



HAL
open science

Suboptimal Filtering over Sensor Networks with Random Communication

Aneel Tanwani

► **To cite this version:**

Aneel Tanwani. Suboptimal Filtering over Sensor Networks with Random Communication. IEEE Transactions on Automatic Control, 2022, 67 (10), pp.5456 - 5463. 10.1109/TAC.2021.3116180 . hal-03106998v3

HAL Id: hal-03106998

<https://laas.hal.science/hal-03106998v3>

Submitted on 29 Sep 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Suboptimal Filtering over Sensor Networks with Random Communication

Aneel Tanwani

Abstract—The problem of filter design is considered for continuous-time linear stochastic systems using distributed sensors. Each sensor unit, represented by a node in an undirected and connected graph, collects some information about the state and communicates its own estimate with the neighbors. It is stipulated that this communication between sensor nodes connected by an edge is time-sampled randomly and for each edge, the sampling process is an independent Poisson counter. Our proposed filtering algorithm for each sensor node is a stochastic hybrid system: It comprises a continuous-time differential equation, and at random time instants when communication takes place, each sensor node updates its state estimate based on the information received by its neighbors. In this setting, we compute the expectation of the error covariance matrix for each unit which is governed by a matrix differential equation. To study the asymptotic behavior of these covariance matrices, we show that if the gain matrices are appropriately chosen and the mean sampling rate is large enough, then the error covariances practically converge to a constant matrix.

Index Terms—Stochastic hybrid system; distributed estimation and filtering; graph theory; random communication; heterogeneous agents.

I. INTRODUCTION

Modern control systems often involve networks of nodes with data acquisition, processing and communication abilities, which has led to growing interest in the area of distributed filtering. The research in this area basically aims at computing an estimate of the state trajectory by combining information from several sources, each of which has some partial information about the state. A conceptual sketch of the distributed filtering architectures is depicted in Figure 1. In this layout, instead of using the centralized measurements y , there are N sensor units which measure different components of these centralized measurements, y_1, \dots, y_N , and each sensor unit computes an estimate of the state based on the partial measurements while exchanging information with its neighbors, which are determined by a communication graph. Several design techniques have now emerged in the literature with varying degree of analysis. For discrete-time deterministic systems, the estimation problem in distributed setting is studied in [1], [2]. For continuous-time deterministic systems, one may refer to [3], [4]. The focus in these works is on proposing estimation algorithms and provide design criteria for each unit which uses

minimal information about the centralized system dynamics. In certain papers, distributed state estimation is studied as an application of synchronization of multi-agent systems [5], [6], [7]. Distributed state estimation for deterministic continuous-time systems with communication at discrete times has been studied in [8], [9]. For stochastic systems, we see consensus based approaches for scalar systems in [10]. Distributed filtering with more general probability distributions and application to Gaussian distributions with linear discrete-time systems is studied in [11]. Prior to that, several algorithms for distributed filtering have been proposed in discrete-setting [12], [13].

An important issue from the point of implementation of filters over networks is to make the algorithms compatible with the underlying communication protocol [14]. In communicating messages over the networks, the transmission may break down at some time instants, or the messages may not be transferred at exact scheduled times. An abstract way to model such scenarios is to assume that the underlying protocols transmit messages only at *randomly* drawn discrete time instants. With this motivation, certain works in the literature have studied the problem of stabilization and control with randomly sampled measurements: The reader may refer to [15] for optimal control and the papers [16], [17] for stability analysis of such systems. A recently published book chapter [18] provides an overview of such results and some recent developments. There have been relatively fewer works which have addressed filtering problem in the presence of measurement errors or communication uncertainties. The papers [19], [20], [21] consider a discrete-time linear dynamical system and associate randomness with the transmission times of the output measurements. A different toolset, based on relative entropy, is adopted in [22] to study the stability and convergence of filters under relaxed assumptions on observation channels. For continuous-time dynamical system driven by white noise, centralized continuous-discrete estimators are studied in [23]. Performance bounds for centralized filters of continuous-time systems under random sampling appear in [24].

In this article, we consider filtering problem in distributed setting for continuous-time linear stochastic systems. Our objective is to propose filtering algorithms, and analyze their performance, when the centralized measurements are not available, and the agents can only communicate their own state estimates to their neighbors at time instants determined by a Poisson counter. In particular, we consider agents (or sensor units) which have partial information about the state at all times. This information is used to compute an estimate of the state using a differential equation. Then, at some random time instants, the agents exchange the value of their estimate

A. Tanwani is affiliated with CNRS – LAAS, University of Toulouse, CNRS, 31400 Toulouse, France. Email: aneel.tanwani@laas.fr

This work has been supported by the project CYPHAL, financed by ANR-JST CREST program, with grant number ANR-20-JSTM-0001.

This is a preprint of an article accepted for publication in IEEE Transactions on Automatic Control, October 2022 issue. DOI: 10.1109/TAC.2021.3116180.

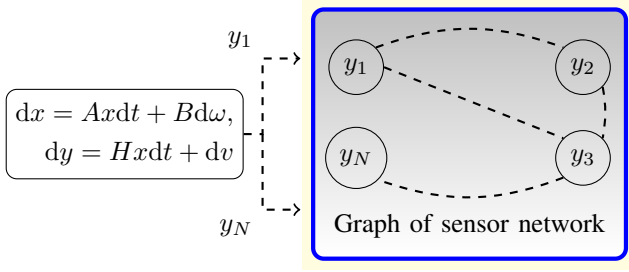


Figure 1: Layout of distributed filters, where dashed lines represent communication at random times.

with their neighboring agents and update their own state based on the information received. In addition to formalizing this intuition, we develop some tools for analysis of such filtering algorithms. This involves computing the bounds on error covariance matrices of each filter while exchanging information with other agents. Also, we are interested in studying the limiting behavior of these bounds as the formal stability analysis of distributed filters with the aforementioned communication architecture has not received much attention in the literature.

To position the contribution of this article in the existing literature on distributed estimation and filtering, we observe that, for distributed state estimation in deterministic setup [1], [2], [3], [4], [6], one directly analyzes the convergence of the norm of estimation error to the origin and the focus is mostly on design aspect. The error covariance bounds are analyzed in the stochastic setup [10], [11], [12], [13], where the problem is essentially studied in the discrete setting without any loss of information and the focus is mostly on achieving optimality via appropriate filter design. On the other hand, our focus is on analyzing the performance of distributed filters for a continuous-time process, where each filter gets continuous information from the sensors, and some updates at random discrete times from its neighbors. The derivation of the bounds on the expectation of error covariance process with Poisson sampling in distributed setup is a unique feature of our work. The elegant and useful outcome of this analysis is that we observe diffusive coupling among covariance bounds, and the mean sampling rate plays the role of coupling strength. This observation allows us to analyze the asymptotic behavior of the expected error covariance for each filtering unit. However, unlike [12], [13], the asymptotic bounds quantifying the performance of the distributed filters are obtained in suboptimal setting through a *simplistic* choice of injection gains and the update rule for information fusion is not necessarily optimal.

Notation: Let us recall some basic definitions from the literature which are used in the statement and derivation of main results. A function $\chi : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ is said to be of class \mathcal{K} if it is continuous, increasing, and satisfies $\chi(0) = 0$. For the matrices M_1 and M_2 , we define the inner product as $\langle M_1, M_2 \rangle = \text{tr}(M_1^T M_2)$. This choice of inner product allows us to introduce the Frobenius norm, defined as $\|M\| = \sqrt{\langle M, M \rangle} = \sqrt{\text{tr}(M^T M)}$. Throughout this article, we will work with this particular norm for matrices. We recall

that the Cauchy-Schwarz inequality for this choice of norm implies that $|\text{tr}(M_1 M_2)| \leq \|M_1\| \|M_2\|$. For a square matrix $M \in \mathbb{R}^{n \times n}$, the notation $\text{Sym}(M)$ stands for $M + M^T$, and $\sigma_i(M)$ denotes the i -th eigenvalue of M with the convention that $\sigma_1(M) \leq \sigma_2(M) \leq \dots \leq \sigma_n(M)$. The symbol $o(s)$ for $s \rightarrow 0$ denotes a function of $s \in \mathbb{R}$ satisfying $\lim_{s \rightarrow 0} \frac{o(s)}{s} = 0$, and the symbol $O(s)$ for $s \rightarrow 0$ stands for a function of $s \in \mathbb{R}$ satisfying $\lim_{s \rightarrow 0} \frac{O(s)}{s}$ exists and is finite.

