arXiv:1611.00170v3 [math.OC] 19 Sep 2017

Online Maximum Likelihood Estimation of the Parameters of Partially Observed Diffusion Processes

Simone Carlo Surace and Jean-Pascal Pfister

***Abstract*—We revisit the problem of estimating the parameters of a partially observed diffusion process, consisting of a hidden state process and an observed process, with a continuous time parameter. The estimation is to be done online, i.e. the parameter estimate should be updated recursively based on the observation filtration. Here, we use an old but under-exploited representation of the incomplete-data log-likelihood function in terms of the filter of the hidden state from the observations. By performing a stochastic gradient ascent, we obtain a fully recursive algorithm for the time evolution of the parameter estimate. We prove the convergence of the algorithm under suitable conditions regarding the ergodicity of the process consisting of state, filter, and tangent filter. Additionally, our parameter estimation is shown numerically to have the potential of improving suboptimal filters, and can be applied even when the system is not identifiable due to parameter redundancies. Online parameter estimation is a challenging problem that is ubiquitous in fields such as robotics, neuroscience, or finance in order to design adaptive filters and optimal controllers for unknown or changing systems.**

I. Introduction

We consider the following family of partially observed dimensional diffusion process under the probability measure P_θ:

$$dX_t = f(X_t,\theta)dt + g(X_t,\theta)dW_t, \tag{1}$$

$$dY_t = h(X_t,\theta)dt + dV_t, \tag{2}$$

parametrized by $\theta \in \Theta$, where $\Theta \subset \mathbb{R}^p$ is an open subset. The process X_t is called the hidden state or signal process with values in $\mathbb{R}^n$, and Y_t is called the observation process with values in $\mathbb{R}^{n_y}$. In addition, W_t, V_t are independent $\mathbb{R}^{n'}$- and $\mathbb{R}^{n_y}$-valued standard Wiener processes (signal and observation noise) respectively. For all $\theta \in \Theta$ we assume the initial conditions $X_0 \sim p_0(\theta)$ and $Y_0 = 0$ under P_θ, and that $f(\cdot,\theta), g(\cdot,\theta), h(\cdot,\theta)$ are functions from $\mathbb{R}^n \times \mathbb{R}^p$ to $\mathbb{R}^n$, $\mathbb{R}^{n\times n'}$, and $\mathbb{R}^{n_y}$ respectively that satisfy the usual conditions that ensure the existence and uniqueness in probability of strong solutions to Eqs. (1,2) for all $t \geq 0$. Additional regularity conditions for f, g, h in both arguments will be required for the convergence proof.

Simone Carlo Surace and Jean-Pascal Pfister are affiliated to the Institute of Neuroinformatics and Neuroscience Center Zurich, University of Zurich and ETH Zurich, 8057 Zurich, Switzerland. Email: surace@ini.uzh.ch and jpfister@ini.uzh.ch.

This work was supported by the Swiss National Science Foundation, grants PZ00P3_137200 and PP00P3_150637.

This paper is an expanded version of work that was first presented at the 2016 Time Series Workshop at NIPS (Neural Information Processing Systems) conference in Barcelona, Spain.

This setting is familiar in classical filtering theory, where the problem is to find (assuming the knowledge of θ) the conditional distribution of X_t conditioned on the history of observations $\mathcal{F}_t^Y = \sigma\{Y_s, 0 \leq s \leq t\}$. In this paper, we focus on the following parameter estimation problem: assuming that a system with parameter θ_0 generates observations Y_t, we want to estimate θ_0 from $\mathcal{F}_t^Y$ recursively.

It is a fundamental theorem of filtering theory that the *innovation process* I_t, defined by

$$I_t = Y_t - \int_0^t \hat{h}_s(\theta)ds, \quad \hat{h}_s(\theta) = \mathbb{E}_\theta\left[h(X_s,\theta)\middle|\mathcal{F}_s^Y\right], \tag{3}$$

is a $(P_\theta, \mathcal{F}_t^Y)$-Brownian motion. By applying Girsanov's theorem, we can change to a measure $\tilde{P}$ under which Y_t is a $(\tilde{P}, \mathcal{F}_t^Y)$-Brownian motion and thus (statistically) independent of both the hidden state X_t and the parameter θ. The change of measure has a Radon-Nikodym derivative

$$\left.\frac{dP_\theta}{d\tilde{P}}\right|_{\mathcal{F}_t^Y} = \exp\left[\int_0^t \hat{h}_s(\theta)\cdot dY_s - \frac{1}{2}\int_0^t \|\hat{h}_s(\theta)\|^2 ds\right], \tag{4}$$

where $\cdot$ denotes the Euclidean scalar product.

For a detailed exposition of the mathematical background (such as the Girsanov's theorem, changes of measure, or the filtering equation below), we suggest a look at the standard literature on filtering theory, e.g. [1].

This paper is structured in the following way. In the next section, we describe our method of obtaining recursive parameter estimates. Next, in Section III, we prove the almost sure convergence of the recursive parameter estimates to stationary points of the asymptotic likelihood. In Section IV we provide a few numerical examples, including cases where the model is not identifiable and the filter is suboptimal. Finally, in Section V we discuss the similarities and differences to existing methods of recursive parameter estimation.

II. Methods

In this paper, we consider the problem of finding an estimator $\tilde{\theta}_t$ that is $\mathcal{F}_t^Y$-measurable and recursively computable, such as to estimate θ_0 online from the continuous stream of observations. For this task, we propose an approach based on a modification of offline maximum likelihood estimation, and therefore need to compute the likelihood of the observations (also called incomplete-data likelihood) as a function of the model parameters. Since the reference measure $\tilde{P}$ does not depend on θ, we can express the incomplete-data log-likelihood function in terms of the optimal filter as

$$\mathcal{L}_t(\theta) = \log \frac{dP_\theta}{d\tilde{P}}\Big|_{\mathcal{F}_t^Y} = \int_0^t \hat{h}_s(\theta) \cdot dY_s - \frac{1}{2}\int_0^t \|\hat{h}_s(\theta)\|^2 ds. \tag{5}$$

A. *Offline algorithm*

We start by describing an offline method for parameter estimation using the log-likelihood function in Eq. (5), which serves as a basis for the online method.

If we were interested in offline learning, our goal would be to maximize the value of $\mathcal{L}_t(\theta)$ for fixed t. There is a number of methods to solve this optimization problem. Among these, a simple iterative method is the gradient ascent, where an estimate $\tilde{\theta}_k$ at iteration k is updated according to

$$\tilde{\theta}_{k+1} = \tilde{\theta}_k + \gamma_k \partial_\theta \mathcal{L}_t(\theta)\Big|_{\theta=\tilde{\theta}_k}, \tag{6}$$

where $\gamma_k > 0$ is called the *learning rate*, and ∂_θ denotes the gradient with respect to the parameter θ. At each iteration, the derivative of the likelihood function has to be recomputed. From Eq. (5), we obtain

$$\partial_\theta \mathcal{L}_t(\theta) = \int_0^t \left(dY_s - \hat{h}_s(\theta)ds\right)^\top \hat{h}_s^\theta(\theta), \tag{7}$$

where $\cdot^\top$ denotes the matrix transpose and the last factor of the integrand, denoted by

$$\hat{h}_s^\theta(\theta) \doteq \partial_\theta \hat{h}_s(\theta), \tag{8}$$

takes values in the matrices of size $n_y \times p$ and is called the *filter derivative* of h with respect to θ.[1]

In the following, we assume that $\hat{h}_t$ admits a finite-dimensional recursive solution or a finite-dimensional recursive approximation. This means that there is a $\mathcal{F}_t^Y$-adapted process $M_t(\theta)$ with values in $\mathbb{R}^m$ and a mapping $\psi_h : \mathbb{R}^m \times \Theta \to \mathbb{R}^{n_y}$ such that either $\hat{h}_t(\theta) = \psi_h(M_t(\theta), \theta)$ (in the case of an exact solution), or such that the equation holds approximately, i.e. with some bounds (preferably uniform in time) on

$$\mathrm{Var}\left[\hat{h}_t(\theta) - \psi_h(M_t(\theta), \theta)\right].$$

For example, in the linear-Gaussian case and if X_0 has a Gaussian distribution, the optimal filter can be represented in terms of a Gaussian distribution with mean μ_t and variance P_t, i.e. $m = 2$, $M_t = (\mu_t, P_t)$, and for $h_\theta(x) = \theta x$, we have $\psi_h(M_t(\theta), \theta) = \theta\mu_t$. Apart from the linear-Gaussian case [2] just mentioned, finite-dimensional (exact) recursive solutions only exist for a small class of systems, namely the Beneš class and its extensions [3]- [7]. Meanwhile, finite-dimensional recursive *approximations* are available for a large class of systems, but the appropriate choice of approximation is a complex topic in its own right and will not be explored here.

