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. 2021. hal-03106998v1

HAL Id: hal-03106998

https://laas.hal.science/hal-03106998v1

Preprint submitted on 12 Jan 2021 (v1), last revised 29 Sep 2021 (v3)

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 linear stochastic systems using distributed sensors. Each sensor unit, represented by a node in an 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 is time-sampled randomly and the sampling process is assumed to be a 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; heterogenous agents.**

I. Introduction

Filtering, or state estimation, in stochastic dynamical systems is one of the fundamental problems in engineering which has found applications in various disciplines ranging from control theory and signal processing to robotics and image/audio processing. A recently compiled book [1] provides an overview of the developments carried out in the area of filtering theory since its inception. Keeping in mind the current trends of large-scale engineering systems operating over networks, we propose filtering algorithms for estimating the state of the continuous-time systems in a distributed manner where the communication between the sensor nodes only takes place at discrete times, possibly due to network constraints.

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, $y_1, \ldots, y_N$, 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 [2], [3]. For continuous-time deterministic systems, one may refer to [4], [5]. The focus in this 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 [6], [7], [8]. Distributed state estimation for deterministic continuous-time systems with communication at discrete times has been studied in [9], [10]. For stochastic systems, we see convergence based approaches for scalar systems in [11]. Distributed filtering with more general probability distributions and application to Gaussian distributions with linear discrete-time systems is studied in [12]. Prior to that, several algorithms for distributed filtering have been proposed in discrete-setting [13], [14].

An important issue from the point of implementation of filters over networks is to make the algorithms compatible with the underlying communication protocol [15]. 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 [16] for optimal control and the papers [17], [18] for stability analysis of such systems. A recently published book chapter [19] 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 [20], [21], [22] 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 [23] 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 observer proposed in [24]. Some analytical results on the performance of centralized

A. Tanwani is affiliated with CNRS – LAAS, University of Toulouse, CNRS, 31400 Toulouse, France.

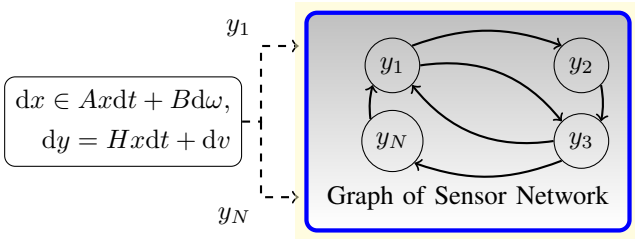

Figure 1: Layout of Distributed Filters

continuous-time systems under random sampling appear in [25].

In this article, we consider filtering problem in distributed setting for continuous-time linear stochastic systems. The particular feature we address is the one where the agents communicate the information between themselves at discrete times which are randomly drawn. 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 with their neighboring agents and update their own state based on the information received. In addition to formalizing this intuition in this paper, 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 [4], [7], [2], [3], [5], one directly analyzes the convergence of the norm of estimation error to the origin. The error covariance bounds are analyzed in the stochastic setup only [11], [12], [13], [14], which is the case here. The papers [11], [12] essentially provide the bounds on error covariance in the discrete-time setup. On the other hand, our main result (Theorem III.1) provides a bound at all time instants for a continuous-time process. Moreover, the derivation and analysis of the expectation of error covariance bound with respect to random (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 allows us to invoke some recent developments on consensus with heterogenous agents to analyze the asymptotic behavior of the expected error covariance for each filtering unit.