II. PROBLEM FORMULATION

Consider the dynamical system

$$dx = Ax dt + B d\omega \quad (1)$$

where $(x(t))_{t \geq 0}$ is an \mathbb{R}^n -valued diffusion process describing the state. Let $(\Omega, \mathcal{F}, \mathbb{P})$ denote the underlying probability space. It is assumed that, for each $t \geq 0$, $(\omega(t))_{t \geq 0}$ is an \mathbb{R}^m -valued standard Wiener process adapted to the filtration $\mathcal{F}_t \subset \mathcal{F}$, with the property that $\mathbb{E}[d\omega(t) d\omega(t)^T] = I_m dt$, for each $t \geq 0$. The matrices $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times m}$ are taken as constant, and the process $(\omega(t))_{t \geq 0}$ does not depend on the state. The solutions of the stochastic differential equation (1) are interpreted in the sense of Itô stochastic integral.

a) Measurements via distributed sensors: The measurements associated with system (1) are obtained from a set of N sensors which are distributed in their localization. Each of these sensors provides a partial measurement about the state described as,

$$dy_i = H_i x dt + dv_i, \quad i = 1, \dots, N, \quad (2)$$

where $H_i \in \mathbb{R}^{p_i \times n}$, and $\sum_{i=1}^N p_i = p$. That is, for each node, $(y_i(t))_{t \geq 0}$ describes an \mathbb{R}^{p_i} -valued continuous-time observation process. In the observation equation (2), $v_i(t)$ is an \mathcal{F}_t -adapted standard Wiener process, taking values in \mathbb{R}^{p_i} , and $\mathbb{E}[dv_i(t) dv_i(t)^T] = V_i dt$, with $V_i \in \mathbb{R}^{p_i \times p_i}$ assumed to be positive definite.

The sensor nodes are connected via a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, \dots, N\}$ is the set of graph nodes, and \mathcal{E} contains all the edges defined by a subset of the pairs (i, j) , $i \neq j$, $i, j \in \mathcal{V}$. We assume that the graph is undirected and connected. The neighbors of a node $i \in \mathcal{V}$ are denoted by $\mathcal{N}_i := \{j \in \mathcal{V} \mid (i, j) \in \mathcal{E}\}$, and we adopt the convention that $i \notin \mathcal{N}_i$. The adjacency matrix $\mathcal{A} := [a_{ij}] \in \{0, 1\}^{N \times N}$ of the graph, which is symmetric, provides the information about which sensor nodes can communicate with each other, that is, if $a_{ij} = 1$ then sensor i and j can communicate, whereas $a_{ij} = 0$ means there is no communication possible between those sensors. The degree of a node $i \in \mathcal{V}$ is defined as $|\mathcal{N}_i|$, that is, the cardinality of the set \mathcal{N}_i . The diagonal matrix $\mathcal{D} = [d_{ii}]$, with $d_{ii} = |\mathcal{N}_i|$ is therefore the degree matrix. We associate a Laplacian \mathcal{L} with this graph, defined as, $\mathcal{L} = \mathcal{D} - \mathcal{A}$. For our purposes, the matrix $\Pi = [\pi_{ij}] \in \mathbb{R}^{N \times N}$, defined as

$$\Pi := I_{N \times N} - \alpha \mathcal{L} \quad (3)$$

with $0 < \alpha \leq \min_{i \in \mathcal{V}} \frac{1}{|\mathcal{N}_i|}$, plays an important role. Note that, by construction, Π is a doubly stochastic matrix, that is,

for each row and each column, the sum of their entries equals one.

b) *Communication Process*: The next main ingredient of our problem formulation is the description of the time instants at which the communication takes place between two sensor nodes connected by an edge. Corresponding to each edge $(i, j) \in \mathcal{E}$, it is stipulated that there is an increasing and divergent sequence $(\tau_k^{ij})_{k \in \mathbb{N}} \subset [0, +\infty[$ with $\tau_0^{ij} := 0$, and the sensor nodes $i, j \in \mathcal{V}$ transmit the value of their state estimate to each other at τ_k^{ij} , $k \in \mathbb{N}$.

In this article, we are interested in the case where the sampling times $(\tau_k^{ij})_{k \in \mathbb{N}}$ are generated *randomly*. Formally, we define

$$N_t^{ij} := \sup\{k \in \mathbb{N} \mid \tau_k^{ij} \leq t\} \quad \text{for } t \geq 0. \quad (4)$$

The map $t \mapsto N_t^{ij}$ increments by 1 at random times, and it provides a description of the number of times the nodes $i, j \in \mathcal{V}$ communicate with each other up to and including time t . Let $\lambda > 0$. The communication times are such that $\tau_{N_t^{ij}}^{ij} \rightarrow +\infty$ almost surely as $t \nearrow +\infty$. Recall that a random process $(N_t^{ij})_{t \geq 0}$ is a *Poisson process of intensity* λ , defined on our probability space, if any one of the following equivalent properties hold:

(P1) $(N_t^{ij})_{t \geq 0}$ is a Markov process taking values in \mathbb{N} , has independent increments, and satisfies $N_0 = 0$, and for $h \searrow 0$ and $t \geq 0$,

$$\mathbb{P}(N_{t+h}^{ij} - N_t^{ij} = k \mid N_t^{ij}) = \begin{cases} 1 - \lambda h + o(h) & \text{if } k = 0, \\ \lambda h + o(h) & \text{if } k = 1, \\ o(h) & \text{if } k \geq 2, \end{cases}$$

where the terms $o(h)$ do not depend on t .

(P2) $(N_t^{ij})_{t \geq 0}$ is a continuous-time random process taking values in \mathbb{N} , having monotone non-decreasing sample paths with increments of 1, and spending an exponentially distributed, with parameter λ , random time in each state k before jumping to $k + 1$ independently for all k .

The communication process between any two neighbors is described by the following hypothesis:

- for each $(i, j) \in \mathcal{E}$, $(N_t^{ij})_{t \geq 0}$ is an independent Poisson process of intensity $\lambda > 0$.

Next, we associate with each node $i \in \mathcal{V}$, the process N_t^i ,

$$N_t^i := \sum_{j \in \mathcal{N}_i} N_t^{ij}$$

so that N_t^i increments by one whenever node $i \in \mathcal{V}$ exchanges information with one of its neighbors. We recall that N_t^i is also a Poisson process. The times at which N_t^i gets incremented are denoted by $\tau_{N_t^i}^i$. We can now introduce the activation set \mathcal{A}_t^i , for $t \geq 0$, which is defined as,

$$\mathcal{A}_t^i := \left\{ j \in \mathcal{N}_i \mid N_t^{ij} - N_{\underline{t}}^{ij} \neq 0; \underline{t} = \tau_{N_{\underline{t}}^i}^i \right\},$$

that is, at times $\tau_{N_t^i}^i$, the set \mathcal{A}_t^i describes the neighbors of agent $i \in \mathcal{V}$ that communicate with it.

c) *Filtering Algorithm*: Based on the aforementioned communication architecture, and the sensor localization, we propose the following algorithm, to be implemented by each sensor node:

$$d\hat{x}_i = (A - L_i H_i) \hat{x}_i dt + L_i dy_i \quad (5)$$

over the interval $[\tau_{N_t^i}^i, \tau_{N_{t+1}^i}^i[$, for each $t \in [0, +\infty[$. Because of the arrival of new information at random times, the estimate \hat{x}_i , $i \in \mathcal{V}$, gets updated at such times, that is, for each node $i \in \mathcal{V}$, we update \hat{x}_i at $t_c = \tau_{N_{t_c}^i}^i$ as follows:

$$\hat{x}_i(t_c^+) = \sum_{j \in \mathcal{A}_{t_c}^i} \pi_{ij} \hat{x}_j(t_c^-) + \left(1 - \sum_{j \in \mathcal{A}_{t_c}^i} \pi_{ij}\right) \hat{x}_i(t_c^-), \quad (6)$$

where π_{ij} are the elements of the matrix Π introduced in (3). Each of these filters is a stochastic hybrid system of the form [25] with continuous evolution described by (5) between sampling times, and the jump rule (6) executed at random sampling instants where we update the estimate \hat{x}_i as the convex combination of itself and its active neighbors. The communication based on the underlying graph results in a particular interconnection of these hybrid systems.

