2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Transport Inspired Particle Filters with Poisson-Sampled Observations in Gaussian Setting

Olga Yufereva Aneel Tanwani

Abstract— **Motivated by the need for developing computationally efficient solutions to filtering problem with limited information, this article develops particle filtering algorithms for continuous-time stochastic processes with time-sampled observation process. The state process is modeled by a continuous-time linear stochastic differential equation driven by Wiener process, and the observation process is a linear mapping of the state with additive Gaussian noise. For practical reasons, we assume that the observations are time-sampled and the underlying sampling process is a Poisson counter. With the aim of developing particle filters for this system, we first propose a mean-field type process which is an observation-driven stochastic differential equation such that the conditional distribution of this process given the observations coincides with the optimal filtering distribution. This model is then used to simulate a collection of particles which are driven only by the sample mean and sample covariance, without simulating the differential equation for the covariance matrix. It is shown that the dynamics of the sample mean and the sample covariance coincide with the optimal ones. An academic example is included for illustration.**

Index Terms— **Continuous-discrete filters; Poisson-sampled observations; Mean-field process; Particle filters.**

I. Introduction

Since the pioneering work of Kalman and Bucy in [1], the problem of filtering in dynamical systems has received considerable attention among the researchers in different communities. The basic idea of computing a statistical estimate of a stochastic process conditioned upon a noisy observation process has found relevance in several applications. The problem has an elegant solution where the optimal filter is characterized by the conditional distribution of the process given the observation process. In the simplest form, with linear dynamics and Gaussian processes, this conditional expectation is Gaussian and one can characterize the filter entirely by describing the evolution of the first moment of the conditional distribution and the covariance of the estimation error conditioned upon the observations. In the nonlinear setting, however, one does not get a finite-dimensional filter in general and it is of interest to study to what extent the techniques developed for the linear case generalize for nonlinear problems. Moreover, for systems with higher dimensions, computing covariance can be burdensome and one looks for alternate methods to approximate the evolution of the error covariance or compute the filtering distribution.

Several techniques have been proposed in the literature with the motivation to provide good approximation of the filtering distribution [2]. Among these, the use of Monte Carlo integration methods for approximating the optimal distribution have gained significant interest in the literature [3], [4]. In the same spirit, particle filters provide a computationally attractive method for approximating the optimal conditional distribution [5]. This later approach is based on simulating the evolution of different particles through differential equations that depend on the associated empirical mean and the empirical error covariance. In continuous-time linear Gaussian setting, it can be shown that, as the number of particles tends to infinity, the empirical mean and the covariance converge to the optimal mean and the covariance of the Kalman-Bucy filter [6], [7]. In fact, the limiting behavior of these particles is described by a McKean–Vlasov type equation, which is also referred as the mean-field process. Put differently, and this is the viewpoint that we adopt later in this article, the particles provide an approximation of this mean-field process and their evolution is coupled to each other through empirical mean and empirical covariances with equal weights. In the literature, this limiting process is seen to be chosen in different ways, e.g., as a modification of the Kalman–Bucy equation [8], or as a non-diffusion equation that is optimal in the measure transportation sense [9].

Apart from developing computationally efficient algorithms, another important research direction is to study the filtering problem with constraints on the information available for computing the optimal distribution. In particular, motivated by the idea of implementing filters subject to observations transmitted over networks through some communication protocols, it is natural to stipulate that the observations arrive at some random time instants [10]. It is of interest to compute the conditional distribution of the state process conditioned upon this discrete observation process [11]. Such a setup for continuous-discrete filters is motivated by the implementation of filters over networks where the observation process is discretized by the presence of a digital channel between the plant and the filter. For certain technical reasons, and to better study the effect of mean sampling rate, the authors stipulate in their previous works that the sampling process is a Poisson counter. In particular, for a system class very close to the one studied in this paper, the authors in [12] propose a continuous-discrete

This work is partially supported by the project CyPhAI, financed by ANR-JST CREST program, with grant number ANR-20-JSTM-0001. The work of O. Yufereva was performed as part of research conducted in the Ural Mathematical Center with the financial support of the Ministry of Science and Higher Education of the Russian Federation (Agreement number 075-02-2023-913).

O. Yufereva and A. Tanwani are affiliated with CNRS – LAAS, University of Toulouse, CNRS, 31400 Toulouse, France (e-mail: oiufereva@laas.fr, aneel.tanwani@cnrs.fr). O. Yufereva is also affiliated with Krasovskii Institute of Mathematics and Mechanics, Ural Branch of RAS, Yekaterinburg, Russia.

filter and analyze the boundedness of error covariance as a function of the mean sampling rate.

Our primary objective in this article is to develop particle filters for continuous-time stochastic processes subject to randomly time-sampled observations. In particular, the state process is modeled by linear continuous-time Ornstein-Uhlenbeck process and the sampling process for the observations is a Poisson counter. In the literature, we do find some variants of particle filters in continuous-discrete setting with different objectives. The paper [13] provides one (and possibly the first) such example, where the authors use mollifiers in the particle equations to smoothen the dynamics, but no statements about the limiting process are provided. The paper [14] develops particle filters for nonlinear systems using the time-discretization procedure as a part of the derivation and studies convergence as the length of the sampling interval converges to zero. Building on the work of [14], the authors in [15] propose a particle filter involving resets in the estimate obtained by the filter, where the reset value is computed by solving a differential equation on a different time-scale.