[1] Here and in the sequel, we use the convention that the gradient operator adds a covariant dimension to the tensor field it acts on. For example, $\partial_\theta \mathcal{L}_t(\theta)$ takes values that are covectors (row vectors), and the gradient of $\hat{h}_t(\theta)$, which has values in $\mathbb{R}^{n_y}$, wrt. θ, is a $(n_y \times p)$–matrix ($\mathbb{R}^{n_y} \otimes \mathbb{R}^{p*}$–tensor) valued process which we denote by $\hat{h}_t^\theta(\theta)$.

We merely mention a few standard approximation schemes: extended and unscented Kalman filters [8], [9], projection or assumed-density filters [10], [11], particle filters [12], and particle filters without weights [13]- [15].

Given a finite-dimensional representation of the filter, a corresponding representation of the filter derivative may be formally defined by differentiation with respect to θ:

$$\hat{h}_t^\theta(\theta) \simeq \partial_\theta \psi_h(M_t(\theta), \theta) + \partial_M \psi_h(M_t(\theta), \theta) M_t^\theta(\theta), \tag{9}$$

where ∂_M denotes the gradient wrt. the first argument of ψ_h and $M_t^\theta(\theta)$ denotes the $(m \times p)$–matrix valued derivative of the process $M_t(\theta)$. For the system in Eqs. (1,2) and for a large class of exact and approximate filters, $M_t(\theta)$ admits a stochastic differential equation (SDE) of the form

$$\begin{aligned} dM_t(\theta) = \mathcal{R}(\theta, M_t(\theta))dt + \mathcal{S}(\theta, M_t(\theta))dY_t \\ + \mathcal{T}(\theta, M_t(\theta))dB_t, \end{aligned} \tag{10}$$

where $\mathcal{R}, \mathcal{S}$, and $\mathcal{T}$ go to $\mathbb{R}^m$, $\mathbb{R}^{m\times n_y}$, and $\mathbb{R}^{m\times m'}$ respectively, and B_t is an m'-dimensional Brownian motion that is independent of $\mathcal{F}_t^{X,Y}$ (e.g. independent noise in particle filters). By differentiating wrt. θ, we find the corresponding SDE for $M_t^\theta(\theta)$

$$\begin{aligned} dM_t^\theta(\theta) = \mathcal{R}'(M_t(\theta), M_t^\theta(\theta), \theta)dt \\ + \mathcal{S}'(M_t(\theta), M_t^\theta(\theta), \theta)dY_t \\ + \mathcal{T}'(M_t(\theta), M_t^\theta(\theta), \theta)dB_t, \end{aligned} \tag{11}$$

where the tensor fields $\mathcal{R}', \mathcal{S}', \mathcal{T}'$ are given by

$$\begin{aligned} \mathcal{R}'(M_t(\theta), M_t^\theta(\theta), \theta) = \partial_\theta \mathcal{R}(M_t(\theta), \theta) \\ + \partial_M \mathcal{R}(M_t(\theta), \theta) M_t^\theta(\theta), \end{aligned} \tag{12}$$

and analogously for $\mathcal{S}$ and $\mathcal{T}$. In Section IV, we will present several examples of both exact and approximate filters for which these calculations will be made explicit.

B. *Online algorithm*

Instead of integrating the gradient of the log-likelihood function up to time t, a stochastic gradient ascent uses the integrand of the gradient of the log-likelihood (evaluated with the current parameter estimate) to update the parameter estimate online as new data is reaching the observer. The time-dependent parameter estimate $\tilde{\theta}_t$ is the $\mathbb{R}^p$-valued stochastic process that is the solution to the SDE

$$d\tilde{\theta}_t = \begin{cases} \gamma_t \tilde{h}_t^{\theta\top}\left(dY_t - \tilde{h}_t dt\right), & \tilde{\theta} \in \Theta, \\ 0, & \tilde{\theta} \notin \Theta, \end{cases} \tag{13}$$

where $\gamma_t > 0$ is a time-dependent learning rate. The processes $\tilde{h}_t$ and $\tilde{h}_t^\theta$ are the filter and filter derivative respectively, recursively updated with the instantaneous parameter estimate $\tilde{\theta}_t$. In terms of the finite-dimensional representation, they are computed as follows:

$$\tilde{h}_t = \psi_h(\tilde{M}_t, \tilde{\theta}_t), \tag{14}$$

$$\tilde{h}_t^\theta = \partial_\theta \psi_h(\tilde{M}_t, \tilde{\theta}_t) + \partial_M \psi_h(\tilde{M}_t, \tilde{\theta}_t)\tilde{M}_t^\theta, \tag{15}$$

where $\tilde{M}_t$ and $\tilde{M}_t^\theta$ evolve according to the SDEs

$$d\tilde{M}_t = \mathcal{R}(\tilde{M}_t, \tilde{\theta}_t)dt + \mathcal{S}(\tilde{M}_t, \tilde{\theta}_t)dY_t + \mathcal{T}(\tilde{M}_t, \tilde{\theta}_t)dB_t, \quad (16)$$

$$d\tilde{M}_t^\theta = \mathcal{R}'(\tilde{M}_t, \tilde{M}_t^\theta, \tilde{\theta}_t)dt + \mathcal{S}'(\tilde{M}_t, \tilde{M}_t^\theta, \tilde{\theta}_t)dY_t + \mathcal{T}'(\tilde{M}_t, \tilde{M}_t^\theta, \tilde{\theta}_t,)dB_t. \quad (17)$$

III. Convergence analysis

In this section, we will prove, under suitable conditions, the convergence of the parameter estimation algorithm given above. Our analysis will closely follow the treatments of related problems by [16] (discrete-time), as well as [17] (continuous-time but fully observed).

The proof of the main convergence result (Therorem 1) relies on the ergodicity of the joint process consisting of the hidden state, filter, and filter derivative. For the purposes of the proof, it is sufficient to assume the conclusions of Propositions 1 and 2, which summarize the main ingredients that will appear in the proof. We will give sufficient conditions (Conditions 1-3 below) that can be verified in applications. From those conditions, proofs for Propositions 1 and 2 are given in Appendices A and B. They rely heavily on theory developed by Veretennikov and Pardoux (see [18]–[20]).

Conditions 1-3 require modifications and regularizations to be applied to most existing filtering algorithms that in many cases are not strictly necessary (i.e. the conclusions of Propositions 1 and 2 are correct even though Conditions 1-3 are not). The theory corresponding to [18]–[20] under milder conditions that are satisfied without these modifications has, to our knowledge, not yet been developed.