Notation: Let us recall some basic definitions from the literature which are used in statement and derivation of main results. A function $\chi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ 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 $\langle M_1, M_2 \rangle = \mathrm{tr}(M_1^\top M_2)$. This choice of inner product allows us to introduce the Frobenius norm, defined as $\|M\| = \sqrt{\langle M, M \rangle} = \sqrt{\mathrm{tr}(M^\top M)}$. 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 $|\,\mathrm{tr}(M_1 M_2)| \leq \|M_1\|\,\|M_2\|$. For a matrix $M \in \mathbb{R}^{n\times n}$, $\sigma_i(M)$ denotes the i-th eigenvalue of M, and we adopt the convention that $\sigma_1(M) \geq \sigma_2(M) \geq \cdots \geq \sigma_n(M)$. The symbol $o(s)$ for $s \to 0$ is a notation for any function of $s \in \mathbb{R}$ satisfying $\lim_{s\to 0} \frac{o(s)}{s} = 0$, and the symbol $O(s)$ for $s \to 0$ stands for any function of $s \in \mathbb{R}$ satisfying $\lim_{s\to 0} \frac{O(s)}{s}$ exists and is finite.

II. Problem Formulation

Consider the dynamical system

$$\mathrm{d}x = Ax\,\mathrm{d}t + B\,\mathrm{d}\omega \tag{1}$$

where $(x(t))_{t\geq 0}$ is an $\mathbb{R}^n$-valued diffusion process describing the state. Let $(\Omega, \mathcal{F}, \mathsf{P})$ 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 $\mathsf{E}[\mathrm{d}\omega(t)\,\mathrm{d}\omega(t)^\top] = I_m$, 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,

$$\mathrm{d}y_i = H_i x\,\mathrm{d}t + \mathrm{d}v_i, \quad i = 1, \ldots, N, \tag{2}$$

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 $\mathsf{E}[\mathrm{d}v_i(t)\,\mathrm{d}v_i(t)^\top] = V_i \in \mathbb{R}^{p_i\times p_i}$, with V_i assumed to be positive definite.

The sensor nodes are connected via a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, \ldots, N\}$ 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 $\mathcal{N}_i$ and we adopt the convention that $i \notin \mathcal{N}_i$. The adjacency matrix $\mathcal{A} := [\alpha_{ij}] \in \{0, 1\}^{N\times N}$ of the graph, which is symmetric, provides the information about which sensor units

can communicate with each other, that is, if $\alpha_{ij} = 1$ then sensor i and j can communicate, whereas $\alpha_{ij} = 0$ 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

$$\Pi := I_{N\times N} - \epsilon\mathcal{L} \tag{3}$$

where $0 < \epsilon \leq \min_{i\in\mathcal{V}} \frac{1}{|\mathcal{N}_i|}$ 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 time instants at which the communication takes place between the sensor nodes. It is assumed that there exists a monotone strictly increasing and divergent sequence $(\tau_k)_{k\in\mathbb{N}} \subset [0, +\infty[$ with $\tau_0 := 0$, and

- at each time instant τ_k, *all* sensor nodes $i \in \mathcal{V}$ transmit the value of their state estimate to their neighbors $\mathcal{N}_i$.

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

$$N_t := \sup\{k \in \mathbb{N} \mid \tau_k \leq t\} \quad \text{for } t \geq 0 \tag{4}$$

and stipulate in addition that

- $(N_t)_{t\geq 0}$ is a continuous-time stochastic process satisfying $\tau_{N_t} \xrightarrow[t\nearrow+\infty]{} +\infty$ almost surely.

The map $t \mapsto N_t$ increments by 1 at random times, and it provides a description of the number of samples up to and including time t. Let $\lambda > 0$. Recall that a random process $(N_t)_{t\geq 0}$ is a *Poisson process of intensity* λ defined on our probability space if any one of the following equivalent properties hold:

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

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

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

(P2) $(N_t)_{t\geq 0}$ 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 each k.

