HAL
open science

Error Bounds for Locally Optimal Distributed Filters with Random Communication Graphs

Aneel Tanwani

To cite this version:

Aneel Tanwani. Error Bounds for Locally Optimal Distributed Filters with Random Communication Graphs. IEEE Conference on Decision and Control (CDC), Dec 2021, Austin, Texas, United States. 10.1109/CDC45484.2021.9682851 . hal-03476850

HAL Id: hal-03476850

https://laas.hal.science/hal-03476850

Submitted on 30 Dec 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.

Error Bounds for Locally Optimal Distributed Filters with Random Communication Graphs

Aneel Tanwani

***Abstract*— We consider the problem of analyzing the performance of distributed filters for continuous-time linear stochastic systems under certain information constraints. We associate an undirected and connected graph with the measurements of the system, where the nodes have access to partial measurements in continuous time. Each node executes a locally optimally filter based on the available measurements. In addition, a node communicates its estimate to a neighbor at some randomly drawn discrete time instants, and these activation times of the graph edges are governed by independent Poisson counters. When a node gets some information from its neighbor, it resets its state using a convex combination of the available information. Consequently, each node implements a filtering algorithm in the form of a stochastic hybrid system. We derive bounds on expected value of error covariance for each node, and show that they converge to a common value for each node if the mean sampling rates for communication between nodes are large enough.**

***Index Terms*— Stochastic hybrid system; distributed filtering; graph theory; random communication.**

I. Introduction

When dealing with systems over networks, the information available for decision making is constrained by the communication channels and the underlying interconnections. Recent advancements in networked systems have led to the motivation for distributed estimation and filtering problems. In this paper, we consider the analysis and design of filters under certain information constraints induced by (a) distributed measurements described by a graph; and (b) frequency of communication between the graph nodes. In the literature on stochastic systems, the filter design, or analysis, has received continued attention. An important question addressed in these works is to quantify the performance of the filters by analyzing the bounds on the evolution of error covariance matrix [1]. 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 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 [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 consensus 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]. On the other hand, some recent dissertations have addressed the problem of distributed optimization and control over random graphs in discrete-time setting [15], [16].

An important issue from the point of implementation of filters over networks is to make the algorithms compatible with the underlying communication protocol. 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 [17] for optimal control and the papers [18], [19] for stability analysis of such systems. A recently published book chapter [20] 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 [21], [22], [23] 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 [24] 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 is proposed in [25]. Some analytical results on the performance of centralized continuous-time systems under random sampling appear in [26].

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. Compared to above mentioned references, the derivation and analysis of the expectation of error covariance bounds with respect to random (Poisson)

A. Tanwani is affiliated with CNRS – LAAS, University of Toulouse, CNRS, 31400 Toulouse, France (e-mail: aneel.tanwani@laas.fr).

This work is sponsored by the project CyPhAI, financed by ANR-JST CREST program with grant number ANR-20-JSTM-0001.

sampling in distributed setup is a unique feature of our work. The results presented in this article build on our recent work in [27]. For each edge describing the communication link between two sensor nodes, we model communication times by an independent Poisson counter, and derive the error covariance bounds quantifying the performance of these individual filters. In comparison to [27], instead of using the constant linear injection gains in the filters, we use time-varying gains that depend on error covariance bounds which makes the filters nonlinear. This motivates the use of term "locally optimal" in the title since we choose injection gains to minimize error covariance between two communication times. Due to this change in filter dynamics, even though the results remain qualitatively similar, the proofs of the main results in this paper require some additional technicalities.

II. Locally Optimal Distributed Filtering

Let us begin with the description of the system class and the formulation of the distributed filtering problem studied in this paper. In the process, we describe the graph structure representing the interconnection between sensor nodes, and the sampling process at which the connected nodes (or the neighbors) exchange information.

A. System Class

We consider dynamical systems modeled by linear stochastic differential equations of the form

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

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 a zero mean $\mathbb{R}^m$-valued standard Wiener process adapted to the filtration $\mathcal{F}_t \subset \mathcal{F}$, with the property that $\mathsf{E}[\mathrm{d}\omega(t)\,\mathrm{d}\omega(t)^\top] = I_m \mathrm{d}t$, 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 with (A, B) controllable, 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.