A. Assumptions and conditions

Definition 1: We say that a function $G : \mathbb{R}^d \times \Theta \to \mathbb{R}$ has the polynomial growth property (PGP) if there are $q, K > 0$ such that for all $\theta \in \Theta$,

$$|G(x, \theta)| \leq K(1 + ||x||^q). \quad (18)$$

Let $\mathbb{G}^d$ be the function space defined by all functions $G : \mathbb{R}^d \times \Theta \to \mathbb{R}$ such that

(a) $G(\cdot, \theta) \in C(\mathbb{R}^d)$,
(b) $G(x, \cdot) \in C^2(\Theta)$,
(c) $\partial_\theta G(x, \cdot)$ and $\partial_\theta^2 G(x, \cdot)$ are Hölder continuous with exponent $\alpha > 0$.

Let $\mathbb{G}_c^d$ be the subset consisting of all $G \in \mathbb{G}^d$ that are centered, i.e. $\int_{\mathbb{R}^d} G(x, \theta)\mu_\theta(dx) = 0$. Let $\bar{\mathbb{G}}^d$ be the subset consisting of all $G \in \mathbb{G}^d$ such that G and its first two derivatives wrt. θ satisfy the PGP. ■

Condition 1: There is a finite-dimensional representation (exact or approximate) of the filter. Let $D = n + m + mp$ and $\mathcal{X}_t$ be the process with values in $\mathbb{R}^D$ defined by concatenating the state X_t, the filter representation and all the filter derivatives as follows:

$$\mathcal{X}_t(\theta) = \mathcal{C}(X_t, M_t(\theta), M_t^\theta(\theta)) \doteq \Big(X_{t,1}, ..., X_{t,n}, M_{t,1}(\theta), ..., M_{t,m}(\theta), M_{t,1}^{\theta_1}(\theta), ..., M_{t,m}^{\theta_1}(\theta), ..., M_{t,1}^{\theta_p}(\theta), ..., M_{t,m}^{\theta_p}(\theta)\Big)^\top. \quad (19)$$

Here we introduced the mapping $\mathcal{C} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times p} \to \mathbb{R}^D$ that performs the concatenation. Under P_{θ_0}, $\mathcal{X}_t(\theta)$ is a diffusion process with SDE

$$d\mathcal{X}_t(\theta) = \Phi(\mathcal{X}_t(\theta), \theta)dt + \Sigma(\mathcal{X}_t(\theta), \theta)d\mathcal{B}_t. \quad (20)$$

Here, $\mathcal{B}_t$ is the D'-dimensional standard Wiener process defined by

$$\mathcal{B}_t = (W_t, V_t, B_t). \quad (21)$$

Similarly, $\tilde{\mathcal{X}}_t \doteq \mathcal{C}(X_t, \tilde{M}_t, \tilde{M}_t^\theta)$ is a diffusion process with SDE

$$d\tilde{\mathcal{X}}_t = \Phi(\tilde{\mathcal{X}}_t, \tilde{\theta}_t)dt + \Sigma(\tilde{\mathcal{X}}_t, \tilde{\theta}_t)d\mathcal{B}_t, \quad (22)$$

consisting of the state as well as the filter and filter derivatives integrated with the online parameter estimate.

We impose the following conditions on the functions Φ and Σ, which go from $\mathbb{R}^D \times \Theta$ to $\mathbb{R}^D$ and $\mathbb{R}^{D\times D'}$ respectively, where $D' = n' + n_y + m'$:

(a) There are $R, \alpha, C > 0$ such that for all $\theta \in \Theta$ and all $||x|| \geq R$, $\Phi(x, \theta) \cdot x \leq -C||x||^\alpha$,
(b) $\exists 0 < \lambda < \Lambda < \infty$ such that $\forall x \in \mathbb{R}^D, \theta \in \Theta : \lambda I \leq \frac{1}{2}\Sigma(x, \theta)\Sigma^\top(x, \theta) \leq \Lambda I$,
(c) $\Phi(\cdot, \theta), \Sigma(\cdot, \theta) \in C_b^2(\mathbb{R}^D)$ for all $\theta \in \Theta$ and $\Phi(x, \cdot), \Sigma(x, \cdot) \in C_b^1(\Theta)$ for all $x \in \mathbb{R}^D$, and all existing derivatives are Hölder continuous with some exponent $\alpha \in (0, 1)$.

In addition, we impose the following constraints on the initial measures $\mu_{\theta,0}$ and $\tilde{\mu}_0$ of $\mathcal{X}_0(\theta)$ and $\tilde{\mathcal{X}}_0$, as well as the function ψ_h giving the filter estimate of h:

(d) For all $q > 0$, $\int_{\mathbb{R}^D} ||x||^q \mu_{\theta,0}(dx) < \infty$ and $\int_{\mathbb{R}^D} ||x||^q \tilde{\mu}_0(dx) < \infty$,
(e) The function ψ_h satisfies the PGP and is in $\mathbb{G}^m$. ■

Let the functions l, F, H, which go from $\mathbb{R}^D \times \Theta$ to $\mathbb{R}$, $\mathbb{R}^p$, and $\mathbb{R}^{n_y\times p}$ respectively, be defined as

$$l(x, \theta) = \psi_h(M, \theta) \cdot \left[h(X, \theta_0) - \tfrac{1}{2}\psi_h(M, \theta)\right], \quad (23)$$

$$F(x, \theta) \doteq H(x, \theta)^\top \left[h(X, \theta_0) - \psi_h(M, \theta)\right], \quad (24)$$

$$H(x, \theta) \doteq \partial_\theta \psi_h(M, \theta) + \partial_M \psi_h(M, \theta)M'. \quad (25)$$

With these definitions, the parameter SDE takes the form

$$d\tilde{\theta}_t = \begin{cases} \gamma_t F(\tilde{\mathcal{X}}_t, \tilde{\theta}_t)dt + \gamma_t H(\tilde{\mathcal{X}}_t, \tilde{\theta}_t)dV_t, & \tilde{\theta}_t \in \Theta \\ 0, & \tilde{\theta}_t \notin \Theta \end{cases}. \quad (26)$$

Condition 2: The function $F(x, \cdot)$ is in $\bar{\mathbb{G}}$ (component-wise). In addition, $l(x, \theta)$ and $H(x, \theta)$ have the PGP (component-wise). ■

Lastly, the following condition on the learning rate is imposed:

Condition 3: $\int_0^\infty \gamma_t dt = \infty$, $\int_0^\infty \gamma_t^2 dt = 0$, and there is a $r > 0$ such that $\lim_{t\to\infty} \gamma_t^2 t^{1/2+2r} = 0$. ■

Remark: With the exception of 1 (b), all of these conditions seem to be reasonable and fulfilled by a broad class of exact and approximate finite-dimensional filtering algorithms. Condition 1 (b), however, is too strict; a degenerate noise covariance matrix occurs in a large number of cases because the state and observation noise are the only noise sources driving all the components of $\mathcal{X}_t$. In order to lift the degeneracy, a small amount of noise can be added to make the diffusion matrix quadratic. Despite this work-around, it would be more desirable to replace Condition 1 (b) by a milder condition (akin to the one used in [21]) that is satisfied by a broad class of filtering algorithms. However, the theory that would allow the full range of results below to be proven under this milder condition has, to our knowledge, not been developed yet. In the cases where the conclusions to Propositions 1 and 2 can be checked directly, this is of no consequence.

B. Results

The consequences of the aforementioned conditions that will be relevant for the convergence analysis are summarized in the following two propositions.

Proposition 1 (Ergodicity): Assume condition 1 and let $\theta \in \Theta$. Then we have:

(i) The process $\mathcal{X}_t(\theta)$ is ergodic under P_{θ_0}, with a unique invariant probability measure μ_θ on $(\mathbb{R}^D, \mathbb{B}_D)$, where $\mathbb{B}_D$ is the Borel σ-algebra on $\mathbb{R}^D$.