In contrast to these aforementiond works, our focus is on developing online particle filters which update their estimate whenever a new measurement from the observation process arrives. In particular, these particles are continuous-discrete: the continuous dynamics describe the evolution between two measurements and the discrete dynamics governed by a Poisson counter provides an update rule whenever the observation is updated. We, therefore, propose a continuous-discrete limiting mean-field process and show that the distribution of this process conditioned upon the discrete observations coincides with the optimal distribution of the Kalman-Bucy filter. We use this process equation to describe a system of interacting articles which are driven by empirical mean and the empirical covariance. These particles are shown to be consistent with the proposed limiting process. This kind of analysis sets up the ground work for studying the performance of these filters as a function of the mean sampling rate.

II. Problem Setup

We consider continuous-time stochastic dynamical systems described by

$$\mathrm{d}x_t = Ax_t\mathrm{d}t + G\,\mathrm{d}\omega_t \tag{1}$$

where $(x_t)_{t\geq 0}$ is an $\mathbb{R}^n$-valued diffusion process describing the state. We let $(\Omega, \mathcal{F}, \mathsf{P})$ denote the underlying probability space. It is assumed that, for each $t \geq 0$, ω_t is an $\mathbb{R}^m$-valued standard Wiener process adapted to the filtration $\mathcal{F}_t \subset \mathcal{F}$, with the property that $\mathsf{E}[\mathrm{d}\omega\,\mathrm{d}\omega^\top] = I_{m\times m}\mathrm{d}t$, for each $t \geq 0$. The matrix $A \in \mathbb{R}^{n\times n}$ and the matrix $G \in \mathbb{R}^{n\times m}$ are assumed to be constant.

For the filtering problem, we associate a discrete observation process with system (1), which essentially provides some noisy partial measurements of the state process 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 $(\tau_k)_{k\in\mathbb{N}}$ taking values in $\mathbb{R}_{\geq 0}$ which denote the time instants at which the measurements are available for computing the statistical estimate of the state process. We introduce the process $(N_t)_{t\geq 0}$ defined as

$$N_t := \sup\big\{k \in \mathbb{N} \;\big|\; \tau_k \leq t\big\} \quad \text{for } t \in \mathbb{R}, \tag{2}$$

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

$$y_{\tau_{N_t}} = C\,x(\tau_{N_t}) + \nu_{\tau_{N_t}}, \qquad t \geq 0, \tag{3}$$

where $\nu_{\tau_k} \sim \mathcal{N}(0, V_{\tau_k})$, that is, ν_{τ_k} is a mean zero Gaussian process with variance V_{τ_k}, which is assumed to be symmetric positive definite for each $k \in \mathbb{Z}_+$.

Define $Y_t = y(\max_{\tau_k\leq t}\{\tau_k\})$, that is a stochastic process with sample paths that are piece-wise constant right-continuous function with probability one. In fact, all the stochastic processes in this paper are considered to be càdlàg (continue à droite avec limite à gauche), and hence their sample paths are also càdlàg. Let $\mathcal{Y}_t$ denote the natural sigma-algebra generated by this piece-wise continuous output $Y_{[0,t]}$. For the filtering problem, we are interested in the càdlàg stochastic process $\widehat{x}_t$, called the estimator, that minimizes the mean-square error, i.e. $\arg\min \mathsf{E}\left[|x_t - \widehat{x}_t|^2 \mid \mathcal{Y}_t\right]$. The optimal (Bayes) solution to this minimization problem is the conditional distribution of the process x_t given the observation process, that is,

$$\mathrm{Law}(x_t|\mathcal{Y}_t). \tag{4}$$

Due to the linearity, and Gaussian noise, this distribution is Gaussian and it suffices to have only its first and second moments. However, the evolution of the mean $\widehat{x}_t := \mathsf{E}[x_t \mid \mathcal{Y}_t]$, depends on the evolution of the error covariance matrix $P_t := \mathsf{E}[(x_t - \widehat{x}_t)(x_t - \widehat{x}_t)^\top|\mathcal{Y}_t]$ and the computation of P_t is often costly for higher dimensional systems.

On the other hand, using the particle approach, it is possible to compute (an approximation) of the conditional distribution without solving the differential equation for the evolution of covariance matrix. This can be done by finding a process S_t that satisfies

$$\mathrm{Law}(S_t \mid \mathcal{Y}_t) \sim \mathrm{Law}(x_t \mid \mathcal{Y}_t) \tag{5}$$

and describing a system of particles whose limiting behavior (as the number of particles tend to infinity) converges to S_t. The computational advantage of this approach is that the simulation of particles, in general, is more efficient and potentially applicable in nonlinear systems as well (although there are very few instances of formal analysis in nonlinear setting). As an example of approximating a process using the particles, one can take for instance a result from [16, Theorem 1.4, p.172] where it is shown that if a process S_t satisfies a simple version of McKean–Vlasov type equation

$$\mathrm{d}S_t = \mathrm{d}B_t + \Big(\int b(S_t, \bar{S})\mu_t(\mathrm{d}\bar{S})\Big)\mathrm{d}t,$$