The centralized output measurement associated to the process (1) is of the form

$$\mathrm{d}y = Hx\,\mathrm{d}t + \mathrm{d}v \tag{2}$$

where $H \in \mathbb{R}^{p\times n}$ is a constant matrix, with (A, H) being observable, and $(v(t))_{t\geq 0}$ is a zero mean $\mathbb{R}^p$-valued standard Wiener process. The conventional filtering problem, with initial state having Gaussian distribution, deals with constructing a mean-square estimate of the state x_t, denoted by $\widehat{x}_t$ so that $\mathsf{E}[|x_t - \widehat{x}_t|^2 \,|\, (\mathrm{d}y(s))_{s\leq t}]$ is minimized. The optimal estimate which achieves this minimum value is $\mathsf{E}[x_t \,|\, (\mathrm{d}y(s))_{s\leq t}]$ and is computed recursively using a Kalman-Bucy filter. For the problem studied in this paper, it is assumed that the centralized measurements are not available and we address the filtering problem with similar assumptions on system data, but under different information constraints which are described next.

B. Information Structure

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,

$$\mathrm{d}y_i = H_i x\,\mathrm{d}t + \mathrm{d}v_i, \quad i = 1, \ldots, N, \tag{3}$$

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 (3), $v_i(t)$ is a zero mean $\mathcal{F}_t$-adapted standard Wiener process, taking values in $\mathbb{R}^{p_i}$, and $\mathsf{E}[\mathrm{d}v_i(t)\,\mathrm{d}v_i(t)^\top] = V_i \mathrm{d}t$, with $V_i \in \mathbb{R}^{p_i\times p_i}$ assumed to be positive definite. The optimal filter which minimizes the mean square estimation error conditioned upon the information available through the measurements $\{\mathrm{d}y_i(s) | s \leq t\}$ is,

$$\mathrm{d}\widehat{x}_i(t) = A\widehat{x}_i(t)\mathrm{d}t + P_i(t)H_i^\top V_i^{-1}(\mathrm{d}y_i(t) - H_i\widehat{x}_i(t)\mathrm{d}t) \tag{4a}$$

$$\dot{P}_i = AP_i + P_iA^\top - P_iH_i^\top V_i^{-1}H_iP_i + BB^\top, \tag{4b}$$

with $\widehat{x}_i(0) = \mathsf{E}[x(0)]$, and $P_i(t)$ is exactly the error covariance $\mathsf{E}[(x_i(t) - \widehat{x}_i(t))(x_i(t) - \widehat{x}_i(t))^\top \,|\, \mathrm{d}y_i(s), s \leq t]$ if $P_i(0) = \mathsf{E}[(x_i(0) - \widehat{x}_i(0))(x_i(0) - \widehat{x}_i(0))^\top]$.

a) Communication Graph: The sensor nodes are connected via a graph $\mathcal{G} = (\mathscr{V}, \mathscr{E})$, where $\mathscr{V} = \{1, \ldots, N\}$ is the set of graph nodes, and $\mathscr{E}$ contains all the edges defined by a subset of the pairs (i, j), $i \neq j$, $i, j \in \mathscr{V}$. We assume that the graph is undirected and connected. The neighbors of a node $i \in \mathscr{V}$ 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 nodes 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 $i \in \mathscr{V}$ 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 nonnegative matrix $\Pi = [\pi_{ij}] \in \mathbb{R}^{N\times N}$, defined as

$$\Pi := I_N - \epsilon\mathcal{L} \tag{5}$$

where $0 < \epsilon \leq \min_{i\in\mathscr{V}} \frac{1}{|\mathcal{N}_i|}$ 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) Random Sampling: 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 $(i, j) \in \mathscr{E}$, 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

- for each $(i,j) \in \mathscr{E}$, the sensor nodes $i,j \in \mathscr{V}$ 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

$$N_t^{ij} := \sup\{k \in \mathbb{N} \mid \tau_k^{ij} \le t\} \quad \text{for } t \ge 0 \tag{6}$$

and assume in addition that, tor each $(i,j) \in \mathscr{E}$, $(N_t^{ij})_{t\ge 0}$ is a continuous-time stochastic process such that $\tau_{N_t^{ij}}^{ij} \to +\infty$ almost surely as $t \to +\infty$. 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 communicate with each other up to and including time t. For the sake of computational tractability, it is stipulated that