(ii) The process $\frac{1}{t}\mathcal{L}_t(\theta)$ converges in probability to $\tilde{\mathcal{L}}(\theta)$, which is given by

$$\tilde{\mathcal{L}}(\theta) = \int_{\mathbb{R}^D} l(x,\theta)\mu_\theta(dx), \tag{27}$$

where $(X, M, M') = \mathcal{C}^{-1}(x)$ are the components of x.

(iii) Let $\mathcal{A}_\mathcal{X}$ be the infinitesimal generator of $\mathcal{X}_t(\theta)$ and let $G \in \mathbb{G}$ such that $\int_{\mathbb{R}^D} G(x,\theta)\mu_\theta(dx) = 0$. Then the Poisson equation $\mathcal{A}_\mathcal{X} v(x,\theta) = G(x,\theta)$ has a unique solution $v(x,\theta)$ that lies in $\mathbb{G}$, with $v(\cdot,\theta) \in C^2(\mathbb{R}^D)$. Moreover, if $G \in \bar{\mathbb{G}}$, then $v \in \bar{\mathbb{G}}$ and also $\partial_x\partial_\theta v$ has the PGP.

(iv) For all $q > 0$, $\mathbb{E}[||\tilde{\mathcal{X}}_t||] < \infty$ and there is a $K > 0$ such that for t large enough,

$$\forall\theta \in \Theta \quad \mathbb{E}_{\theta_0}\left[\sup_{s\le t} ||\mathcal{X}_s(\theta)||^q\right] \le K\sqrt{t}, \tag{28}$$

$$\mathbb{E}_{\theta_0}\left[\sup_{s\le t} ||\tilde{\mathcal{X}}_s||^q\right] \le K\sqrt{t}. \tag{29}$$

Proof: See Appendix A. ■

Proposition 2 (Regularity of the asymptotic likelihood): Assume conditions 1-3 and let $\theta \in \Theta$. Then we have:

(i) The asymptotic likelihood function $\tilde{\mathcal{L}}(\theta)$ is in $C^2(\Theta)$, and the gradient g and Hessian $\mathcal{H}$ of the asymptotic likelihood are given in terms of the invariant measure μ_θ and its derivative ν_θ as

$$g(\theta) \doteq \partial_\theta\tilde{\mathcal{L}}(\theta) = \int_{\mathbb{R}^D} F(x,\theta)\mu_\theta(dx), \tag{30}$$

$$\begin{aligned}\mathcal{H}(\theta) \doteq \partial_\theta^2\tilde{\mathcal{L}}(\theta) = &\int_{\mathbb{R}^D} \partial_\theta F(x,\theta)\mu_\theta(dx) \\ &+ \int_{\mathbb{R}^D} F(x,\theta)\nu_\theta(dx),\end{aligned} \tag{31}$$

(ii) The function $G(x,\theta) \doteq F(x,\theta) - g(\theta)$ is in $\bar{\mathbb{G}} \cap \mathbb{G}_c$.

(iii) For any $q > 0$ and $\theta \in \Theta$ there is a constant $K_q > 0$ such that

$$\int_{\mathbb{R}^D} (1 + ||x||^q)\,\mu_\theta(dx) \le K_q. \tag{32}$$

(iv) Let the finite signed measures $\nu_{\theta,i} = \partial_{\theta_i}\mu_\theta$, $i = 1, ..., p$ and $|\nu_{\theta,i}(dx)|$ be their variation. For any $q > 0$ and $\theta \in \Theta$ there is a constant $K'_q > 0$ such that

$$\int_{\mathbb{R}^D} (1 + ||x||^q)\,|\nu_{\theta,i}(dx)| \le K'_q. \tag{33}$$

(v) There is a constant $C > 0$ such that

$$\tilde{\mathcal{L}}(\theta) + \|g(\theta)\| + \|\mathcal{H}(\theta)\| \le C. \tag{34}$$

Proof: See Appendix B. ■

Now we can formulate our main result. Its proof relies on several lemmas that are given in Appendix C.

Theorem 1 (main theorem): Assume conditions 1-3 and let $\tilde{\theta}_0 \in \Theta$. Then, with probability one

$$\lim_{t\to\infty} \left\|g(\tilde{\theta}_t)\right\| = 0 \quad \text{or} \quad \tilde{\theta}_t \to \partial\Theta. \tag{35}$$

Proof: See Appendix D. ■

norm. MSE	w/o learning	w/ learning	ground truth	optimal
Linear model	0.99	0.29	0.28	0.28
Bimodal model	0.56	0.20	0.32	0.18

TABLE I
SUMMARY OF THE PERFORMANCE RESULTS OF FIG. 2 AND 5. THE AVERAGE MSE VALUES WITH LEARNING ARE TAKEN FROM THE LAST THIRD OF THE TRIAL WHERE PERFORMANCE HAS CONVERGED. NOTE THAT THE NUMBERS IN THE LEFT COLUMN (WITHOUT LEARNING) DEPEND ON THE INITIAL PARAMETER ESTIMATES. THE INITIAL PARAMETERS WE USED ARE FOUND IN THE MAIN TEXT AS WELL AS IN THE CAPTIONS TO FIG. 1 AND 4. 'GROUND TRUTH' MEANS THAT THE FILTER IS RUN WITH THE GROUND TRUTH PARAMETERS, WHICH ACHIEVES OPTIMAL PERFORMANCE IN THE LINEAR CASE (BECAUSE THE KALMAN-BUCY FILTER IS EXACT), BUT NOT IN THE NONLINEAR CASE. IN THE LATTER, OPTIMAL PERFORMANCE IS ESTIMATED BY USING A PARTICLE FILTER.

IV. EXAMPLES AND NUMERICAL VALIDATION

Here, we consider two different example filtering problems and show explicitly how the parameter learning rules are derived. We also study the numerical performance of the learning method. Since under suitable conditions on the decay of the learning rate, convergence is guaranteed by the results in the preceding section, we do not study this case. Instead, we study whether the method also converges with constant learning rate, i.e. when violating Condition 3. A constant learning rate is a sensible choice when the system parameters are expected to change.

All numerical experiments use the Euler-Maruyama method to integrate the SDEs. We evaluate the performance of the learned filter by the mean squared error (MSE), normalized by the variance of the hidden process.

A. One-dimensional Kalman-Bucy filter (linear filtering problem)

We shall first consider the simple case of the linear filtering problem, for which it is possible to obtain an exact finite-dimensional filter as well as exact expressions for the asymptotic likelihood. Here, we have a three-dimensional parameter vector $\theta = (a, \sigma, w)$, where $a, \sigma > 0$ and $w \in \mathbb{R}$, and we have

$f(x,\theta) = -ax$, $g(x,\theta) = \sigma$ and $h(x,\theta) = wx$, such that the filtering problem reads

$$dX_t = -aX_t dt + \sigma dW_t, \qquad dY_t = wX_t dt + dV_t. \quad (36)$$

Assuming a Gaussian initialization, i.e. $X_0 \sim \mathcal{N}(0, \sigma^2/2a)$, the optimal filter has a Gaussian distribution with mean μ_t and variance P_t (the Kalman-Bucy filter [2]). This is a two-dimensional representation with $M_t(\theta) = (\mu_t(\theta), P_t(\theta))^\top$, which can be expressed as

$$dM_t(\theta) = \begin{pmatrix} -a\mu_t(\theta) - w^2\mu_t(\theta)P_t(\theta) \\ \sigma^2 - 2aP_t(\theta) - w^2 P_t(\theta)^2 \end{pmatrix} dt + \begin{pmatrix} wP_t(\theta) \\ 0 \end{pmatrix} dY_t. \quad (37)$$

We have $\psi_h(M_t(\theta), \theta) = w\mu_t(\theta)$.