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:

$$\mathrm{d}\widehat{x}_i(t) = (A - L_iH_i)\widehat{x}_i(t)\mathrm{d}t + L_i\mathrm{d}y_i(t) \tag{5}$$

over the interval $[\tau_k, \tau_{k+1}[$, for each $k \in \mathbb{N}$. At sampling times τ_k, $k \in \mathbb{N}$, when the sensor node i receives the information from its neighbors, we update the state as follows:

$$\widehat{x}_i^+ = \widehat{x}_i^- + \epsilon \sum_{j\in\mathcal{N}_i} (\widehat{x}_j^- - \widehat{x}_i^-) \tag{6a}$$

$$= \pi_{ii}\widehat{x}_i^- + \sum_{j\in\mathcal{N}_i} \pi_{ij}\widehat{x}_j^-, \tag{6b}$$

where we recall that $0 < \epsilon \leq \min_{i\in\mathcal{V}} \frac{1}{|\mathcal{N}_i|}$ and $\pi_{ij} \in [0, 1]$ are the elements of the doubly stochastic matrix Π introduced in (3). Each of these filters is a stochastic hybrid system of the form [26] with continuous evolution described by (5) between sampling times, and the jump rule (6) executed at random sampling instants where we update the estimate $\widehat{x}_i$ as the convex combination of itself and its 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 to study the performance of the distributed filters proposed in (5)–(6) under the aforementioned communication architecture. 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, $\mathcal{Y}_t^i := \{(\mathrm{d}y_i(s), \widehat{x}_j(\tau_{N_s})) \mid s \leq t, j \in \mathcal{N}_i\}$, then our objectives are:

- Compute an upper bound on the expectation (with respect to sampling process) of the error covariance matrices $\mathsf{E}[\mathsf{E}[(x(t) - \widehat{x}_i(t))(x(t) - \widehat{x}_i(t))^\top \mid \mathcal{Y}_t^i]]$, 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:

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

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

Theorem III.1. *Consider system* (1) *with distributed measurements* (2) *and the corresponding hybrid filters* (5), (6) *linked together by an undirected and connected graph. If*

the communication between graph nodes nodes takes place at random times generated by a Poisson process of intensity $\lambda > 0$, then the following items hold:

1) *For each $\lambda > 0$, and $i = 1, \cdots, N$, it holds that*

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

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

$$\dot{\mathcal{P}}_i = (A - L_iC_i)\mathcal{P}_i + \mathcal{P}_i(A - L_iC_i)^\top + BB^\top + L_iV_iL_i^\top + \lambda \sum_{j\in\mathcal{N}_i} \pi_{ij}(\mathcal{P}_j - \mathcal{P}_i). \quad (9)$$

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

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

$$0 = (A - LH)\, S + S\,(A - LH)^\top + BB^\top + \frac{1}{N}\sum_{i=1}^{N} L_iV_iL_i^\top \quad (11)$$

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

$$\limsup_{t\to\infty} \|\mathcal{P}_i(t) - S\| \leq \sqrt{N}\,\frac{\sigma_1(R)}{\sigma_n(R)}\sqrt{\chi\left(\frac{1}{\lambda\sigma_2(\mathcal{L}) - C_1}\right)}\,(C_3\overline{\sigma}_G + C_2S)\,, \quad (12)$$

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

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, $e_i := x - \widehat{x}_i$. Over an interval $[\tau_k, \tau_{k+1}[$, we have the differential equation:

$$\mathrm{d}e_i = (A - L_iH_i)\mathrm{d}e_i + L_i\mathrm{d}v_i + B\mathrm{d}\omega \quad (13a)$$

and at sampling time τ_k, it holds that

$$e_i(\tau_k^+) = \pi_{ii}e_i(\tau_k^-) + \sum_{j\in\mathcal{N}_i} \pi_{ij}e_j(\tau_k^-). \quad (13b)$$

Error covariance bound on estimation error $\mathsf{E}[e_ie_i^\top]$ is described in the following statement.