Remark II.1. In our filtering algorithm (5), (6), we are choosing the gains L_i to be constant. In optimal filtering, such as Kalman filtering, one uses the information of the evolution of the covariance matrices to update the gains. Our choice is clearly suboptimal but for this paper, we focus only on constant gains to keep the presentation simple.

d) *Problem statement*: The basic problem studied in this paper is the performance of the distributed filters proposed in (5)–(6) when the communication across each edge is described by an independent Poisson counter. The design parameters are the gain matrices L_i , and we want to study the role of mean sampling rate λ in quantifying the performance of the system. More precisely, if we let \mathcal{Y}_t^i denote the information available to sensor node $i \in \mathcal{V}$ up till time $t \in [0, +\infty[$, that is, $\mathcal{Y}_t^i := \{(dy_i(s), \hat{x}_j(\tau_{N_s^j}^j)) \mid s \leq t, j \in \mathcal{N}_i\}$, then our objectives are:

- Compute an upper bound on the expectation (with respect to sampling process) of the error covariance matrices $\mathbb{E}[\mathbb{E}[(x(t) - \hat{x}_i(t))(x(t) - \hat{x}_i(t))^T \mid \mathcal{Y}_t^i]]$, for $t \geq 0$.
- Provide a criteria for boundedness of the expected error covariance in terms of the gain matrices, and the mean sampling rate.

III. MAIN RESULT

We now address the objectives outlined in the previous section. To do so, we consider the injection gains $L_i \in \mathbb{R}^{n \times p_i}$, and introduce the matrices $L \in \mathbb{R}^{n \times p}$ and $H \in \mathbb{R}^{p \times n}$ as follows:

$$L := \frac{1}{N} [L_1 \mid \dots \mid L_N], \quad H = \begin{bmatrix} H_1 \\ \vdots \\ H_N \end{bmatrix}. \quad (7)$$

Consequently, we see that $A - LH = A - \frac{1}{N} \sum_{i=1}^N L_i H_i$.

Theorem III.1. Consider system (1) with distributed measurements (2) and the corresponding hybrid filters (5), (6) coupled by an undirected and connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Suppose that the communication across an edge $(i, j) \in \mathcal{E}$ takes place at random times described by an independent Poisson process $(N_t^{ij})_{t \geq 0}$ of intensity $\lambda > 0$, then the following items hold:

1) For each $\lambda > 0$, and $i = 1, \dots, N$, we have

$$\mathbb{E}[\mathbb{E}[(x(t) - \hat{x}_i(t))(x(t) - \hat{x}_i(t))^\top | \mathcal{Y}_t^i] \leq \mathcal{P}_i(t), \quad (8)$$

where the matrix-valued function $\mathcal{P}_i : [0, \infty[\rightarrow \mathbb{R}^{n \times n}$ satisfies the differential equation

$$\begin{aligned} \dot{\mathcal{P}}_i &= (A - L_i H_i) \mathcal{P}_i + \mathcal{P}_i (A - L_i H_i)^\top + B B^\top + L_i V_i L_i^\top \\ &\quad + \lambda \sum_{j \in \mathcal{N}_i} \pi_{ij} (\mathcal{P}_j - \mathcal{P}_i). \end{aligned} \quad (9)$$

2) If the gains L_i are chosen such that the matrix $(A - LH)$ is Hurwitz, which in turn implies that there exist a symmetric positive definite matrix R and a symmetric positive semidefinite matrix S satisfying

$$R(A - LH) + (A - LH)^\top R \leq -I_n \quad (10)$$

$$0 = (A - LH)S + S(A - LH)^\top + B B^\top + \frac{1}{N} \sum_{i=1}^N L_i V_i L_i^\top \quad (11)$$

then for every $\lambda > 0$ sufficiently large, the corresponding solution of (9) satisfies

$$\begin{aligned} \limsup_{t \rightarrow \infty} \|\mathcal{P}_i(t) - S\| &\leq \\ \sqrt{N} \frac{\sigma_n(R)}{\sigma_1(R)} \sqrt{\chi \left(\frac{1}{\lambda \alpha \sigma_2(\mathcal{L}) - C_1} \right)} &\quad (C_3 \bar{\sigma}_G + C_2 \|S\|), \end{aligned} \quad (12)$$

for some nonnegative constants C_1, C_2, C_3 , while $\bar{\sigma}_G := \max_{1 \leq i \leq N} \|B B^\top + L_i V_i L_i^\top\|$, and χ is a class \mathcal{K} function satisfying $\chi(s) = O(s)$ as $s \rightarrow 0$.

The proof of Theorem III.1 involves several intermediate results, and is carried out in the remainder of this section within various subsections.

A. Analysis for fixed sampling times

The first step in the proof of Theorem III.1 is to obtain a bound on the covariance of estimation error resulting from (5), (6) for a given value of sampling times. To do so, it is useful to introduce the dynamics for the state estimation error, $e_i := x - \hat{x}_i$. For $i \in \mathcal{V}$, over an interval $[\tau_{N_i}^i, \tau_{N_i+1}^i[$, we have the differential equation:

$$de_i = (A - L_i H_i) de_i + L_i dv_i + B d\omega \quad (13a)$$

and at communication times $t_c = \tau_{N_i}^i$, it holds that

$$e_i(t_c^+) = \left(1 - \sum_{j \in \mathcal{A}_{t_c}^i} \pi_{ij}\right) e_i(t_c^-) + \sum_{j \in \mathcal{A}_{t_c}^i} \pi_{ij} e_j(t_c^-). \quad (13b)$$

Error covariance bound on estimation error is described in the following statement.

Proposition III.2. Consider system (1) and the filter (5), (6). For a node $i \in \mathcal{V}$, consider the sequence $\{\tau_k^i\}_{k \in \mathbb{N}}$ to be fixed. Let the process $P_i : [0, \infty[\rightarrow \mathbb{R}^{n \times n}$, with $P_i(0) \geq \mathbb{E}[(x_0 - \hat{x}_i(0))(x_0 - \hat{x}_i(0))^\top]$, be defined by

$$\frac{dP_i}{dt} = (A - L_i H_i) P_i + P_i (A - L_i H_i)^\top + B B^\top + L_i V_i L_i^\top, \quad (14a)$$

for $\tau_k^i \leq t < \tau_{k+1}^i$, for each $k \in \mathbb{N}$, and at $t_c = \tau_k^i$, let

$$P_i(t_c^+) := \left(1 - \sum_{j \in \mathcal{A}_{t_c}^i} \pi_{ij}\right) P_i(t_c^-) + \sum_{j \in \mathcal{A}_{t_c}^i} \pi_{ij} P_j(t_c^-). \quad (14b)$$

Then, it holds that, for each $t \geq 0$,

$$\mathbb{E}[\mathbb{E}[(x(t) - \hat{x}_i(t))(x(t) - \hat{x}_i(t))^\top | \mathcal{Y}_t^i] \leq P_i(t). \quad (15)$$