Let us first calculate the asymptotic log-likelihood. It follows from the above that $P_t(\theta)$ (and its derivatives with respect to θ) will tend to a unique steady state given by

$$P_\infty(\theta) = \frac{1}{w^2}\left(\sqrt{a^2 + w^2\sigma^2} - a\right). \quad (38)$$

By initializing the filter with this steady-state value, the representation can be made one-dimensional, i.e.

$$dM_t(\theta) = \left(-a\mu_t(\theta) - w^2\mu_t(\theta)P_\infty(\theta)\right) dt + wP_\infty(\theta)dY_t. \quad (39)$$

The process $\mathcal{X}_t(\theta)$ consisting of X_t, $\mu_t(\theta)$, and the filter derivatives $\mu_t^a(\theta), \mu_t^\sigma(\theta), \mu_t^w(\theta)$, therefore admits the SDE representation

$$d\mathcal{X}_t(\theta) = A\mathcal{X}_t(\theta)dt + B\begin{pmatrix} dW_t \\ dV_t \end{pmatrix}, \quad (40)$$

with matrices

$$A = \begin{pmatrix} -a_0 & 0 & 0 & 0 & 0 \\ ww_0P & -a - w^2P & 0 & 0 & 0 \\ ww_0P^a & -1 - w^2P^a & -a - w^2P & 0 & 0 \\ ww_0P^\sigma & -w^2P^\sigma & 0 & -a - w^2P & 0 \\ w_0(P + wP^w) & -w^2P^w & 0 & 0 & -a - w^2P \end{pmatrix}, \quad (41)$$

and

$$B = \begin{pmatrix} \sigma_0 & 0 \\ 0 & wP \\ 0 & wP^a \\ 0 & wP^\sigma \\ 0 & P + wP^w \end{pmatrix}, \quad (42)$$

where P is a short-hand for $P_\infty(\theta)$ and P^a etc. are partial derivatives of $P_\infty(\theta)$.

The process $\mathcal{X}_t(\theta)$ is ergodic, and its unique invariant probability measure is multivariate Gaussian with zero mean and covariance matrix K given by the solution to

$$BB^\top + AK + KA^\top = 0. \quad (43)$$

In terms of this, the asymptotic log-likelihood reads

$$\begin{aligned} \tilde{\mathcal{L}}(\theta) &= ww_0K_{12} - \frac{1}{2}w^2K_{22} \\ &= \frac{P_\infty(\theta)w^2\sigma_0^2w_0^2(2a + P_\infty(\theta)w^2)}{4a_0(a + P_\infty(\theta)w^2)(a + a_0 + P_\infty(\theta)w^2)} \\ &\quad - \frac{P_\infty(\theta)^2w^4}{4(a + P_\infty(\theta)w^2)} \end{aligned} \quad (44)$$

With suitable boundaries of the parameter space, all the conclusions of Propositions 1 and 2 can be fulfilled.

This model is non-identifiable from the observations. The set of critical points of the asymptotic likelihood is characterized by

$$\partial_\theta\tilde{\mathcal{L}}(\theta) = 0 \Leftrightarrow \theta = \left(a_0, \sigma, \frac{w_0\sigma_0}{\sigma}\right)^\top, \quad \sigma > 0, \quad (45)$$

i.e. convergence can be guaranteed to one of these points only, and not to the ground truth parameters $\theta_0 = (a_0, \sigma_0, w_0)^\top$. The model becomes identifiable if either σ_0 or w_0 is known. Alternatively, one may choose a parametrization for which X_t has unit variance (i.e. $\sigma = \sqrt{2a}$).

Let us now derive the parameter update equations. The filtering equations for the mismatched filter, expressed in terms of the online parameter estimates, read

$$d\mu_t = -\tilde{a}_t\mu_t dt + \tilde{w}_tP_t(dY_t - \tilde{w}_t\mu_t dt), \quad \mu_0 = 0, \quad (46)$$

$$dP_t = \left(\tilde{\sigma}_t^2 - 2\tilde{a}_tP_t - \tilde{w}_t^2P_t^2\right) dt, \quad P_0 = \frac{\tilde{\sigma}_0^2}{2\tilde{a}_0}, \quad (47)$$

where the initialization of P_0 reflects the prior belief of the variance of X_0 based on the initial parameter estimates.

The online parameter update equations read

$$d\tilde{a}_t = \gamma_a\tilde{a}_t\tilde{w}_t\mu_t^a\left(dY_t - \tilde{w}_t\mu_t dt\right), \quad (48)$$

$$d\tilde{\sigma}_t = \gamma_\sigma\tilde{\sigma}_t\tilde{w}_t\mu_t^\sigma\left(dY_t - \tilde{w}_t\mu_t dt\right), \quad (49)$$

$$d\tilde{w}_t = \gamma_w\tilde{w}_t\left(\mu_t + \tilde{w}_t\mu_t^w\right)\left(dY_t - \tilde{w}_t\mu_t dt\right). \quad (50)$$

In order to prevent sign changes of the parameters we chose time-dependent learning rates that are proportional to the parameters ($\tilde{a}_t$ has to stay non-negative because the filter equations turn unstable otherwise; for $\tilde{\sigma}_t$ and $\tilde{w}_t$ it is because of identifiability, i.e. the signs of σ and w are not identifiable from $\mathcal{F}_t^Y$). Here, we introduced the filter derivatives μ_t^a, μ_t^σ and μ_t^w of the mean, which, together with the filter derivatives of the variance, satisfy the coupled system of SDEs

$$\begin{aligned} d\mu_t^a &= -\left[\mu_t + \left(\tilde{a}_t + \tilde{w}_t^2P_t\right)\mu_t^a + \tilde{w}_t^2\mu_tP_t^a\right] dt \\ &\quad + \tilde{w}_tP_t^a dY_t, \end{aligned} \quad (51)$$

$$dP_t^a = -\left[2P_t + 2\left(\tilde{a}_t + \tilde{w}_t^2P_t\right)P_t^a\right] dt, \quad (52)$$

$$\begin{aligned} d\mu_t^\sigma &= -\left[\left(\tilde{a}_t + \tilde{w}_t^2P_t\right)\mu_t^\sigma + \tilde{w}_t^2\mu_tP_t^\sigma\right] dt \\ &\quad + \tilde{w}_tP_t^\sigma dY_t, \end{aligned} \quad (53)$$

$$dP_t^\sigma = \left[2\tilde{\sigma}_t - 2\left(\tilde{a}_t + \tilde{w}_t^2P_t\right)P_t^\sigma\right] dt, \quad (54)$$

$$\begin{aligned} d\mu_t^w &= -\left[2\tilde{w}_t\mu_tP_t + \left(\tilde{a}_t + \tilde{w}_t^2P_t\right)\mu_t^w\right] dt \\ &\quad - \tilde{w}_t^2\mu_tP_t^w dt + \left[P_t + \tilde{w}_tP_t^w\right] dY_t, \end{aligned} \quad (55)$$

$$dP_t^w = -\left[2\tilde{w}_tP_t^2 + 2\left(\tilde{a}_t + \tilde{w}_t^2P_t\right)P_t^w\right] dt, \quad (56)$$

$$\mu_0^a = \mu_0^\sigma = \mu_0^w = 0, \quad (57)$$

$$P_0^a = -\frac{\tilde{\sigma}_0^2}{2\tilde{a}_0^2}, \quad P_0^\sigma = \frac{\tilde{\sigma}_0}{\tilde{a}_0}, \quad P_0^w = 0. \quad (58)$$

The right-hand sides of the filter derivative equations and the initial conditions of the filter derivatives are obtained from the corresponding equations of the filtered mean and variance and their initial conditions by differentiating with respect to each of the parameters (see Section II for details).