where $\mu_t(\mathrm{d}\bar{S})$ is the law of S_t, B_t is a standard Brownian motion, the function b is bounded and Lipschitz continuous and initial distribution is given, then S_t can be approximated by a system of M interacting particles S_t^i, $i = 1, \dots, M$ with the dynamics

$$\mathrm{d}S_t^i = \mathrm{d}B_t^i + \frac{1}{M}\sum_{j=1}^{M} b(S_t^i, S_t^j)\mathrm{d}t, \quad i = 1, \dots, M$$

and with the corresponding initial distributions. In particular, when $M \to \infty$, each S_t^i approaches a process which is an independent copy of the process S_t. For the filtering problem, with the observation process Z_t given in the form $\mathrm{d}Z_t = Cx_t\mathrm{d}t + V\mathrm{d}\omega_t$, the authors in [9] show that the property $\mathrm{Law}(S_t \mid \mathcal{Y}_t) \sim \mathrm{Law}(x_t \mid \mathcal{Y}_t)$ holds if S_t satisfies the following McKean–Vlasov type equation

$$\begin{aligned}\mathrm{d}S_t = AS_t\mathrm{d}t + K_{1,t}(S_t - \widehat{S}_t)\mathrm{d}t \\ + K_{2,t}\left(\mathrm{d}Z_t - \frac{CS_t + C\widehat{S}_t}{2}\mathrm{d}t\right), \quad (6)\end{aligned}$$

where $\widehat{S}_t = \mathsf{E}[S_t \mid Z_{[0,t]}]$, and $K_{1,t}$, $K_{2,t}$ are certain gain matrices. The authors in [9] also show that this evolution rule minimizes certain cost associated with the transportation of probability measures. The evolution equations for the individual particles are then obtained from (6) by replacing $\widehat{S}_t$ with the empirical mean.

The basic problem studied in this paper is to design particle filters for the continuous-time system (1) with the discrete observation process (3), and this is done in following steps:

- Find a process s_t such that $\mathsf{E}(s_t \mid \mathcal{Y}_t) \sim \mathsf{E}(x_t \mid \mathcal{Y}_t)$.
- Describe a system of particles s_t^i, $i = 1, \dots, M$ coupled to each other via empirical mean and empirical variance, such that, each s_t^i represents an independent copy of s_t when $M \to \infty$.
- Show that the empirical mean of the particles is consistent with the optimal solution to the filtering problem.

The first step basically corresponds to describing a mean-field filtering process for system (1), (3) which is addressed in Section IV. The corresponding particle system, discussed in Section V, provides an approximation of this mean field process when the number of particles is large enough.

III. Poisson Process Integration

A. *Differential equation with Poisson processes*

In this paper, we work under the assumption that the observation process is driven by a Poisson counter. Consequently, we use integration with respect to this Poisson process to define the filtering equations and it is important to develop a chain rule to describe the derivative of functions of the processes governed by such differential equations.

Definition 3.1: Fix $\lambda > 0$. A family of random variables $(N_t)_{t\geq 0}$ with values in $\mathbb{Z}_+$ is called a Poisson process of intensity λ if

1) $N_0 = 0$;
2) its increments are independent, i.e. $(N_{t_1} - N_{t_0}), \dots, (N_{t_n} - N_{t_{n-1}})$ are independent for all $0 = t_1 < \dots < t_n$;
3) for all $t \geq 0$ the following infinitesimal property holds

$$\begin{aligned}&\mathsf{P}(N_{t+\delta} - N_t = 0) = 1 - \lambda\delta + o(\delta),\\ &\mathsf{P}(N_{t+\delta} - N_t = 1) = \lambda\delta + o(\delta),\\ &\mathsf{P}(N_{t+\delta} - N_t \geq 2) = o(\delta).\end{aligned}$$

For equivalent definitions, we refer to [17, Sec. 2.3] and more properties of N_t are provided in [18, Ch. I, Sec. 3]. Note that each sample path of a Poisson process is a piecewise constant function and almost all of them are such that the difference between a value N_t and its left limit $N_{t^-} = \lim_{\delta\downarrow 0} N_{t-\delta}$ equals either 0 or 1 for all $t \geq 0$:

$$\Delta N_t := (N_t - N_{t^-}) \in \{0, 1\}.$$

Corollary 3.1: Let $(N_t)_{t\geq 0}$ be a Poisson counter and let f and g be Lipschitz continuous functions. A process e, which sample paths are defined as follows

$$\mathrm{d}e_t/\mathrm{d}t = f(e_t) \quad \text{for } \Delta N_t = 0 \qquad (7a)$$

$$e_t = e_{t^-} + g(e_{t^-})\,\Delta N_t \quad \text{for } \Delta N_t \neq 0 \qquad (7b)$$

is a càdlàg process.

In our case, it is convenient to introduce the following differential notation for such processes:

$$\mathrm{d}e_t = f(e_t)\mathrm{d}t + g(e_t)\mathrm{d}N_t. \qquad (8)$$

An $\mathbb{R}^n$-valued càdlàg process $(e_t)_{t\geq 0}$ is called a solution of (8) if, for almost all sample path of Poisson process $(N_t)_{t\geq 0}$, it satisfies (7). Such a solution is uniquely defined when $f : \mathbb{R}^n \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}^n$ are Lipschitz continuous functions.