Proof. For the sake of conciseness, we will denote the expectation on the left-hand side of (15) by $\mathbb{E}[e_i e_i^\top]$. We first show that, for each $t \in [\tau_k^i, \tau_{k+1}^i[$, $k \in \mathbb{N}$,

$$\begin{aligned} \frac{d\mathbb{E}[e_i(t) e_i(t)^\top]}{dt} &= (A - L_i H_i) \mathbb{E}[e_i(t) e_i(t)^\top] \\ &\quad + \mathbb{E}[e_i(t) e_i(t)^\top] (A - L_i H_i)^\top + B B^\top + L_i V_i L_i^\top \end{aligned}$$

and that, at the communication instants, $t_c = \tau_k^i$,

$$\begin{aligned} \mathbb{E}[e_i(t_c^+) e_i(t_c^+)^\top] &\leq \left(1 - \sum_{j \in \mathcal{A}_{t_c}^i} \pi_{ij}\right) \mathbb{E}[e_i(t_c^-) e_i(t_c^-)^\top] \\ &\quad + \sum_{j \in \mathcal{A}_{t_c}^i} \pi_{ij} \mathbb{E}[e_j(t_c^-) e_j(t_c^-)^\top]. \end{aligned}$$

The bound in (15) then holds by observing that $P_i(t)$ satisfies (14a) over $[\tau_k^i, \tau_{k+1}^i[$, and (14b) at $t = \tau_k^i$, with $\mathbb{E}[e_i(0) e_i(0)^\top] \leq P_i(0)$.

a) *Continuous part:* We first show that for $t \in [\tau_k^i, \tau_{k+1}^i[$, we have the equality $\frac{d\mathbb{E}[e_i(t) e_i(t)^\top]}{dt} = \frac{dP_i(t)}{dt}$. This is a rather classical derivation, and as an example, one may refer to the arguments in [24] for details. On this time interval, let us consider the Itô process (13), and the function $v : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ given by, $\mathbb{R}^n \ni e \mapsto v(e) = e e^\top \in \mathbb{R}^{n \times n}$. Consequently, $(v \circ e_i(t))_{t \geq 0}$ is an Itô process. By applying Itô's differential chain rule, and following the calculations as in [24, Section 3], we obtain

$$\begin{aligned} \frac{d\mathbb{E}[v(e_i(t))]}{dt} &= (A - L_i H_i) \mathbb{E}[e_i(t) e_i(t)^\top] + \\ &\quad \mathbb{E}[e_i(t) e_i(t)^\top] (A - L_i H_i)^\top + B B^\top + L_i V_i L_i^\top. \end{aligned}$$

b) *Jump part:* Computing the bound on error covariance at sampling times basically requires us to compute bounds on covariance of a random variable described by the convex combination of different random variables. This is based on the following lemma:

Lemma III.3. Let m be a positive integer, and let $x_1, \dots, x_m \in \mathbb{R}^n$. If $z := \sum_{j=1}^m \gamma_j x_j$ for some $\gamma_j \in [0, 1]$, $\sum_{j=1}^m \gamma_j = 1$, then

$$zz^\top \leq \sum_{j=1}^m \gamma_j x_j x_j^\top. \quad (16)$$

Proof of Lemma III.3. Let \mathcal{S}_+^n denote the set of positive semi-definite matrices in $\mathbb{R}^{n \times n}$, and consider the function $f : \mathbb{R}^n \rightarrow \mathcal{S}_+^n$, so that, for $x \in \mathbb{R}^n$, $f(x) = xx^\top$. We consider the epigraph of f , denoted by $\text{epi}(f)$, and defined as,

$$\text{epi}(f) = \{(M, x) \in \mathcal{S}_+^n \times \mathbb{R}^n \mid M \geq f(x)\}.$$

We claim that $\text{epi}(f) \subset \mathcal{S}_+^n \times \mathbb{R}^n$ is convex, that is, if (M_1, \bar{x}_1) and (M_2, \bar{x}_2) belong to $\text{epi}(f)$, then for each $\gamma \in [0, 1]$, it holds that $(\gamma M_1 + (1-\gamma)M_2, \gamma \bar{x}_1 + (1-\gamma)\bar{x}_2) \in \text{epi}(f)$. To see this, we first observe that, for a fixed $\gamma \in [0, 1]$,

$$\begin{aligned} & \left(\sqrt{\frac{1-\gamma}{\gamma}} \gamma \bar{x}_1 - \sqrt{\frac{\gamma}{1-\gamma}} (1-\gamma) \bar{x}_2 \right) \\ & \left(\sqrt{\frac{1-\gamma}{\gamma}} \gamma \bar{x}_1 - \sqrt{\frac{\gamma}{1-\gamma}} (1-\gamma) \bar{x}_2 \right)^\top \geq 0 \end{aligned}$$

and hence

$$\gamma(1-\gamma)(\bar{x}_1 \bar{x}_2^\top + \bar{x}_2 \bar{x}_1^\top) \leq \frac{1-\gamma}{\gamma} \gamma^2 \bar{x}_1 \bar{x}_1^\top + \frac{\gamma}{1-\gamma} (1-\gamma)^2 \bar{x}_2 \bar{x}_2^\top.$$

Using this last inequality, we get

$$\begin{aligned} f(\gamma \bar{x}_1 + (1-\gamma)\bar{x}_2) &= \gamma^2 \bar{x}_1 \bar{x}_1^\top + \gamma(1-\gamma)(\bar{x}_1 \bar{x}_2^\top + \bar{x}_2 \bar{x}_1^\top) \\ & \quad + (1-\gamma)^2 \bar{x}_2 \bar{x}_2^\top \\ &\leq \gamma^2 \left(1 + \frac{1-\gamma}{\gamma} \right) \bar{x}_1 \bar{x}_1^\top + (1-\gamma)^2 \left(1 + \frac{\gamma}{1-\gamma} \right) \bar{x}_2 \bar{x}_2^\top, \\ &= \gamma \bar{x}_1 \bar{x}_1^\top + (1-\gamma) \bar{x}_2 \bar{x}_2^\top \\ &\leq \gamma M_1 + (1-\gamma) M_2, \end{aligned}$$

which implies that $\text{epi}(f)$ is convex. Coming back to the statement of the lemma, we observe that, for each $j = 1, \dots, m$, $(f(x_j), x_j) \in \text{epi}(f)$ and since we have shown that this set is convex, it follows that

$$\sum_{j=1}^m \gamma_j (f(x_j), x_j) = \left(\sum_{j=1}^m \gamma_j f(x_j), \sum_{j=1}^m \gamma_j x_j \right)$$

also belongs to $\text{epi}(f)$. By construction, we therefore have

$$f \left(\sum_{j=1}^m \gamma_j x_j \right) \leq \sum_{j=1}^m \gamma_j f(x_j)$$

which yields the desired inequality by recalling the definition of f . \square

At communication times, $t_c = \tau_k^i$, the statement of Lemma III.3 directly yields an upper bound on $\mathbb{E}[e_i(t_c^+) e_i(t_c^+)^\top]$ by recalling the definition of $e(t_c^+)$ in (13b), and using the linearity of the expectation operator. The proof of Proposition III.2 is therefore complete. \square

B. Expectation of error covariance with respect to sampling process

In Proposition III.2, we computed a bound on the error covariance for a fixed sequence of time instants at which measurements are received by node $i \in \mathcal{V}$. That is, $P_i(t)$ is an upper bound on $\mathbb{E}[(x(t) - \hat{x}_i(t))(x(t) - \hat{x}_i(t))^\top \mid \mathcal{Y}_t^i]$ along one particular sample path as the realization of $P_i(\cdot)$ depends on the observed sampling times. It is thus of interest to compute the expected value of $P_i(\cdot)$ along all possible sample paths generated by the communication processes $(N_t^{ij})_{t \geq 0}$, $j \in \mathcal{N}_i$. We do so next under the assumption that $(N_t^{ij})_{t \geq 0}$ is a Poisson process of intensity λ .