First, we investigated one of the cases where the model is identifiable, i.e. the parameter w was assumed to be known and we set $\tilde{w}_0 = w_0 = 3$ and $\gamma_w = 0$. The performance of the algorithm is visualized in Fig. 1 where the learning process is shown in a single trial, and in Fig. 2, where we show trial-averaged learning curves for the MSE and the parameter estimates. For both figures, the ground truth parameters were set to $a_0 = 1$, $\sigma_0 = 2$, and the initial parameter estimates were $\tilde{a}_0 = 10$ and $\tilde{\sigma}_0 = \sqrt{0.2}$, making for a strongly mismatched model that produces an MSE close to 1 without learning, i.e. with all learning rates set to zero. With constant learning rates $\gamma_a = \gamma_\sigma = 0.03$, the filter performance can be improved to almost optimal performance within a time-frame of $T = 1000$, after which the parameter estimates approach the ground truth. The log-likelihood function is not globally concave, but it has a single global maximum (see Fig. 3).

Next, we looked at the non-identifiable case where all three parameters have to be learned. Depending on the initial conditions, the performance does not always reach optimal performance within this time-frame, and parameter estimates do not necessarily converge to the ground truth. However, Fig. 3 shows that the filter error can be dramatically reduced within a time-frame of $T = 3000$ for all initial parameter estimates that we tested. This holds even for initial parameter estimates that lead to an initial MSE > 1.

B. Bimodal state and linear observation model with (approximate) projection filter

Consider the following system with four positive parameters (a, b, σ, w):

$$dX_t = X_t \left(a - bX_t^2\right) dt + \sigma dW_t, \tag{59}$$
$$dY_t = wX_t dt + dV_t. \tag{60}$$

In this problem the hidden state X_t has a bimodal stationary distribution with modes at $x = \pm\sqrt{a/b}$. Since the observation model is linear like in Section IV-A, the parameter learning rules are expressed in terms of the posterior mean $\mu_t = \hat{X}_t$ as

$$d\tilde{a}_t = \gamma_a \tilde{a}_t \tilde{w}_t \mu_t^a \left(dY_t - \tilde{w}_t \mu_t dt\right), \tag{61}$$
$$d\tilde{b}_t = \gamma_b \tilde{b}_t \tilde{w}_t \mu_t^b \left(dY_t - \tilde{w}_t \mu_t dt\right), \tag{62}$$
$$d\tilde{\sigma}_t = \gamma_\sigma \tilde{\sigma}_t \tilde{w}_t \mu_t^\sigma \left(dY_t - \tilde{w}_t \mu_t dt\right), \tag{63}$$
$$d\tilde{w}_t = \gamma_w \tilde{w}_t \left(\mu_t + \tilde{w}_t \mu_t^w\right) \left(dY_t - \tilde{w}_t \mu_t dt\right), \tag{64}$$

We have made the learning rules proportional to the parameters in order to prevent sign changes, i.e. to guarantee that all parameters remain positive. In contrast to the linear model in Section IV-A, the filtering problem is not exactly solvable. We use the projection filter on the manifold of Gaussian densities introduced by [11], or equivalently, the Gaussian assumed density filter (ADF) in Stratonovich calculus. The mean μ_t and variance P_t of the Gaussian approximation to the filter evolve as

$$\begin{aligned} d\mu_t = &\left[\tilde{a}_t \mu_t - \tilde{b}_t \mu_t^3 - \left(3\tilde{b}_t + \tilde{w}_t^2\right) \mu_t P_t\right] dt \\ &+ \tilde{w}_t P_t dY_t, \quad \mu_0 = 0, \end{aligned} \tag{65}$$
$$dP_t = \left[\tilde{\sigma}_t^2 + \left(2\tilde{a}_t - \tilde{w}_t^2 P_t^2 - 6\tilde{b}_t(\mu_t^2 + P_t)\right) P_t\right] dt, \tag{66}$$

where the initial variance as a function of the initial parameter estimates is the variance of the stationary distribution obtained by solving the time-independent Fokker-Planck equation $\mathcal{A}^\dagger = 0$.

$$P_0 = \Gamma\left(\tilde{a}_0, \tilde{b}_0, \tilde{\sigma}_0\right) = \frac{\int_{-\infty}^{\infty} x^2 e^{\tilde{\sigma}_0^{-2}\left(\tilde{a}_0 x^2 - \frac{1}{2}\tilde{b}_0 x^4\right)} dx}{\int_{-\infty}^{\infty} e^{\tilde{\sigma}_0^{-2}\left(\tilde{a}_0 x^2 - \frac{1}{2}\tilde{b}_0 x^4\right)} dx}. \tag{67}$$

By differentiating Eqs. (65,66) with respect to the parameters, we obtain the following equations for the filter derivatives:

$$d\mu_t^a = \left[\mu_t + \alpha_t \mu_t^a + \beta_t P_t^a\right] dt + \tilde{w}_t P_t^a dY_t, \tag{68}$$
$$dP_t^a = \left[2P_t + A_t \mu_t^a + B_t P_t^a\right] dt, \tag{69}$$
$$\begin{aligned} d\mu_t^b = &\left[-\mu_t \left(\mu_t^2 + 3P_t\right) + \alpha_t \mu_t^b + \beta_t P_t^b\right] dt \\ &+ \tilde{w}_t P_t^a dY_t, \end{aligned} \tag{70}$$
$$dP_t^b = \left[-6P_t \left(\mu_t^2 + P_t\right) + A_t \mu_t^b + B_t P_t^b\right] dt, \tag{71}$$
$$d\mu_t^\sigma = \left[\alpha_t \mu_t^\sigma + \beta_t P_t^\sigma\right] dt + \tilde{w}_t P_t^\sigma dY_t, \tag{72}$$
$$dP_t^\sigma = \left[2\tilde{\sigma}_t + A_t \mu_t^\sigma + B_t P_t^\sigma\right] dt, \tag{73}$$
$$\begin{aligned} d\mu_t^w = &\left[-2\tilde{w}_t \mu_t P_t + \alpha_t \mu_t^w + \beta_t P_t^w\right] dt \\ &+ \left[P_t + \tilde{w}_t P_t^w\right] dY_t, \end{aligned} \tag{74}$$
$$dP_t^w = \left[-2\tilde{w}_t P_t^2 + A_t \mu_t^w + B_t P_t^w\right] dt, \tag{75}$$

$$\mu_0^a = \mu_0^b = \mu_0^\sigma = \mu_0^w = 0, \tag{76}$$
$$P_0^a = \frac{\partial}{\partial \tilde{a}_0} \Gamma\left(\tilde{a}_0, \tilde{b}_0, \tilde{\sigma}_0\right), \tag{77}$$
$$P_0^b = \frac{\partial}{\partial \tilde{b}_0} \Gamma\left(\tilde{a}_0, \tilde{b}_0, \tilde{\sigma}_0\right), \tag{78}$$
$$P_0^\sigma = \frac{\partial}{\partial \tilde{\sigma}_0} \Gamma\left(\tilde{a}_0, \tilde{b}_0, \tilde{\sigma}_0\right), \quad P_0^w = 0, \tag{79}$$

where we introduced the following auxiliary processes

$$\alpha_t = \tilde{a}_t - \tilde{w}_t^2 P_t - 3\tilde{b}_t \left(\mu_t^2 + P_t\right), \tag{80}$$
$$\beta_t = -\left(\tilde{w}_t^2 + 3\tilde{b}_t\right) \mu_t, \tag{81}$$
$$A_t = -12\tilde{b}_t \mu_t P_t, \tag{82}$$
$$B_t = 2\tilde{a}_t - 2\tilde{w}_t^2 P_t - 6\tilde{b}_t \left(\mu_t^2 + 2P_t\right). \tag{83}$$

We numerically tested the learning algorithm for this nonlinear model by simulating a system with $a_0 = 4$, $b_0 = 3$, $\sigma_0 = 1$ and $w_0 = 2$, leading to a variance $\mathrm{Var}(X_t) = 1.17$. Initial parameter estimates were set to a permutation of the ground truth, i.e. $\tilde{a}_0 = 1$, $\tilde{b}_0 = 2$, $\tilde{\sigma}_0 = 3$ and $\tilde{w}_0 = 4$ and the simulations lasted $T = 2000$ (due to the longer time-scale compared to the linear model) with a time-step of $dt = 10^{-3}$. In Fig. 4 we show a an example of the learning process.