To design a filter, we need the chain rule for such processes.

Proposition 3.2 (Chain rule): Let $f : \mathbb{R}^n \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}^n$ be Lipschitz continuous, and let $(e_t)_{t\geq 0}$ be a solution of (8). If $\psi\colon \mathbb{R}^n \to \mathbb{R}$ is a differentiable function, then

$$\begin{aligned}\mathrm{d}\psi(e_t) = (\nabla\psi(e_t))^\top f(e_t)\mathrm{d}t \\ + \big(\psi\big(e_t + g(e_t)\big) - \psi(e_t)\big)\,\mathrm{d}N_t. \quad (9)\end{aligned}$$

Proof: Let us consider only such sample paths of the Poisson process N that $\Delta N_t \in \{0, 1\}$ for all $t > 0$. Then, we can check (9) separately for $\{t > 0 \mid \Delta N_t = 0\}$ and for $\{t > 0 \mid \Delta N_t = 1\}$.

In the first case, we have that $\mathrm{d}e_t/\mathrm{d}t = f(e_t)$ by definition. Plugging this and (9) in (7a) for the process $\psi(e_t)$ we get the equality and so the case $\Delta N_t = 0$ is checked.

In the second case, we should notice that each finite interval $[0, t']$ contains only a finite set of the discontinuity points of N_t, say $\{\tau_i\}_{i=1,\dots,k} = \{t \in]0, t'] \mid \Delta N_t \neq 0\}$. Consider an arbitrary $i \in \{1, \dots, k\}$. For the process e_t we have $e_{\tau_i} = e_{\tau_i^-} + g(e_{\tau_i^-})\Delta N_{\tau_i}$ by definition. Using this and (9) to check (7b) for the process $\psi(e_t)$ we get that its right-hand side is

$$\psi(e_{\tau_i^-}) + \Big(\psi\big(\big(e_{\tau_i^-} + g(e_{\tau_i^-})\Delta N_{\tau_i}\big)\big) - \psi(e_{\tau_i^-})\Big)\Delta N_{\tau_i}$$

which equals $\psi(e_{\tau_i})$, the left-hand side of (7b), as $\Delta N_{\tau_i} = 1$. This completes the proof. ■

B. Filters with discrete observations

Coming to the question of optimal filter for continuous system (1) with discrete observations (3), we first recall the filter design based on computing the covariance matrix exactly via the differential equations. For an arbitrary strictly increasing real-valued sequence $(\tau_k)_{k\in\mathbb{Z}_+}$, an optimal filter is designed in [11, Th. 7.1]. If we specify a sequence $(\tau_k)_{k\in\mathbb{Z}_+}$ so that it corresponds to the arrival times of a Poisson process, we can reformulate the result from [11, Th. 7.1] as follows.

Theorem 3.3: Suppose that the process $(N_t)_{t\geq 0}$ in (2) is a Poisson counter. The optimal filter for the system (1), (3) is a Gaussian process for which the mean $\widehat{x}_t$ satisfies

$$\mathrm{d}\widehat{x}_t = A\widehat{x}_t\mathrm{d}t + K_t(y_t - C\widehat{x}_t)\mathrm{d}N_t, \tag{10}$$

and the covariance matrix P_t satisfies

$$\mathrm{d}P_t = (AP_t + P_tA^\top + GG^\top)\mathrm{d}t - K_tCP_t\mathrm{d}N_t \tag{11}$$

where $K_t = P_tC^\top(CP_tC^\top + V_t)^{-1}$.

In the above result, one makes the observation that the optimal conditional distribution is Gaussian for each realization of $(N_t)_{t\geq 0}$ despite the fact that the mean and covariance are discontinuous along each sample path.

IV. Continuous-Discrete Mean-Field Process

In contrast to the filter proposed in Theorem 3.3, we now turn our attention to the main topic of the paper, that is, compute the (approximate) filtering distribution using the particle approach which would alleviate the burden of solving the covariance equation explicitly.

On the first glance, one may try the "continuous-discrete version" of the process (6) for Poisson-sampled observation case, and define the process S_t as follows:

$$\begin{aligned}\mathrm{d}S_t = AS_t\mathrm{d}t &+ K_{1,t}(S_t - \widehat{S}_t)\mathrm{d}t \\ &+ K_{2,t}\left(y_t - \frac{CS_t + C\widehat{S}_t}{2}\right)\mathrm{d}N_t\end{aligned} \tag{12}$$

for some gain matrices $K_{1,t}$ and $K_{2,t}$. It turns out that, for such processes, $\mathrm{Law}(S_t \mid \mathcal{Y}_t)$ is not necessarily equivalent to $\mathrm{Law}(x_t \mid \mathcal{Y}_t)$. Retaining the structure of the continuous part of the transport inspired equation (12), we allow for the update term to be a linear combination of the process itself and the conditional mean. Thus, we propose the following continuous-discrete model for the mean-field process:

$$\boxed{\begin{aligned}\mathrm{d}s_t &:= As_t\mathrm{d}t + \frac{1}{2}GG^\top Q_t^{-1}(s_t - \widehat{s}_t)\mathrm{d}t \\ &\quad + \Big(L_t(y_t - C\widehat{s}_t) - \Xi_t(s_t - \widehat{s}_t)\Big)\mathrm{d}N_t, \\ \widehat{s}_t &:= \mathsf{E}[s_t \mid \mathcal{Y}_t], \\ Q_t &:= \mathsf{E}[(s_t - \widehat{s}_t)(s_t - \widehat{s}_t)^\top \mid \mathcal{Y}_t],\end{aligned}} \tag{13}$$

with initial conditions $s_0 \sim x_0$. Here, for each $t \geq 0$, L_t is a gain matrix, and the matrix $\Xi_t \in \mathbb{R}^{n\times n}$ will be specified shortly. The special case of $\Xi_t = \frac{1}{2}L_tC$ brings us in the form (12), but this particular choice does not provide a consistent posterior distribution.