Proposition III.2. *Consider system* (1) *and the filter* (5), (6) *with $\{\tau_k\}_{k\in\mathbb{N}}$ fixed. Let the process $P_i : [0, \infty[\to \mathbb{R}^{n\times n}$, with $P_i(0) \geq \mathsf{E}[(x_0 - \widehat{x}_i(0))(x_0 - \widehat{x}_i(0))^\top]$, be defined by*

$$\frac{\mathrm{d}P_i}{\mathrm{d}t} = (A - L_iH_i)P_i + P_i(A - L_iH_i)^\top + BB^\top + L_iV_iL_i^\top, \quad (14a)$$

for $\tau_k \leq t < \tau_{\tau_k+1}$, and let

$$P_i(\tau_k^+) = \pi_{ii}P_i(\tau_k^-) + \sum_{j\in\mathcal{N}_i} \pi_{ij}P_j(\tau_k^-) \quad (14b)$$

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

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

Proof. We first show that, for each $t \in [\tau_k, \tau_{k+1}[$,

$$\frac{\mathrm{d}\mathsf{E}[e_i(t)e_i(t)^\top]}{\mathrm{d}t} \leq (A - L_iH_i)\mathsf{E}[e_i(t)e_i(t)^\top] + \mathsf{E}[e_i(t)e_i(t)^\top](A - L_iH_i)^\top + BB^\top + L_iV_iL_i^\top$$

and that, at the sampling instants

$$\mathsf{E}[e_i(\tau_k^+)e_i(\tau_k^+)^\top] \leq \pi_{ii}\mathsf{E}[e_i(\tau_k^-)e_i(\tau_k^-)^\top] + \sum_{j\in\mathcal{N}_i} \pi_{ij}\mathsf{E}[e_j(\tau_k^-)e_j(\tau_k^-)^\top].$$

The bound in (15) then holds by observing that $P_i(t)$ satisfies (14a) over $[\tau_k, \tau_{k+1}[$, and (14b) at $t = \tau_k$, with $\mathsf{E}[e_0e_0^\top] \leq P_i(0)$.

a) Continuous part: We first show that for $t \in [\tau_k, \tau_{k+1}[$, we have the inequality $\frac{\mathrm{d}\mathsf{E}[e_i(t)e_i(t)^\top]}{\mathrm{d}t} = \frac{\mathrm{d}P_i(t)}{\mathrm{d}t}$. This is a rather classical derivation, and as an example, one may refer to the arguments in [25] for details. On this time interval, let us consider the Itô process (13), and the function $v : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ given by, $\mathbb{R}^n \ni e \mapsto v(e) = ee^\top \in \mathbb{R}^{n\times n}$. 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 [25, Section 3], we obtain

$$\frac{\mathrm{d}\mathsf{E}[v(e_i(t))]}{\mathrm{d}t} = (A - L_iH_i)\mathsf{E}\big[e_i(t)e_i(t)^\top\big] + \mathsf{E}\big[e_i(t)e_i(t)^\top\big](A - L_iH_i)^\top + BB^\top + L_iV_iL_i^\top.$$

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 $x_1, \ldots, x_m \in \mathbb{R}^n$. If $z := \sum_{j=1}^m \gamma_jx_j$ for some $\gamma_j \in [0, 1]$, $\sum_{j=1}^m \gamma_j = 1$, then*

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

Proof. Let $\mathcal{S}_+^n$ denote the set of positive semidefinite matrices in $\mathbb{R}^{n\times n}$, and consider the function $f : \mathbb{R}^n \to \mathcal{S}_+^n$, so that, for

$x \in \mathbb{R}^n$, $f(x) = xx^\top$. We consider the epigraph of f, denoted by $\mathrm{epi}(f)$, and defined as,

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

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

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

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 $\mathrm{epi}(f)$ is convex. Coming back to the statement of the lemma, we observe that, for each $j = 1, \ldots, m$, $(f(x_j), x_j) \in \mathrm{epi}(f)$ 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 $\mathrm{epi}(f)$. By construction, we therefore have

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

which yields the desired inequality be recalling the definition of f. □

The statement of Lemma III.3 directly yields an upper bound on $\mathsf{E}[e_i(\tau_k^+)e_i(\tau_k^+)^\top]$. Recalling the definition of $e(\tau_k^+)$ in (13b), and using the linearity of the expectation operator, we get

$$\begin{aligned} \mathsf{E}[e_i(\tau_k^+)e_i(\tau_k^+)^\top] &\leq \pi_{ii}\mathsf{E}[e_i(\tau_k^-)e_i(\tau_k^-)^\top] \\ &\quad + \sum_{j \in \mathcal{N}_i} \pi_{ij}\mathsf{E}[e_j(\tau_k^-)e_j(\tau_k^-)^\top]. \end{aligned}$$