To compute the expected value of P_i with respect to sampling times, we introduce the operator $\mathcal{L}_{\text{inf}} : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ defined as

$$\mathcal{L}_{\text{inf}}(Q) := \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \left(\mathbb{E}[P_i(t+\varepsilon) \mid P_i(t) = Q] - Q \right). \quad (17)$$

We will see that this operator describes infinitesimal evolution of the expected value of the error covariance process and hence, (9) and the proof of first item in Theorem III.1, are a direct consequence of the following result.

Proposition III.4. Consider the process $(P_i(t))_{t \geq 0}$ given by (14) with the processes $(N_t^{ij})_{t \geq 0}$ of intensity $\lambda > 0$, $j \in \mathcal{N}_i$. Let $\mathcal{P}_i(t) := \mathbb{E}[P_i(t) \mid P_0]$ denote the expected value at time $t \geq 0$ with respect to distribution associated with the process $N_t^i := \sum_{j \in \mathcal{N}_i} N_t^{ij}$. Then, (9) holds.

Proof. We first note that $(P_i(t))_{t \geq 0}$ is Markovian because the future of $P_i(t)$ depends only on the last sampling instant $\tau_{N_t^i}^i$. According to Dynkin's formula,

$$\mathcal{P}_i(t) = \mathbb{E}[P_i(t)] = P_i(0) + \mathbb{E} \left[\int_0^t \mathcal{L}_{\text{inf}}(P_i(s)) ds \right].$$

Since \mathcal{L}_{inf} is a linear operator, we get $\dot{\mathcal{P}}_i(t) = \mathcal{L}_{\text{inf}}(\mathcal{P}_i(t))$. To show (9), we thus compute $\mathcal{L}_{\text{inf}}(Q)$ for a given matrix Q in the remainder of the proof. Recalling the definition of \mathcal{L}_{inf} , we observe that

$$\begin{aligned} \mathbb{E}[P_i(t+\varepsilon) \mid P_i(t) = Q] &= \mathbb{E}[(P_i(t+\varepsilon)) (1_{\{N_{t+\varepsilon}^i = N_t^i\}} \\ & \quad + 1_{\{N_{t+\varepsilon}^i = 1 + N_t^i\}} + 1_{\{N_{t+\varepsilon}^i - N_t^i \geq 2\}}) \mid P_i(t)]. \end{aligned} \quad (18)$$

We now compute the conditional probability distribution of $(P_i(t+\varepsilon))$ for small $\varepsilon > 0$ given $(P_i(t))$. We note that N_t^i is a Poisson process of intensity $d_{ii}\lambda$, where d_{ii} is the degree of node $i \in \mathcal{V}$. Since the communication times are independent of the process $(P_i(t))_{t \geq 0}$, by definition of the Poisson process we have, for $\varepsilon \downarrow 0$, $\mathbb{P}(N_{t+\varepsilon}^i - N_t^i = 0 \mid N_t^i, P_i(t)) = 1 - d_{ii}\lambda\varepsilon + o(\varepsilon)$, $\mathbb{P}(N_{t+\varepsilon}^i - N_t^i = 1 \mid N_t^i, P_i(t)) = d_{ii}\lambda\varepsilon + o(\varepsilon)$, $\mathbb{P}(N_{t+\varepsilon}^i - N_t^i \geq 2 \mid N_t^i, P_i(t)) = o(\varepsilon)$. Using these expressions we develop (18) for $\varepsilon \downarrow 0$ as

$$\begin{aligned} & \mathbb{E}[P_i(t+\varepsilon) \mid P_i(t) = Q] \\ &= \mathbb{E}[P_i(t+\varepsilon) (1_{\{N_{t+\varepsilon}^i = N_t^i\}} + 1_{\{N_{t+\varepsilon}^i = 1 + N_t^i\}}) \mid P_i(t)] + o(\varepsilon) \\ &= \mathbb{E}[P_i(t+\varepsilon) \mid P_i(t), N_{t+\varepsilon}^i = N_t^i] \cdot (1 - d_{ii}\lambda\varepsilon + o(\varepsilon)) \\ & \quad + \mathbb{E}[P_i(t+\varepsilon) \mid P_i(t), N_{t+\varepsilon}^i = 1 + N_t^i] (d_{ii}\lambda\varepsilon) + o(\varepsilon). \end{aligned} \quad (19)$$

The two significant terms on the right-hand side of (19) are now computed separately. For the event $N_{t+\varepsilon}^i = N_t^i$, given $P_i(t) = Q$, \dot{P}_i is governed by (14a), so we have for $\varepsilon \downarrow 0$,

$$P_i(t + \varepsilon) = P_i(t) + \varepsilon \dot{P}_i(t) + o(\varepsilon) = Q + \varepsilon [(A - L_i H_i)Q + Q(A - L_i H_i)^\top + BB^\top + L_i V_i L_i^\top] + o(\varepsilon),$$

leading to the first term on the right-hand side of (19) having the estimate

$$\begin{aligned} & \mathbb{E}[P_i(t + \varepsilon) \mid P_i(t), N_{t+\varepsilon}^i = N_t^i] \cdot (1 - d_{ii}\lambda\varepsilon + o(\varepsilon)) \\ &= Q + \varepsilon ((A - L_i H_i)Q + Q(A - L_i H_i)^\top + BB^\top + L_i V_i L_i^\top) \\ & \quad - (d_{ii}\lambda\varepsilon)Q + o(\varepsilon), \end{aligned} \quad (20)$$

for $\varepsilon \downarrow 0$. Concerning the second term on the right-hand side of (19), we observe that $\mathbb{P}(N_{t+\varepsilon}^i = 1 + N_t^i \mid N_t^i)$ is equivalent to

$$\sum_{j \in \mathcal{N}_i} \mathbb{P}(N_{t+\varepsilon}^{ij} - N_t^{ij} = 1 \mid N_t^{ij}) \prod_{k \neq j} \mathbb{P}(N_{t+\varepsilon}^{ik} - N_t^{ik} = 0 \mid N_t^{ik}).$$

Moreover, conditional on $N_{t+\varepsilon}^{ij} = 1 + N_t^{ij}$, $j \in \mathcal{N}_i$, the probability distribution of $\tau_{N_{t+\varepsilon}^{ij}}^{ij}$ is uniform over $[t, t + \varepsilon]$ by definition of the Poisson process [26, Theorem 2.3.7]. We introduce $\theta \in [0, 1[$ such that $\tau_{N_{t+\varepsilon}^{ij}}^{ij} = t + \theta\varepsilon$; then θ is uniformly distributed on $[0, 1[$ given $N_{t+\varepsilon}^{ij} = 1 + N_t^{ij}$. We now write the right-hand side of (14a) more compactly using the map $\bar{\psi}$,

$$Q \mapsto \bar{\psi}(Q) = (A - L_i H_i)Q + Q(A - L_i H_i)^\top + BB^\top + L_i V_i L_i^\top$$

and (14b) using the map ψ_t^j , $Q \mapsto \psi_t^j(Q) := (1 - \pi_{ij})Q + \pi_{ij}P_j(t)$. We thus have, conditioned on the event $N_{t+\varepsilon}^{ij} = 1 + N_t^{ij}$, $P_i(t) = Q$,

$$P_i(\tau_{N_{t+\varepsilon}^{ij}}^{ij}) = P_i(t + \theta\varepsilon) = \psi_t^j(P_i(t + \theta\varepsilon)).$$

The above expressions then lead to, conditioned on the same event, and for $\varepsilon \downarrow 0$,

$$\begin{aligned} P_i(t + \varepsilon) &= P_i(t + \theta\varepsilon) + (1 - \theta)\varepsilon \dot{P}_i(t + \theta\varepsilon) + o(\varepsilon) \\ &= P_i(t + \theta\varepsilon) + (1 - \theta)\varepsilon \bar{\psi}(P_i(t + \theta\varepsilon)) + o(\varepsilon) \\ &= \psi_{t+\theta\varepsilon}^j(P_i(t) + \theta\varepsilon \bar{\psi}(P_i(t)) + o(\varepsilon)) \\ & \quad + (1 - \theta)\varepsilon \bar{\psi}(P_i(t + \theta\varepsilon)) + o(\varepsilon) \\ &= \psi_t^j(P_i(t)) + \theta O(\varepsilon) + O(\varepsilon) + o(\varepsilon). \end{aligned}$$