To show that a different choice of the matrix valued process Ξ_t provides a posterior distribution consistent with the optimal filter, we first derive the differential equations for the conditional mean $\widehat{s}_t$ and the conditional error covariance Q_t.

Proposition 4.1: Let L_t and Ξ_t be càdlàg $\mathcal{Y}_t$-measurable processes. The system (13) then yields

$$\mathrm{d}\widehat{s}_t = A\widehat{s}_t\mathrm{d}t + L_t(y_t - C\widehat{s}_t)\mathrm{d}N_t, \tag{14}$$

$$\begin{aligned}\mathrm{d}Q_t &= \left(AQ_t + Q_tA^\top + GG^\top\right)\mathrm{d}t \\ &\quad + \left(\Xi_tQ_t\Xi_t^\top - \Xi_tQ_t - Q_t\Xi_t^\top\right)\mathrm{d}N_t.\end{aligned} \tag{15}$$

Proof: Employing $\mathcal{Y}_t$-measurability of L_t and Ξ_t, we immediately obtain that

$$\mathrm{d}\widehat{s}_t = A\widehat{s}_t\mathrm{d}t + L_t(y_t - C\widehat{s}_t)\mathrm{d}N_t.$$

Next, we consider the error variable $e_t := s_t - \widehat{s}_t$, and observe that

$$\mathrm{d}e_t := \mathrm{d}(s_t - \widehat{s}_t) = Ae_t\mathrm{d}t + \frac{1}{2}GG^\top Q_t^{-1}e_t\mathrm{d}t - \Xi e_t\mathrm{d}N_t.$$

Using the chain rule from Proposition 3.2, we get

$$\begin{aligned}\mathrm{d}e_te_t^\top &= \Big(Ae_te_t^\top + \frac{1}{2}GG^\top Q_t^{-1}e_te_t^\top \\ &\quad + e_te_t^\top A^\top + \frac{1}{2}e_te_t^\top(GG^\top Q_t^{-1})^\top\Big)\mathrm{d}t \\ &\quad + \left((e_t - \Xi_te_t)(e_t - \Xi_te_t)^\top - e_te_t^\top\right)\mathrm{d}N_t.\end{aligned}$$

It follows that

$$\begin{aligned}\mathrm{d}Q_t &= \left(AQ_t + Q_tA^\top + GG^\top\right)\mathrm{d}t \\ &\quad + \left(\Xi_tQ_t\Xi_t^\top - \Xi_tQ_t - Q_t\Xi_t^\top\right)\mathrm{d}N_t,\end{aligned}$$

and hence (15) holds. ■

We now use the result of Proposition 4.1 to show that an appropriate choice of L_t and Ξ_t indeed yields a conditional distribution that is consistent with optimal filtering distribution.

Theorem 4.2: Consider system (13) with initial condition $s_0 \sim x_0$. Suppose that $L_t = Q_tC^\top(CQ_tC^\top + V_t)^{-1}$ and Ξ_t satisfies

$$\begin{aligned}(\Xi_t - I)&Q_t(\Xi_t^\top - I) \\ &= Q_t - Q_tC^\top(CQ_tC^\top + V_t)^{-1}CQ_t\end{aligned} \tag{16}$$

then $\mathrm{Law}(s_t \mid \mathcal{Y}_t) \sim \mathrm{Law}(x_t \mid \mathcal{Y}_t)$.

Proof: We first claim that the conditional distribution described by $\mathrm{Law}(s_t \mid \mathcal{Y}_t)$ is Gaussian. Indeed, it follows from Corollary 3.1, that the evolution of the process s_t is described by multiplying the initial normal distribution with certain matrices. It therefore suffices to show that the conditional mean and conditional variance of the process s_t coincides with the solution of (10) and (11), respectively.

The proposed L_t and Ξ_t are such that

$$\Xi_tQ_t\Xi_t^\top - \Xi_tQ_t - Q_t\Xi_t^\top = -L_tCQ_t$$

so the dynamics of (14) and (15) coincide with the dynamics of (10) and (11), respectively. Moreover, for the initial distribution $s_0 \sim x_0$, it holds that $\widehat{s}_0 = \widehat{x}_0 = \mathsf{E}[x_0]$, and similarly, $Q_0 = P_0$. The desired solution therefore follows from the uniqueness of solutions to the differential equations (14) and (15). ∎

So, to compute s_t, one needs to solve the quadratic equation (16) only countably many times almost surely, at the renewal times of the Poisson process, that is, whenever the observation process updates. We will provide some remarks about the solvability of the quadratic equation (16) in the next section.