The proof of Proposition III.2 is therefore complete. □

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. That is, $P_i(t)$ is an upper bound on $\mathsf{E}[(x(t) - \widehat{x}_i(t))(x(t) - \widehat{x}_i(t))^\top \,|\, \mathcal{Y}_t^i]$ 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 sampling process $(N_t)_{t\geq 0}$. We do so next under the assumption that $(N_t)_{t\geq 0}$ is a Poisson process of intensity λ.

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

$$\mathcal{L}_{\inf}(Q) := \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\Big(\mathsf{E}[P_i(t+\varepsilon) \,|\, P_i(t) = Q] - Q\Big). \quad (17)$$

We will see that this operator describes infintesimal evolution of the expected value of the error covariance process, and hence (9) is a direct consequence of the following result.

Proposition III.4. *Consider the process $(P_i(t))_{t\geq 0}$ given by* (14) *with the sampling process $(N_t)_{t\geq 0}$ of intensity $\lambda > 0$. Let $\mathcal{P}_i(t) := \mathsf{E}[P_i(t) \,|\, P_0]$ denote the expected value of the covariance process at time $t \geq 0$ with respect to sampling process. Then,* (9) *holds.*

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

$$\mathcal{P}_i(t) = \mathsf{E}[P_i(t)] = P_i(0) + \mathsf{E}\Big[\int_0^t \mathcal{L}_{\inf}(P_i(s))\Big]\mathrm{d}s.$$

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

$$\begin{aligned} \mathsf{E}\big[P_i(t+\varepsilon) \,\big|\, P_i(t) = Q\big] &= \mathsf{E}\big[\big(P_i(t+\varepsilon)\big)\big(1_{\{N_{t+\varepsilon}=N_t\}} \\ &\quad + 1_{\{N_{t+\varepsilon}=1+N_t\}} + 1_{\{N_{t+\varepsilon}-N_t\geq 2\}}\big) \,\big|\, P(t)\big]. \end{aligned} \quad (18)$$

We now compute the conditional probability distribution of $\big(P_i(t+\varepsilon)\big)$ for small $\varepsilon > 0$ given $\big(P_i(t)\big)$. Since the sampling process is independent of the process $\big(P_i(t)\big)_{t\geq 0}$, by definition of the sampling (Poisson) process we have, for $\varepsilon \downarrow 0$, $\mathsf{P}\big(N_{t+\varepsilon} - N_t = 0 \,\big|\, N_t, P_i(t)\big) = 1 - \lambda\varepsilon + o(\varepsilon)$, $\mathsf{P}\big(N_{t+\varepsilon} - N_t = 1 \,\big|\, N_t, P_i(t)\big) = \lambda\varepsilon + o(\varepsilon)$, $\mathsf{P}\big(N_{t+\varepsilon} - N_t \geq 2 \,\big|\, N_t, P_i(t)\big) = o(\varepsilon)$. Using these expressions we develop (18) further for $\varepsilon \downarrow 0$ as

$$\begin{aligned} &\mathsf{E}\big[P_i(t+\varepsilon) \,\big|\, P_i(t) = Q\big] \\ &= \mathsf{E}\big[P_i(t+\varepsilon)\big(1_{\{N_{t+\varepsilon}=N_t\}} + 1_{\{N_{t+\varepsilon}=1+N_t\}}\big) \,\big|\, P_i(t)\big] + o(\varepsilon) \\ &= \mathsf{E}\big[P_i(t+\varepsilon) \,\big|\, P_i(t), N_{t+\varepsilon} = N_t\big] \cdot \big(1 - \lambda\varepsilon + o(\varepsilon)\big) \\ &\quad + \mathsf{E}\big[P_i(t+\varepsilon) \,\big|\, P_i(t), N_{t+\varepsilon} = 1 + N_t\big]\big(\lambda\varepsilon\big) + o(\varepsilon). \end{aligned} \quad (19)$$