One can repeat the above procedure for each $j \in \mathcal{N}_i$, so that, for $\varepsilon \downarrow 0$,

$$\begin{aligned} & \mathbb{E}[P_i(t + \varepsilon) \mid P_i(t) = Q, N_{t+\varepsilon}^i = 1 + N_t^i] \mathbb{P}(N_{t+\varepsilon}^i = 1 + N_t^i) \\ &= \sum_{j \in \mathcal{N}_i} \int_0^1 \left(\mathbb{E}[\psi_t^j(Q)] + \theta O(\varepsilon) + O(\varepsilon) + o(\varepsilon) \right) d\theta \cdot (\lambda\varepsilon) \\ &= \sum_{j \in \mathcal{N}_i} \left(\mathbb{E}[\psi_t^j(Q)] + O(\varepsilon) \right) \cdot (\lambda\varepsilon) \\ &= \sum_{j \in \mathcal{N}_i} (\lambda\varepsilon) \mathbb{E}[\psi_t^j(Q)] + o(\varepsilon). \end{aligned} \quad (21)$$

Substituting (20) and (21) in (19), we obtain

$$\begin{aligned} & \mathbb{E}[P_i(t + \varepsilon) \mid P_i(t) = Q] = Q - (\lambda\varepsilon) \sum_{j \in \mathcal{N}_i} (Q - \mathbb{E}[\psi_t^j(Q)]) + o(\varepsilon) \\ & \quad + \varepsilon \left((A - L_i H_i)Q + Q(A - L_i H_i)^\top + BB^\top + L_i V_i L_i^\top \right). \end{aligned}$$

Substituting these expressions in (17), we see that for each $Q \in \mathbb{R}^{n \times n}$,

$$\begin{aligned} \mathcal{L}_{\text{inf}}(Q) &= (A - L_i H_i)Q + Q(A - L_i H_i)^\top + BB^\top + L_i V_i L_i^\top \\ & \quad + \lambda (\mathbb{E}[\psi_t^j(Q)] - Q), \end{aligned}$$

which, upon recalling the definition of ψ_t^j , leads to the desired expression in (9). \square

C. Asymptotic behavior

The second item in Theorem III.1 relates to the asymptotic behavior of the coupled differential equations (9). Each of these equations is linear in the state, driven by a constant term which corresponds to the noise level in the system and sensors, and are interconnected by diffusive coupling which corresponds to the mean sampling rate λ . Such systems in the literature are studied under the framework of heterogenous multi-agent systems since the dynamics of \mathcal{P}_i are different for each $i \in \mathcal{V}$. In contrast to homogenous agents, consensus in heterogenous agents is not possible in general. However, one can get the states of all the agents close to desired accuracy by increasing the coupling strength λ .

For the proof of item 2) in Theorem III.1, we introduce the notation,

$$A_i := A - L_i H_i, \quad \text{and} \quad \bar{A} := A - LH = \frac{1}{N} \sum_{i=1}^N A_i.$$

Furthermore, let

$$F_i(\mathcal{P}_i) := A_i \mathcal{P}_i + \mathcal{P}_i A_i^\top, \quad \text{and} \quad G_i := BB^\top + L_i V_i L_i^\top$$

so that, $\dot{\mathcal{P}}_i = F_i(\mathcal{P}_i) + G_i$. Also, we let

$$\mathcal{P} := \begin{bmatrix} \mathcal{P}_1 \\ \vdots \\ \mathcal{P}_N \end{bmatrix}, \quad F(\mathcal{P}) := \begin{bmatrix} F_1(\mathcal{P}_1) \\ \vdots \\ F_N(\mathcal{P}_N) \end{bmatrix}, \quad G := \begin{bmatrix} G_1 \\ \vdots \\ G_N \end{bmatrix}.$$

Then, we can write

$$\begin{aligned} \dot{\mathcal{P}} &= F(\mathcal{P}) + G + \lambda(\Pi \otimes I_n)\mathcal{P} - \lambda(I_N \otimes I_n)\mathcal{P} \\ &= F(\mathcal{P}) + G - \lambda\alpha(\mathcal{L} \otimes I_n)\mathcal{P}, \end{aligned}$$

where we recall that $\Pi = I_N + \alpha\mathcal{L}$. For an undirected connected graph with Laplacian \mathcal{L} , we can find a matrix U such that $U\mathcal{L}U^\top = \text{diag}(0, \Lambda)$. Let v_{ℓ_1} denote the normalized eigenvector corresponding to the eigenvalue 0 of the Laplacian, so that

$$v_{\ell_1}^\top = \frac{1}{\sqrt{N}} [1 \quad 1 \quad \cdots \quad 1] \in \mathbb{R}^{1 \times N}, \quad \text{and} \quad v_{\ell_1}^\top \mathcal{L} = 0.$$

There exists a matrix $\tilde{U} \in \mathbb{R}^{N \times (N-1)}$, with $\tilde{U}^\top \tilde{U} = I_{N-1}$, $\tilde{U}^\top \cdot \mathbf{1}_N = 0$, such that

$$U = \begin{bmatrix} v_{\ell_1}^\top \\ \tilde{U}^\top \end{bmatrix}, \quad U^{-1} = U^\top = [v_{\ell_1}, \tilde{U}^\dagger]$$

where $\tilde{U}^\dagger \in \mathbb{R}^{N \times (N-1)}$ is the matrix satisfying

$$\tilde{U}^\top \tilde{U}^\dagger = I_{N-1}, \quad \text{and} \quad 1_N^\top \tilde{U}^\dagger = 0.$$

Introduce the coordinate transformation

$$\mathcal{Q} = \frac{1}{\sqrt{N}}(U \otimes I_n)\mathcal{P} = \begin{bmatrix} \frac{1}{\sqrt{N}}(1_N^\top \otimes I_n) \\ \frac{1}{\sqrt{N}}(\tilde{U}^\top \otimes I_n) \end{bmatrix} \mathcal{P} =: \begin{bmatrix} \mathcal{Q}_1 \\ \tilde{\mathcal{Q}} \end{bmatrix}$$

and the inverse of this transformation gives,

$$\begin{aligned} \mathcal{P} &= \sqrt{N}(U \otimes I_n)^{-1}\mathcal{Q} = \sqrt{N}(U^{-1} \otimes I_n)\mathcal{Q} \\ &= \sqrt{N} \begin{bmatrix} (v_{l_1} \otimes I_n) & (\tilde{U}^\dagger \otimes I_n) \end{bmatrix} \mathcal{Q} \\ &= (1_N \otimes I_n)\mathcal{Q}_1 + \sqrt{N}(\tilde{U}^\dagger \otimes I_n)\tilde{\mathcal{Q}}. \end{aligned}$$