V. Particle Filter

We now use the process equations (13) to define a system of particles. The inconvenient aspect of (13) is that the differential equation depends on the conditional mean and conditional covariance of the process. The use of particles allows us to replace these two terms with empirical mean and empirical covariance respectively.

For $M \geq n$ and $t \geq 0$, let M-particle system be defined as

$$\mathrm{d}s_t^i = As_t^i \mathrm{d}t + \frac{1}{2} GG^\top (Q_t^{\mathrm{e}})^{-1}(s_t^i - \widehat{s}_t^{\,\mathrm{e}})\mathrm{d}t + \Big(L_t^{\mathrm{e}}(y_t - C\widehat{s}_t^{\mathrm{e}}) - \Xi_t^{\mathrm{e}}(s_t^i - \widehat{s}_t^{\mathrm{e}})\Big)\mathrm{d}N_t, \tag{17a}$$

for $i = 1, \dots, M$, where $\widehat{s}_t^{\mathrm{e}}$ denotes the empirical mean

$$\widehat{s}_t^{\mathrm{e}} = \frac{1}{M}\sum_{i=1}^{M} s_t^i, \tag{17b}$$

and Q_t^{e} describes the empirical covariance

$$Q_t^{\mathrm{e}} = \frac{1}{M}\sum_{i=1}^{M}(s_t^i - \widehat{s}_t^{\mathrm{e}})(s_t^i - \widehat{s}_t^{\mathrm{e}})^\top. \tag{17c}$$

The gain L_t^{e} is defined as a function of the empirical mean and covariance as

$$L_t^{\mathrm{e}} = Q_t^{\mathrm{e}} C^\top (CQ_t^{\mathrm{e}} C^\top + V_t)^{-1}, \tag{17d}$$

and, similar to (16), the matrix Ξ_t^{e} satisfies

$$(\Xi_t^{\mathrm{e}} - I)Q_t^{\mathrm{e}}((\Xi_t^{\mathrm{e}})^\top - I) = Q_t^{\mathrm{e}} - Q_t^{\mathrm{e}} C^\top (CQ_t^{\mathrm{e}} C^\top + V_t)^{-1} CQ_t^{\mathrm{e}}. \tag{17e}$$

One readily observes that, for each particle, the only differential equation to be simulated is (17a), and the terms $\widehat{s}_t^{\mathrm{e}}$ and Q_t^{e} appearing in this equation are defined via static coupling with other particles.

A. Consisteny of the particle filter

We now show that the particle filter (17) is indeed consistent with the optimal filter of Theorem 3.3.

Proposition 5.1: The evolution of the sample mean $\widehat{s}_t^{\mathrm{e}}$ and the sample covariance Q_t^{e} satisfies the following equations

$$\mathrm{d}\widehat{s}_t^{\mathrm{e}} = A\widehat{s}_t^{\mathrm{e}} + L_t^{\mathrm{e}}(y_t - C\widehat{s}_t^{\mathrm{e}})\mathrm{d}N_t, \tag{18a}$$

$$\mathrm{d}Q_t^{\mathrm{e}} = \left(AQ_t^{\mathrm{e}} + Q_t^{\mathrm{e}} A^\top + GG^\top\right)\mathrm{d}t - L_t^{\mathrm{e}} CQ_t^{\mathrm{e}} \mathrm{d}N_t. \tag{18b}$$

where L_t^{e} satisfies (17d).

Proof: The differential equation for $\widehat{s}_t^{\mathrm{e}}$ can be obtained by taking the time derivative of the equation $\widehat{s}_t^{\mathrm{e}} = \frac{1}{M}\sum_{i=1}^{M} s_t^i$. Repeating the arguments of the proof of Proposition 4.1 and using $Q_t^{\mathrm{e}} = \frac{1}{M}\sum_{i=1}^{M}(s_t^i - \widehat{s}_t^{\mathrm{e}})(s_t^i - \widehat{s}_t^{\mathrm{e}})^\top$, we obtain that

$$\mathrm{d}Q_t^{\mathrm{e}} = \left(AQ_t^{\mathrm{e}} + Q_t^{\mathrm{e}} A^\top + GG^\top\right)\mathrm{d}t + \left(\Xi_t^{\mathrm{e}} Q_t^{\mathrm{e}} (\Xi_t^{\mathrm{e}})^\top - \Xi_t^{\mathrm{e}} Q_t^{\mathrm{e}} - Q_t^{\mathrm{e}}(\Xi_t^{\mathrm{e}})^\top\right)\mathrm{d}N_t.$$

Then we notice that (17e) is equivalent to

$$\Xi_t^{\mathrm{e}} Q_t^{\mathrm{e}} (\Xi_t^{\mathrm{e}})^\top - \Xi_t^{\mathrm{e}} Q_t^{\mathrm{e}} - Q_t^{\mathrm{e}}(\Xi_t^{\mathrm{e}})^\top = L_t^{\mathrm{e}} CQ_t^{\mathrm{e}}$$

which leads to (18b). ∎

The result of Proposition 5.1 therefore shows that the differential equations for empirical mean $\widehat{s}_t^{\mathrm{e}}$ and covariance Q_t^{e} are the same as those given in Theorem 3.3. To build on this result and show that s_t^{e} and Q_t^{e} are indeed consistent with $\widehat{x}_t$ and P_t, we also need consistency of initial conditions. Assuming that $s_0 \sim x_0$, we see that a realization of s_0^{e} approximates $\mathsf{E}[x_0]$ if the number of particles M is large enough. In the same way, for large enough M, a realization of Q_0 approximates P_0. Formal analysis of this approximation and studying the effect of sampling rate on the asymptotic behavior of the resulting error is a topic of further research.