- For each $(i,j) \in \mathscr{E}$, $(N_t^{ij})_{t\ge 0}$ is an independent *Poisson process of intensity* $\lambda_{ij} > 0$. That is, $(N_t^{ij})_{t\ge 0}$ is a Markov process taking values in $\mathbb{N}$, has independent increments, and satisfies $N_0^{ij} = 0$, and for $h \searrow 0$ and $t \ge 0$,

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

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

Because of the arrival of new information at random times, the estimate $\widehat{x}_i$, $i \in \mathscr{V}$, gets updated. To describe this update rule, we associate with each node $i \in \mathscr{V}$, 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 $i \in \mathscr{V}$ exchanges information with any of its neighbor. We recall that N_t^i is also a Poisson process of intensity $\lambda_i := \sum_{j\in\mathcal{N}_i} \lambda_{ij}$. The times at which N_t^i gets incremented are denoted by $\tau_{N_t^i}$. We can now introduce the activation set $\mathcal{A}_t^i$,

$$\mathcal{A}_t^i := \left\{ j \in \mathcal{N}_i \,\middle|\, N_t^{ij} - N_{\underline{t}}^{ij} \ne 0, \underline{t} = \tau_{N_t^i - 1} \right\},$$

so that, at communication times $t_c = \tau_{N_t^i}$, the set $\mathcal{A}_{t_c}^i$ describes the neighbors of node $i \in \mathscr{V}$ that communicate their estimate to node $i \in \mathscr{V}$. Consequently, at $t_c = \tau_{N_t^i}$, we update the state estimate as follows:

$$\widehat{x}_i(t_c^+) = \sum_{j\in\mathcal{A}_{t_c}^i} \pi_{ij}\widehat{x}_j(t_c^-) + \Big(1 - \sum_{j\in\mathcal{A}_{t_c}^i} \pi_{ij}\Big)\widehat{x}_i(t_c^-), \tag{7}$$

where π_{ij} are the elements of the matrix Π introduced in (5). If $e_i := x - \widehat{x}_i$ denotes the estimation error, then because of this update rule, it is observed that,

$$e_i(t_c^+)e_i^\top(t_c^+) \le \Big(1 - \sum_{j\in\mathcal{A}_{t_c}^i} \pi_{ij}\Big) e_i(t_c^-)e_i^\top(t_c^-) + \sum_{j\in\mathcal{A}_{t_c}^i} \pi_{ij} e_i(t_c^-)e_i^\top(t_c^-)$$

which is a direct consequence of the following lemma, whose proof appears in [27, Lemma III.3]:

Lemma 1. *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 $\gamma_j \in [0,1]$, then*

$$zz^\top \le \sum_{j=1}^m \gamma_j x_j x_j^\top. \tag{8}$$

C. Summary of Filtering Algorithm

So far, we have specified the information available to each sensor node and a filtering algorithm, (4), (7), which uses this available information. If $\mathcal{Y}_t^i$ denotes the information available to sensor node $i \in \mathscr{V}$ up till time $t \in [0, +\infty[$, then we can write $\mathcal{Y}_t^i = \{(\mathrm{d}y_i(s), \widehat{x}_j(\tau_{N_s^i})) \mid s \le t, j \in \mathcal{N}_i\}$. Our goal is to quantify the performance of these distributed filters by computing a bound on expected value of the error covariance matrices, that is, $\mathsf{E}[\mathsf{E}[(x(t) - \widehat{x}_i(t))(x(t) - \widehat{x}_i(t))^\top \mid \mathcal{Y}_t^i]]$, for $t \ge 0$. Here, the outer expectation is with respect to the random update times, and the inner expectation is with respect to the noise process in the state and output equation. The estimate computed by each node $i \in \mathscr{V}$ is obtained by executing the following steps:

- Integrate (4a) and (4b) over the interval $[\tau_{N_t^i}, \tau_{N_t^i+1}[$,
- At $t_c = \tau_{N_t^i}$, reset the state $\widehat{x}_i$ via (7), and set

$$P_i(t_c^+) = \big(1 - \sum_{j\in\mathcal{A}_{t_c}^i} \pi_{ij}\big)P_i(t_c^-) + \sum_{j\in\mathcal{A}_{t_c}^i} \pi_{ij}P_j(t_c^-) \tag{9}$$

  with $P_i(0) \ge \mathsf{E}[(x(0) - \widehat{x}_i(0))(x(0) - \widehat{x}_i(0))]$.