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

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

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

$$\begin{aligned}&\mathsf{E}\big[P_i(t+\varepsilon) \,\big|\, P_i(t), N_{t+\varepsilon} = N_t\big] \cdot \big(1 - \lambda\varepsilon + o(\varepsilon)\big)\\ &= Q + \varepsilon((A - L_iH_i)Q + Q(A - L_iH_i)^\top + BB^\top + L_iV_iL_i^\top)\\ &\quad - (\lambda\varepsilon)Q + o(\varepsilon).\end{aligned} \tag{20}$$

for $\varepsilon \downarrow 0$. Concerning the second term on the right-hand side of (19), we observe that conditional on $N_{t+\varepsilon} = 1 + N_t$, the probability distribution of $\tau_{N_{t+\varepsilon}}$ is [27, Theorem 2.3.7] uniform over $[t, t+\varepsilon[$ by definition of the sampling (Poisson) process. We introduce $\theta \in [0, 1[$ such that $\tau_{N_{t+\varepsilon}} = t + \theta\varepsilon$; then θ is uniformly distributed on $[0, 1[$ given $N_{t+\varepsilon} = 1 + N_t$. We now write the right-hand side of (14a) more compactly using the map $\overline{\psi}$,

$$Q \mapsto \overline{\psi}(Q) = (A - L_iH_i)Q + Q(A - L_iH_i)^\top + BB^\top + L_iV_iL_i^\top$$

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

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

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

$$\begin{aligned}P_i(t+\varepsilon) &= P_i(t+\theta\varepsilon) + (1-\theta)\varepsilon\dot{P}_i(t+\theta\varepsilon) + o(\varepsilon)\\ &= P_i(t+\theta\varepsilon) + (1-\theta)\varepsilon\overline{\psi}\big(P_i(t+\theta\varepsilon)\big) + o(\varepsilon)\\ &= \psi_{t+\theta\varepsilon}(P_i(t) + \theta\varepsilon\overline{\psi}(P_i(t)) + o(\varepsilon))\\ &\qquad + (1-\theta)\varepsilon\overline{\psi}\big(P_i(t+\theta\varepsilon)\big) + o(\varepsilon)\\ &= \psi_t(P_i(t)) + \theta O(\varepsilon) + O(\varepsilon) + o(\varepsilon).\end{aligned}$$

Therefore, for $\varepsilon \downarrow 0$,

$$\begin{aligned}&\mathsf{E}\big[P_i(t+\varepsilon) \,\big|\, P_i(t) = Q, N_{t+\varepsilon} = 1 + N_t\big] \cdot (\lambda\varepsilon)\\ &= \int_0^1 \Big(\psi_t(Q) + \theta O(\varepsilon) + O(\varepsilon) + o(\varepsilon)\Big)\,\mathrm{d}\theta \cdot (\lambda h)\\ &= \Big(\psi_t(Q) + O(h)\Big) \cdot (\lambda h) = (\lambda h)\psi_t(Q) + o(\varepsilon).\end{aligned} \tag{21}$$

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

$$\begin{aligned}&\mathsf{E}\big[P_i(t+\varepsilon) \,\big|\, P_i(t) = Q\big] = Q - (\lambda\varepsilon)\big(Q - \psi_t(Q)\big) + o(\varepsilon)\\ &+ \varepsilon\Big((A - L_iH_i)Q + Q(A - L_iH_i)^\top + BB^\top + L_iV_iL_i^\top\Big).\end{aligned}$$

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

$$\begin{aligned}\mathcal{L}_{\inf}(Q) &= (A - L_iH_i)Q + Q(A - L_iH_i)^\top + BB^\top + L_iV_iL_i^\top\\ &\quad + \lambda\big(\psi_t(Q) - Q\big),\end{aligned}$$

which, upon recalling the definition of ψ_t, leads to the desired expression in (9). □

This also completes the proof of first item in the statement of Theorem III.1.