Let \mathcal{S} denote the solution of the differential equation $\dot{\mathcal{S}} = \bar{A}\mathcal{S} + \mathcal{S}\bar{A}^\top + BB^\top + \frac{1}{N} \sum_{i=1}^N L_i V_i L_i^\top$. We introduce the variable \mathcal{E} to denote the difference between the mean value of \mathcal{P}_i , $i = 1, \dots, N$, and \mathcal{S} , that is,

$$\mathcal{E} := \mathcal{Q}_1 - \mathcal{S} = \frac{1}{N} \sum_{i=1}^N \mathcal{P}_i - \mathcal{S},$$

and it is observed that

$$\begin{aligned} \dot{\mathcal{E}} &= \dot{\mathcal{Q}}_1 - \dot{\mathcal{S}} = \frac{1}{N} \sum_{i=1}^N \dot{\mathcal{P}}_i - \dot{\mathcal{S}} \\ &= \bar{A}\mathcal{E} + \mathcal{E}\bar{A}^\top + \frac{1}{N} \left(\sum_{i=1}^N F_i \left(\sqrt{N}(\tilde{U}_i^\dagger \otimes I_n)\tilde{\mathcal{Q}} \right) \right) \end{aligned} \quad (22)$$

where \tilde{U}_i^\dagger denotes the i -th row of \tilde{U}^\dagger . Also, we can write

$$\begin{aligned} \dot{\tilde{\mathcal{Q}}} &= -\lambda\alpha(\Lambda \otimes I_n)\tilde{\mathcal{Q}} + \frac{1}{\sqrt{N}}(\tilde{U}^\top \otimes I_n)(F(\mathcal{P}) + G) \\ &= -\lambda\alpha(\Lambda \otimes I_n)\tilde{\mathcal{Q}} + \frac{1}{\sqrt{N}}(\tilde{U}^\top \otimes I_n) \left[F((1_N \otimes I_n)\mathcal{Q}_1 \right. \\ &\quad \left. + \sqrt{N}(\tilde{U}^\dagger \otimes I_n)\tilde{\mathcal{Q}}) + G \right] \\ &= -\lambda\alpha(\Lambda \otimes I_n)\tilde{\mathcal{Q}} + (\tilde{U}^\top \otimes I_n)F((\tilde{U}^\dagger \otimes I_n)\tilde{\mathcal{Q}}) \\ &\quad + \frac{1}{\sqrt{N}}(\tilde{U}^\top \otimes I_n)(F((1_N \otimes I_n)(\mathcal{E} + \mathcal{S})) + G). \end{aligned} \quad (23)$$

For stability analysis of the process \mathcal{P} , we equivalently analyze the stability of (22) and (23). Towards this end, we recall the matrix R that satisfies (10), and introduce the functions

$$\begin{aligned} \mathcal{V}_1(\mathcal{E}) &:= \frac{1}{2} \text{tr}(R^{1/2}\mathcal{E}R\mathcal{E}R^{1/2}), \\ \mathcal{V}_2(\tilde{\mathcal{Q}}) &:= \frac{1}{2} \text{tr}(\tilde{\mathcal{Q}}\tilde{\mathcal{Q}}^\top). \end{aligned}$$

To analyze the evolution of \mathcal{V}_1 along the solutions of (22), we compute the bound on the derivative of \mathcal{V}_1 . Using the cyclic property of the trace operator, we get

$$\begin{aligned} \dot{\mathcal{V}}_1 &= \frac{1}{2} \text{tr} \left[\text{Sym}((R\bar{A} + \bar{A}^\top R)\mathcal{E}R\mathcal{E}) \right] \\ &\quad + \frac{1}{2N} \text{tr} \left[\text{Sym} \left(R\mathcal{E}R \sum_{i=1}^N F_i(\sqrt{N}(\tilde{U}_i^\dagger \otimes I_n)\tilde{\mathcal{Q}}) \right) \right] \\ &\leq -\text{tr}(\mathcal{E}R\mathcal{E}) \end{aligned}$$

$$\begin{aligned} &+ \frac{1}{2\sqrt{N}} \text{tr} \left[\sum_{i=1}^N \text{Sym} \left(R\mathcal{E}R F_i((\tilde{U}_i^\dagger \otimes I_n)\tilde{\mathcal{Q}}) \right) \right] \\ &\leq -\sigma_1(R)\|\mathcal{E}\|^2 + C_0\sqrt{N}\|\mathcal{E}\|\|\tilde{\mathcal{Q}}\| \end{aligned}$$

where we used the bounds on the trace of products from [27], $\text{tr}(\mathcal{E}R\mathcal{E}) = \text{tr}(R\mathcal{E}^2) \geq \sigma_1(R)\text{tr}(\mathcal{E}^2)$, and $C_0 \geq 0$ is a constant satisfying

$$\max_{1 \leq i \leq N} \frac{1}{2} \text{tr} \left[\text{Sym}(R\mathcal{E}R(F_i((\tilde{U}_i^\dagger \otimes I_n)\tilde{\mathcal{Q}}))) \right] \leq C_0\|\mathcal{E}\|\|\tilde{\mathcal{Q}}\|.$$

On the other hand, using similar techniques, we can bound the derivative of \mathcal{V}_2 along the solutions of (23) as follows:

$$\begin{aligned} \dot{\mathcal{V}}_2 &\leq -\lambda\alpha\sigma_2(\mathcal{L})\text{tr}(\tilde{\mathcal{Q}}\tilde{\mathcal{Q}}^\top) + C_1\text{tr}(\tilde{\mathcal{Q}}\tilde{\mathcal{Q}}^\top) \\ &\quad + C_2\|\mathcal{E}\| \cdot \|\tilde{\mathcal{Q}}\| + C_2\|\mathcal{S}\| \cdot \|\tilde{\mathcal{Q}}\| + C_3\bar{\sigma}_G \cdot \|\tilde{\mathcal{Q}}\| \end{aligned}$$

where we recall that $\bar{\sigma}_G = \max_{1 \leq i \leq N} \|G_i\|$, and the positive scalars $C_1, C_2, C_3 \geq 0$ are chosen to satisfy

$$\begin{aligned} 0.5 \text{tr} \left[\text{Sym}(\tilde{\mathcal{Q}}^\top(\tilde{U}^\top \otimes I_n)(F((\tilde{U}^\dagger \otimes I_n)\tilde{\mathcal{Q}}))) \right] &\leq C_1\text{tr}(\tilde{\mathcal{Q}}\tilde{\mathcal{Q}}^\top) \\ 0.5 \text{tr} \left[\text{Sym}(\tilde{\mathcal{Q}}^\top(\tilde{U}^\top \otimes I_n)(F((1_N \otimes I_n)\mathcal{E}))) \right] &\leq C_2\sqrt{N}\|\mathcal{E}\| \cdot \|\tilde{\mathcal{Q}}\| \\ 0.5 \text{tr} \left[\text{Sym}(\tilde{\mathcal{Q}}^\top(\tilde{U}^\top \otimes I_n)(F((1_N \otimes I_n)\mathcal{S}))) \right] &\leq C_2\sqrt{N}\|\mathcal{S}\| \cdot \|\tilde{\mathcal{Q}}\| \\ 0.5 \text{tr} \left[\text{Sym}(\tilde{\mathcal{Q}}^\top(\tilde{U}^\top \otimes I_n)G) \right] &\leq C_3\sqrt{N}\bar{\sigma}_G\|\tilde{\mathcal{Q}}\|. \end{aligned}$$

For the combined dynamical system (22), (23), we now consider the proper, positive definite Lyapunov function

$$\mathcal{V}(\mathcal{E}, \mathcal{Q}) = \mathcal{V}_1(\mathcal{E}) + \mathcal{V}_2(\tilde{\mathcal{Q}})$$

and observe that

$$\begin{aligned} \dot{\mathcal{V}} &\leq -\sigma_1(R)\|\mathcal{E}\|^2 - (\lambda\alpha\sigma_2(\mathcal{L}) - C_1)\|\tilde{\mathcal{Q}}\|^2 \\ &\quad + (C_0\sqrt{N} + C_2)\|\mathcal{E}\|\|\tilde{\mathcal{Q}}\| + (C_2\|\mathcal{S}\| + C_3\bar{\sigma}_G)\|\tilde{\mathcal{Q}}\|. \end{aligned}$$

Applying Lemma A.1 from Appendix, we readily obtain,

$$\begin{aligned} \limsup_{t \rightarrow \infty} \|\mathcal{E}(t)\|^2 + \|\tilde{\mathcal{Q}}(t)\|^2 &\leq \frac{1}{\sigma_1^2(R)} \limsup_{t \rightarrow \infty} \mathcal{V}(\mathcal{E}(t), \mathcal{Q}(t)) \\ &\leq \frac{\sigma_n^2(R)}{\sigma_1^2(R)} \left(C_3\bar{\sigma}_G + C_2 \limsup_{t \rightarrow \infty} \|\mathcal{S}(t)\| \right)^2 \chi \left(\frac{1}{\lambda\alpha\sigma_2(\mathcal{L}) - C_1} \right). \end{aligned}$$

With \bar{A} Hurwitz and \mathcal{S} satisfying (11), we have $\lim_{t \rightarrow \infty} \mathcal{S}(t) = \mathcal{S}$, which leads to the following bound for each $i \in \mathcal{V}$,

$$\begin{aligned} \limsup_{t \rightarrow \infty} \|\mathcal{P}_i(t) - \mathcal{S}\| &= \limsup_{t \rightarrow \infty} \|\mathcal{P}_i(t) - \mathcal{S}(t)\| \\ &\leq \limsup_{t \rightarrow \infty} \sqrt{N} \sqrt{\|\mathcal{E}(t)\|^2 + \|\tilde{\mathcal{Q}}(t)\|^2} \\ &\leq \sqrt{N} \frac{\sigma_n(R)}{\sigma_1(R)} \sqrt{\chi \left(\frac{1}{\lambda\alpha\sigma_2(\mathcal{L}) - C_1} \right)} (C_3\bar{\sigma}_G + C_2\|\mathcal{S}\|) \end{aligned}$$

for some class \mathcal{K} function χ such that $\chi(s) = O(s)$ as $s \rightarrow 0$. Hence, we obtain (12) and this completes the proof of Theorem III.1.