As a first step in obtaining the desired bounds, one immediately observes that, for each $i \in \mathscr{V}$, if we fix the times at which node $i \in \mathscr{V}$ communicates with its neighbors, then

$$\mathsf{E}[(x(t) - \widehat{x}_i(t))(x(t) - \widehat{x}_i(t))^\top \mid \mathcal{Y}_t^i] \le P_i(t), \quad t \ge 0 \tag{10}$$

where P_i is described by (4b), (9). The remaining task therefore is to compute expectation with respect to the distributions assigned to the times at which information between sensor nodes takes place.

III. Main Results

The basic problem studied in this paper is the performance of the distributed filters proposed in the previous section. In particular, we want to relate the mean sampling rates λ_{ij}, corresponding to the edges $(i,j) \in \mathscr{E}$, with the bounds on the error covariance. As our first main result, we compute an upper bound on the expectation (with respect to sampling process N_t^i) 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 $t \ge 0$.

Theorem 2. *Consider system* (1) *with distributed measurements* (3) *and the corresponding hybrid filters* (4), (7), (9) *linked together by an undirected and connected graph $\mathcal{G} = (\mathscr{V}, \mathscr{E})$. For an edge $(i,j) \in \mathscr{E}$, if the communication between nodes $i,j \in \mathscr{V}$ takes place at random times generated by a Poisson process of intensity $\lambda_{ij} > 0$, then for each $i = 1, \cdots, N$, it holds that*

$$\mathsf{E}[\mathsf{E}[(x(t) - \widehat{x}_i(t))(x(t) - \widehat{x}_i(t))^\top \mid \mathcal{Y}_t^i]] \le \mathcal{P}_i(t), \tag{11}$$

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

$$\dot{\mathcal{P}}_i = A\mathcal{P}_i + \mathcal{P}_i A^\top - \mathcal{P}_i H_i^\top V_i^{-1} H_i \mathcal{P}_i + BB^\top + \sum_{j\in\mathcal{N}_i} \lambda_{ij}\pi_{ij}(\mathcal{P}_j - \mathcal{P}_i). \quad (12)$$

The proof of Theorem 2 is very similar to the proof of [27, Theorem III.1, item 1)] and is omitted due to space constraints. The result of Theorem 2 provides a bound on the evolution of error covariance for each node in terms of a differential equation. These equations are quadratic (and hence nonlinear), driven by a constant term which corresponds to the noise level in the system, and are interconnected by some coupling term. 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 $i \in \mathscr{V}$. 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. The next result relates to the asymptotic behavior of the coupled differential equations (12).

Theorem 3. *For* $i = 1,\dots,N$*, consider the matrix-valued equations* (12) *and assume that* $\lambda_{ij} = \lambda$ *is the same for each* $(i,j) \in \mathscr{E}$*. Let* $S \in \mathbb{R}^{n\times n}$ *be symmetric positive semidefinite matrix satisfying*

$$0 = AS + SA^\top - \frac{1}{N} S \left(\sum_{i=1}^{N} H_i^\top V_i^{-1} H_i \right) S + BB^\top. \quad (13)$$

Then for every $\delta > 0$*, there exists* $\lambda > 0$ *sufficiently large, such that the corresponding solution of* (12) *satisfies*[1]

$$\limsup_{t\to\infty} \|\mathcal{P}_i(t) - S\| \le \delta. \quad (14)$$

The proof of Theorem 3 is carried out in Section IV. To conclude this section, we provide some remarks about our main results.

Remark 1. The injection gains used in (4) over an interval $[\tau_{N_t^i}, \tau_{N_t^i+1}[$ do not need any information about how the other filters in the network choose their gains. Moreover, they minimize the value of P_i in (4b) over the class of linear time-varying gains. The latter statement follows from the fact that

$$(A-L_iH_i)P_i + P_i(A_i-L_iH_i)^\top + L_iV_iL_i^\top = (A-\overline{L}_iH_i)P_i + P_i(A_i-\overline{L}_iH_i)^\top + \overline{L}_iV_i\overline{L}_i^\top - (\overline{L}_i - L_i)V_i(\overline{L}_i - L_i)^\top \quad (15)$$

for any constant matrix $\overline{L}_i \in \mathbb{R}^{n\times p_i}$, and $L_i = P_iH_i^\top V_i^{-1}$. This shows that the filters (4) perform better than the constant linear gains proposed by the author in [27].

