Error covariance bounds for suboptimal filters with Lipschitzian drift and Poisson-sampled measurements

Aneel Tanwani, Olga Yufereva

Error Covariance Bounds for Suboptimal Filters with Lipschitzian Drift and Poisson-Sampled Measurements

Aneel Tanwani Olga Yufereva

Abstract

We design suboptimal filters for a class of continuous-time nonlinear stochastic systems when the measurements are assumed to arrive randomly at discrete times under a Poisson distribution. The proposed filter is a dynamical system with a differential equation and a reset map which updates the estimate whenever a new measurement is received. We analyze the performance of the proposed filter by computing the expected value of the error covariance which is described by a differential equation. We study throughly the conditions under which the error covariance remains bounded, which depend on the system data and the mean sampling rate associated with the measurement process. We also study the particular cases when the error covariance is seen to decrease with the increase in the sampling rate. For the particular case of linear filters, we can also compare the error covariance bounds with the case when the measurements are continuously available.

1 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 [5] provides an overview of the developments carried out in the area of filtering theory since its inception. While the problem is relatively well-studied for linear dynamical systems with closed form solutions, the analysis and implementation of nonlinear filters has proven to be a rather challenging problem. For this reason, many research works have focused on providing relaxed versions of optimal filters for nonlinear systems, which are more amenable for implementation. However, from analysis perspective, relatively fewer works exist which rigorously quantify the performance of nonlinear filters. For these reasons, nonlinear filtering continues to be an active area of research where the search of effective implementation algorithms and analysis methods is of interest for may applications [8, 11, 37, 33, 39, 32, 17]. This article aims at making a contribution in the area of nonlinear filtering for continuous-time systems where the amount of information provided by the output measurements is not only limited by continuous white noise, but also by sampling in time; In particular, it is stipulated that the output measurement arrive randomly at some discrete time instants only. With this additional uncertainty due to random sampling of the output measurement, we study the problem of obtaining bounds on the covariance of the state estimation error for certain class of nonlinear dynamical systems with our proposed filtering algorithms.

The filtering problem aims at minimizing the expected value of squared estimation error at each time conditioned upon the measurements available up to that point. Under some mild assumptions, it is seen that the optimal estimator, minimizing the aforementioned function, is the expected value of the state variable conditioned upon the last available measurement. The key element in computing this estimate is thus to get the probability density function of the state conditioned upon the system output. Pioneering works related to describing the evolution of density function using partial differential equations have been carried out in [30, 20, 40, 10], see also the historical overview [27] and the book [22]. Analysis, or implementation, of such equations is of course a different, and possibly a difficult, task. With relevance to this article, the results on analyzing the performance of the filters by analyzing the estimation error are the most pertinent ones. In this regard, the works [12, 28] provide bounds on error covariance under certain structural assumption on the system dynamics. We impose similar assumptions

This is a preprint of a published article: A. Tanwani and O. Yufereva. Error Covariance Bounds for Suboptimal Filters with Lipschitzian Drift and Poisson-Sampled Measurements. Automatica 122, 2020. DOI: 10.1016/j.automatica.2020.109280

on our nonlinear system to propose a suboptimal filter and analyze its performance under randomly sampled measurements.

More specifically, the problem formulation studied in this paper is motivated by the need of controlling dynamical systems over networks [13]. A common feature of the communication across networks is that the underlying protocols allow transmission of measurements only at randomly chosen discrete time instants [35]. Certain works in the literature have studied this problem of stabilization and control with randomly sampled measurements: The reader may refer to [1] for optimal control and the papers [14, 2] for stability conditions. A recently published book chapter [34] 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 [23, 29, 15] 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, one can adopt the structure of continuous-discrete observer proposed in [16]. However, such filters also suffer from the problem of implementation.

There are not many works that analyze the performance of *continuous-time* nonlinear filters in the presence of communication errors, or measurement errors in general, which is the topic of this paper. For continuous-time linear systems, filtering problem with randomly sampled measurements has been studied in [25] and [26], where the authors compute a conditional density function and use it to get the desired expectation. In this article, instead, we focus on nonlinear dynamics with a certain structure and propose a filter directly for estimating the expected value of the state. This filter is described by a piecewise deterministic Markov process [6, 7] where the randomness is due to sampling times of the measurements. To analyze the performance of the proposed filter, it is seen that the resulting error dynamics are described by an Itô process with resets at random times. Using stochastic differential rule [18], we first compute an upper bound on the error covariance matrix for a given sequence of sampling times. Assuming that the sampling process is Poisson, we compute the expected value of the resulting bound which is a matrix differential equation. In the resulting analysis of this differential equation, we are aided by the tools appearing in [38] and [9]. Such methods have also found their utility in linear filtering problems [3].

The remainder of the paper is organized as follows: The dynamical system and the observation process is described in Section 2. In Section 3, we propose a filter which estimates the mean value of the state trajectory for fixed sampling times, and compute the bounds on resulting error covariance. The expected value of these bounds with respect to sampling times are derived in Section 4. We then study the evolution of the mean covariance bound with respect to time in Section 5, and provide conditions in Section 6 and Section 7 under which the error covariance stays bounded. Finally, in Section 8, we study dependence with respect to the sampling rate before concluding the article with some directions of future research, and an appendix which collects some known results from the literature used in our derivations.

2 Problem Formulation

2.1 System Class

We consider stochastic nonlinear systems of the form

$$\mathrm{d}x_t = f(x_t)\,\mathrm{d}t + G(x_t)\,\mathrm{d}\omega_t \tag{1a}$$

$$\mathrm{d}z_t = h(x_t)\,\mathrm{d}t + \mathrm{d}\eta_t, \tag{1b}$$

where $(x_t)_{t\geqslant 0}$ is an $\mathbb{R}^n$-valued diffusion process describing the state, and $(z_t)_{t\geqslant 0}$ describes an $\mathbb{R}^p$-valued continuous-time observation process. Let $(\Omega, \mathscr{F}, \mathsf{P})$ denote the underlying probability space. It is assumed that, for each $t \geqslant 0$, ω_t is an $\mathbb{R}^m$-valued standard Wiener process adapted to the filtration $\mathscr{F}_t \subset \mathscr{F}$, with the property that

$\mathsf{E}[\mathrm{d}\omega\,\mathrm{d}\omega^\top] = I_{m\times m}$, for each $t \geqslant 0$. Also, in the observation equation (1b), η_t is an $\mathscr{F}_t$-adapted standard Wiener process, taking values in $\mathbb{R}^p$, and $\mathsf{E}[\mathrm{d}\eta_t\,\mathrm{d}\eta_t^\top] = V_t \in \mathbb{R}^{p\times p}$, with V_t assumed to be positive definite. The drift term $f : \mathbb{R}^n \to \mathbb{R}^n$, the matrix of diffusion coefficients $G : \mathbb{R}^n \to \mathbb{R}^{n\times m}$, and the output map $h : \mathbb{R}^n \to \mathbb{R}^p$ are assumed to be continuous. It is also assumed that the processes $(\omega_t)_{t\geqslant 0}$, $(\eta_t)_{t\geqslant 0}$ are independent of each other, and do not depend on the state either. The solutions of the stochastic differential equation are given by

$$x_t - x_0 = \int_0^t f(x_s)\,\mathrm{d}s + \int_0^t G(x_s)\,\mathrm{d}\omega_s$$

where the last integral is interpreted in the sense of Itô, and is referred to as Itô stochastic integral. Further details about Itô's integral and its connection to stochastic differential equations can be found in the books [18, 22], or in a recent article [19] dedicated to the memory of K. Itô.

The conventional filtering problem deals with constructing a mean-square estimate of the state x_t, denoted by $\widehat{x}_t$ so that $\mathsf{E}[|x_t - \widehat{x}_t|^2 \,|\, (\mathrm{d}z_s)_{s\leqslant t}]$ is minimized. The optimal estimate which achieves this minimum value is $\mathsf{E}[x_t \,|\, (\mathrm{d}z_s)_{s\leqslant t}]$ and it is typically computed under the premise that the output measurements are available continuously at all time instants. But here we work under some constraints on the availability of the output measurements.

2.2 Observation Process

Our goal is to study the state estimation problem when the output measurements are available only at random times. The motivation to work with randomly time-sampled measurements comes from several applications, such as, communication over networks which allow information packets to be sent at some discrete randomly distributed time instants. Thus, we consider a monotone nondecreasing sequence $(\tau_n)_{n\in\mathbb{N}}$ taking values in $\mathbb{R}_{\geqslant 0}$ which denote the time instants at which the measurements are available for estimation. We introduce the process N_t defined as

$$N_t := \sup\big\{n \in \mathbb{N} \,\big|\, \tau_n \leqslant t\big\} \quad \text{for } t \in \mathbb{R},$$

and it is assumed that $(N_t)_{t\geqslant 0}$ is a Poisson stochastic process independent of the noise and the state processes. The discretized, and noisy, observation process is thus defined as[1]

$$y_{\tau_{N_t}} = h(x(\tau_{N_t})) + v_{\tau_{N_t}}, \qquad t \geqslant 0.$$

Our goal is to construct the estimate $\widehat{x}_t$, which minimizes the mean square estimation error, using the observations $\mathscr{Y}_t := \{y_{\tau_k} \,|\, k \leqslant N_t\}$. For a certain class of nonlinear functions f and h in (1), we will derive filtering equations to construct an estimate, and provide conditions on the sampling process which guarantee bounded covariance of the estimation error.

2.3 Problem Context

With time-sampled measurements, it is natural to consider the filtering problem:

$$\min_{\widehat{x}_t} \; \mathsf{E}\big[|x_t - \widehat{x}_t|^2 \,\big|\, \{y_{\tau_k}\}_{k\leqslant N_t}\big] \tag{2}$$

for each $t \geqslant 0$. It has been shown in [16] that the solution to the minimization problem (2) is given by

$$\widehat{x}_t = \mathsf{E}[x_t \,|\, \{y_{\tau_k}\}_{k\leqslant N_t}], \tag{3}$$

[1] Equation (1b) is formally equivalent to $y_t = h(x_t) + v_t$, with the identifications $y_t \sim \frac{\mathrm{d}z_t}{\mathrm{d}t}$ and $v_t \sim \frac{\mathrm{d}\eta_t}{\mathrm{d}t}$, so that v_t is a white Gaussian process, $v_t \sim N(0, V_{\tau_{N_t}})$; see [16, Chapter 4] for further details.

that is, we need to compute the expected value of x_t conditioned upon the last available output measurements. Because of the discrete nature of the output measurements, it is seen that, between two consecutive sampling instants τ_{N_t} and τ_{N_t+1}, the conditional expectation is obtained by computing the conditional probability $\varrho(x_t|\{y_{\tau_k}\}_{k\leqslant N_t})$. Between two successive measurements, we have

$$\varrho(x_t|\{y_{\tau_k}\}_{k\leqslant N_t}) = \varrho\,(x_t|y_{\tau_{N_t}}), \quad \tau_{N_t} \leqslant t < \tau_{N_t+1},$$

where $\varrho\,(x_t|y_{\tau_{N_t}})$ is obtained by the same evolution rule which governs the propagation of the probability density function of the state process $(x_t)_{t\geqslant 0}$, that is, the Fokker-Planck operator with the initial condition $\varrho\,(x_{\tau_{N_t}}|y_{\tau_{N_t}})$, see for example [4]. When a new measurement arrives, the conditional probability density function gets updated, and can be computed using Bayes' rule [16, Theorem 6.1]. From these equations for conditional distribution, one can arrive at the following equations to describe the evolution of the estimate in (3):

$$\frac{\mathrm{d}\widehat{x}_t}{\mathrm{d}t} = \widehat{f}(x_t) = \mathsf{E}[f(x_t)\,|\,\mathscr{Y}_t], \quad \tau_{N_t} \leqslant t < \tau_{N_t+1} \tag{4a}$$

$$\widehat{x}_{\tau_{N_t}} = \frac{\mathsf{E}[x_{\tau_{N_t}}\varrho(y_{\tau_{N_t}}|x_{\tau_{N_t}})\,|\,\mathscr{Y}_{\tau_{N_t-1}}]}{\mathsf{E}[\varrho(y_{\tau_{N_t}}|x_{\tau_{N_t}})\,|\,\mathscr{Y}_{\tau_{N_t-1}}]}. \tag{4b}$$

Similarly, the equation for the covariance of the estimation error conditioned upon the discrete measurements, $P_t := \mathsf{E}[\,|x_t-\widehat{x}_t|^2\,|\,\{y_{\tau_k}\}_{k\leqslant N_t}]$, is given by

$$\frac{\mathrm{d}P_t}{\mathrm{d}t} = (\mathsf{E}[x_t f^\top(x_t)\,|\,\mathscr{Y}_t] - \widehat{x}_t\widehat{f}^\top(x_t)) + (\mathsf{E}[f(x_t)x_t^\top\,|\,\mathscr{Y}_t] - \widehat{f}(x_t)\widehat{x}_t^\top) + \mathsf{E}[G(x_t)G(x_t)^\top\,|\,\mathscr{Y}_t] \tag{5a}$$

$$P_{\tau_{N_t}} = \frac{\mathsf{E}[x_{\tau_{N_t}}x_{\tau_{N_t}}^\top\varrho(y_{\tau_{N_t}}|x_{\tau_{N_t}})\,|\,\mathscr{Y}_{\tau_{N_t-1}}]}{\mathsf{E}[\varrho(y_{\tau_{N_t}}|x_{\tau_{N_t}})\,|\,\mathscr{Y}_{\tau_{N_t-1}}]} - \widehat{x}_{\tau_{N_t}}\widehat{x}_{\tau_{N_t}}^\top. \tag{5b}$$

In general, solving these equations is not possible. In fact, as noted in [16, Chapter 6], "the right-hand sides of (4) and (5) involve expectations that require the whole conditional density for their evaluation." Thus, the computation of the estimate (first moment) and the error covariance (the second moment) depend on all the higher order moments of the conditional density $\varrho(x_t|\mathscr{Y}_t)$, which remain unknown. To bypass aforementioned difficulties, we work with certain assumptions on the system dynamics which allow us to obtain a suboptimal estimate independently of the higher order moments. We analyze the performance of this suboptimal filter by getting bounds on the resulting error covariance which provides an upper bound on (2). For this choice of $\widehat{x}_t$, the value of $\mathsf{E}[|x_t-\widehat{x}_t|^2\,|\,\mathscr{Y}_t]$ depends on the random sampling times. We then compute the expectation with respect to sampling times and hence obtain a bound on $\mathsf{E}[\mathsf{E}[|x_t-\widehat{x}_t|^2\,|\,\mathscr{Y}_t]]$.

2.4 Structural Assumption on System Nonlinearities

To achieve the performance of the nonlinear filters, we introduce a structural assumption on the nonlinearities in the model (1).

(A1) There exist matrices $A\in\mathbb{R}^{n\times n}$, $C\in\mathbb{R}^{p\times n}$, the constants a,c such that for all $\delta\in\mathbb{R}^n$

$$|f(x+\delta)-f(x)-A\delta| \leqslant a\,|\delta|, \tag{6a}$$

$$|h(x+\delta)-h(x)-C\delta| \leqslant c|\delta|, \tag{6b}$$

and the pair (A,C) is detectable. Moreover, the diffusion term is constant, that is, $G(x)=G\in\mathbb{R}^{n\times m}$ for every $x\in\mathbb{R}^n$, with (A,G) stabilizable.

Our goal is to work out the filtering equations in the previous section for this particular class of nonlinear systems, and analyze the covariance of the estimation error as a function of the parameters in the model and the observation process.

Remark 2.1. In other words, the nonlinearities, that we consider, are to be seen as the perturbations in the linear terms and the size of these perturbations is handled by the constants a and c. If we write $f(x) = Ax + f_{\mathrm{nl}}(x)$, and $h(x) = Cx + h_{\mathrm{nl}}(x)$, then Assumption **(A1)** requires that f_{nl} is Lipschitz continuous with modulus $a \geqslant 0$, and that h_{nl} is Lipschitz continuous with modulus $c \geqslant 0$. This allows us to cover drift terms with at most linear growth, and output measurements (from sensors) with bias and saturation. Assumption **(A1)**, however, rules out nonlinearities in diffusion coefficient matrix G, and considering such nonlinearities is a topic of further investigations.

Remark 2.2. It is possible to consider time-dependence in nonlinearities f, g, h and the covariances of the noise. As a result, the matrices A, C, G, the constants a, c, and the covariance matrices $\mathsf{E}[\mathrm{d}\omega_t \mathrm{d}\omega_t^\top]$, $\mathsf{E}[\mathrm{d}\eta_t \mathrm{d}\eta_t^\top]$ can also be time-dependent. Such generalities can be addressed within the framework of this paper. In particular, the derivation of bounds in Proposition 3.1, Proposition 3.2, and Proposition 4.1 could be easily generalized without much differences. However, the results of Section 6, Section 7, and Section 8, where we analyze the stability of the differential equations associated with the bounds of the error covariance, require more careful analysis for time-varying systems.

3 Suboptimal Filtering

In what follows, we consider the dynamical system (1) under the assumption **(A1)**. As already mentioned, the exact filtering equations are not easy to analyse. Thus, we propose a filter structure, which is clearly a suboptimal solution and analyze the performance of this particular filter.