IV. CONCLUSIONS AND PERSPECTIVES

We considered the problem of distributed filtering over undirected and connected graphs. The sensor nodes exchange information about their estimate, and the communication between the sensor nodes is driven by independent Poisson processes. The algorithm implemented by each node is in the form of a stochastic hybrid system. We derive bounds on the covariance of the estimation error for each node, and show practical convergence for large enough sampling rates.

There are some immediate research directions that emanate from this work. Building on Theorem III.1, it is interesting to investigate algorithms which minimize the covariance bounds under different information patterns. Another interesting direction is to study a broader class of random processes for the communication between sensor nodes. While the memoryless Poisson counter treated in this paper results in diffusive coupling between the dynamics of expected error covariances, it remains to be seen what kind of interconnection is obtained from other communication processes.

APPENDIX

The following lemma has been used in the proof of Theorem III.1.

Lemma A.1. Consider a function $W : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$, and a function $g : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$, such that

$$W(t, r, s) \leq -a_1 r^2 + a_2 r s - a_3 s^2 + g(t)s$$

for some positive scalars $a_1, a_2, a_3 > 0$. Then, there exists $c > 0$ such that

$$W(t, r, s) \leq -c(r^2 + s^2), \quad \text{if } r^2 + s^2 \geq g^2(t)\chi\left(\frac{1}{a_3}\right)$$

where χ is a class \mathcal{K} function and $\chi(s) = O(s)$ as $s \rightarrow 0$.

The proof of Lemma A.1 is very similar to the proof of [5, Lemma 2], where one can also find the exact expressions for the constant c , and the function χ .

ACKNOWLEDGEMENTS

The author would like to thank Hyungbo Shim for indicating several references on consensus problems in multi-agent systems with heterogeneous agents.

REFERENCES

- [1] S. Park and N. Martins, "Design of distributed LTI observers for state omniscience," *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 561–576, 2017.
- [2] A. Mitra and S. Sundaram, "Distributed observers for LTI systems," *IEEE Transactions on Automatic Control*, vol. 63, no. 11, pp. 3689–3704, 2018.
- [3] W. Han, H. Trentelman, Z. Wang, and Y. Shen, "A simple approach to distributed observer design for linear systems," *IEEE Transactions on Automatic Control*, vol. 64, no. 1, pp. 329–336, 2019.
- [4] L. Wang and A. Morse, "A distributed observer for a time-invariant linear system," *IEEE Transactions on Automatic Control*, vol. 63, no. 7, pp. 2123–2130, 2018.
- [5] J. Kim, J. Yang, H. Shim, J.-S. Kim, and J. Seo, "Robustness of synchronization of heterogeneous agents by strong coupling and a large number of agents," *IEEE Transactions on Automatic Control*, vol. 61, no. 10, pp. 3096–3102, 2016.
- [6] J. Lee and H. Shim, "A tool for analysis and synthesis of heterogeneous multi-agent systems under rank-deficient coupling," *Automatica*, vol. 117, 2020.
- [7] E. Panteley and A. Loria, "Synchronization and dynamic consensus of heterogeneous networked systems," *IEEE Transactions on Automatic Control*, vol. 62, no. 8, pp. 3758–3773, 2017.
- [8] F. Dorfler, F. Pasqualetti, and F. Bullo, "Continuous-time distributed observers with discrete communication," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 2, pp. 296–304, 2013.
- [9] Y. Li, S. Phillips, and R. Sanfelice, "Robust distributed estimation for linear systems under intermittent information," *IEEE Transactions on Automatic Control*, vol. 63, no. 4, pp. 973–988, 2017.
- [10] 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, 2008.
- [11] G. Battistelli and L. Chisci, "Kullback–Leibler average, consensus on probability densities, and distributed state estimation with guaranteed stability," *Automatica*, vol. 50, pp. 707–718, 2014.
- [12] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *Proc. 46th IEEE Conf. on Decision and Control*, 2007, pp. 5492–5498.
- [13] —, "Kalman-Consensus Filter: Optimality, stability, and performance," in *Proc. 48th IEEE Conf. on Decision and Control*, 2009, pp. 7036–7042.
- [14] B. Andrievsky, A. Matveev, and A. Fradkov, "Control and estimation under information constraints: Toward a unified theory of control, computation and communications," *Automation and Remote Control*, vol. 71, no. 4, pp. 572–633, 2010.
- [15] M. Adès, P. Caines, and R. Malhamé, "Stochastic optimal control under Poisson-distributed observations," *IEEE Transactions on Automatic Control*, vol. 45, no. 1, pp. 3–13, 2000.
- [16] J. Hespanha and A. Teel, "Stochastic impulsive systems driven by renewal processes," in *Proc. 17th International Symposium on Mathematical Theory of Networked Systems*, 2006, pp. 606–618.
- [17] D. Antunes, J. Hespanha, and C. Silvestre, "Volterra integral approach to impulsive renewal systems: Application to networked control," *IEEE Transactions on Automatic Control*, vol. 57, no. 3, pp. 607–619, 2012.
- [18] A. Tanwani, D. Chatterjee, and D. Liberzon, "Stabilization of continuous-time deterministic systems under random sampling: Overview and recent developments," in *Uncertainty in Complex Networked Systems*, T. Başar, Ed. Switzerland: Springer Nature, 2018, pp. 209–246.
- [19] A. Matveev and A. Savkin, "The problem of state estimation via asynchronous communication channels with irregular transmission times," *IEEE Transactions on Automatic Control*, vol. 48, no. 4, pp. 670–676, 2003.
- [20] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, 2004.
- [21] M. Huang and S. Dey, "Stability of Kalman filtering with Markovian packet losses," *Automatica*, vol. 43, pp. 598–607, 2007.
- [22] C. McDonald and S. Yüksel, "Stability of non-linear filters, observability and relative entropy," in *Proc. of 56th Annual Allerton Conference on Communication, Control, and Computing*, 2018, pp. 110–114.
- [23] A. Jazwinski, *Stochastic processes and filtering theory*. New York: Dover Publications, Inc., 2007.
- [24] A. Tanwani and O. Yufereva, "Error covariance bounds for suboptimal filters with Lipschitzian drift and Poisson-sampled measurements," *Automatica*, vol. 122, 2020.
- [25] A. Teel, "Lyapunov conditions certifying stability and recurrence for a class of stochastic hybrid systems," *Annual Reviews in Control*, vol. 37, pp. 1–24, 2013.
- [26] Y. Suhov and M. Kelbert, *Probability and Statistics by Example. II*. Cambridge University Press, Cambridge, 2008, Markov chains: A primer in random processes and their applications.
- [27] Y. Fang, K. Loparo, and X. Feng, "Inequalities for the trace of matrix product," *IEEE Transactions on Automatic Control*, vol. 39, no. 12, pp. 2489–2490, 1994.