Remark 2. In Theorem 3, we basically study convergence of the differential equations (12) which contain quadratic nonlinearities. In general, such nonlinearities result in semi-global convergence, that is, the solutions converge starting from initial conditions in a compact set. However, because of the minimum property described in Remark 1, we get global convergence with no restrictions on the initial condition. However, the convergence is not necessarily asymptotic, but only up to a neighborhood of a fixed point, which we often call practical convergence. The practical aspect of the convergence is unavoidable since we allow different noise covariance levels for each filter.

Remark 3. In the formulation of Theorem 2, since we associate a different Poisson process to each link, the nodes communicate with each other at different times. However, we associate the same sampling rate λ with each edge (i,j). The motivation for doing so is that, when we write the collective dynamics for each node, the last term in (12) is written as a scalar multiple of the Laplacian. If we assume that the process N_t^{ij} associated with edge (i,j) has intensity λ_{ij}, then the arguments required for establishing practical convergence are more involved and are not carried out in this paper. One would expect that if each $\lambda_{ij} > 0$ is large enough, then we do get practical convergence.

IV. Proof of Theorem 3

For the proof of Theorem 3, we introduce the following notation:

$$L_i(t) := \mathcal{P}_i(t)H_i^\top V_i^{-1}, \quad A_i(t) := A - L_i(t)H_i$$

where we will drop the time argument in the sequel. Furthermore, let

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

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

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

Then, we can write

$$\begin{aligned} \dot{\mathcal{P}} &= F(\mathcal{P}) + G(\mathcal{P}) + \lambda(\Pi \otimes I_n)\mathcal{P} - \lambda(I_N \otimes I_n)\mathcal{P} \\ &= F(\mathcal{P}) + G(\mathcal{P}) - \lambda\epsilon(\mathcal{L} \otimes I_n)\mathcal{P}, \end{aligned} \quad (16)$$

where $\otimes$ denotes the Kronecker product, and we recalled the definition of Π to write $\Pi - I_N = -\epsilon\mathcal{L}$.

Proposition 4. *The solutions of* (16) *stay bounded over the interval* $[0,+\infty[$.

The proof of Proposition 4 is based on the observation made in Remark 1 combined with the fact that the convergence of solutions of (16), with time-invariant L_i, has been established in author's work [27, Theorem III.1]. In the sequel, we establish the practical convergence of the solutions of (16). The analysis is based on rewriting the dynamics in new coordinates and providing a Lyapunov function in those coordinates.

[1] When taking the norm of a matrix, we refer to Frobenius norm.

To describe the coordinate transformation, we recall that, 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 to the eigenvalue $\sigma_1(\mathcal{L}) = 0$, 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 $\widetilde{U}^\dagger \in \mathbb{R}^{N\times(N-1)}$ 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,

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

Let $\mathcal{S}$ denote the solution of the differential equation

$$\begin{aligned}\dot{\mathcal{S}} &= A\mathcal{S} + \mathcal{S}A^\top - \frac{1}{N}\mathcal{S}\left(\sum_{i=1}^N H_i^\top V_i^{-1} H_i\right)\mathcal{S} + BB^\top \\ &= A\mathcal{S} + \mathcal{S}A^\top - \mathcal{S}H^\top\widetilde{V}^{-1}H\mathcal{S} + BB^\top\end{aligned}$$

where we observe that $\frac{1}{N}\left(\sum_{i=1}^N H_i^\top V_i^{-1}H_i\right) = H^\top\widetilde{V}^{-1}H$. Under the assumption that (A,H) is observable and (A,B) is controllable, it holds that $\limsup_{t\to\infty}\mathcal{S}(t) = S$, where S is positive semidefinite matrix satisfying (13). In particular, $\mathcal{S}(t)$ is uniformly bounded for $t\geq 0$. We introduce the variable $\mathcal{E}$ to denote the difference between the mean value of $\mathcal{P}_i$, $i = 1,\ldots,N$, 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