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. Practical stabilization of heterogenous agents has been studied in [6], [8] and here we base our analysis on similar concepts adapted to linear matrix-valued processes. For the proof of item 2) in Theorem III.1, we introduce the notation,

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

Furthermore, let

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

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

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

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

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

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

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

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

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

Introduce the coordinate transformation

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

and the inverse of this transformation gives,

$$\begin{aligned}\mathcal{P} &= \sqrt{N}(U\otimes I_n)^{-1}\mathcal{Q} = \sqrt{N}(U^{-1}\otimes I_n)\mathcal{Q}\\ &= \sqrt{N}\begin{bmatrix}(v_{\ell_1}\otimes I_n) & (\widetilde{U}^\dagger\otimes I_n)\end{bmatrix}\mathcal{Q}\\ &= (1_N\otimes I_n)\mathcal{Q}_1 + \sqrt{N}(\widetilde{U}^\dagger\otimes I_n)\widetilde{\mathcal{Q}}.\end{aligned}$$

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

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

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

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

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

$$\begin{aligned}V_1(\mathcal{E}) &:= \frac{1}{2}\operatorname{tr}(R^{1/2}\mathcal{E}R\mathcal{E}R^{1/2}),\\ V_2(\widetilde{\mathcal{Q}}) &:= \frac{1}{2}\operatorname{tr}(\widetilde{\mathcal{Q}}\widetilde{\mathcal{Q}}^\top).\end{aligned}$$

To analyze the evolution of V_1 along the solutions of (22), we compute the bound on the derivative of V_1. Using the cyclic property of the trace operator, we get

$$\begin{aligned}\dot{V}_1 &= \frac{1}{2}\operatorname{tr}\left[\operatorname{Sym}((R\overline{A}+\overline{A}^\top R)\mathcal{E}R\mathcal{E})\right]\\ &\quad + \frac{1}{2N}\operatorname{tr}\left[\operatorname{Sym}\left(R\mathcal{E}R\sum_{i=1}^N F_i(\sqrt{N}(\widetilde{U}_i^\dagger\otimes I_n)\widetilde{\mathcal{Q}})\right)\right]\\ &\le -\operatorname{tr}(\mathcal{E}R\mathcal{E})\\ &\quad + \frac{1}{2\sqrt{N}}\operatorname{tr}\left[\sum_{i=1}^N \operatorname{Sym}\left(R\mathcal{E}RF_i((\widetilde{U}_i^\dagger\otimes I_n)\widetilde{\mathcal{Q}})\right)\right]\\ &\le -\sigma_n(R)\|\mathcal{E}\|^2 + C_0\sqrt{N}\,\|\mathcal{E}\|\,\|\widetilde{\mathcal{Q}}\|\end{aligned}$$

where we used the bounds on the trace of products from [28], $\operatorname{tr}(\mathcal{E}R\mathcal{E}) = \operatorname{tr}(R\mathcal{E}^2) \ge \sigma_n(R)\operatorname{tr}(\mathcal{E}^2)$, and $C_0 \ge 0$ is a constant satisfying

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

On the other hand, using similar techniques, we can bound the derivative of V_2 along the solutions of (23) as follows:

$$\begin{aligned}\dot{V}_2 &\le -\lambda\,\sigma_2(\mathcal{L})\operatorname{tr}(\widetilde{\mathcal{Q}}\widetilde{\mathcal{Q}}^\top) + C_1\operatorname{tr}(\widetilde{\mathcal{Q}}\widetilde{\mathcal{Q}}^\top)\\ &\quad + C_2\|\mathcal{E}\|\cdot\|\widetilde{\mathcal{Q}}\| + C_2\|\mathcal{S}\|\cdot\|\widetilde{\mathcal{Q}}\| + C_4\overline{\sigma}_G\cdot\|\widetilde{\mathcal{Q}}\|\end{aligned}$$