3.1 Filtering Equations

For given values of randomly sampled time instants $\{\tau_{N_t}\}_{t\geqslant 0}$, or in other words, a fixed sample path of $(N_t)_{t\geqslant 0}$, we define the estimate $\widehat{x}_t$ using the filter (7), whose sample paths are càdlàg functions; a function $\phi : [0,\infty) \to \mathbb{R}^n$ is a càdlàg function if $\lim_{s\searrow t} \phi_s = \phi_t$ for each $t \in [0,\infty)$, and $\phi_{t^-} := \lim_{s\nearrow t} \phi_s$ exists for each $t \in (0,\infty)$.

$$\mathrm{d}\widehat{x}_t = f(\widehat{x}_t)\,\mathrm{d}t \tag{7a}$$

$$\widehat{x}_{\tau^+_{N_t}} = \widehat{x}_{\tau^-_{N_t}} + K_{\tau_{N_t}}(y_{\tau_{N_t}} - h(\widehat{x}_{\tau^-_{N_t}})). \tag{7b}$$

where $\widehat{x}_t$ is the state estimate, and $K_{\tau_k} \in \mathbb{R}^{n\times p}$, $k \in \mathbb{N}$, are the injection gains that need to be designed appropriately. The estimate obtained from (7) is not necessarily a solution to the optimization problem (2) and, in general, not optimal for minimizing the mean square estimation error. However, by analyzing the performance of the suboptimal filter in (7), we can get an upper bound on the minimal mean square estimation error. Because of the randomness in sampling times, it is noted that the filter (7) is a piecewise-deterministic Markov process [6], or more generally, a stochastic hybrid system of the form considered in [36].

Hence, for the particular choice of filter in (7), we are interested in computing bounds on the expectation (with respect to the sampling process) of the error covariance and find conditions (involving the design parameter K_{τ_k}, $k \in \mathbb{N}$) so that the resulting covariance is bounded. The dependency of the error covariance on the sampling rate is also analyzed for specific cases.

3.2 Error Covariance Bounds with Arbitrary Injection Gains

The following statement describes a bound on the covariance of estimation error resulting from (7) for a given value of sampling times.

Proposition 3.1. *Consider system* (1) *under assumption* **(A1)** *and the filter* (7) *with* $\{\tau_k\}_{k\in\mathbb{N}}$ *fixed. Let the process* P_t, *with* $P_0 := \mathsf{E}[(x_0 - \widehat{x}_0)(x_0 - \widehat{x}_0)^\top]$, *be defined by*

$$\frac{\mathrm{d}P_t}{\mathrm{d}t} = AP_t + P_tA^\top + aP_t + a\,\mathrm{tr}(P_t)I_{n\times n} + GG^\top, \tag{8a}$$

for $\tau_{N_t} \leqslant t < \tau_{N_t+1}$, *and for* $k = N_t$, *let*

$$P_{\tau_k^+} = P_{\tau_k^-} + cP_{\tau_k^-} - K_{\tau_k} CP_{\tau_k^-} - P_{\tau_k^-} C^\top K_{\tau_k}^\top + K_{\tau_k}\Big((1+c)CP_{\tau_k^-}C^\top + (2c+c^2)\operatorname{tr}(P_{\tau_k^-})I_{p\times p} + V\Big)K_{\tau_k}^\top. \tag{8b}$$

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

$$\mathsf{E}[(x_t - \widehat{x}_t)(x_t - \widehat{x}_t)^\top \mid \mathscr{Y}_t] \leqslant P_t. \tag{9}$$

Proof. We denote the state estimation error by $e_t := x_t - \widehat{x}_t$, so that the resulting error dynamics are given by

$$\begin{aligned} \mathrm{d}e_t &= (f(x_t) - f(\widehat{x}_t))\mathrm{d}t + G\mathrm{d}\omega_t \\ &=: \widetilde{f}(x_t, \widehat{x}_t)\mathrm{d}t + G\mathrm{d}\omega_t \end{aligned} \tag{10a}$$

for $t \in [\tau_{N_t}, \tau_{N_t+1}[$, and for each $k = N_t$,

$$\begin{aligned} e_{\tau_k^+} &= x_{\tau_k^+} - \widehat{x}_{\tau_k^+} = x_{\tau_k^-} - \widehat{x}_{\tau_k^-} - K_{\tau_k}(y_{\tau_k} - h(\widehat{x}_{\tau_k^-})) \\ &= e_{\tau_k^-} - K_{\tau_k}(y_{\tau_k} - h(\widehat{x}_{\tau_k^-})). \end{aligned} \tag{10b}$$

Under the small nonlinearity assumption **(A1)**, we can write $f(x) = Ax + f_{\mathrm{nl}}(x)$ for some Lipschitz function $f_{\mathrm{nl}}(\cdot)$ with modulus a, so that

$$\widetilde{f}(x_t, \widehat{x}_t) = f(x_t) - f(\widehat{x}_t) = Ae + f_{\mathrm{nl}}(x_t) - f_{\mathrm{nl}}(\widehat{x}_t),$$

where $f_{\mathrm{nl}}(\cdot)$ satisfies the bound, $|f_{\mathrm{nl}}(x_t) - f_{\mathrm{nl}}(\widehat{x}_t)| \leqslant a|e_t|$ for every $x_t, \widehat{x}_t \in \mathbb{R}^n$. In what follows, we denote $\mathsf{E}[(x_t - \widehat{x}_t)(x_t - \widehat{x}_t)^\top \mid \mathscr{Y}_t]$ by $\mathsf{E}[e_t e_t^\top]$, and show that, for each $t \in [\tau_{N_t}, \tau_{N_t+1}[$,

$$\frac{\mathrm{d}\mathsf{E}[e_t e_t^\top]}{\mathrm{d}t} \leqslant A\mathsf{E}[e_t e_t^\top] + \mathsf{E}[e_t e_t^\top]A^\top + a\,\mathsf{E}[e_t e_t^\top] + a\operatorname{tr}(\mathsf{E}[e_t e_t^\top])\, I_{n\times n} + GG^\top$$

and that, at the sampling instants

$$\mathsf{E}[e_{\tau_k^+} e_{\tau_k^+}^\top] - \mathsf{E}[e_{\tau_k^-} e_{\tau_k^-}^\top] \leqslant P_{\tau_k^+} - P_{\tau_k^-}.$$

The bound in (9) then holds by observing that P_t satisfies (8a) over $[\tau_{N_t}, \tau_{N_t+1}[$, and (8b) at $t = \tau_{N_t}$, with $P_0 = \mathsf{E}[e_0 e_0^\top]$.

Continuous part Let us first show that for every $t \in [\tau_{N_t}, \tau_{N_t+1}[$ we have the inequality $\frac{\mathrm{d}\mathsf{E}[e_t e_t^\top]}{\mathrm{d}t} \leqslant \frac{\mathrm{d}P_t}{\mathrm{d}t}$. On this time interval, let us consider the Itô process in (10a):

$$\mathrm{d}e_t = \widetilde{f}(x_t, \widehat{x}_t)\mathrm{d}t + G\mathrm{d}\omega_t,$$

and the function $v : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ given by, $\mathbb{R}^n \ni e \mapsto v(e) = ee^\top \mathbb{R}^{n\times n}$, where an element of the matrix $v(e)$ is denoted by $v_{k,l}(e) = e^k e^l$. Consequently, $(v \circ e_t)_{t\geqslant 0}$ is an Itô process, and applying Itô's differential chain rule, we obtain

$$\mathrm{d}v_{k,l}(e_t) = \sum_{i=1}^n \frac{\partial v_{k,l}(e_t)}{\partial e_t^i}\mathrm{d}e_t^i + \frac{1}{2}\sum_{i,j=1}^n \frac{\partial^2 v_{k,l}(e_t)}{\partial e_t^i \partial e_t^j}\mathrm{d}e_t^i \mathrm{d}e_t^j.$$

Substituting $\mathrm{d}e_t^i = \widetilde{f}_i(x_t, \widehat{x}_t)\mathrm{d}t + \sum_{k=1}^m G_{i,k}(\mathrm{d}\omega_t)_k$ and using Itô's multiplication table [18, Page 154], we have $(\mathrm{d}\omega_t)_i \cdot (\mathrm{d}\omega_t)_j = \delta_{ij}\mathrm{d}t$, $\mathrm{d}t \cdot \mathrm{d}t = (\mathrm{d}\omega_t)_i \cdot \mathrm{d}t = \mathrm{d}t \cdot (\mathrm{d}\omega_t)_i = 0$, and hence

$$\begin{aligned} \mathrm{d}v_{k,l}(e_t) &= \sum_{i=1}^n \frac{\partial v_{k,l}(e_t)}{\partial e_t^i}\widetilde{f}_i(x_t, \widehat{x}_t)\mathrm{d}t + \frac{1}{2}\sum_{i,j=1}^n \frac{\partial^2 v_{k,l}(e_t)}{\partial e_t^i \partial e_t^j}\sum_{k_i,k_j=1}^m \delta_{k_i,k_j} G_{i,k_i} G_{j,k_j}\mathrm{d}t + \sum_{i=1}^n\sum_{j=1}^m \frac{\partial v_{k,l}(t,e_t)}{\partial e_t^i}G_{i,j}(\mathrm{d}\omega_t)_j \\ &= \sum_{i=1}^n \frac{\partial v_{k,l}(e_t)}{\partial e_t^i}\widetilde{f}_i(x_t, \widehat{x}_t)\mathrm{d}t + \frac{1}{2}\sum_{i,j=1}^n \frac{\partial^2 v_{k,l}(e_t)}{\partial e_t^i \partial e_t^j}\langle G_i, G_j\rangle\mathrm{d}t + \sum_{i=1}^n\sum_{j=1}^m \frac{\partial v_{k,l}(t,e_t)}{\partial e_t^i}G_{i,j}(\mathrm{d}\omega_t)_j, \end{aligned}$$

where $\langle G_i, G_j\rangle$ is the inner product of i-th and j-th rows of the matrix G. We are interested in computing $\frac{\mathrm{d}\mathsf{E}[v(e_t)]}{\mathrm{d}t}$, and we have the following expression for its coordinates:

$$\frac{\mathrm{d}\mathsf{E}[v_{k,l}(e_t)]}{\mathrm{d}t} = \lim_{\varepsilon\to 0}\frac{\mathsf{E}[v_{k,l}(e_{t+\varepsilon})]-v_{k,l}(e_t)}{\varepsilon} = \mathsf{E}\bigg[\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\mathrm{d}v(e_s)\bigg]$$

Substituting the expression for $\mathrm{d}v(e_s)$, and using the fact that ω_s is a Wiener process, we get

$$\frac{\mathrm{d}\mathsf{E}[v_{k,l}(e_t)]}{\mathrm{d}t} = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\mathsf{E}\bigg[\sum_{i=1}^n\frac{\partial v_{k,l}(e_s)}{\partial e_s^i}\widetilde{f}_i(x_s,\widehat{x}_s)+\frac{1}{2}\sum_{i,j=1}^n\frac{\partial^2 v_{k,l}(e_s)}{\partial e_s^i\partial e_s^j}\langle G_i,G_j\rangle\bigg]\mathrm{d}s$$

and hence, recalling that, $v_{k,l}(e_t)=e_t^k e_t^l$,

$$\frac{\mathrm{d}\mathsf{E}[v_{k,l}(e_t)]}{\mathrm{d}t} = \mathsf{E}\big[(e_l\widetilde{f}_k(x_t,\widehat{x}_t)+e_k\widetilde{f}_l(x_t,\widehat{x}_t))\big]+\frac{1}{2}\left(\langle G_k,G_l\rangle+\langle G_l,G_k\rangle\right).$$

Thus,

$$\begin{aligned}\frac{\mathrm{d}\mathsf{E}[v(e_t)]}{\mathrm{d}t} &= \mathsf{E}\big[\widetilde{f}(x_t,\widehat{x}_t)e_t^\top + e_t\widetilde{f}(x_t,\widehat{x}_t)^\top\big]+GG^\top\\ &= \mathsf{E}\big[Ae_te_t^\top+\widetilde{f}_{\mathrm{nl}}(x_t,\widehat{x}_t)e_t^\top+e_t(Ae_t)^\top\big]+\mathsf{E}\big[e_t\widetilde{f}_{\mathrm{nl}}(x_t,\widehat{x}_t)^\top\big]+GG^\top\\ &\leqslant A\mathsf{E}\big[e_te_t^\top\big]+\mathsf{E}\big[e_te_t^\top\big]A^\top+a\,\mathsf{E}\big[e_te_t^\top\big]+a\mathsf{E}\big[|e_t|^2\big]I_{n\times n}+GG^\top\\ &= A\mathsf{E}\big[e_te_t^\top\big]+\mathsf{E}\big[e_te_t^\top\big]A^\top+a\,\mathsf{E}\big[e_te_t^\top\big]+a\,\mathrm{tr}(\mathsf{E}\big[e_te_t^\top\big])\,I_{n\times n}+GG^\top,\end{aligned}$$

where we used the fact that $\widetilde{f}(x_t,\widehat{x}_t)=Ae_t+\widetilde{f}_{\mathrm{nl}}(x_t,\widehat{x}_t)$, with $\widetilde{f}_{\mathrm{nl}}(x_t,\widehat{x}_t):=f_{\mathrm{nl}}(x_t)-f_{\mathrm{nl}}(\widehat{x}_t)$, and the last inequality was obtained using the bound[2]

$$e\widetilde{f}_{\mathrm{nl}}^\top+\widetilde{f}_{\mathrm{nl}}e^\top\leqslant a\,ee^\top+a|e|^2I_{n\times n}$$

and the fact that $|e|^2=\mathrm{tr}\,(ee^\top)$.

Jump part To analyze the jump in error dynamics at time instants when the new measurements are received, we first observe that, under **(A1)**, $h(x)=Cx+h_{\mathrm{nl}}(x)$, where $h_{\mathrm{nl}}(x)$ satisfies the bound, $|h_{\mathrm{nl}}(x-e)-h_{\mathrm{nl}}(x)|\leqslant c|e|$, for all $x,e\in\mathbb{R}^n$. Since x_t is continuous, recalling that $e_t=x_t-\widehat{x}_t$ and (10b), it follows that, for $k=N_t$,

$$\begin{aligned}e_{\tau_k^+}-e_{\tau_k^-} &= -(\widehat{x}_{\tau_k^+}-\widehat{x}_{\tau_k^-}) = -K_{\tau_k}(y_{\tau_k}-h(\widehat{x}_{\tau_k^-}))\\ &= -K_{\tau_k}(Cx_{\tau_k}+h_{\mathrm{nl}}(x_{\tau_k^-})+v_{\tau_k}-C\widehat{x}_{\tau_k^-}-h_{\mathrm{nl}}(\widehat{x}_{\tau_k^-}))\\ &= -K_{\tau_k}(Ce_{\tau_k^-}+\widetilde{y}_{\tau_k^-}+v_{\tau_k})\end{aligned}$$

where $\widetilde{y}_{\tau_k^-}:=h_{\mathrm{nl}}(x_{\tau_k^-})-h_{\mathrm{nl}}(\widehat{x}_{\tau_k^-})$. With this relation for jump in error value at sampling instants, we can compute the change in the error covariance at the sampling instants. We do so by noting that

$$\begin{aligned}\mathsf{E}[e_{\tau_k^+}e_{\tau_k^+}^\top] &= \mathsf{E}\Big[\Big(e_{\tau_k^-}-K_{\tau_k}(Ce_{\tau_k^-}+\widetilde{y}_{\tau_k^-}+v_{\tau_k})\Big)\cdot\Big(e_{\tau_k^-}-K_{\tau_k}(Ce_{\tau_k^-}+\widetilde{y}_{\tau_k^-}+v_{\tau_k})\Big)^\top\Big]\\ &= \mathsf{E}[e_{\tau_k^-}e_{\tau_k^-}^\top]-\mathsf{E}[e_{\tau_k^-}(Ce_{\tau_k^-}+\widetilde{y}_{\tau_k^-}+v_{\tau_k})^\top K_{\tau_k}^\top+K_{\tau_k}(Ce_{\tau_k^-}+\widetilde{y}_{\tau_k^-}+v_{\tau_k})e_{\tau_k^-}^\top]\\ &\quad+\mathsf{E}[K_{\tau_k}(Ce_{\tau_k^-}+\widetilde{y}_{\tau_k^-}+v_{\tau_k})(Ce_{\tau_k^-}+\widetilde{y}_{\tau_k^-}+v_{\tau_k})^\top K_{\tau_k}^\top].\end{aligned}\tag{11}$$

[2] It follows from the fact that, for each $z,e\in\mathbb{R}^n$, and $f_{\mathrm{nl}}\in\mathbb{R}^n$ satisfying $|f_{\mathrm{nl}}|\leqslant a|e|$, we have that, $z^\top(e\widetilde{f}_{\mathrm{nl}}^\top+\widetilde{f}_{\mathrm{nl}}e^\top)z\leqslant 2\,|z^\top e|\cdot|z^\top\widetilde{f}_{\mathrm{nl}}|\leqslant 2\,a\,|z^\top e|\cdot|z|\,|e|$
$\leqslant a\,z^\top ee^\top z+a|e|^2z^\top z.$

We will now get a bound on the second and third terms on the right-hand side. To do so, we observe that, for each $K \in \mathbb{R}^{n\times p}$, $\tilde{y} \in \mathbb{R}^p$, $e \in \mathbb{R}^n$, such that $|\tilde{y}| \leqslant c|e|$, we have

$$-K\tilde{y}e^\top - e\tilde{y}^\top K^\top \leqslant cee^\top + c|e|^2 KK^\top. \tag{12}$$

This indeed follows from the fact that, for each $z \in \mathbb{R}^n$, $z^\top(-K\tilde{y}e^\top - e\tilde{y}^\top K^\top)z \leqslant 2\,|z^\top e| \cdot |z^\top K\tilde{y}| \leqslant 2c\,|z^\top e| \cdot |K^\top z|\,|e| \leqslant c\,z^\top ee^\top z + c|e|^2 z^\top KK^\top z$. Applying the inequality (12), the second term on the right-hand side of (11) can be bounded as follows:

$$\begin{aligned}
&-\mathsf{E}[e_{\tau_k^-}(Ce_{\tau_k^-} + \tilde{y}_{\tau_k^-} + v_{\tau_k})^\top K_{\tau_k}^\top - K_{\tau_k}(Ce_{\tau_k^-} + \tilde{y}_{\tau_k^-} + v_{\tau_k})e_{\tau_k^-}^\top] \\
&\leqslant -\mathsf{E}[e_{\tau_k^-}e_{\tau_k^-}^\top]C^\top K_{\tau_k}^\top - K_{\tau_k}C\mathsf{E}[e_{\tau_k^-}e_{\tau_k^-}^\top] + c\,\mathsf{E}[e_{\tau_k^-}e_{\tau_k^-}^\top] + c\,|e|^2 K_{\tau_k}K_{\tau_k}^\top \\
&= -\mathsf{E}[e_{\tau_k^-}e_{\tau_k^-}^\top]C^\top K_{\tau_k}^\top - K_{\tau_k}C\mathsf{E}[e_{\tau_k^-}e_{\tau_k^-}^\top] + c\,\mathsf{E}[e_{\tau_k^-}e_{\tau_k^-}^\top] + c\,\operatorname{tr}(\mathsf{E}[e_{\tau_k^-}e_{\tau_k^-}^\top])K_{\tau_k}K_{\tau_k}^\top,
\end{aligned}$$

where we used the fact that $|e_{\tau_k^-}|^2 = \operatorname{tr}(\mathsf{E}[e_{\tau_k^-}e_{\tau_k^-}^\top])$.