B. Solvability of the design parameters

The injection gains L_t^{e} and Ξ_t^{e} in our particle system (17a) only apply at times of observation updates and are obtained from (17d) and (17e), resepctively. In these equations, we note that Q_t^{e} gets updated due to reset in s_t and $\widehat{s}_t^{\mathrm{e}}$ with the arrival of new observations. Thus, the key step for the simulation of the particles is to solve the algebraic equation (17e) at renewal times of the Poisson process. Solvability of this equation can be guaranteed and carried out in following steps:

Step 1: We can find an invertible matrix E_t such that $E_t E_t^\top = Q_t^{\mathrm{e}}$. This is always possible because, for the number of particles M large enough, Q_t^{e} is symmetric and positive definite. Using the eigenvalue decomposition, we can choose E_t to be the matrix of orthonormal eigenvectors multiplied by a diagonal matrix with square root of the eigenvalues.

Step 2: We then observe that $(Q_t^{\mathrm{e}} - L_t^{\mathrm{e}} CQ_t^{\mathrm{e}})$ is positive semi-definite. This follows from the fact that, we can write:

$$(Q_t^{\mathrm{e}} - L_t^{\mathrm{e}} CQ_t^{\mathrm{e}}) = (I - L_t^{\mathrm{e}} C)Q_t^{\mathrm{e}}(I - L_t^{\mathrm{e}} C)^\top + L_t^{\mathrm{e}} V_t L_t^{\mathrm{e}\top}$$

Step 3: Next, we compute the matrix F_t such that $F_t F_t^\top = (Q_t^{\mathrm{e}} - L_t^{\mathrm{e}} CQ_t^{\mathrm{e}})$. Due to the fact that $(Q_t^{\mathrm{e}} - L_t^{\mathrm{e}} CQ_t^{\mathrm{e}})$ is symmetric positive semi-definite, such a matrix F_t always exists.

Thus, using eigenvalue decomposition of the matrices Q_t^{e} and $(Q_t^{\mathrm{e}} - L_t^{\mathrm{e}} CQ_t^{\mathrm{e}})$, we have an analytic solution. That is, for each renewal time $t \geq 0$, we let $X_t = \Xi_t - I$, then the quadratic equation becomes

$$X_t E_t E_t^\top X_t^\top = F_t F_t^\top$$

for which the solution is $X_t = F_t E_t^{-1}$, that is, $\Xi_t = I + X_t$.

VI. Simulations

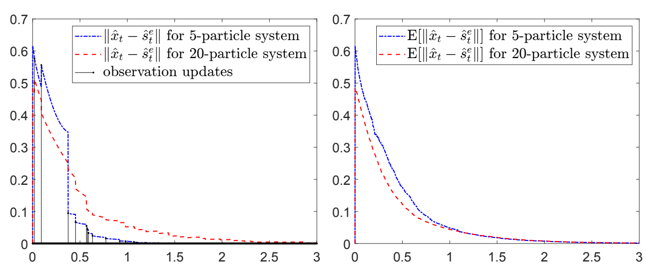
We will demonstrate the results of Section IV and Section V through an academic example. Consider the linear stochastic system

$$\mathrm{d}x_t = Ax_t\mathrm{d}t + G\mathrm{d}\omega_t \tag{19a}$$

$$y_{\tau_k} = Cx_{\tau_k} + \nu_{\tau_k}, \tag{19b}$$

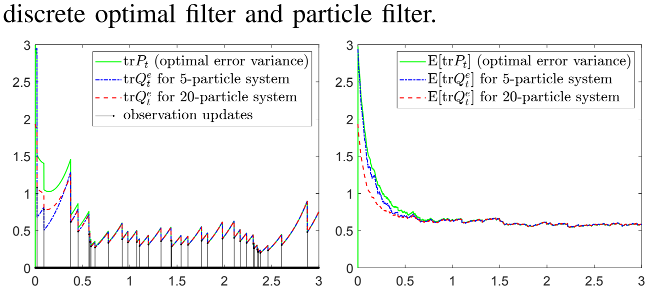
where $A = \begin{bmatrix} 0 & 3 & 1 \\ 2 & -2 & 1 \\ -2 & 1 & -3 \end{bmatrix}$, $C = \begin{bmatrix} 1 & -1 & 2 \\ 1 & 0 & 1 \end{bmatrix}$, $G = \begin{bmatrix} 0.5 & 0.5 & 0.5 \end{bmatrix}^\top$, and ν_t is normally distributed with mean $(0,0)^\top$ and the constant variance $V = \begin{bmatrix} 0.5 & 0.1 \\ 0.1 & 0.5 \end{bmatrix}$.