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

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

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

$$V(\mathcal{E},\mathcal{Q}) = V_1(\mathcal{E}) + V_2(\mathcal{Q})$$

and observe that

$$\begin{aligned}\dot{V} &\le -\sigma_n(R)\|\mathcal{E}\|^2 - (\lambda\sigma_2(\mathcal{L}) - C_1)\|\widetilde{\mathcal{Q}}\|^2\\ &\quad + (C_0\sqrt{N} + C_2)\|\mathcal{E}\|\,\|\widetilde{\mathcal{Q}}\| + (C_2\|\mathcal{S}\| + C_3\overline{\sigma}_G)\,\|\widetilde{\mathcal{Q}}\|.\end{aligned}$$

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

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

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

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

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

IV. Conclusions and Perspectives

We considered the problem of distributed estimation over undirected and connected graphs. The sensor nodes exchange information about their estimate, and the the communication between the sensor nodes is driven by a Poisson process. The algorithm implemented by each node is in the form of a stochastic hybrid system. We obtain pointwise and asymptotic bounds on the covariance of the estimation error for each sensor node, and show that these covariances converge practically to a constant matrix if the mean sampling rate is large enough.

There are some immediate research directions that emanate from this work. Going through the proof of Theorem III.1, it would be interesting to investigate if the bounds on $\|\mathcal{P}_i - S\|$ can be tightened. More specifically, it needs to be checked if instead of practical convergence, one can achieve asymptotic convergence for some fixed value of λ. In general, asymptotic convergence in distributed estimation is ensured without any sensor noises, so that the agents share a common equilibrium point but it remains to be seen if some structural design allows us to achieve asymptotic convergence despite different noise covariances in sensor units.

Another interesting direction is, to investigate a broader class of random processes for the communication between agents. While the memoryless Poisson counter treated in this paper results in diffusive coupling between the dynamics of 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* $W : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$, *and a function* $g : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, *such that*

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

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

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

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

The proof of Lemma A.1 is very similar to the proof of [6, 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 heterogenous agents.

References

[1] D. Crisan and B. Rozovskii, *The Oxford Handbook of Nonlinear Filtering*. Oxford University Press, 2011.

[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] 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.

[4] 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.

[5] 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.

[6] 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.

[7] J. Lee and H. Shim, "A tool for analysis and synthesis of heterogeneous multi-agent systems under rank-deficient coupling," *Automatica*, vol. 117, 2020.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *Proceedings of 46th IEEE Conf. on Decision and Control (CDC)*, 2007.

[14] ——, "Kalman-Consensus Filter : Optimality, stability, and performance," in *Proceedings of 48th IEEE Conf. on Decision and Control (CDC)*, 2009.

[15] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 138 – 162, 2007.

[16] M. Adès, P. E. Caines, and R. P. Malhamé, "Stochastic optimal control under Poisson-distributed observations," *IEEE Transactions on Automatic Control*, vol. 45, no. 1, pp. 3–13, 2000.

[17] J. P. Hespanha and A. R. Teel, "Stochastic impulsive systems driven by renewal processes," in *Proc. 17th International Symposium on Mathematical Theory of Networked Systems*, 2006, pp. 606–618.

[18] D. J. Antunes, J. P. Hespanha, and C. J. 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.

[19] 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.

[20] A. S. Matveev and A. V. 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.

[21] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, 2004.

[22] M. Huang and S. Dey, "Stability of Kalman filtering with Markovian packet losses," *Automatica*, vol. 43, pp. 598–607, 2007.

[23] 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.

[24] A. Jazwinski, *Stochastic processes and filtering theory*. New York: Dover Publications, Inc., 2007.

[25] A. Tanwani and O. Yufereva, "Error covariance bounds for suboptimal filters with Lipschitzean drift and Poisson-sampled measurements," *Automatica*, vol. 122, 2020.

[26] 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.

[27] 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.

[28] 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.