In this case, the sub-optimality of the Gaussian approximation inherent in the projection filter allows the filter error (MSE) to be lower with learning than with the ground truth parameters in the absence of learning, getting close to the performance of the optimal filter. This is shown in Fig. 5 in terms of trial-averaged learning curves. The normalized MSE with learning decreases within the time frame of $T = 2000$ and converges below the MSE for the projection filter with fixed parameters set to the ground truth. The optimal performance was estimated by running a particle filter with prior importance function, resampling at every time-step, 1000 particles and parameters set to the ground truth [22].

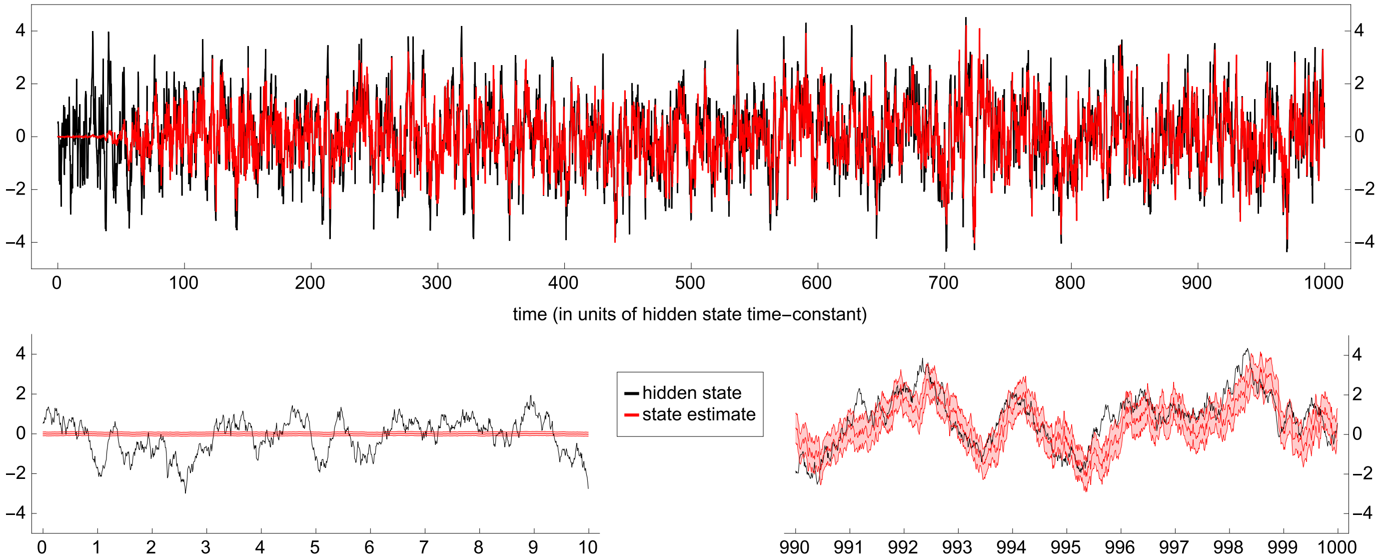

Fig. 1. **Online learning and filtering in the linear model**. The hidden state X_t (black) and Kalman-Bucy state estimate μ_t (red, shaded region shows μ_t $\pm$ one standard deviation $\sqrt{P_t}$, c.f. Eqs. (46,47)) are shown for the linear model of Section IV-A with parameters $a_0 = 1$, $\sigma_0 = 2$, $w_0 = 3$. The time-step is $dt = 10^{-3}$, initial parameter estimates are $\tilde{a}_0 = 10$, $\tilde{\sigma}_0 = \sqrt{0.2}$, $\tilde{w}_0 = 3$ (i.e. the parameter w_0 is known), and the learning rates are $\gamma_a = \gamma_\sigma = 0.03$ and $\gamma_w = 0$. Top: the entire learning period of $T = 1000$ shows a gradual improvement of the performance of the filter. Bottom left: during the first 10 seconds, the model is still strongly mismatched. Bottom right: during the last 10 seconds, the filter optimally tracks the hidden state.

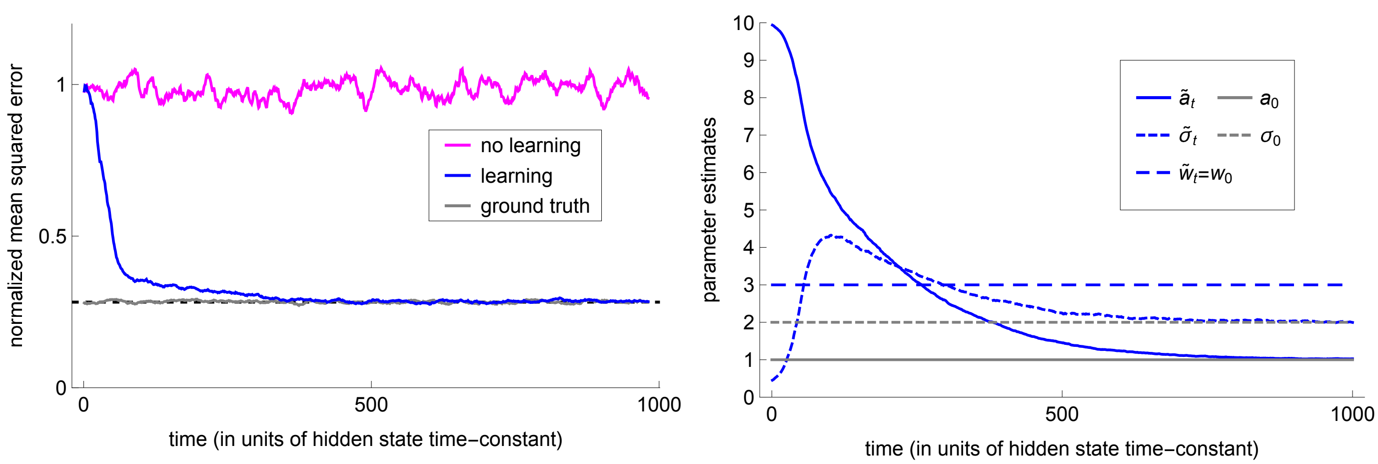

Fig. 2. **Online learning and filtering in the linear model**. The time evolution of the MSE and parameter estimates are shown for the linear model of Section IV-A (see Fig. 1 caption for details). Left: the moving average of the normalized MSE (time window of 20 seconds) shows how the learning algorithm leads to a gradual improvement of the performance of the filter, which eventually reaches the performance of an optimal Kalman-Bucy filter with ground truth parameters. The black, dashed line shows the theoretical result for the performance of the Kalman-Bucy filter. Right: the parameter estimates for the unknown parameters converge to the ground truth parameters. All curves are trial-averaged ($N = 100$ trials).

V. Related approaches

In this section, we attempt to review similar approaches for online maximum likelihood estimation, and their relations to our method. We note that most of the literature on this topic is formulated for discrete-time systems, and we realize that the list of reviewed works is not exhaustive. Some of the approaches for Hidden Markov Models (HMMs) discussed here are also surveyed in more detail in [23]- [25].

A. Recursive maximum-likelihood approaches

This work is the continuous-time analogue of the online stochastic gradient ascent algorithm of [16], [26] for HMMs. The behavior of the algorithm is analyzed by casting it in the Robbins-Monro framework of stochastic approximations. We used a similar approach to studying convergence in Section III. More recently, the convergence of discrete-time stochastic gradient algorithms for parameter estimation was studied under more general conditions [27]. To our knowledge, it is an open problem to obtain a similarly general result for continuous-time models such as the one in this paper.