For the last term on the right-hand side of (11), we observe that, for each $C \in \mathbb{R}^{p\times n}$, $e \in \mathbb{R}^n$, and $\tilde{y} \in \mathbb{R}^p$ satisfying $|\tilde{y}| \leqslant c|e|$, we have

$$Ce\tilde{y}^\top + \tilde{y}e^\top C^\top \leqslant c\,Cee^\top C^\top + c|e|^2 I_{p\times p}. \tag{13}$$

This, in turn, leads to

$$\begin{aligned}
&\mathsf{E}[K_{\tau_k}(Ce_{\tau_k^-} + \tilde{y}_{\tau_k^-} + v_{\tau_k})(Ce_{\tau_k^-} + \tilde{y}_{\tau_k^-} + v_{\tau_k})^\top K_{\tau_k}^\top] \\
&= \mathsf{E}[K_{\tau_k}(Ce_{\tau_k^-}e_{\tau_k^-}^\top C^\top + Ce_{\tau_k^-}\tilde{y}_{\tau_k^-}^\top + \tilde{y}_{\tau_k^-}e_{\tau_k^-}^\top C^\top + \tilde{y}_{\tau_k^-}\tilde{y}_{\tau_k^-}^\top + v_{\tau_k}(Ce_{\tau_k^-} + \tilde{y}_{\tau_k^-})^\top + (Ce_{\tau_k^-} + \tilde{y}_{\tau_k^-})v_{\tau_k}^\top + v_{\tau_k}v_{\tau_k}^\top)K_{\tau_k}^\top] \\
&= K_{\tau_k}\mathsf{E}[Ce_{\tau_k^-}e_{\tau_k^-}^\top C^\top + Ce_{\tau_k^-}\tilde{y}_{\tau_k^-}^\top + \tilde{y}_{\tau_k^-}e_{\tau_k^-}^\top C^\top + \tilde{y}_{\tau_k^-}\tilde{y}_{\tau_k^-}^\top + v_{\tau_k}(Ce_{\tau_k^-} + \tilde{y}_{\tau_k^-})^\top + (Ce_{\tau_k^-} + \tilde{y}_{\tau_k^-})v_{\tau_k}^\top + v_{\tau_k}v_{\tau_k}^\top]K_{\tau_k}^\top \\
&\leqslant K_{\tau_k}\Big(C\mathsf{E}[e_{\tau_k^-}e_{\tau_k^-}^\top]C^\top + \mathsf{E}[c\,Ce_{\tau_k^-}e_{\tau_k^-}^\top C^\top + c|e_{\tau_k^-}|^2 I_{p\times p} + c^2|e_{\tau_k^-}|^2] + V\Big)K_{\tau_k}^\top \\
&= K_{\tau_k}\Big(C\mathsf{E}[\tilde{x}_{\tau_k^-}\tilde{x}_{\tau_k^-}^\top]C^\top + c\,C\mathsf{E}[e_{\tau_k^-}e_{\tau_k^-}^\top]C^\top + (c+c^2)\operatorname{tr}(\mathsf{E}[e_{\tau_k^-}e_{\tau_k^-}^\top])I_{p\times p} + V\Big)K_{\tau_k}^\top.
\end{aligned}$$

Plugging this last bound in (11), we get

$$\begin{aligned}
\mathsf{E}[e_{\tau_k^+}e_{\tau_k^+}^\top] \leqslant \mathsf{E}[e_{\tau_k^-}e_{\tau_k^-}^\top] - \mathsf{E}[e_{\tau_k^-}e_{\tau_k^-}^\top]C^\top K_{\tau_k}^\top - K_{\tau_k}C\mathsf{E}[e_{\tau_k^-}e_{\tau_k^-}^\top] + c\,\mathsf{E}[e_{\tau_k^-}e_{\tau_k^-}^\top] + K_{\tau_k}\Big((1+c)C\mathsf{E}[e_{\tau_k^-}e_{\tau_k^-}^\top]C^\top \\
+ (2c+c^2)\operatorname{tr}(\mathsf{E}[e_{\tau_k^-}e_{\tau_k^-}^\top]) + V\Big)K_{\tau_k}^\top
\end{aligned}$$

which is the desired statement. □

3.3 Minimizing Error Covariance with Suboptimal Filters

In the result of Proposition 3.1, we computed bounds on the estimation error covariance for arbitrary choice of gains. However, by choosing the gains in a certain dynamic manner such that K in (7b) depends on P_t obtained from equations of form (8a)-(8b), we can minimize P_t and $\mathsf{E}[e_t e_t^\top]$. For the chosen class of filters, this optimal choice of dynamic gain and the resulting covariance bound are indicated in the next result.

Proposition 3.2. *Consider system* (1) *under assumption* **(A1)** *and the filter* (7) *with* $\{\tau_k\}_{k\in\mathbb{N}}$ *given. If the output injection gains* K_{τ_k}, $k \in \mathbb{N}$, *in* (7b) *are chosen as*

$$K_{\tau_k} = P_{\tau_k^-}^\star C^\top M_{P_{\tau_k}}^{-1} \tag{14}$$

where $P^\star$ is obtained from

$$\frac{\mathrm{d}P_t^\star}{\mathrm{d}t} = AP_t^\star + P_t^\star A^\top + aP_t^\star + a\,\mathrm{tr}(P_t^\star)I + GG^\top, \tag{15a}$$

$$P_{\tau_k^+}^\star = (1+c)P_{\tau_k^-}^\star - P_{\tau_k^-}^\star C^\top M_{P_{\tau_k}}^{-1} C P_{\tau_k^-}^\star \tag{15b}$$

and $M_{P_{\tau_k}}$ is defined as[3]

$$M_{P_{\tau_k}} := (1+c)CP_{\tau_k^-}^\star C^\top + (2c+c^2)\,\mathrm{tr}\,(P_{\tau_k^-}^\star)I_{p\times p} + V, \tag{16}$$

then, for each $t \geqslant 0$, it holds that

$$\mathsf{E}[(x_t - \widehat{x}_t)(x_t - \widehat{x}_t)^\top] \leqslant P_t^\star \leqslant P_t \tag{17}$$

where P_t is obtained as a solution of (8).

The proof of this statement follows from Proposition 3.1. It is noted that K_{τ_k} only affects the bound P_t at the jump times via (8b). The right-hand side of (8b) is a quadratic convex function of K_{τ_k} which is minimized by choosing K_{τ_k} as in (14).

4 Expectation of Error Covariance

In the previous section, we computed a bound on the error covariance for a fixed sequence of time instants at which measurements are received. That is, P_t is an upper bound on $\mathsf{E}[(x_t-\widehat{x}_t)(x_t-\widehat{x}_t)^\top \,|\, \mathscr{Y}_t]$ along one particular sample path as the realization of P_t depends on the observed sampling times. It was initially stipulated that the random sampling instants are governed by a Poisson process. It is thus of interest to compute the expected value of P_t along all possible sample paths generated by Poisson sampling process.

Toward this end, it is noted that the evolution of P_t in (8) is governed by a piecewise deterministic process, where P_t is obtained from the differential equation (8a) between two sampling instants, and then at the sampling times, P_t is reset according to the equation (8b).

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

$$\mathscr{L}_{\mathrm{id}}(Q) = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\Big(\mathsf{E}[P(t+\varepsilon)\,|\,P(t)=Q] - Q\Big). \tag{18}$$

We will see that this operator defines the evolution of the expected value of the error covariance process, and it is computed in the following result:

Proposition 4.1. *Consider the process $(P_t)_{t\geqslant 0}$ given by* (8) *with $K_{\tau_k} = K \in \mathbb{R}^{n\times p}$ for each $k \in \mathbb{N}$, and assume that the sampling process $(N_t)_{t\geqslant 0}$ has intensity $\lambda > 0$. Let $\mathscr{P}_t := \mathsf{E}[P_t \,|\, P_0]$ denote the expected value of the covariance process at time $t \geqslant 0$. Then, it holds that*

$$\dot{\mathscr{P}}_t = \mathscr{L}_{\mathrm{id}}(\mathscr{P}_t) = (A-\lambda KC)\mathscr{P}_t + \mathscr{P}_t(A-\lambda KC)^\top + \Pi_a(\mathscr{P}_t) + GG^\top + \lambda\, c\mathscr{P}_t + \lambda K M_{\mathscr{P}_t} K^\top, \tag{19}$$

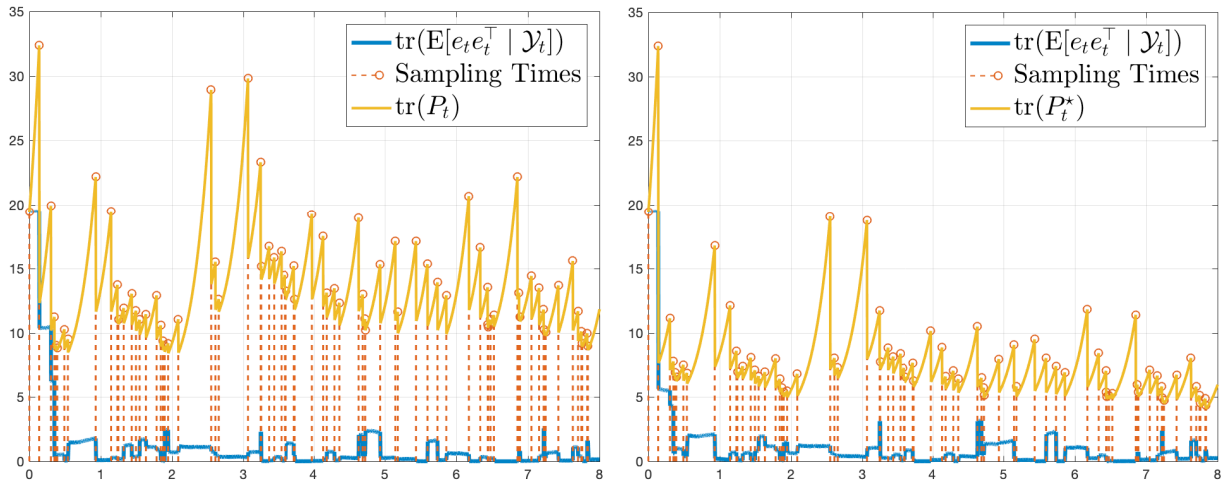
where the linear operator $\Pi_a : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ is defined as

$$\Pi_a(P) := a\,P + a\,\mathrm{tr}(P)I, \tag{20}$$

and the matrix $M_{\mathscr{P}_t} \in \mathbb{R}^{p\times p}$ is given by

$$M_{\mathscr{P}_t} := (1+c)C\mathscr{P}_t C^\top + (2c+c^2)\,\mathrm{tr}\,(\mathscr{P}_t)I_{p\times p} + V. \tag{21}$$

[3] For a given matrix P, the notation M_P used in (16) will be used extensively in the remainder of this article where we will change the matrix appearing in the subscript. With V assumed to be symmetric positive definite, M_P is a symmetric positive definite matrix, whenever P is symmetric positive semidefinite.



(a) Constant gain $K = [0.5\ 0.5]^\top$.

(b) Dynamic gain $K_t = P_t^\star C^\top M_{P_t^\star}^{-1}$.

Figure 1: Plot of $\mathrm{tr}(e_t e_t^\top)$ (blue curve) and $\mathrm{tr}(P_t)$ (yellow curve) along one sample path (of sampling times) for two choices of K_{τ_k} with $\lambda = 10$.

In particular, if we choose $K_{\tau_k} = P^\star_{\tau_k^-} C^\top M^{-1}_{P_{\tau_k}}$, *then the expected value of the error covariance process governed by* (15) *is*

$$\dot{\mathscr{P}}_t^\star = A\mathscr{P}_t^\star + \mathscr{P}_t^\star A^\top + \Pi_a(\mathscr{P}_t^\star) + GG^\top + \lambda c \mathscr{P}_t^\star - \lambda \mathscr{P}_t^\star C^\top M^{-1}_{\mathscr{P}_t^\star} C \mathscr{P}_t^\star. \tag{22}$$

Before providing the proof of this result, we present two examples for an illustration of the results presented so far.

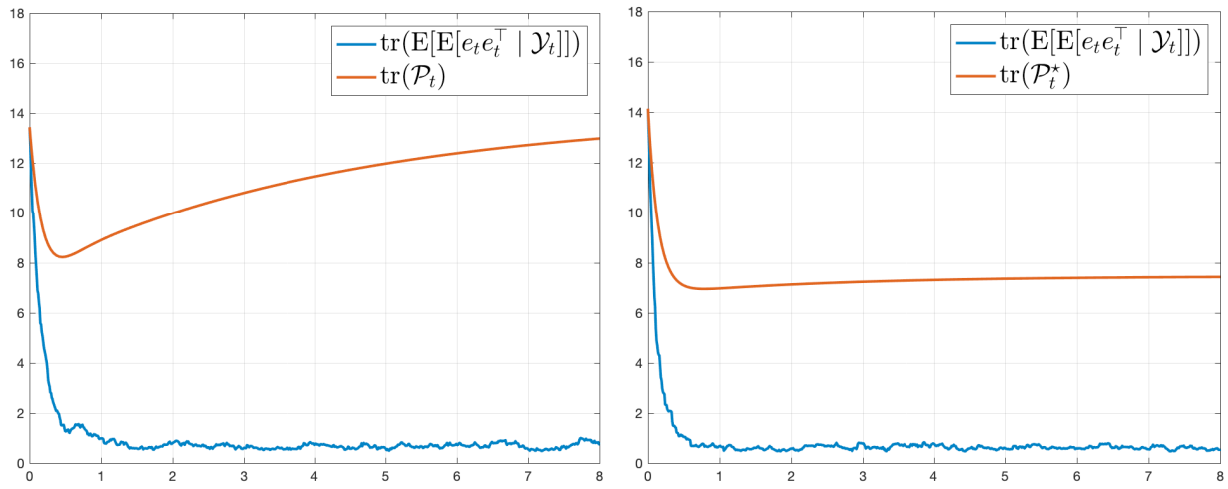
Example 1. Consider the nonlinear system described by the equations

$$\mathrm{d}x_{1,t} = \mathrm{sat}(x_{2,t})\mathrm{d}t + \mathrm{d}\omega_t, \quad \mathrm{d}x_{2,t} = (x_{1,t} - x_{2,t})\mathrm{d}t + \mathrm{d}\omega_t$$

with the measurement process $\mathrm{d}z_t = (1 + x_{1,t})\mathrm{d}t + \mathrm{d}\eta_t$. The noise processes are normalized as $\mathrm{d}\omega_t \sim \mathcal{N}(0,1)$ and $\mathrm{d}\eta_t \sim \mathcal{N}(0,1)$. The function $\mathrm{sat}(\cdot)$ represents the standard saturation function, that is, for each $z \in \mathbb{R}$, $\mathrm{sat}(z) = z - f_{\mathrm{nl}}(z)$, where f_{nl} is a Lipschitz function described as, $f_{\mathrm{nl}}(z) = z - 1$, if $z \geqslant 1$, $f_{nl}(z) = 0$ if $z \in [-1, 1]$, and $f_{\mathrm{nl}}(z) = z + 1$, if $z \leqslant -1$. Hence, by letting $A = \left[\begin{smallmatrix} 0 & 1 \\ 1 & -1 \end{smallmatrix}\right]$, $C = [1\ 0]$, $G = [1\ 1]^\top$, it is seen that Assumption **(A1)** holds with $a = 1$, and $c = 0$. Associating a Poisson process of intensity λ, the corresponding measurement equations are given by, $y_{\tau_{N_t}} = 1 + x_{1,\tau_{N_t}} + v_{\tau_{N_t}}$, where we take $v_{\tau_{N_t}} \sim \mathcal{N}(0,1)$ for each $t \geqslant 0$. For appropriately chosen $k_{1,\tau_{N_t}}, k_{2,\tau_{N_t}} \in \mathbb{R}$, the proposed filter is then described by

$$\begin{aligned} \mathrm{d}\widehat{x}_{1,t} &= \mathrm{sat}(\widehat{x}_{2,t})\mathrm{d}t, & \widehat{x}_{1,\tau^+_{N_t}} &= \widehat{x}_{1,\tau^-_{N_t}} + k_{1,\tau_{N_t}}(y_{\tau_{N_t}} - \widehat{x}_{1,\tau^-_{N_t}}) \\ \mathrm{d}\widehat{x}_{2,t} &= (\widehat{x}_{1,t} - \widehat{x}_{2,t})\mathrm{d}t, & \widehat{x}_{2,\tau^+_{N_t}} &= \widehat{x}_{2,\tau^-_{N_t}} + k_{2,\tau_{N_t}}(y_{\tau_{N_t}} - \widehat{x}_{1,\tau^-_{N_t}}). \end{aligned}$$

The simulation results for this example with two different choices of $K_{\tau_{N_t}} := [k_{1,\tau_{N_t}}\ k_{2,\tau_{N_t}}]^\top$ are reported in Figure 1 and Figure 2. For one particular choice of randomly determined sampling times, Figure 1a provides an illustration of (9) in Proposition 3.1, where we plot $\mathrm{tr}(\mathsf{E}[e_t e_t^\top \,|\, \mathscr{Y}_t])$ and the bound $\mathrm{tr}(P_t)$ obtained from solving (8), with constant gain $K_{\tau_{N_t}} = [0.5\ 0.5]^\top$ for each $t \geqslant 0$. Similarly, Figure 1b is obtained by choosing $K_{\tau_{N_t}} = P^\star_{\tau_{N_t}} C^\top M^{-1}_{P^\star_{\tau_{N_t}}}$ and is an illustration of inequality (17) stated in Proposition 3.2. In Figure 2, we compute expectation with respect to the sampling process. Figure 2a plots $\mathrm{tr}(\mathsf{E}[\mathsf{E}[ee^\top \,|\, \mathscr{Y}_t]])$, and its upper bound $\mathrm{tr}(\mathscr{P}_t)$ obtained from (19) with constant gain $K_{\tau_{N_t}} = [0.5\ 0.5]^\top$ for each $t \geqslant 0$, and choosing the intensity of Poisson sampling process



(a) Constant gain $K = [0.5\ 0.5]^\top$.