$$\begin{aligned}\dot{\mathcal{E}} &= \dot{\mathcal{Q}}_1 - \dot{\mathcal{S}} = \frac{1}{N}\sum_{i=1}^N \dot{\mathcal{P}}_i - \dot{\mathcal{S}} \\ &= A\mathcal{E} + \mathcal{E}A^\top - \mathcal{E}H^\top\widetilde{V}^{-1}H\mathcal{E} + g_1(\mathcal{E},\mathcal{S},\widetilde{\mathcal{Q}}) \qquad (17)\end{aligned}$$

for some matrix-valued function $g_1(\cdot)$. Due to the boundedness of $\mathcal{P}_i$ and $\mathcal{S}$, it can be shown that $\|g_1(\mathcal{E},\mathcal{S},\widetilde{\mathcal{Q}})\| \leq M_1\|\widetilde{\mathcal{Q}}\|$ for some $M_1\geq 0$. Also, we can write

$$\begin{aligned}\dot{\widetilde{\mathcal{Q}}} &= -\lambda\epsilon(\Lambda\otimes I_n)\widetilde{\mathcal{Q}} + \frac{1}{\sqrt{N}}(\widetilde{U}^\top\otimes I_n)\left(F(\mathcal{P}) + G(\mathcal{P})\right) \\ &= -\lambda\epsilon(\Lambda\otimes I_n)\widetilde{\mathcal{Q}} + \frac{1}{\sqrt{N}}(\widetilde{U}^\top\otimes I_n)\Big[(F+G)\cdot \\ &\qquad ((1_N\otimes I_n)\mathcal{Q}_1 + \sqrt{N}(\widetilde{U}^\dagger\otimes I_n)\widetilde{\mathcal{Q}})\Big] \\ &= -\lambda\epsilon(\Lambda\otimes I_n)\widetilde{\mathcal{Q}} + g_2(\mathcal{E},\mathcal{S},\widetilde{\mathcal{Q}}), \qquad (18)\end{aligned}$$

for some matrix-valued function $g_2(\cdot)$. Once again, due to the boundedness of the solutions, we can find a constant $M_2\geq 0$ such that $\|g_2(\mathcal{E},\mathcal{S},\widetilde{\mathcal{Q}})\| \leq M_2\|\mathcal{E}\| + M_3$.

For stability analysis of the process $\mathcal{P}$, we equivalently analyze the stability of (17) and (18). Towards this end, we recall that, if the matrix S that satisfies (13) then $(A - SH^\top\widetilde{V}^{-1}H)$ is a Hurwitz matrix, and there exists a positive definite matrix R such that

$$(A - SH^\top\widetilde{V}^{-1}H)R + R(A - SH^\top\widetilde{V}^{-1}H)^\top \leq -I. \qquad (19)$$

Let us introduce the functions

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

To analyze the evolution of $\mathcal{V}_1$ along the solutions of (17), we compute the bound on the derivative of $\mathcal{V}_1$. Using the cyclic property of the trace operator, we get

$$\begin{aligned}\dot{\mathcal{V}}_1 &= \frac{1}{2}\operatorname{tr}\left[\operatorname{Sym}((R\overline{A} + \overline{A}^\top R)\mathcal{E}R\mathcal{E})\right] \\ &\qquad + \operatorname{tr}\left[\operatorname{Sym}\left(R\mathcal{E}Rg_1(\mathcal{E},\mathcal{S},\widetilde{\mathcal{Q}})\right)\right] \\ &\leq -\operatorname{tr}(\mathcal{E}R\mathcal{E}) + C_0\,\|\mathcal{E}\|\,\|g_1(\mathcal{E},\mathcal{S},\widetilde{\mathcal{Q}})\| \\ &\leq -\sigma_1(R)\|\mathcal{E}\|^2 + C_0\,M_1\,\|\mathcal{E}\|\,\|\widetilde{\mathcal{Q}}\|\end{aligned}$$

where we used the bounds on the trace of products from [28], $\operatorname{tr}(\mathcal{E}R\mathcal{E}) = \operatorname{tr}(R\mathcal{E}^2) \geq \sigma_1(R)\operatorname{tr}(\mathcal{E}^2)$, with $\sigma_1(R)$ being the smallest eigenvalue of R and $C_0\geq 0$ is some constant.

