. 4, pp. 1300–1316, 2018.

[10] X. He, W. Xue, X. Zhang, and H. Fang, “Distributed filtering for uncertain systems under switching sensor networks and quantized communications,” Automatica, vol. 114, p. 108842, 2020.

[11] P. Duan, Z. Duan, G. Chen, and L. Shi, “Distributed state estimation for uncertain linear systems: A regularized least-squares approach,” Automatica, vol. 117, p. 109007, 2020.

[12] W. Li, Z. Wang, D. W. C. Ho, and G. Wei, “On boundedness of error covariances for Kalman consensus filtering problems,” IEEE Transactions on Automatic Control, vol. 65, no. 6, pp. 2654–2661, 2020.

[13] S. P. Talebi, S. Werner, V. Gupta, and Y.-F. Huang, “On stability and convergence of distributed filters,” IEEE Signal Processing Letters, vol. 28, pp. 494–498, 2021.

[14] M. Kamgarpour and C. J. Tomlin, “Convergence properties of a decentralized Kalman filter,” in 2008 47th IEEE Conference on Decision and Control, 2008, pp. 3205–3210.

[15] S. Battilotti, F. Cacace, and M. d’Angelo, “A stability with optimality analysis of consensus-based distributed filters for discrete-time linear systems,” Automatica, vol. 129, p. 109589, 2021.

[16] B. D. O. Anderson and J. B. Moore, Optimal filtering. Englewood Cliffs, N.J: Prentice-Hall, 1979.

[17] W. Li, G. Wei, D. W. C. Ho, and D. Ding, “A weighted uniform detectability for sensor networks,” IEEE Transactions on Neural Networks and Learning Systems, vol. 29, no. 11, pp. 5790–5796, 2018.

[18] R. A. Horn and C. R. Johnson, Matrix Analysis. New York: Cambridge: Cambridge University Press, 1985.

[19] L. Xiao and S. P. Boyd, “Fast linear iterations for distributed averaging,” Systems and Control Letters, vol. 53, no. 1, pp. 65–78, 2004.

[20] G. Battistelli, L. Chisci, and D. Selvi, “A distributed Kalman filter with event-triggered communication and guaranteed stability,” Automatica, vol. 93, pp. 75–82, 2018.

[21] B. Gharesifard and J. Cortes, “Distributed strategies for generating weight-balanced and doubly stochastic digraphs,” European Journal of Control, vol. 18, no. 6, pp. 539–557, 2012.
[22] Z. Duan, L. Huang, and Z.-P. Jiang, “On the effects of redundant control inputs in discrete-time systems,” in Proceedings of the 31st Chinese Control Conference. IEEE, 2012, pp. 165–170.

[23] X. He, W. Xue, and H. Fang, “Consistent distributed state estimation with global observability over sensor network,” Automatica, vol. 92, pp. 162–172, 2018.

[24] S. P. Talebi and S. Werner, “Distributed Kalman filtering and control through embedded average consensus information fusion,” IEEE Transactions on Automatic Control, vol. 64, no. 10, pp. 4396–4403, 2019.

[25] S. P. Talebi, S. Werner, and D. P. Mandic, “Quaternion-valued distributed filtering and control,” IEEE Transactions on Automatic Control, vol. 65, no. 10, pp. 4246–4257, 2020.

[26] J. Yi, L. Chai, and J. Zhang, “Average consensus by graph filtering: New approach, explicit convergence rate, and optimal design,” IEEE Transactions on Automatic Control, vol. 65, no. 1, pp. 191–206, 2020.

[27] D. A. Dowler, “Bounding the norm of matrix powers,” 2013.

[28] T. Kailath, A. H. Sayed, and B. Hassibi, Linear estimation. Prentice Hall, 2000.