(b) Dynamic gain $K_t = P_t^\star C^\top M_{P_t^\star}^{-1}$.

Figure 2: Experimental value (blue curve) and theoretical upper bound (red curve) for expectation (with respect to sampling process) of trace of error covariance with $\lambda = 10$ for two choices of K.

$\lambda = 10$. In Figure 2b, we choose $K_{\tau_{N_t}} = P^\star_{\tau_{N_t}} C^\top M^{-1}_{P^\star_{\tau_{N_t}}}$, with the the theoretical upper bound $\mathrm{tr}(\mathscr{P}_t^\star)$ obtained from (22). We indeed observe that $\mathrm{tr}(\mathscr{P}_t^\star)$ has a smaller value than $\mathrm{tr}(\mathscr{P}_t)$, for each $t \geqslant 0$. It is not obvious, for what choices of gains $K_{\tau_{N_t}}$, we should get finite values of our bounds; such questions are addressed in subsequent sections.

Example 2. Consider the linear stochastic system

$$\mathrm{d}x_t = Ax_t \mathrm{d}t + G\mathrm{d}\omega_t \tag{23a}$$
$$\mathrm{d}z_t = Cx_t \mathrm{d}t + \mathrm{d}\eta_t. \tag{23b}$$

where $\mathrm{d}z_t$ is the measurement process, and the processes ω_t, η_t satisfy the same hypothesis as in system (1). If (A, C) is detectable and (A, G) is stabilizable, then Assumption **(A1)** holds with $a = c = 0$. The output observed at randomly sampled time instants is described by $y_{\tau_{N_t}} = Cx_{\tau_{N_t}} + v_{\tau_{N_t}}$, where for each $t \geqslant 0$, $v_{\tau_{N_t}} \sim \mathcal{N}(0, V)$. The filter (7) takes the form

$$\mathrm{d}\widehat{x}_t = A\widehat{x}_t\,\mathrm{d}t, \qquad \widehat{x}_{\tau^+_{N_t}} = \widehat{x}_{\tau^+_{N_t}} + K_{\tau_{N_t}}(y_{\tau_{N_t}} - h(\widehat{x}_{\tau^-_{N_t}})).$$

Inequality (9) in Proposition 3.1 holds, where P_t satisfies

$$\dot{P}_t = AP_t + P_tA^\top + GG^\top,$$

for $\tau_{N_t} \leqslant t < \tau_{N_t+1}$, and for $k = N_t$,

$$P_{\tau_k^+} = P_{\tau_k^-} - K_{\tau_k} C P_{\tau_k^-} - P_{\tau_k^-} C^\top K_{\tau_k}^\top + K_{\tau_k}\left(CP_{\tau_k^-}C^\top + V\right)K_{\tau_k}^\top.$$

By choosing, $K_{\tau_k} = P_{\tau_k^-} C^\top M^{-1}_{P_{\tau_k}}$, for each $k \in \mathbb{N}$, the covariance bound is minimized and the resulting equations are given by (15) by setting $a = c = 0$.

Assuming that the sampling process is Poisson of intensity λ, and the gain K is constant, the expected value of P_t is

$$\dot{\mathscr{P}}_t = (A - \lambda KC)\mathscr{P}_t + \mathscr{P}_t(A - \lambda KC)^\top + GG^\top + \lambda K\left(C\mathscr{P}_t C^\top + V\right)K^\top.$$

However, by choosing $K_{\tau_k} = P^\star_{\tau_k^-} C^\top M^{-1}_{P_{\tau_k}}$, the expected value of the error covariance process governed by (15) in linear case is

$$\dot{\mathscr{P}}_t^\star = A\mathscr{P}_t^\star + \mathscr{P}_t^\star A^\top + GG^\top - \lambda \mathscr{P}_t^\star C^\top \left(C\mathscr{P}_t C^\top + V\right)^{-1} C\mathscr{P}_t^\star.$$

Thus, in linear case, an upper bound on the expectation (with respect to sampling process) of error covariance is obtained by solving a nonlinear differential equations. The boundedness of the solutions of equations in Proposition 4.1 are analyzed in later sections. Here, we end this section with the proof of Proposition 4.1.

Proof of Proposition 4.1. We first note that $\big(P_t\big)_{t\geqslant 0}$ is Markovian because the future of P_t depends only on the last sampling instant τ_{N_t}. According to Dynkin's formula,

$$\mathscr{P}_t = \mathsf{E}[P_t] = P_0 + \mathsf{E}\Big[\int_0^t \mathscr{L}_{\mathrm{id}}(P_s)\Big]\mathrm{d}s.$$

Since $\mathscr{L}_{\mathrm{id}}$ is a linear operator, we readily get $\dot{\mathscr{P}}_t = \mathscr{L}_{\mathrm{id}}(\mathscr{P}_t)$. To show (19), we thus compute $\mathscr{L}_{\mathrm{id}}(Q)$ for a given matrix Q in the remainder of the proof. Recalling the definition of $\mathscr{L}_{\mathrm{id}}$, we observe that

$$\mathsf{E}\big[P(t+\varepsilon)\,\big|\,P_t = Q\big] = \mathsf{E}\Big[\big(P(t+\varepsilon)\big)\big(1_{\{N_{t+\varepsilon}=N_t\}} + 1_{\{N_{t+\varepsilon}=1+N_t\}} + 1_{\{N_{t+\varepsilon}-N_t\geqslant 2\}}\big)\,\Big|\,P(t)\Big]. \tag{24}$$

We now compute the conditional probability distribution of $\big(P(t+\varepsilon)\big)$ for small $\varepsilon > 0$ given $\big(P_t\big)$. Since the sampling process is independent of the process $\big(P_t\big)_{t\geqslant 0}$, by definition of the sampling (Poisson) process we have, for $\varepsilon \downarrow 0$, $\mathsf{P}\big(N_{t+\varepsilon}-N_t = 0\,\big|\,N_t, P_t\big) = 1-\lambda\varepsilon+o(\varepsilon)$, $\mathsf{P}\big(N_{t+\varepsilon}-N_t = 1\,\big|\,N_t, P_t\big) = \lambda\varepsilon+o(\varepsilon)$, $\mathsf{P}\big(N_{t+\varepsilon}-N_t \geqslant 2\,\big|\,N_t, P_t\big) = o(\varepsilon)$. Using these expressions we develop (24) further for $\varepsilon \downarrow 0$ as

$$\begin{aligned}\mathsf{E}\big[P(t+\varepsilon)\,\big|\,P_t = Q\big] &= \mathsf{E}\big[P(t+\varepsilon)\big(1_{\{N_{t+\varepsilon}=N_t\}} + 1_{\{N_{t+\varepsilon}=1+N_t\}}\big)\,\big|\,P_t\big] + o(\varepsilon)\\ &= \mathsf{E}\big[P(t+\varepsilon)\,\big|\,P_t, N_{t+\varepsilon} = N_t\big]\cdot\big(1-\lambda\varepsilon+o(\varepsilon)\big) + \mathsf{E}\big[P(t+\varepsilon)\,\big|\,P_t, N_{t+\varepsilon} = 1+N_t\big]\big(\lambda\varepsilon\big) + o(\varepsilon).\end{aligned} \tag{25}$$

The two significant terms on the right-hand side of (25) are now computed separately. For the event $N_{t+\varepsilon} = N_t$, given $P(t) = Q$, $\dot{P}$ is governed by (8a), so we have for $\varepsilon \downarrow 0$,

$$\begin{aligned}P(t+\varepsilon) &= P(t) + \varepsilon\dot{P}(t) + o(\varepsilon)\\ &= Q + \varepsilon(AQ + QA^\top + \Pi_a(Q) + GG^\top) + o(\varepsilon),\end{aligned}$$

leading to the first term on the right-hand side of (25) having the estimate

$$\begin{aligned}&\mathsf{E}\big[P(t+\varepsilon)\,\big|\,P(t), N_{t+\varepsilon} = N_t\big]\cdot\big(1-\lambda\varepsilon+o(\varepsilon)\big)\\ &= Q + \varepsilon(AQ + QA^\top + \Pi_a(Q) + GG^\top) - (\lambda\varepsilon)Q + o(\varepsilon),\end{aligned} \tag{26}$$

for $\varepsilon \downarrow 0$. Concerning the second term on the right-hand side of (25), we observe that conditional on $N_{t+\varepsilon} = 1+N_t$, the probability distribution of $\tau_{N_{t+\varepsilon}}$ is [31, 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 (8a) more compactly using the map ψ_1,

$$Q \mapsto \psi_1(Q) = AQ + QA^\top + \Pi_a(Q) + GG^\top$$

and (8b) using the map ψ_2,

$$Q \mapsto \psi_2(Q) = (1+c)Q - KCQ - QC^\top K + K\big((1+c)CQC^\top + (2c+c^2)\operatorname{tr}QI_{p\times p} + V\big)K.$$

We thus have, conditioned on the event $N_{t+\varepsilon} = 1+N_t, P_t = Q$,

$$P(\tau_{N_{t+\varepsilon}}) = P(t+\theta\varepsilon) = \psi_2(P(t+\theta\varepsilon)^-).$$

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

$$\begin{aligned}P(t+\varepsilon) &= P(t+\theta\varepsilon) + (1-\theta)\varepsilon\dot{P}(t+\theta\varepsilon) + o(\varepsilon)\\ &= P(t+\theta\varepsilon) + (1-\theta)\varepsilon\psi_1\big(P(t+\theta\varepsilon)\big) + o(\varepsilon)\\ &= \psi_2(P(t) + \theta\varepsilon\psi_1(P(t)) + o(\varepsilon)) + (1-\theta)\varepsilon\psi_1\big(P(t+\theta\varepsilon)\big) + o(\varepsilon)\\ &= \psi_2(P(t)) + \theta O(\varepsilon) + O(\varepsilon) + o(\varepsilon).\end{aligned}$$

Therefore, for $\varepsilon \downarrow 0$,

$$\begin{aligned}\mathsf{E}\big[P(t+\varepsilon)\,\big|\,P_t = Q, N_{t+\varepsilon} = 1+N_t\big]\cdot(\lambda\varepsilon) &= \int_0^1 \Big(\psi_2(Q) + \theta O(\varepsilon) + O(\varepsilon) + o(\varepsilon)\Big)\mathrm{d}\theta \cdot (\lambda h)\\ &= \Big(\psi_2(Q) + O(h)\Big)\cdot(\lambda h) = (\lambda h)\psi_2(Q) + o(\varepsilon). \end{aligned} \tag{27}$$

Substituting (26) and (27) in (25), we obtain

$$\mathsf{E}\big[P(t+\varepsilon)\,\big|\,P(t) = Q\big] = Q - (\lambda\varepsilon)\big(Q - \psi_2(Q)\big) + \varepsilon\Big(AQ + QA^\top + \Pi_a(Q) + GG^\top\Big) + o(\varepsilon).$$

Substituting these expressions in (18), we see that for each $Q \in \mathbb{R}^{n\times n}$,

$$\mathscr{L}_{\mathrm{id}}(Q) = AQ + QA^\top + \Pi_a(Q) + GG^\top + \lambda\big(\psi_2(Q) - Q\big),$$

which, upon recalling the definition of ψ_2, leads to the desired expression in (19).

The derivation of (22) follows the same recipe: the only difference arises in the definition of the function ψ_2 which now represents the right-hand side of (15b). □

Remark 4.2. In (22), if we introduce the variable $\mathscr{K}_t := \mathscr{P}_t^\star C^\top M_{\mathscr{P}_t^\star}^{-1}$, then (22) is equivalently written as

$$\dot{\mathscr{P}}_t^\star = (A - \lambda\mathscr{K}_t C)\mathscr{P}_t^\star + \mathscr{P}_t^\star(A - \lambda\mathscr{K}_t C)^\top + \Pi_a(\mathscr{P}_t^\star) + GG^\top + \lambda\, c\mathscr{P}_t^\star + \lambda\mathscr{K}_t M_{\mathscr{P}_t^\star}\mathscr{K}_t^\top.$$

This last equation resembles (19), the only difference being that a constant K is replaced by a time-varying term $\mathscr{K}_t$.

5 Evolution of Error Covariance

We are interested in studying the solutions of the equation (19) and (22), which characterize the expected value of the error covariance resulting from the suboptimal nonlinear filter when the available measurements are Poisson distributed. In particular, we want to study conditions on the system dynamics and the sampling rates which guarantee boundedness of the solution. In doing so, the following algebraic equation also plays an important role:

$$0 = A\overline{\mathscr{P}} + \overline{\mathscr{P}}A^\top + \Pi_a(\overline{\mathscr{P}}) + \lambda c\overline{\mathscr{P}} + GG^\top - \lambda\overline{\mathscr{P}}C^\top M_{\overline{\mathscr{P}}}^{-1} C\overline{\mathscr{P}}. \tag{28}$$

In the remainder of this section, we provide a statement of the existence of solution to the differential equation (22), and then study boundedness of its solution and connections to steady state algebraic equations in the next section.

To study Carathédory solutions to the differential equation (19), we introduce the notation,

$$\Psi(\mathscr{P}, M_{\mathscr{P}}, K) := A_{\lambda K}\mathscr{P} + \mathscr{P}A_{\lambda K}^\top + \Pi(\mathscr{P}) + GG^\top + \lambda K^\top M_{\mathscr{P}} K. \tag{29}$$

where $M_{\mathscr{P}}$ is defined as in (21), and we take

$$A_{\lambda K} := (A - \lambda KC) \tag{30a}$$

$$\Pi(\mathscr{P}) := \Pi_a(\mathscr{P}) + \lambda\, c\,\mathscr{P} = a\mathscr{P} + a\,\mathrm{tr}(\mathscr{P})I + \lambda c\mathscr{P}. \tag{30b}$$

It is readily seen that the particular choice of $K = \mathscr{K} = \mathscr{P}C^\top M_{\mathscr{P}}^{-1}$ minimizes $\Psi(\mathscr{P}, M_{\mathscr{P}}, K)$ with respect to the ordering in the cone of positive semidefinite matrices. Indeed,

$$\begin{aligned}\Psi(\mathscr{P}, M_{\mathscr{P}}, \mathscr{K}) &= (A - \lambda\mathscr{K}C)\mathscr{P} + \mathscr{P}(A - \lambda\mathscr{K}C)^\top + \Pi(\mathscr{P}) + GG^\top + \lambda\mathscr{K}M_{\mathscr{P}}\mathscr{K}^\top\\ &= (A - \lambda KC)\mathscr{P} + \mathscr{P}(A - \lambda KC)^\top + \Pi(\mathscr{P}) + GG^\top + \lambda K M_{\mathscr{P}} K^\top - \lambda(K - \mathscr{K})M_{\mathscr{P}}(K - \mathscr{K})^\top\\ &= \Psi(\mathscr{P}, M_{\mathscr{P}}, K) - \lambda(K - \mathscr{K})M_{\mathscr{P}}(K - \mathscr{K})^\top\\ &\leqslant \Psi(\mathscr{P}, M_{\mathscr{P}}, K). \end{aligned} \tag{31}$$

We can apply similar arguments to prove the following lemma which compares the solutions of equations (19) and (22).