On the other hand, using similar techniques, we can bound the derivative of $\mathcal{V}_2$ along the solutions of (18) as follows:

$$\begin{aligned}\dot{\mathcal{V}}_2 &\leq -\lambda\epsilon\,\sigma_2(\mathcal{L})\operatorname{tr}(\widetilde{\mathcal{Q}}\widetilde{\mathcal{Q}}^\top) + C_1\,\operatorname{tr}(\widetilde{\mathcal{Q}}\widetilde{\mathcal{Q}}^\top) \\ &\qquad + C_2M_2\,\|\mathcal{E}\|\,\|\widetilde{\mathcal{Q}}\| + C_3M_3\,\|\widetilde{\mathcal{Q}}\|\end{aligned}$$

for some appropriate positive scalars $C_1, C_2, C_3 \geq 0$. For the combined dynamical system (17), (18), we now consider the proper, positive definite Lyapunov function

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

and observe that

$$\begin{aligned}\dot{\mathcal{V}} &\leq -\sigma_n(R)\|\mathcal{E}\|^2 - (\lambda\epsilon\sigma_2(\mathcal{L}) - C_1)\|\widetilde{\mathcal{Q}}\|^2 \\ &\qquad + (C_0M_1 + C_2M_2)\|\mathcal{E}\|\,\|\widetilde{\mathcal{Q}}\| + C_3M_3\,\|\widetilde{\mathcal{Q}}\|.\end{aligned}$$

To get a bound on $V(\mathcal{E},\widetilde{\mathcal{Q}})$, bet us recall the following lemma:

Lemma 5 ([6]). *Consider a function* $W : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}$*, 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*

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

where χ *is a continuous nondecreasing function and* $\chi(s) = O(s)$ *as* $s \to 0$*.*

Applying Lemma 5, it follows that for every $\delta > 0$, there exists $\lambda > 0$ such that

$$\limsup_{t\to\infty} \|\mathcal{E}(t)\|^2 + \|\widetilde{\mathcal{Q}}(t)\|^2 \leq \frac{\delta^2}{\sqrt{N}}$$

For S satisfying (13), we have $\lim_{t\to\infty} \mathcal{S}(t) = S$, which leads to the following bound for each $i \in \mathscr{V}$,

$$\begin{aligned}\limsup_{t\to\infty} \|\mathcal{P}_i(t) - S\| &= \limsup_{t\to\infty} \|\mathcal{P}_i(t) - \mathcal{S}(t)\| \\ &\leq \limsup_{t\to\infty} \sqrt{N}\sqrt{\|\mathcal{E}(t)\|^2 + \|\widetilde{\mathcal{Q}}(t)\|^2} \leq \delta.\end{aligned}$$

Hence, we obtain (14) and this completes the proof of Theorem 3.

V. 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 the communication between the sensor nodes, along each edge, is driven by an independent 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.

A possible extension is to generalize the main results for a broader class of random processes for the communication between nodes. 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.

References

[1] A. Taghvaei and P. G. Mehta, "Error analysis of the stochastic linear feedback particle filter," in *Proc. of 57th IEEE Conference on Decision and Control*, 2018, pp. 7194–7199.

[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 *Proc. of 46th IEEE Conf. on Decision and Control*, 2007.

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

[15] B. Touri, "Product of random stochastic matrices and distributed optimization," Ph.D. dissertation, Department of Electrical and Computer Engineering, Univ. of Illinois at Urbana-Champaign, 2011.

[16] S. Etesami, "Potential-based analysis of social, communication, and distributed networks," Ph.D. dissertation, Department of Electrical and Computer Engineering, Univ. of Illinois at Urbana-Champaign, 2015.

[17] 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.

[18] 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.

[19] 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.

[20] 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. Springer Nature, 2018, pp. 209–246.

[21] 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.

[22] 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.

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

[24] 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.

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

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

[27] A. Tanwani, "Suboptimal filtering over sensor networks with random communication," *IEEE Transactions on Automatic Control*, 2022, DOI: 10.1109/TAC.2021.3116180.

[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.