To measure the effectiveness of the particle system, we compare it with the optimal estimator. For a simulated and fixed path of the state x_t and the observation noise ν_t, we compare $\widehat{x}_t$ and $\widehat{s}^{\mathrm{e}}_t$. For the first moment, we consider the time plot of $\|\widehat{x}_t - \widehat{s}^{\mathrm{e}}_t\|$. While $\widehat{x}_t$ is defined by the differential equation (10), to compute $\widehat{s}^{\mathrm{e}}_t$ we simulate particles s^i_t and take their empirical mean. Figure 1a shows a realization of this first moment process with observation updates at discrete times. We fix Poisson intensity $\lambda = 10$. To get a more general picture we simulate 200 sample paths of the Poisson process, look at the time evolution of $\|\widehat{x}_t - \widehat{s}^{\mathrm{e}}_t\|$, and take their pointwise (in time) average, denoted as $\mathsf{E}\|\widehat{x}_t - \widehat{s}^{\mathrm{e}}_t\|$ in Figure 1b. Thus, Figure 1b presents the evolution of the expectation (with respect to the sampling process) of the first moment of the difference between $\widehat{x}_t$ and $\widehat{s}^{\mathrm{e}}_t$.



(a) One realization behavior. (b) Mean over 200 sample paths.

Fig. 1: Evolution of the first moment of the continuous-discrete optimal filter and particle filter.



(a) One realization behavior. (b) Mean over 200 sample paths.

Fig. 2: Evolution of the second moment of the continuous-discrete optimal filter and particle filter.

In a similar way, we compare the evolution of second moments as well. To compute P_t, we need to solve the matrix differential equation (11) and Q^{e}_t is computed as in (17c). The corresponding simulations are reported in Figure 2, where we compare $\mathrm{trace}(P_t)$ with $\mathrm{trace}(Q^{\mathrm{e}}_t)$ for a single realization in Fig. 2a, and over multiple realizations of sampling process in Fig. 2b.

VII. Conclusions

We considered the problem of designing particle filters for continuous-time linear stochastic systems with discrete observation process. The proposed mean-field type process is a continuous-discrete differential equation whose posterior distribution coincides with optimal conditional distribution. We then sample this process to design a system of interacting particles which are coupled through empirical mean and covariance. This system of particles is shown to be consistent with the optimal Kalman–Bucy filter in the sense that the mean and covariance satisfy the same differential equation.

As a topic of further investigation, we are analyzing quantitatively the approximation error between the empirical mean and the optimal mean, as well as the difference between empirical covariance and the optimal covariance. It remains to be seen how the approximation error varies with the number of particles. It is also interesting to see how this approximation is affected by the mean sampling rate of the underlying observation process.

References

[1] R. Kalman and R. Bucy, "New results in linear filtering and prediction theory," *Journal of Basic Engineering*, vol. 83, no. 1, pp. 95–108, 1961.

[2] B. Anderson and J. Moore, *Optimal Filtering*. Dover Publications, 1979.

[3] A. Doucet, N. De Freitas, N. J. Gordon *et al.*, *Sequential Monte Carlo methods in practice*. Springer, 2001, vol. 1, no. 2.

[4] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman filter: Particle filters for tracking applications*. Artech house, 2003.

[5] Z. Chen, "Bayesian filtering: From Kalman filters to particle filters, and beyond," *Statistics*, vol. 182, no. 1, pp. 1–69, 2003.

[6] G. Evensen, *Data Assimilation: The Ensemble Kalman Filter*, 2nd ed. Springer Science & Business Media, 2009.

[7] A. N. Bishop and P. Del Moral, "On the mathematical theory of ensemble (linear-Gaussian) Kalman–Bucy filtering," *Math. Control, Signals and Systems*, 2023.

[8] ——, "On the stability of matrix-valued Riccati diffusions," *Elec. J. of Probability*, vol. 24, no. 84, pp. 1–40, 2019.

[9] A. Taghvaei and P. G. Mehta, "An optimal transport formulation of the linear feedback particle filter," in *2016 American Control Conference (ACC)*. IEEE, 2016, pp. 3614–3619.

[10] J. P. Hespanha, "Modeling and analysis of networked control systems using stochastic hybrid systems," *Annual Reviews in Control*, vol. 38, no. 2, pp. 155 – 170, 2014.

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

[12] A. Tanwani and O. Yufereva, "Error covariance bounds for suboptimal filters with Lipschitzian drift and Poisson-sampled measurements," *Automatica*, vol. 122, p. 109280, 2020.

[13] K. Bergemann and S. Reich, "A mollified ensemble Kalman filter," *Q.J. Royal Meteorological Society*, vol. 136, no. 651, pp. 1636–1643.

[14] T. Yang, P. G. Mehta, and S. P. Meyn, "Feedback particle filter," *IEEE Transactions on Automatic Control*, vol. 58, no. 10, pp. 2465–2480, 2013.

[15] X. Chen, X. Luo, J. Shi, and S. S.-T. Yau, "General convergence result for continuous-discrete feedback particle filter," *International Journal of Control*, vol. 95, no. 11, pp. 2972–2986, 2022.

[16] A.-S. Sznitman, "Topics in propagation of chaos," *Ecole d'été de probabilités de Saint-Flour XIX—1989*, vol. 1464, pp. 165–251, 1991.

[17] Y. Suhov and M. Kelbert, *Probability and statistics by example: volume 2, Markov chains: a primer in random processes and their applications*. Cambridge University Press, 2008, vol. 2.

[18] P. E. Protter, *Stochastic differential equations*. Springer, 2005.