Lemma 5.1. *Let $\mathscr{P}_t^\star$ be the solution of*

$$\dot{\mathscr{P}}_t^\star = \Psi(\mathscr{P}_t^\star, M_{\mathscr{P}_t^\star}, \mathscr{P}_t^\star C^\top M_{\mathscr{P}_t^\star}^{-1}), \quad \mathscr{P}_0^\star = \mathscr{P}_0,$$

and let $\mathscr{P}_t$ be the solution of

$$\dot{\mathscr{P}}_t = \Psi(\mathscr{P}_t, M_{\mathscr{P}_t}, K), \quad \mathscr{P}_0 \geqslant 0 \text{ given},$$

where $M_{\mathscr{P}_t^\star}$ and $M_{\mathscr{P}_t}$ are defined as in (16). Then, it holds that

$$\mathscr{P}_t^\star \leqslant \mathscr{P}_t, \quad \forall\, t \geqslant 0.$$

Proof. Notice that for $\widetilde{\mathscr{P}}(t) = \mathscr{P}_t - \mathscr{P}_t^\star$ the minimum property (31) implies

$$\begin{aligned}\dot{\widetilde{\mathscr{P}}} &= \Psi(\mathscr{P}_t, M_{\mathscr{P}_t}, K) - \Psi(\mathscr{P}_t^\star, M_{\mathscr{P}_t^\star}, \mathscr{P}_t^\star C^\top M_{\mathscr{P}_t^\star}^{-1}) \\ &\geqslant \Psi(\mathscr{P}_t, M_{\mathscr{P}_t}, K) - \Psi(\mathscr{P}_t^\star, M_{\mathscr{P}_t^\star}, K) \\ &= A_{\lambda K}\widetilde{\mathscr{P}}_t + \widetilde{\mathscr{P}}_t A_{\lambda K}^\top + \Pi(\widetilde{\mathscr{P}}_t) + \lambda K M_{\widetilde{\mathscr{P}}_t} K^\top,\end{aligned}$$

where $M_{\widetilde{\mathscr{P}}} = M_{\mathscr{P}_t} - M_{\mathscr{P}_t^\star}$ is consistent and depends linearly on $\widetilde{\mathscr{P}}_t$. So, we have

$$\dot{\widetilde{\mathscr{P}}} = A_{\lambda K}\widetilde{\mathscr{P}} + \widetilde{\mathscr{P}} A_{\lambda K}^\top + \Pi(\widetilde{\mathscr{P}}) + \lambda K M_{\widetilde{\mathscr{P}}} K^\top + R_t,$$

for some $R_t \geqslant 0$. Since $\widetilde{\mathscr{P}}(0) = 0$, for $t \geqslant 0$, $\widetilde{\mathscr{P}}_t$ is a solution of the following Volterra equation

$$\widetilde{\mathscr{P}}_t = \int_0^t \mathrm{e}^{sA_{\lambda K}}(\Pi(\widetilde{\mathscr{P}}_s) + \lambda K^\top M_{\widetilde{\mathscr{P}}_s} K)\mathrm{e}^{sA_{\lambda K}^\top}\mathrm{d}s.$$

The matrix on the right-hand side is positive semidefinite, and hence $\mathscr{P}_t \geqslant \mathscr{P}_t^\star$, for each $t \geqslant 0$. □

A direct consequence of this observation is that the right-hand side of (19) is minimized by choosing K_{τ_k} as in (14). Even though the perturbation Π is linear in $\mathscr{P}$, the aforementioned choice of K makes the right-hand side of (19) a nonlinear function in $\mathscr{P}$, which is different from the quadratic nonlinearity seen in Riccati differential equations. We are thus interested in knowing whether there exists a solution to the resulting equation. To this end, we will focus our attention on the initial value problem (22) with $\mathscr{P}(0) = \mathscr{P}_0$, where $\mathscr{P}_0$ is a positive semidefinite matrix.

Proposition 5.2. *For each symmetric positive semidefinite matrix $\mathscr{P}_0 \in \mathbb{R}^{n\times n}$ and a given $T > 0$, there exists a unique absolutely continuous solution $\mathscr{P} : [0, T] \to \mathbb{R}^{n\times n}$ in the class of symmetric positive definite matrices such that $\mathscr{P}(0) = \mathscr{P}_0$ and the differential equation (22) holds for almost every $t \in [0, T]$.*

The proof of this result is given in Appendix A. In the infinite-time horizon case we have a unique continuous solution $\mathscr{P} : [0, +\infty) \to \mathbb{R}^{n\times n}$ in the class of symmetric positive semidefinite matrices satisfying (22).

6 Conditions for Bounded Covariance

In this section, we study the conditions under which the error covariance $\mathscr{P}_t$ stays bounded for all $t \geqslant 0$. In doing so, we recall that the error covariance can be compactly represented as $\dot{\mathscr{P}}_t = \Psi(\mathscr{P}_t, M_{\mathscr{P}_t}, K)$ where Ψ is

defined in (29). For the purposes of this section, it is convenient to write Ψ in an alternate form. To do so, we let $\Pi_c(P) := cP + cKCPC^\top K^\top + (2c+c^2)\operatorname{tr}(P)KK^\top$, and let

$$\Pi_\lambda^K(P) := \Pi_a(P) + \lambda\Pi_c(P),$$

then we can write

$$\begin{aligned}
\Psi(\mathscr{P}, M_{\mathscr{P}}, K) &= A\mathscr{P} + \mathscr{P}A^\top + \Pi_\lambda^K(\mathscr{P}) + GG^\top + \lambda KVK^\top + \lambda\big[KC\mathscr{P}C^\top K^\top - KC\mathscr{P} - \mathscr{P}C^\top K^\top\big] \\
&= \left(A - \frac{\lambda}{2}I\right)\mathscr{P} + \mathscr{P}\left(A - \frac{\lambda}{2}I\right)^\top + \Pi_\lambda^K(\mathscr{P}) + GG^\top \\
&\quad + \lambda KVK^\top + \lambda\big[\mathscr{P} - \mathscr{P}C^\top K^\top - KC\mathscr{P} + KC\mathscr{P}C^\top K^\top\big] \\
&= \left(A - \frac{\lambda}{2}I\right)\mathscr{P} + \mathscr{P}\left(A - \frac{\lambda}{2}I\right)^\top + \Pi_\lambda^K(\mathscr{P}) + GG^\top \\
&\quad + \lambda KVK^\top + \lambda\big[(I-KC)\mathscr{P}(I-KC)^\top\big].
\end{aligned}$$

More compactly, by letting $A_\lambda := \left(A - \frac{\lambda}{2}I\right)$, $J_K := (I - KC)$, we get

$$\Psi(\mathscr{P}, M_{\mathscr{P}}, K) := A_\lambda\mathscr{P} + \mathscr{P}A_\lambda^\top + \Pi_\lambda^K(\mathscr{P}) + GG^\top + \lambda KVK^\top + \lambda J_K\mathscr{P}J_K^\top. \tag{32}$$

Based on this representation, we can write the solution $\mathscr{P}_t$ to equation (19) as follows:

$$\mathscr{P}_t = \mathrm{e}^{A_\lambda t}\mathscr{P}_0\mathrm{e}^{A_\lambda^\top t} + \int_0^t \mathrm{e}^{A_\lambda(t-s)}\Big[\Pi_\lambda^K(\mathscr{P}_s) + \lambda J_K\mathscr{P}_s J_K^\top + GG^\top + \lambda KVK^\top\Big]\mathrm{e}^{A_\lambda^\top(t-s)}\,\mathrm{d}s.$$

Naturally, to study the bounds on $\mathscr{P}_t$ for large t, we consider the operator $\mathscr{F}_\lambda^K : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ defined as:

$$\mathscr{F}_\lambda^K(P) := \int_0^\infty \mathrm{e}^{A_\lambda t}\left[\Pi_\lambda^K(P) + \lambda J_K P J_K^\top\right]\mathrm{e}^{A_\lambda^\top t}\,\mathrm{d}t. \tag{33}$$

This dependence is stated in the following result:

Theorem 6.1. *Consider the matrix valued function $\mathscr{P}_t$ described by (19), and the following statements:*

(S1) *All the eigenvalues of the operator $\mathscr{F}_\lambda^K$, defined in (33), lie inside the unit disk of the complex plane, that is,*

$$\rho(\mathscr{F}_\lambda^K) < 1$$

where $\rho(\cdot)$ denotes the spectral radius of its argument.

(S2) *There exists a positive semidefinite solution $\mathscr{Q}$ to the algebraic equation*

$$0 = A_\lambda\mathscr{Q} + \mathscr{Q}A_\lambda^\top + \Pi_\lambda^K(\mathscr{Q}) + GG^\top + \lambda KVK^\top + \lambda(I-KC)\mathscr{Q}(I-KC)^\top. \tag{34}$$

Statement **(S1)** *implies that* **(S2)** *holds. If (A, G) is controllable, then* **(S2)** *implies that*

$$\lim_{t\to\infty}\mathscr{P}_t = \mathscr{Q}, \tag{35}$$

where $\mathscr{P}_t$ is the solution of (19) and $\mathscr{Q}$ is obtained from **(S2)**.

Proof. The implication that **(S2)** follows from **(S1)** is a consequence of [9, Theorem 3.1: (d) ⇒ (b)].

The implication that **(S2)** implies (35) can be obtained as a result of Lemma C.2 given in Appendix C. To apply this result, we need to show that $(A_\lambda, \overline{G})$ is controllable, where $\overline{G}$ is such that $\lambda KVK^\top + GG^\top = \overline{G}\,\overline{G}^\top$. This indeed follows from the fact that (A, G) is controllable. To see this, we use Lemma C.1 and show that (A_λ, G) is controllable. To prove the later statement, let us proceed by contradiction. If (A_λ, G) is not controllable, there exists $v \in \mathbb{R}^n$ such that $v^\top(A - 0.5\lambda I) = (\mu - 0.5\lambda)v^\top$ and $v^\top G = 0$, where μ is an eigenvalue of A. This implies that $v^\top A = \mu v^\top$ and $v^\top G = 0$, meaning (A, G) is not controllable, which is a contradiction. □

We now prove a similar result for the nonlinear differential equation (22).

Theorem 6.2. *Consider the matrix valued function $\mathscr{P}_t$ described by the nonlinear equation differential equation (22), and the following statement:*

(S3) *There exists a positive semidefinite solution $\mathscr{Q}$ to the algebraic equation*

$$0 = A\mathscr{Q} + \mathscr{Q}A^\top + \Pi_a(\mathscr{Q}) + \lambda\, c\, \mathscr{Q} - \lambda \mathscr{Q}C^\top M_{\mathscr{Q}}^{-1} C\mathscr{Q} + GG^\top \tag{36}$$

1. *If (A, G) is stabilizable and there exists K such that* **(S1)** *holds, then* **(S3)** *holds and $(A - \lambda\mathscr{Q}C^\top M_{\mathscr{Q}}^{-1})$ is a Hurwitz matrix.*

2. *If (A, G) is controllable,* **(S3)** *implies that,*

$$\lim_{t\to\infty} \mathscr{P}_t = \mathscr{Q}, \tag{37}$$

where $\mathscr{P}_t$ solves (22) and $\mathscr{Q}$ is obtained from **(S3)**.

Proof. To prove the first item, we first choose K_1 such that **(S1)** holds. Therefore, there exists a positive semidefinite matrix $\mathscr{Q}_1$ that solves the matrix equation (34), and $\mathscr{Q}_1$ satisfies the fixed point equation, $\mathscr{Q}_1 = \mathscr{F}_\lambda^{K_1}(\mathscr{Q}_1) + \mathfrak{G}_{K_1}$, where $\mathscr{F}_\lambda^{K_1}(\cdot)$ is defined as in (33), and $\mathfrak{G}_{K_1} = \int_0^\infty \mathrm{e}^{A_\lambda t}(GG^\top + \lambda K_1 V K_1^\top)\mathrm{e}^{A_\lambda^\top t}\mathrm{d}t$. Due to linearity of $\mathscr{F}_\lambda^{K_1}$, it readily follows that $\mathscr{Q}_1 = \mathscr{F}_\lambda^{K_1}(\mathscr{Q}_1) + \mathfrak{G}_{K_1} = (\mathscr{F}_\lambda^{K_1})^i(\mathscr{Q}_1) + \sum_{j=0}^{i-1}(\mathscr{F}_\lambda^{K_1})^j(\mathfrak{G}_{K_1})$, for each $i \in \mathbb{N}$. Further, since the spectral radius of $\mathscr{F}_\lambda^{K_1}$ is strictly less than one, there exist $\gamma \geqslant 0$ and $0 < \delta < 1$ such that $\|(\mathscr{F}_\lambda^{K_1})^i\| \leqslant \gamma\delta^i$ for each $i \in \mathbb{N}$. Consequently, let

$$\mathscr{G}_{K_1,i_1}(\mathscr{Q}) := (\mathscr{F}_\lambda^{K_1})^{i_1}(\mathscr{Q}) + \sum_{j=0}^{i_1-1}(\mathscr{F}_\lambda^{K_1})^j(\mathfrak{G}_{K_1})$$

where i_1 is chosen such that $\gamma\delta^{i_1} < 1$. Since $\mathscr{G}_{K_1,i_1}$ is a contraction mapping, there is a unique solution of the equation $\mathscr{Q} = \mathscr{G}_{K_1,i_1}(\mathscr{Q})$, which we have denoted by $\mathscr{Q}_1$. Further, let $K_2 := M_{\mathscr{Q}_1}^{-1} C^\top \mathscr{Q}_1$, $\mathscr{Q}_2 = \mathscr{G}_{K_2,i_1}(\mathscr{Q}_2)$, where $\mathscr{Q}_2$ can be obtained by successive approximations:

$$\mathscr{Q}_2^{(1)} = 0, \quad \mathscr{Q}_2^{(k+1)} = \mathscr{F}_\lambda^{K_2}(\mathscr{Q}_2^{(k)}) + \mathfrak{G}_{K_2};$$

Indeed, $\mathscr{G}_{K_2,i_1}(\mathscr{Q}_2)$ is well-defined since A_λ is Hurwitz. Notice that the sequence of $\mathscr{Q}_2^{(k)}$ is nondecreasing and nonnegative.

Recall that $\mathscr{Q}_1$ satisfies $\Psi(\mathscr{Q}_1, M_{\mathscr{Q}_1}, K_1) = 0$, so that, from (31) and our choice of K_2, we have

$$0 = -\Psi(\mathscr{Q}_1, M_{\mathscr{Q}_1}, K_1) \leqslant -\Psi(\mathscr{Q}_1, M_{\mathscr{Q}_1}, K_2)$$

which leads to (recall that $J_{K_2} = (I - K_2 C)$),

$$\begin{aligned} 0 &\leqslant -\int_0^\infty \mathrm{e}^{A_\lambda t}\big(A_\lambda\mathscr{Q}_1 + \mathscr{Q}_1 A_\lambda^\top + J_{K_2}\mathscr{Q}_1 J_{K_2}^\top + \Pi_\lambda^{K_2}(\mathscr{Q}_1) + GG^\top + \lambda K_2 V K_2^\top\big)\mathrm{e}^{A_\lambda^\top t}\mathrm{d}t \\ &\leqslant \mathscr{Q}_1 - \int_0^\infty \mathrm{e}^{A_\lambda t}\big(\Pi_\lambda^{K_2}(\mathscr{Q}_1) + J_{K_2}\mathscr{Q}_1 J_{K_2}^\top\big)\mathrm{e}^{A_\lambda^\top t}\mathrm{d}t - \mathfrak{G}_{K_2}. \end{aligned}$$

Let us show by induction that $\mathscr{Q}_2^{(k)} \leqslant \mathscr{Q}_1$ for each k. Firstly it holds for $k = 1$, and for $k > 1$,

$$\begin{aligned} \mathscr{Q}_2^{(k+1)} &= \mathscr{F}_\lambda^{K_2}(\mathscr{Q}_2^{(k)}) + \mathfrak{G}_{K_2} \\ &= \int_0^\infty \mathrm{e}^{tA_\lambda}\big(\Pi_\lambda^{K_2}(\mathscr{Q}_2^{(k)}) + \lambda J_{K_2}(\mathscr{Q}_2^{(k)})J_{K_2}^\top\big)\mathrm{e}^{tA_\lambda^\top}\mathrm{d}t + \mathfrak{G}_{K_2} \\ &\leqslant \mathscr{Q}_1 - \int_0^\infty \mathrm{e}^{tA_\lambda}\Big(\Pi_\lambda^{K_2}(\mathscr{Q}_1 - \mathscr{Q}_2^{(k)}) + \lambda J_{K_2}(\mathscr{Q}_1 - \mathscr{Q}_2^{(k)})J_{K_2}^\top\Big)\mathrm{e}^{tA_\lambda^\top}\mathrm{d}t. \end{aligned}$$

Since the second-term on the right-hand side is negative definite, for each $k \in \mathbb{N}$, $\mathcal{Q}_2^{(k)}$ is no greater than $\mathcal{Q}_1$ and consequently $0 \leqslant \mathcal{Q}_2 \leqslant \mathcal{Q}_1$.

This procedure can be repeated to construct the sequences $\{K_i\}$ and $\{\mathcal{Q}_i\}$ similarly. Since $0 \leqslant \mathcal{Q}_{i+1} \leqslant \mathcal{Q}_i$ there exists $\overline{\mathcal{Q}} = \lim_{i\to\infty} \mathcal{Q}_i$, and moreover, for $M_{\overline{\mathcal{Q}}} = (1+c)C\overline{\mathcal{Q}}C^\top + (2c+c^2)\operatorname{tr}(\overline{\mathcal{Q}})I_{n\times n} + V$, we obtain, $\overline{K} = \lim_{i\to\infty} K_i = M_{\overline{\mathcal{Q}}}^{-1} C^\top \overline{\mathcal{Q}}$, and get $\Psi(\overline{\mathcal{Q}}, M_{\overline{\mathcal{Q}}}, \overline{K}) = 0$ where $\overline{\mathcal{Q}} \geqslant 0$.

To complete the proof of first item, we next show that $A - \lambda\overline{K}C$ is stable. It is seen that $\overline{\mathcal{Q}}$ satisfies the equation

$$\overline{\mathcal{Q}} = \int_0^\infty \mathrm{e}^{sA_{\lambda\overline{K}}}\Big(\Pi(\overline{\mathcal{Q}}) + GG^\top + \lambda\overline{K}M_{\overline{\mathcal{Q}}}\overline{K}^\top\Big)\mathrm{e}^{sA_{\lambda\overline{K}}^\top}\mathrm{d}s \geqslant \int_0^\infty \mathrm{e}^{sA_{\lambda\overline{K}}} GG^\top \mathrm{e}^{sA_{\lambda\overline{K}}^\top}\,\mathrm{d}s.$$

Existence of positive semidefinite $\overline{\mathcal{Q}}$ implies that the integral on the right-hand side is bounded and since (A, G) is assumed to be stabilizable, it follows that $A_{\lambda\overline{K}}$ is Hurwitz.

For the proof of second item in Theorem 6.2, we show that there are $\overline{\mathcal{P}}_t$ and $\underline{\mathcal{P}}_t$ such that $\underline{\mathcal{P}}_t \leqslant \mathcal{P}_t \leqslant \overline{\mathcal{P}}_t$ and both converge to $\mathcal{Q}$. Set $K_{\mathcal{Q}} = \mathcal{Q}C^\top M_{\mathcal{Q}}^{-1}$, $A_{\mathcal{Q}} = A - \lambda K_{\mathcal{Q}} C$, where $A_{\mathcal{Q}}$ is Hurwitz stable by the first item. Denote by $\overline{\mathcal{P}}_t$ the solution of $\dot{\overline{\mathcal{P}}}_t = \Psi(\overline{\mathcal{P}}_t, M_{\overline{\mathcal{P}}_t}, K_{\mathcal{Q}}) = A_\lambda\overline{\mathcal{P}}_t + \overline{\mathcal{P}}_t A_\lambda^\top + \Pi_\lambda^{K_{\mathcal{Q}}}(\overline{\mathcal{P}}_t) + GG^\top + \lambda K_{\mathcal{Q}} V K_{\mathcal{Q}}{}^\top + \lambda(I - K_{\mathcal{Q}}C)\overline{\mathcal{P}}_t(I - K_{\mathcal{Q}}C)^\top$, with $\overline{\mathcal{P}}_0 = \mathcal{P}_0 \geqslant 0$. By the minimum property (Lemma 5.1) $\mathcal{P}_t \leqslant \overline{\mathcal{P}}_t$. Let $\overline{G}_{\mathcal{Q}}$ be such that $\lambda K_{\mathcal{Q}} V K_{\mathcal{Q}}^\top + GG^\top = \overline{G}_{\mathcal{Q}}\overline{G}_{\mathcal{Q}}^\top$. Using the same reasoning as in the proof of Theorem 6.1, it follows from Lemma C.1 that $(A_\lambda, \overline{G}_{\mathcal{Q}})$ is controllable. Then, applying Lemma C.2, we have $\overline{\mathcal{P}}_t \to \mathcal{Q}$ as $t \to +\infty$.

Next, consider the function $\underline{\mathcal{P}}_t$ obtained by solving the equation, $\dot{\underline{\mathcal{P}}}_t = \Psi(\underline{\mathcal{P}}_t, M_{\underline{\mathcal{P}}_t}, \underline{\mathcal{P}}_t C^\top M_{\underline{\mathcal{P}}_t}^{-1})$, with $\underline{\mathcal{P}}(0) = 0$. Due to zero initial condition, $\underline{\mathcal{P}}_t$ enjoys a monotonicity property which is stated in the following lemma and its proof appears in Appendix B.

Lemma 6.3. *The solution $\mathcal{P}_t$ of the following Cauchy problem*

$$\frac{\mathrm{d}\mathcal{P}_t}{\mathrm{d}t} = \Psi(\mathcal{P}_t, M_{\mathcal{P}_t}, \mathcal{P}_t C^\top M_{\mathcal{P}_t}^{-1}), \quad \mathcal{P}_0 = 0$$

is monotone nondecreasing for $t \in [0, +\infty)$.

Notice that the limit $\lim_{t\to+\infty} \underline{\mathcal{P}}_t$ exists as $\underline{\mathcal{P}}_t$ is monotone function by Lemma 6.3 and this limit is bounded because of minimal property in Lemma 5.1 and Theorem 6.1. Moreover, this means that $\frac{\mathrm{d}\underline{\mathcal{P}}}{\mathrm{d}t}$ and $\frac{\mathrm{d}^2\underline{\mathcal{P}}}{\mathrm{d}t^2}$ are bounded too. As $\underline{\mathcal{P}}_t \geqslant 0$ the equality $\lim_{t\to+\infty} \underline{\mathcal{P}}_t = \int_0^{+\infty} \frac{\mathrm{d}\underline{\mathcal{P}}_t}{\mathrm{d}t}\mathrm{d}t$ implies that $\frac{\mathrm{d}\underline{\mathcal{P}}_t}{\mathrm{d}t} \to 0$ as $t \to \infty$. Since $\mathcal{Q}$ is a unique solution of the algebraic equation, it follows that $\underline{\mathcal{P}}_t \to \mathcal{Q}$ as $t \to \infty$.

It remains to prove that $\underline{\mathcal{P}}_t \leqslant \mathcal{P}_t$. Let $K_{\mathcal{P}_t} := \mathcal{P}_t C^\top M_{\mathcal{P}_t}^{-1}$ and consider $\widetilde{\mathcal{P}}_t := \mathcal{P}_t - \underline{\mathcal{P}}_t$, so that $\widetilde{\mathcal{P}}_0 = \mathcal{P}_0$, and

$$\begin{aligned}
\frac{\mathrm{d}\widetilde{\mathcal{P}}_t}{\mathrm{d}t} &= \Psi(\mathcal{P}_t, M_{\mathcal{P}_t}, K_{\mathcal{P}_t}) - \Psi(\underline{\mathcal{P}}_t, M_{\underline{\mathcal{P}}_t}, \underline{\mathcal{P}}_t C^\top M_{\underline{\mathcal{P}}_t}^{-1}) \\
&\geqslant \Psi(\mathcal{P}_t, M_{\mathcal{P}_t}, K_{\mathcal{P}_t}) - \Psi(\underline{\mathcal{P}}_t, M_{\underline{\mathcal{P}}_t}, K_{\mathcal{P}_t}) \\
&= (A - \lambda K_{\mathcal{P}_t} C)\widetilde{\mathcal{P}} + \widetilde{\mathcal{P}}(A - \lambda K_{\mathcal{P}_t} C)^\top + \Pi(\widetilde{\mathcal{P}}_t) + \lambda K_{\mathcal{P}_t} M_{\widetilde{\mathcal{P}}_t} K_{\mathcal{P}_t}.
\end{aligned}$$

This leads to

$$\begin{aligned}
\widetilde{\mathcal{P}}_t &= \Phi(t,0)\mathcal{P}_0\Phi^\top(t,0) + \int_0^t \Phi(t,s)\big[\Pi(\widetilde{\mathcal{P}}_s) + \lambda K_{\mathcal{P}_s} M_{\widetilde{\mathcal{P}}_s} K_{\mathcal{P}_s}^\top\big]\Phi^\top(t,s)\,\mathrm{d}s \\
&\geqslant \int_0^t \Phi(t,s)\big[\Pi(\widetilde{\mathcal{P}}_s) + \lambda K_{\mathcal{P}_s} M_{\widetilde{\mathcal{P}}_s} K_{\mathcal{P}_s}^\top\big]\Phi^\top(t,s)\,\mathrm{d}s,
\end{aligned}$$

where $\Phi(t,s)$ is the fundamental matrix associated with $A-\lambda K_{\mathscr{P}_t}C$. From this expression[4], we obtain $\widetilde{\mathscr{P}_t} \geqslant 0$ and hence $\underline{\mathscr{P}}_t \leqslant \mathscr{P}_t$. □

7 Sufficient Conditions for Boundedness

In this section, we study some particular cases which provide sufficient conditions to guarantee the boundedness of the error covariance matrix, based on the results given in the previous section.

7.1 Eigenvalue Assignment Condition

Since the conventional design of the state estimators provide conditions for bounded estimation error based on the eigenvalues of certain matrix, we rewrite the conditions from previous section as an eigenvalue test.

Proposition 7.1. *For a given $\lambda > 0$ and $K \in \mathbb{R}^{n\times p}$, consider the matrices,*[5]

$$\begin{aligned}\mathscr{A}_{\lambda K} &:= (I \otimes (A-\lambda KC)^\top) + ((A-\lambda KC)^\top \otimes I)\\ &\quad + \lambda((KC)^\top \otimes (KC)^\top)\\ \Sigma_a &:= a(I\otimes I + \operatorname{vec}(I)\cdot\operatorname{vec}(I)^\top)\\ \Sigma_c &:= c(I\otimes I + (KC)^\top\otimes(KC)^\top)\\ &\quad + (2c+c^2)\cdot(\operatorname{vec}(I)\cdot\operatorname{vec}(I)^\top)\cdot(I\otimes KK^\top)\end{aligned}$$

and the following statement,

(S4) *The matrix $\mathscr{A}_{\lambda K}+\Sigma_a+\lambda\Sigma_c \in \mathbb{R}^{n^2\times n^2}$ is Hurwitz.*

Then **(S4)** *implies* **(S2)** *stated in Theorem 6.1.*

Proof. To prove this result, we will make use of [9, Theorem 3.1(ii)]. Using this result, we just need to show that, if **(S4)** holds, then the matrix-valued function $\mathscr{R} : [0,\infty) \to \mathbb{R}^{n\times n}$ satisfying the differential equation $\dot{\mathscr{R}} = \Sigma(\mathscr{R})$ has the property that $\|\mathscr{R}_t\|$ converges to zero, as $t\to\infty$, where $\Sigma(\mathscr{R}) = A_\lambda^\top\mathscr{R} + \mathscr{R}A_\lambda + \lambda J_K^\top\mathscr{R}J_K + \Pi^*_{\lambda,K}(\mathscr{R})$ and $\Pi^*_{\lambda,K}$ denotes the adjoint[6] of Π^K_λ. In particular,

$$\Pi^*_{\lambda,K}(\mathscr{R}) = a\mathscr{R} + a\operatorname{tr}(\mathscr{R})I + \lambda c\mathscr{R} + \lambda cC^\top K^\top\mathscr{R}KC + \lambda(2c+c^2)\operatorname{tr}(KK^\top\mathscr{R})I.$$

We now write the matrix differential equation in vectorial form, so that, $\operatorname{vec}(\dot{\mathscr{R}}) = \operatorname{vec}(\Sigma(\mathscr{R}))$, and using the properties of the Kronecker product of the matrices, we observe that[7]

$$\operatorname{vec}(\Sigma(R)) = (I\otimes A^\top)\operatorname{vec}(R) + (A^\top\otimes I)\operatorname{vec}(R) - \lambda(I\otimes I)\operatorname{vec}(R) + \lambda(J_K^\top\otimes J_K^\top)\operatorname{vec}(R) + \operatorname{vec}(\Pi^*_{\lambda,K}(\mathscr{R})). \tag{38}$$

Recalling the definition of Σ_a and Σ_c, the last term in (38) can be written as, $\operatorname{vec}(\Pi^*_{\lambda,K}(R)) = \Sigma_a\operatorname{vec}(R) + \lambda\Sigma_c\operatorname{vec}(R)$. Also, since $J_K = (I-KC)$, the second to last term in (38) is rewritten as $(I-KC)^\top\otimes(I-KC)^\top = (I\otimes I) - ((KC)^\top\otimes I) - (I\otimes(KC)^\top) + ((KC)^\top\otimes(KC)^\top)$ so that, $\operatorname{vec}(\Sigma(\mathscr{R})) = (\mathscr{A}_{\lambda K}+\Sigma_a+\lambda\Sigma_c)\operatorname{vec}(\mathscr{R})$. Clearly, if the matrix $(\mathscr{A}_{\lambda K}+\Sigma_a+\lambda\Sigma_c)$ is Hurwitz, then $\operatorname{vec}(\mathscr{R}_t)$ converges to 0, as $t\to\infty$, and, in particular, $\|\mathscr{R}_t\|$ converges to 0 as well. □

[4] If $Q_t \geqslant \int_0^t \Phi(t,s)(\Pi(Q_s)+\lambda K_s M_{Q_s}K_s)\Phi^\top(t,s)\mathrm{d}s$, then $Q_t \geqslant 0$.

[5] The symbol $\otimes$ denotes the matrix Kronecker product.

[6] For a linear operator $\Pi : \mathbb{R}^{n\times n}\to\mathbb{R}^{n\times n}$, its adjoint is a linear operator $\Pi^* : \mathbb{R}^{n\times n}\to\mathbb{R}^{n\times}$ that satisfies $\langle\Pi(M_1),M_2\rangle = \langle M_1,\Pi^*(M_2)\rangle$, for all matrices $M_1, M_2 \in \mathbb{R}^{n\times n}$. Here, the inner product over the space of matrices is defined as $\langle M_1,M_2\rangle = \operatorname{tr}(M_1^\top M_2)$.

[7] For given matrices A,B,C,D of appropriate dimensions, we have: 1) $(A\otimes B)(C\otimes D) = AC\otimes BD$, and 2) If $ABC = D$, then $\operatorname{vec}(D) = (C^\top\otimes A)\operatorname{vec}(B)$.

7.2 A sufficient condition with C square and invertible

In the particular case, when C is square and invertible, one can provide certain conditions under which a particular choice of the gain matrix K bounds the error covariance. This is captured by the following result:

Proposition 7.2. *Consider the system* (1) *and the filter* (7). *Assume that $C \in \mathbb{R}^{n\times n}$ is invertible and take $K = C^{-1}$. Let*

$$\alpha := \int_0^\infty \mathrm{e}^{(A-0.5\lambda I)t}\mathrm{e}^{(A-0.5\lambda I)^\top t}\mathrm{d}t \tag{39}$$

Then, the error covariance is bounded if

$$a(n+1) + 2\lambda c + \lambda n(2c+c^2)\,\|C^{-2}\| < \frac{1}{\alpha}. \tag{40}$$

Proof. Based on the result of Theorem 6.1, it suffices to show that (40) leads to the spectral radius of $\mathscr{F}_\lambda^K$ being strictly less than 1. By letting $K = C^{-1}$, it is seen that $\mathscr{F}_\lambda^K(P) = \int_0^\infty \mathrm{e}^{A_\lambda t}\Pi_\lambda^K(P)\mathrm{e}^{A_\lambda t}\,\mathrm{d}t$. We can now introduce the bound

$$-\|P\|\,\Pi_\lambda^K(I) \leqslant \Pi_\lambda^K(P) \leqslant \Pi_\lambda^K(I)\,\|P\|$$

so that, using the definition of α in (39), $\|\mathscr{F}_\lambda^K(P)\| \leqslant \alpha\,\|\Pi_\lambda^K(I)\|\,\|P\|$. Condition (40) ensures that $\alpha\|\Pi_\lambda^K\| < 1$ because

$$\Pi_\lambda^K(I) = a(1+\mathrm{tr}(I))I + 2\lambda cI + \lambda(2c+c^2)\,\mathrm{tr}(I)C^{-2}.$$

As a result, $\|\mathscr{F}_\lambda^K\| < 1$ and in particular, the spectral radius of $\mathscr{F}_\lambda^K$ is less than 1. □

In the proof of Proposition 7.2, we basically showed that, if C is invertible and the nonlinearities are small enough, then for certain choice of K and λ such that

$$\|\mathscr{F}_\lambda^K\| < 1 \tag{41}$$

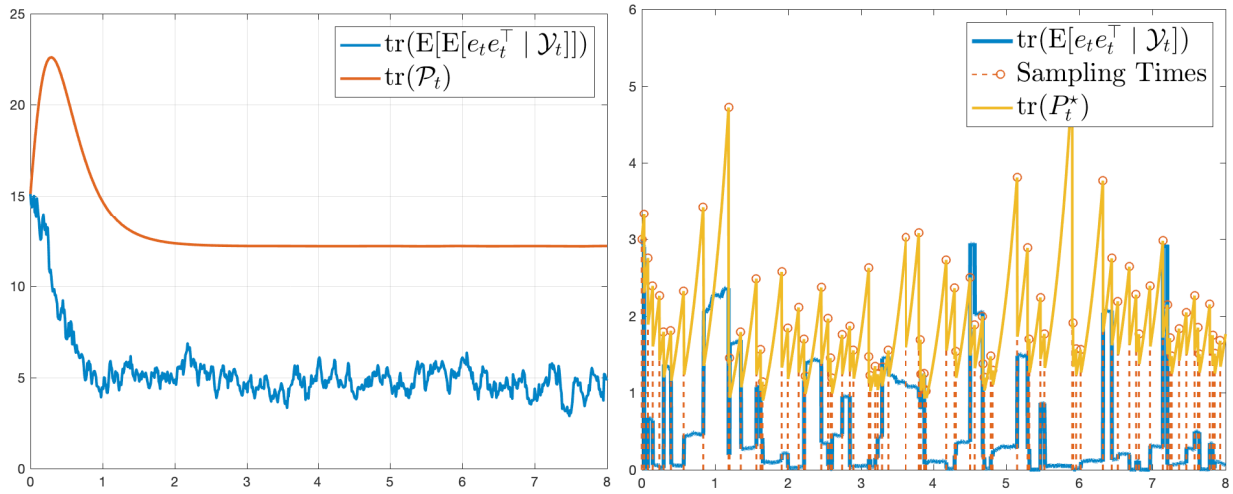
it follows that the spectrum of $\mathscr{F}_\lambda^K$ lies within the unit disk of the complex plane, that is, the statement **(S1)** holds. However, if C is not invertible, then even for linear systems ($a = 0$ and $c = 0$), there may not exist λ and K such that (41) holds. This is shown by the following example.

Example 3. Consider the dynamical system (1) with $f(t,x) = Ax$ and $h(x) = Cx$, $G = I_{2\times 2}$, where $A = \left[\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right]$; $C = [\,1\ 0\,]$. With this example, even if there are no nonlinearities, that is, $a, c = 0$, it is seen that, for any $\lambda > 0$, there does not exist a $K \in \mathbb{R}^2$ such that (41) holds for this system. To see this, observe that (41) translates to finding a matrix K such that, $\lambda\left|\int_0^\infty e^{t(A-0.5\lambda I)}J_K J_K^\top e^{t(A-0.5\lambda I)^\top}\mathrm{d}t\right| < 1$. We let $K = \begin{bmatrix}k_1\\ k_2\end{bmatrix}$, for some $k_1, k_2 \in \mathbb{R}$. However, with straightforward computations involving matrix exponential, we see that, the integral is finite for each $\lambda > 2$, and

$$e^{t(A-0.5\lambda I)}J_K J_K^\top e^{t(A-0.5\lambda I)^\top} = \begin{bmatrix} * & * \\ * & e^{(1-0.5\lambda)t}(1+k_2^2)e^{(1-0.5\lambda)t}\end{bmatrix}.$$

Clearly, with $\lambda > 2$, $\lambda\int_0^\infty e^{(1-0.5\lambda)t}(1+k_2^2)e^{(1-0.5\lambda)t} > 1$, and hence the condition (41) does not hold true.

However, it is possible to check that for each $\lambda > 2$, we can find K such that **(S4)** holds true. For simulation purposes, we choose $\lambda = 10$ and see that $K = [1\ 2]^\top$ satisfies the condition **(S4)**, where we take $\Sigma_a = \Sigma_c = 0$. The simulation results are reported in Figure 3. Moreover, since eigenvalues of a matrix are continuous functions of its entries, the matrix $\mathscr{A}_{\lambda K} + \Sigma_a + \lambda\Sigma_c$ also has eigenvalues in open left-half of the complex plane, for small enough values of a and c.



(a) Expectation with respect to sampling times with constant gain $K = [1\ 2]^\top$.

(b) One sample path with bound minimizing gain.

Figure 3: Trace of error covariance and their theoretical upper bounds for Example 3 with $\lambda = 10$.

8 Dependence on Sampling Rate

A natural question for analyzing the performance of our proposed estimator is to study the dependence of error covariance with respect to the sampling rate λ of the Poisson measurement process. This question has not been straightforward for the class of systems studied in this paper, that is, the systems with (f,h) Lipschitz continuous. However, some constructive statements can be derived if we assume that f,h are linear, that is, $f(x) = Ax$ and $h(x) = Cx$, so that the system is described as in Example 2 using equation (23).

8.1 A lower bound on sampling rate

We first provide a necessary lower bound on the sampling rate which ensures that the error covariance remains bounded for all times. In what follows, we denote an eigenvalue of A by $\mu_j(A)$ and its real part by $\Re\mu_j(A)$, $j = 1,\dots,n$.

Proposition 8.1. *Consider system (23) with (A,C) detectable and (A,G) controllable. Then, for each $\lambda < 2\max\{\Re\mu_j(A) \mid j = 1,\dots,n\}$, the solution to (19), with $a = c = 0$, is such that $\lim_{t\to\infty} \mathscr{P}(t) = \infty$.*

Proof. We first look at the differential equation, $\dot{Q}_t = A_\lambda Q_t + Q_t A_\lambda^\top + GG^\top$ with initial condition $Q_0 = \mathscr{P}_0$. It is clear that Q_t stays bounded if and only if, $0 = A_\lambda \overline{Q} + \overline{Q}A_\lambda^\top + GG^\top$ has a symmetric positive definite solution $\overline{Q}$. With (A,G) controllable, this is the case if and only if A_λ is Hurwitz, that is, $\lambda > 2\max\{\Re\mu_j(A) \mid j = 1,\dots,n\}$. We claim that $Q_t \leqslant \mathscr{P}_t$, for every $t \geqslant 0$. To see this, let $Q_0 = \mathscr{P}_0$. Let $\widetilde{Q}_t = \mathscr{P}_t - Q_t$, then

$$\dot{\widetilde{Q}}_t = \psi_1(\widetilde{Q}_t) + \lambda\big[(I-KC)\mathscr{P}_t(I-KC)^\top + KVK^\top\big]$$

where $\psi_1(\widetilde{Q}_t) = A_\lambda \widetilde{Q}_t + \widetilde{Q}_t A_\lambda^\top + GG^\top$. Because of the zero initial condition, this equation has the solution given by

$$\widetilde{Q}_t = \int_0^t \mathrm{e}^{A_\lambda(t-s)}\big[\lambda(I-KC)\mathscr{P}_s(I-KC)^\top + GG^\top + \lambda KVK^\top\big]\mathrm{e}^{A_\lambda^\top(t-s)}\,\mathrm{d}s \geqslant 0$$

and hence $\widetilde{Q}_t \geqslant 0$, that is, $\mathscr{P}_t \geqslant Q_t$ for each $t \geqslant 0$. As a result, $\mathscr{P}_t$ goes to $+\infty$ when Q_t grows unbounded. □

8.2 Monotonicity relation

After deriving a lower bound which is necessary for boundedness, the next question is to study how this steady state upper bound on the error covariance changes as a function of the sampling rate.

Proposition 8.2. *Consider system* (1) *under* **(A1)** *with* $a = 0$, $c = 0$. *Let* $\lambda_1, \lambda_2 \geqslant 0$ *describe the rate of Poisson measurement process and* $\mathscr{P}_t^{\lambda_1}$, $\mathscr{P}_t^{\lambda_2}$ *denote the corresponding solutions to* (22) *with* $a = c = 0$ *and* $\mathscr{P}^{\lambda_1}(0) = \mathscr{P}^{\lambda_2}(0)$. *For each* $\lambda_2 \geqslant \lambda_1$, *we have*

$$\mathscr{P}_t^{\lambda_2} \leqslant \mathscr{P}_t^{\lambda_1}, \quad \forall\, t \geqslant 0. \tag{42}$$

Proof. For a given λ, the solution to (22) satisfies the equation, $\dot{\mathscr{P}}_t^{\lambda} = \Psi_\lambda(\mathscr{P}_t^{\lambda}, M_{\mathscr{P}_t^{\lambda}}, \mathscr{P}_t^{\lambda} C^\top M_{\mathscr{P}_t^{\lambda}}^{-1})$, where, due to the linearity of the measurement map,

$$\Psi_\lambda(\mathscr{P}_t^{\lambda}, M_{\mathscr{P}_t^{\lambda}}, \mathscr{P}_t^{\lambda} C^\top M_{\mathscr{P}_t^{\lambda}}^{-1}) = A\mathscr{P}_t^{\lambda} + \mathscr{P}_t^{\lambda} A^\top + GG^\top - \lambda \mathscr{P}_t^{\lambda} C^\top M_{\mathscr{P}_t^{\lambda}}^{-1} C \mathscr{P}_t^{\lambda}.$$

It therefore holds that $\frac{\partial \Psi_\lambda}{\partial \lambda} \leqslant 0$ and thus, for each $\lambda_2 \geqslant \lambda_1$, we have $\Psi_{\lambda_2} \leqslant \Psi_{\lambda_1}$, for every $\lambda_2 \geqslant \lambda_1$. In particular, $\dot{\mathscr{P}}_t^{\lambda_2} \leqslant \dot{\mathscr{P}}_t^{\lambda_2}$, for each $t \geqslant 0$ and hence (42) holds. □

8.3 Convergence with large sampling rate

Building up on the monotonicity relation stated in Proposition 8.2, we now compare the error covariance of the sampled system with the error covariance of an estimator with continuous time measurements. Toward this end, let us consider the *optimal* Kalman-Bucy filter for (23) with continuously available output measurements, which is given by

$$\mathrm{d}\widehat{z}_t = A\widehat{z}_t\,\mathrm{d}t + K_t^{\mathrm{cont}}(\mathrm{d}z_t - C\widehat{z}_t\,\mathrm{d}t), \tag{43}$$

with $\widehat{z}_0 = \mathsf{E}[x_0]$. To minimize the error covariance, the injection gain K_t^{cont} for this estimator is obtained as

$$K_t^{\mathrm{cont}} = \mathscr{R}_t C^\top V^{-1} \tag{44a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{R}_t = A\mathscr{R}_t + \mathscr{R}_t A^\top + GG^\top - \mathscr{R}_t C^\top V^{-1} C \mathscr{R}_t. \tag{44b}$$

When $t \to \infty$, the error covariance $\mathsf{E}[|x_t - \widehat{z}_t|^2 \,|\, \mathscr{Y}_t]$ is given by the matrix $\overline{\mathscr{R}}$ that solves

$$0 = A\overline{\mathscr{R}} + \overline{\mathscr{R}} A^\top + GG^\top - \overline{\mathscr{R}} C^\top V^{-1} C \overline{\mathscr{R}}. \tag{45}$$

We are interested in knowing how the steady state error covariance for the estimator using sampled measurements compares with $\overline{\mathscr{R}}$ when we increase the sampling rate λ, and in particular, when we let λ go to ∞.

The next result addresses these issues and to formulate it, we consider the operator $\overline{\mathscr{F}}_\lambda^K$ which is obtained from $\mathscr{F}_\lambda^K$ by setting $\Pi_\lambda^K = 0$ in (33), that is,

$$\overline{\mathscr{F}}_\lambda^K(P) := \lambda \int_0^\infty \mathrm{e}^{tA_\lambda}(I - KC)P(I - KC)^\top \mathrm{e}^{tA_\lambda^\top}\,\mathrm{d}t, \tag{46}$$

with $\rho\left(\overline{\mathscr{F}}_\lambda^K\right)$ denoting its spectral radius. We denote the Neumann series associated with $\overline{\mathscr{F}}_\lambda^K$ by $\mathscr{N}_\lambda$, so that $\mathscr{N}_\lambda := \left(I - \overline{\mathscr{F}}_\lambda^K\right)^{-1} = \sum_{j=0}^\infty (\overline{\mathscr{F}}_\lambda^K)^j$. Thus, if $\rho\left(\overline{\mathscr{F}}_\lambda^K\right) < 1$, there exists ℓ such that $\left\|\left(\overline{\mathscr{F}}_\lambda^K\right)^\ell\right\| < 1$, and we can write

$$\begin{aligned}\mathscr{N}_\lambda &= \left(I + \overline{\mathscr{F}}_\lambda^K + \cdots + \left(\overline{\mathscr{F}}_\lambda^K\right)^{\ell-1}\right)\left(I + \left(\overline{\mathscr{F}}_\lambda^K\right)^\ell + \left(\overline{\mathscr{F}}_\lambda^K\right)^{2\ell} + \cdots\right) \\ &= \left(I + \overline{\mathscr{F}}_\lambda^K + \cdots + \left(\overline{\mathscr{F}}_\lambda^K\right)^{\ell-1}\right)\left(I - \left(\overline{\mathscr{F}}_\lambda^K\right)^\ell\right)^{-1}.\end{aligned}$$

Using triangle inequality[8], it follows that $\|\mathscr{N}_\lambda\| \leqslant \beta_\lambda := \left\|I + \overline{\mathscr{F}}_\lambda^K + \cdots + \left(\overline{\mathscr{F}}_\lambda^K\right)^{\ell-1}\right\| \Big/ \left(1 - \left\|\left(\overline{\mathscr{F}}_\lambda^K\right)^\ell\right\|\right)$.

Theorem 8.3. *Consider system (1) under (A1) with $a = 0$, $c = 0$, and assume that (A, G) is controllable. Let $\overline{\mathscr{R}}$ denote the steady state minimal covariance of the estimator (43) that satisfies (45). Let $K = K^{\mathrm{cont}} = \overline{\mathscr{R}} C^\top V^{-1}$ and assume that the operator $\overline{\mathscr{F}}_\lambda^K$ in (46) satisfies*

$$\rho\left(\overline{\mathscr{F}}_\lambda^K\right) < 1, \tag{47}$$

for each $\lambda > 0$. Then the matrix $\overline{\mathscr{P}}_\lambda \geqslant 0$ that solves (34), and provides an upper bound on the error covariance due to filter (7) with Poisson measurement process of intensity λ, satisfies

$$\overline{\mathscr{P}}_\lambda = \overline{\mathscr{R}} + \mathscr{N}_\lambda\left(\int_0^\infty \mathrm{e}^{A_\lambda t} K\big((1-\lambda)V + \lambda C\overline{\mathscr{R}}C^\top\big)K^\top \mathrm{e}^{A_\lambda^\top t}\,\mathrm{d}t\right). \tag{48}$$

In particular, if $\beta_\infty = \limsup_{\lambda\to\infty} \beta_\lambda$, then

$$\limsup_{\lambda\to\infty} \|\overline{\mathscr{P}}_\lambda\| \leqslant \|\overline{\mathscr{R}}\| + \beta_\infty \|K(C\overline{\mathscr{R}}C^\top - V)K^\top\|. \tag{49}$$

Proof. Consider the solution $\mathscr{P}_t^\lambda$ of (19) with $\Pi_\lambda^K = 0$ and $K := \overline{\mathscr{R}} C^\top V^{-1}$. Let $\mathscr{Q}_t^\lambda = \mathscr{P}_t^\lambda - \overline{\mathscr{R}}$, then $\mathscr{Q}_t^\lambda$ satisfies the following relation

$$\begin{aligned}
\dot{\mathscr{Q}}_t^\lambda &= A_\lambda \mathscr{Q}_t^\lambda + \mathscr{Q}_t^\lambda A_\lambda^\top + (1+\lambda)KVK^\top + \lambda(I-KC)\mathscr{P}_t^\lambda(I-KC)^\top - \lambda\overline{\mathscr{R}} \\
&= A_\lambda \mathscr{Q}_t^\lambda + \mathscr{Q}_t^\lambda A_\lambda^\top + (1+\lambda)KVK^\top - \lambda\overline{\mathscr{R}} + \lambda(I-KC)\mathscr{Q}_t^\lambda(I-KC)^\top + \lambda(I-KC)\overline{\mathscr{R}}(I-KC)^\top \\
&= A_\lambda \mathscr{Q}_t^\lambda + \mathscr{Q}_t^\lambda A_\lambda^\top + KVK^\top - \lambda KVK^\top + \lambda(I-KC)\mathscr{Q}_t^\lambda(I-KC)^\top + \lambda KC\overline{\mathscr{R}}C^\top K^\top
\end{aligned}$$

where the last equality followed from the fact that

$$\begin{aligned}
(I-KC)\overline{\mathscr{R}}(I-KC)^\top - \mathscr{R} &= KC\overline{\mathscr{R}}C^\top K^\top - 2\overline{\mathscr{R}}C^\top V^{-1} C\overline{\mathscr{R}} \\
&= -2KVK^\top + KC\overline{\mathscr{R}}C^\top K^\top.
\end{aligned}$$

Because of condition (47) and Theorem 6.1, we have $\lim_{t\to\infty} \mathscr{P}_t^\lambda(t) = \overline{\mathscr{P}}_\lambda$. Hence, by letting $\overline{\mathscr{Q}}_\lambda := \overline{\mathscr{P}}_\lambda - \overline{\mathscr{R}}$ (which is not necessarily positive definite), we have that $\lim_{t\to\infty} \mathscr{Q}_t^\lambda(t) = \overline{\mathscr{Q}}_\lambda$. Moreover, $\overline{\mathscr{Q}}_\lambda$ solves the following algebraic equation:

$$0 = A_\lambda \overline{\mathscr{Q}}_\lambda + \overline{\mathscr{Q}}_\lambda A_\lambda^\top + KVK^\top - \lambda KVK^\top + \lambda(I-KC)\overline{\mathscr{Q}}_\lambda(I-KC)^\top + \lambda KC\overline{\mathscr{R}}C^\top K^\top$$

Alternatively, the matrix $\overline{\mathscr{Q}}_\lambda$ satisfies the equation,

$$\overline{\mathscr{Q}}_\lambda = \overline{\mathscr{F}}_\lambda^K(\overline{\mathscr{Q}}_\lambda) + \overline{\mathfrak{G}}_\lambda^K,$$

where $\overline{\mathfrak{G}}_\lambda^K := \int_0^\infty \mathrm{e}^{A_\lambda t} K\big((1-\lambda)V + \lambda C\overline{\mathscr{R}}C^\top\big)K^\top \mathrm{e}^{A_\lambda^\top t}\,\mathrm{d}t$. Iterating the fixed point equation, we get

$$\overline{\mathscr{Q}}_\lambda = \left(\overline{\mathscr{F}}_\lambda^K\right)^k(\overline{\mathfrak{G}}_\lambda^K) + \sum_{j=0}^{k-1}\left(\overline{\mathscr{F}}_\lambda^K\right)^j(\overline{\mathfrak{G}}_\lambda^K) \overset{k\to\infty}{=} (I - \overline{\mathscr{F}}_\lambda^K)^{-1}(\overline{\mathfrak{G}}_\lambda^K),$$

and hence we arrive at (48): $\overline{\mathscr{P}}_\lambda = \overline{\mathscr{R}} + (I - \overline{\mathscr{F}}_\lambda^K)^{-1}(\overline{\mathfrak{G}}_\lambda^K)$. The inequality in (49) is obtained by taking the norms and using the bound $\left\|\mathrm{e}^{tA_\lambda}\right\| = \left\|\mathrm{e}^{tA}\cdot \mathrm{e}^{-0.5t\lambda I}\right\| \leqslant \mathrm{e}^{(\|A\|-0.5\lambda)t}$, which results in

$$\limsup_{\lambda\to\infty} \left\|\overline{\mathfrak{G}}_\lambda^K\right\| \leqslant \left\|K(C\overline{\mathscr{R}}C^\top - V)K^\top\right\|,$$

along with the bound $\|\mathscr{N}_\lambda\| \leqslant \beta_\lambda$. □

[8] We use the fact that $\left\|\left(I - \left(\overline{\mathscr{F}}_\lambda^K\right)^\ell\right)^{-1}\right\| \leqslant \|I\| + \left\|\left(\overline{\mathscr{F}}_\lambda^K\right)^\ell\right\| + \left\|\left(\overline{\mathscr{F}}_\lambda^K\right)^\ell\right\|^2 + \cdots = \sum_{j=0}^\infty \left\|\left(\overline{\mathscr{F}}_\lambda^K\right)^\ell\right\|^j = \left(1 - \left\|\left(\overline{\mathscr{F}}_\lambda^K\right)^\ell\right\|\right)^{-1}$.

9 Conclusions

In this paper, we studied the problem of filter design for Lipschitz nonlinear systems under randomly sampled output measurements. We provided filter design equations and derived equations for computing the covariance of the state estimation error resulting from our filter design. Conditions were provided under which the expected error covariance remains bounded, and the dependence on the sampling rate was also studied in some special cases.

To build on this line of work, it could be interesting to address similar questions with other types of distributions on the measurement process as most of the developments in the current work are specific to Poisson processes of constant intensity. Such generalizations could start with considering sampling processes where the rate depends on the state. It would be interesting to see if the tools developed in process allow us to handle a broader system class with state-dependent diffusion coefficient matrix G. One could also think of developing more general tools for studying dependence on the sampling rate, which at the moment is presented only for the linear measurement case. In this regard, one can also draw connections with the results that appear in the literature on cheap control and quantifying filtering error bounds when the coefficient next to diffusion terms in the system dynamics converge to zero [21, 28].

Appendix

A Proof of Proposition 5.2

Proof. To prove the existence of solution to equation (22), we first consider the differential equation (19) with a time-varying $K(\cdot)$. The resulting time-varying differential equation, subject to the initial condition $\mathscr{P}(0)=\mathscr{P}_0$, has the solution

$$\mathscr{P}_t=\varphi(t,\mathscr{P},M_{\mathscr{P}},K):=\Phi(t,0)\mathscr{P}_0\Phi^\top(t,0)+\int_0^t\Phi(t,s)\Big(GG^\top+\Pi(\mathscr{P}_s)+\lambda K_sM_{\mathscr{P}_s}K_s^\top\Big)\Phi^\top(t,s)\mathrm{d}s.$$

For any given $K(\cdot)$ and a positive definite matrix $M_{\mathscr{P}}\in\mathbb{R}^{p\times p}$ (affine in $\mathscr{P}$), we can find a solution of the Volterra equation $\mathscr{P}_t=\varphi(t,\mathscr{P},M_{\mathscr{P}},K)$ by successive approximation. Let us construct a sequence of these solutions P_i for $i\geqslant 1$ such that $M_0=I_{p\times p}$, $K_0=C^\top$, $\mathscr{P}_{i+1}(t)=\varphi(t,\mathscr{P}_{i+1},M_i,K_i)$, $M_{i+1}=(1+c)C\mathscr{P}_{i+1}C^\top+(2c+c^2)\operatorname{tr}P_{i+1}I_{p\times p}+V$, and $K_{i+1}=M_{i+1}^{-1}C^\top\mathscr{P}_{i+1}$. Notice that all $\mathscr{P}_i$ are positive semidefinite, all M_i are positive definite and so invertible, all matrices $K_i,M_i,\mathscr{P}_i$, $i\in\mathbb{N}$, are measurable and bounded with respect to t. Consider $Q_i(t)=\mathscr{P}_i(t)-\mathscr{P}_{i+1}(t)$ with $Q_i(0)=0$. From (19) and (31), we get

$$\begin{aligned}
\dot{Q}_i&=\Psi(\mathscr{P}_i,M_i,K_i)-\Psi(\mathscr{P}_{i+1},M_{i+1},K_{i+1})\\
&\geqslant\Psi(\mathscr{P}_i,M_i,K_i)-\Psi(\mathscr{P}_{i+1},M_{i+1},K_i)\\
&=(A-\lambda K_iC)Q_i+Q_i(A-\lambda K_iC)^\top+\Pi(Q_i)+\lambda K_i(M_i-M_{i+1})K_i^\top\\
&=(A-\lambda K_iC)Q_i+Q_i(A-\lambda K_iC)^\top+\Pi(Q_i)+\lambda(1+c)K_iCQ_iC^\top K_i^\top+\lambda\operatorname{tr}(Q_i)(2c+c^2)K_iK_i^\top\\
&=(A-\lambda K_iC)Q_i+Q_i(A-\lambda K_iC)^\top+\Pi(Q_i)+\lambda K_i\Big((1+c)CQ_iC^\top+\operatorname{tr}(Q_i)(2c+c^2)I_p\Big)K_i^\top
\end{aligned}$$

so there exists a positive semidefinite matrix R such that

$$\dot{Q}_i=(A-\lambda K_iC)Q_i+Q_i(A-\lambda K_iC)^\top+\Pi(Q_i)+\lambda K_i\Big((1+c)CQ_iC^\top+\operatorname{tr}(Q_i)(2c+c^2)I_p\Big)K_i^\top+R.$$

Hence, $Q_i(t)$ can be rewritten as a solution of the equation

$$Q_i(t)=\int_0^t\Phi(t,s)(\Pi(Q_i)+\lambda K_iM_{Q_i}K_i^\top+R)\Phi^\top(t,s)\mathrm{d}s,$$

implying that $Q_i(t) \geqslant 0$ for each $i \geqslant 1$ and $t \in [0,T]$, as it can be calculated by successive approximation with $Q_i(0) = 0$. Consequently, the sequence $\{\mathscr{P}_i(t)\}_{i\in\mathbb{N}}$ is non-increasing for each $t \in [0,T]$. It also means that on the finite segment $[0,T]$, $\mathscr{P}_i(t)$ is uniformly bounded by the number, $\sup\{|\mathscr{P}_1(t)| \mid t \in [0,T]\}$, for every $i \geqslant 1$ and $t \in [0,T]$; consequently $M_i(t)$ and $K_i(t)$ are uniformly bounded as well. Moreover, we let $\mathscr{P}_t = \lim_{i\to\infty} \mathscr{P}_i(t)$, $M_t = \lim_{i\to\infty} M_i(t) = (1+c)C\mathscr{P}_t C^\top + (2c+c^2)\operatorname{tr}(\mathscr{P}_t) I_{n\times n} + V$, $K_t = \lim_{i\to\infty} K_i(t) = \mathscr{P} C^\top M^{-1}$. So, the Volterra equation has the form: $\mathscr{P}_t = \varphi(t, \mathscr{P}, M, K)$ for M_t and K_t defined above, meaning that the chosen solution $\mathscr{P}$ is well-defined and unique. □

B Proof of Lemma 6.3

Let $\mathscr{K}_{\mathscr{P}_t} = \mathscr{P}_t C^\top M_{\mathscr{P}_t}^{-1}$ and let $\Phi(t,s)$ denote the fundamental matrix associated with $A - \lambda \mathscr{K}_{\mathscr{P}_t} C$. Then

$$\mathscr{P}_t = \int_0^t \Phi(t,s)(\Pi(\mathscr{P}_s) + GG^\top + \lambda \mathscr{K}_{\mathscr{P}_s} M_{\mathscr{P}_s} \mathscr{K}_{\mathscr{P}_s}^\top)\Phi^\top(t,s)\mathrm{d}s. \tag{50}$$

Let us fix an arbitrary $\tau \geqslant 0$ and define $\widetilde{\mathscr{K}}_{\mathscr{P}_t} = \mathscr{K}_{\mathscr{P}_{t+\tau}}$. Further, let $\widetilde{\Phi}(t,s)$ be the fundamental matrix associated with $A - \lambda \widetilde{\mathscr{K}}_{\mathscr{P}_t} C$; then $\widetilde{\Phi}(t,s) = \Phi(t+\tau, s+\tau)$. Let $\widetilde{\mathscr{P}}_t$ be the solution of

$$\dot{\widetilde{\mathscr{P}}}_t = \Psi(\widetilde{\mathscr{P}}_t, M_{\widetilde{\mathscr{P}}_t}, \widetilde{\mathscr{K}}_{\mathscr{P}_t}), \quad \widetilde{\mathscr{P}}_0 = 0,$$

then $\mathscr{P}_t \leqslant \widetilde{\mathscr{P}}_t$, for each $t \geqslant 0$, by minimum property. Moreover, we have the following representations of $\widetilde{\mathscr{P}}_{t-\tau}$:

$$\begin{aligned}
\widetilde{\mathscr{P}}_{t-\tau} &= \int_0^{t-\tau} \widetilde{\Phi}(t-\tau,s)(\Pi(\widetilde{\mathscr{P}}_s) + GG^\top + \lambda \widetilde{\mathscr{K}}_{\mathscr{P}_s} M_{\widetilde{\mathscr{P}}_s} \widetilde{\mathscr{K}}_{\mathscr{P}_s}^\top)\widetilde{\Phi}^\top(t-\tau,s)\mathrm{d}s \\
&= \int_0^{t-\tau} \Phi(t,s+\tau)(\Pi(\widetilde{\mathscr{P}}_s) + GG^\top + \lambda \mathscr{K}_{\mathscr{P}_{s+\tau}} M_{\widetilde{\mathscr{P}}_s} \mathscr{K}_{\mathscr{P}_{s+\tau}}^\top)\Phi^\top(t,s+\tau)\mathrm{d}s \\
&= \int_\tau^t \Phi(t,s)(\Pi(\widetilde{\mathscr{P}}_{s-\tau}) + GG^\top + \lambda \mathscr{K}_{\mathscr{P}_s} M_{\widetilde{\mathscr{P}}_{s-\tau}} \mathscr{K}_{\mathscr{P}_s}^\top)\Phi^\top(t,s)\mathrm{d}s \\
&\leqslant \int_0^t \Phi(t,s)(\Pi(\widetilde{\mathscr{P}}_{s-\tau}) + GG^\top + \lambda \mathscr{K}_{\mathscr{P}_s} M_{\widetilde{\mathscr{P}}_{s-\tau}} \mathscr{K}_{\mathscr{P}_s}^\top)\Phi^\top(t,s)\mathrm{d}s.
\end{aligned}$$

The last inequality, combined with (50), yields

$$\mathscr{P}_t - \widetilde{\mathscr{P}}_{t-\tau} \geqslant \int_0^t \Phi(t,s)(\Pi(\mathscr{P}_s - \widetilde{\mathscr{P}}_{s-\tau}) + \lambda \mathscr{K}_{\mathscr{P}_s} M_{(\mathscr{P}_s - \widetilde{\mathscr{P}}_{s-\tau})} K_{\mathscr{P}_s}^\top)\Phi^\top(t,s)\mathrm{d}s$$

and hence we get $\mathscr{P}_s - \widetilde{\mathscr{P}}_{s-\tau} \geqslant 0$, for each $s \geqslant \tau$. By letting $s = t+\tau$, we obtain the desired monotonicity, that is, $\mathscr{P}_t \leqslant \widetilde{\mathscr{P}}_t \leqslant \mathscr{P}_{t+\tau}$, and this completes the proof.

C Auxiliary Results

Lemma C.1 ([38, Lemma 4.1]). *Consider the equation, $GG^\top + KK^\top = \overline{GG}^\top$, and let F be an arbitrary matrix of suitable dimension.*

1. *If (A,G) is controllable, then $(A+KF, \overline{G})$ is controllable.*
2. *If (A,G) is stabilizable, then $(A+KF, \overline{G})$ is stabilizable.*

Lemma C.2 ([38, Lemma 5.1]). *Let (A,G) be controllable, $\mathscr{P}_0 \geqslant 0$ and let $\mathscr{P}$ satisfy*

$$\dot{\mathscr{P}}_t = A\mathscr{P}_t + \mathscr{P}_t A^\top + \Pi(\mathscr{P}_t) + GG^\top, \quad \mathscr{P}(0) = \mathscr{P}_0.$$

If there exists a constant matrix $\mathscr{Q} \geqslant 0$ such that $A\mathscr{Q} + \mathscr{Q}A^\top + \Pi(\mathscr{Q}) + GG^\top = 0$, then $\mathscr{P}_t \to \mathscr{Q}$ as $t \to +\infty$.

References

[1] M. Adès, P. E. Caines, and R. P. Malhamé. Stochastic optimal control under Poisson-distributed observations. *IEEE Transactions on Automatic Control*, 45(1):3–13, 2000.

[2] 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*, 57(3):607 – 619, 2012.

[3] A. Bishop and P. Del Moral. On the stability of Kalman–Bucy diffusion processes. *SIAM Journal on Control and Optimization*, 55(6):4015–4047, 2017.

[4] R. Brockett. Stochastic control: Lecture notes. 2009.

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

[6] M. H. A. Davis. Piecewise-deterministic Markov processes: a general class of nondiffusion stochastic models. *Journal of the Royal Statistical Society. Series B. Methodological*, 46(3):353–388, 1984. With discussion.

[7] M. H. A. Davis. *Markov Models and Optimization*, volume 49 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 1993.

[8] A. Doucet and A.M. Johansen. A tutorial on particle filtering and smoothing: Fifteen years later. In *Handbook of Nonlinear Filtering*, chapter 24, pages 656–704. Oxford University Press, 2011.

[9] M.D. Fragoso, O.L.V. Costa, and C.E. de Souza. A new approach to linearly perturbed Riccati equations arising in stochastic control. *Applied Mathematics and Optimization*, 37:99–126, 1998.

[10] M. Fujisaki, G. Kallianpur, and H. Kunita. Stochastic differential equations for the nonlinear filtering problem. *Osaka J. Math*, 9(1):19–40, 1972.

[11] A. Germani, C. Manes, and P. Palumbo. Linear filtering for bilinear stochastic differential systems with unknown inputs. *IEEE Transactions on Automatic Control*, 47(10):1726–1730, 2002.

[12] A. Gilman and I. Rhodes. Cone-bounded nonlinearities and mean-square bounds–Estimation upper bound. *IEEE Transactions on Automatic Control*, 18(3):260–265, 1973.

[13] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu. A survey of recent results in networked control systems. *Proceedings of the IEEE*, 95(1):138 – 162, 2007.

[14] 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*, pages 606–618. 2006.

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

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

[17] A. D. Kara, N. Saldi, and S. Yüksel. Weak Feller property of non-linear filters. *Systems & Control Letters*, 134:104512, 2019.

[18] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New York, 2nd edition, 1998.

[19] Hiroshi Kunita. Itô's stochastic calculus: Its surprising power for applications. *Stochastic Processes and their Applications*, 120(5):622–652, 2010.

[20] H. Kushner. Dynamical equations for optimal nonlinear filtering. *J. Differential Equations*, 30:209–245, 1967.

[21] H. Kwakernaak and R. Sivan. The maximally achievable accuracy of linear optimal regulators and linear optimal filters. *IEEE Trans. on Automatic Control*, AC-17(1):79–86, 1972.

[22] R. Liptser and A.N. Shiryaev. *Statistics of Random Processes. Part I & II*. Springer-Verlag, Berlin, 2nd edition, 2001.

[23] 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*, 48(4):670–676, 2003.

[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*, pages 110–114, 2018.

[25] M. Micheli. Random sampling of a continuous-time stochastic dynamical system: Analysis, state estimation, and applications. Master's thesis, Dept. of Elec. & Comp. Engg., University of California, Berkley, 2001.

[26] M. Micheli and M.I. Jordan. Random sampling of a continuous-time stochastic dynamical system. In *Proc. of 15th Intl. Symp. Mathematical Theory of Networks and Systems (MTNS)*, 2002.

[27] S.K. Mitter. Filtering and stochastic control: A historical perspective. *IEEE Control Systems Magazine*, 16(3):67–76, 1996.

[28] J. Picard. Efficiency of the extended Kalman filter for nonlinear systems with small noise. *SIAM J. Control & Optimization*, 51(3):843–885, 1991.

[29] 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*, 49(9):1453–1464, 2004.

[30] R.L. Stratonovich. Conditional markov processes. *Theory of Probability & Its Applications*, 5(2):156–178, 1960.

[31] 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.

[32] T. Sutter, A. Ganguly, and H. Koeppl. A variational approach to path estimation and parameter inference of hidden diffusion processes. *J. Machine Learning Research*, 17:1–37, 2016.

[33] A. Taghvaei, J. de Wiljes, P.G. Mehta, and S. Reich. Kalman filter and its modern extensions for the continuous-time nonlinear filtering problem. *ASME. J. Dyn. Sys., Meas., Control*, 140(3):030904–11, 2017.

[34] A. Tanwani, D. Chatterjee, and D. Liberzon. Stabilization of continuous-time deterministic systems under random sampling: Overview and recent developments. In T. Başar, editor, *Uncertainty in Complex Networked Systems*, pages 209–246. Springer Nature, Switzerland, 2018.

[35] A. Tanwani and A. R. Teel. Stabilization with event-driven controllers over a digital communication channel with random transmissions. In *Proc. 56th IEEE Conference on Decision and Control*, pages 6063–6068. 2017.

[36] A.R. Teel. Lyapunov conditions certifying stability and recurrence for a class of stochastic hybrid systems. *Annual Reviews in Control*, 37:1 – 24, 2013.

[37] R. van Handel. Nonlinear filtering and systems theory. In *Proceedings of the 19th International Symposium on Mathematical Theory of Networks and Systems (MTNS)*, pages 2157–2168, 2010.

[38] W.M. Wonham. On a matrix Riccati equation of stochastic control. *SIAM Journal on Control*, 6(4):681–697, 1968.

[39] T. Yang, P. G. Mehta, and S. P. Meyn. Feedback particle filter. *IEEE Transactions on Automatic Control*, 58(10):2465–2480, 2013.

[40] M. Zakai. On the optimal filtering of diffusion processes. *Z. Wahrsch. Verw. Gebiete*, 11(3):230–243, 1969.