B. Prediction error algorithms

Another stochastic approximation scheme is the recursive minimum prediction error scheme (see [16] and [28]) for HMMs. Instead of finding maxima of the likelihood, it finds minima of the average (squared) prediction error, i.e. the error

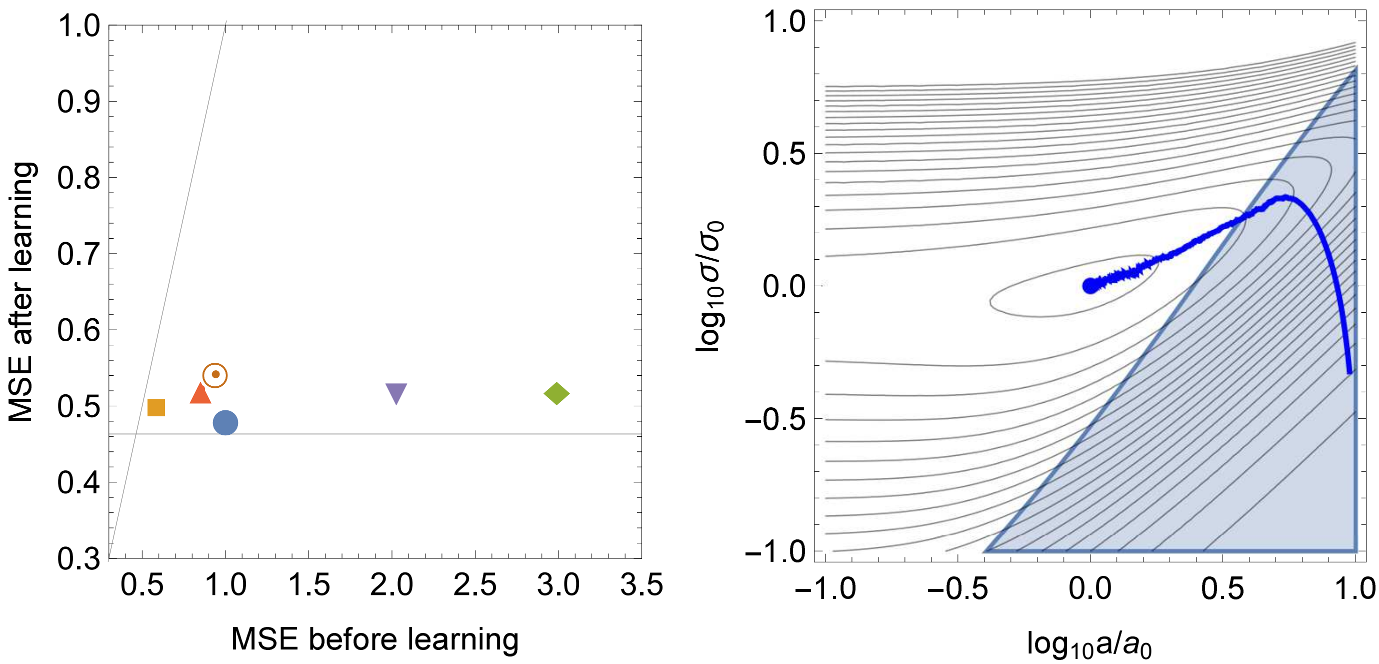

Fig. 3. **Online learning and filtering in the linear model**. Left: the filter error before and after learning is shown for different initial conditions of the parameter estimates. For all initial conditions, the learning algorithm achieves a dramatical improvement of the filter error. The horizontal line shows the error of the optimal filter and the diagonal line corresponds to matching before-learning and after-learning error. The blue dot shows the example from Figs. 1 and 2. Right: the asymptotic log-likelihood function from Eq. (44) in the parameter subspace spanned by a and σ for $w = w_0 = 3$ has a single global maximum near $a = a_0$ and $\sigma = \sigma_0$. The shading shows the region where the function is non-concave, and the blue line is the trial-averaged learning trajectory from Fig. 2.

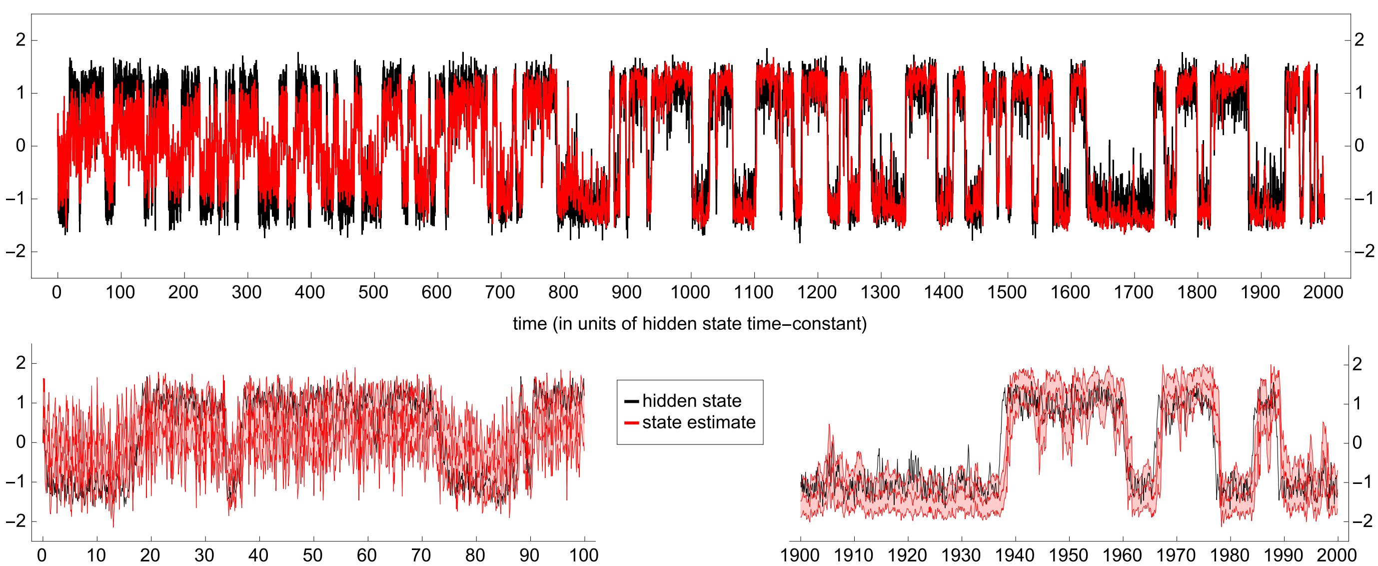

Fig. 4. **Online learning and filtering in the nonlinear model**. The hidden state X_t (black) and mean μ_t of the projection filter are shown for the bimodal model of Section IV-B with parameters $a_0 = 4$, $b_0 = 3$, $\sigma_0 = 1$ and $w_0 = 2$, $\tilde{a}_0 = 1$, $\tilde{b}_0 = 2$, $\tilde{\sigma}_0 = 3$, $\tilde{w}_0 = 4$, $\gamma_a = \gamma_b = \gamma_w = 10^{-1}$ and $\gamma_\sigma = 0.04$. Top: the entire learning period of $T = 2000$ shows an improvement in both step size between the two attractors and the variability within both attractors. Bottom left: during the first 100 seconds, the filter is too sensitive to observations and has an incorrect spacing between attractors. Bottom right: during the last 100 seconds, the filter shows good tracking performance.

between the observations and the predicted observations. In our continuous-time model, the prediction error is given by the infinitesimal pseudo-innovation increment $dY_t - \tilde{h}_t dt$. Formal differentiation of $(dY_t - \tilde{h}_t dt)^2$ with respect to the parameter yields the same parameter update rule as that derived in Section II. While a rigorous analysis has not been done, it seems natural to conjecture that recursive maximum likelihood and recursive minimum prediction error are equivalent in continuous time.

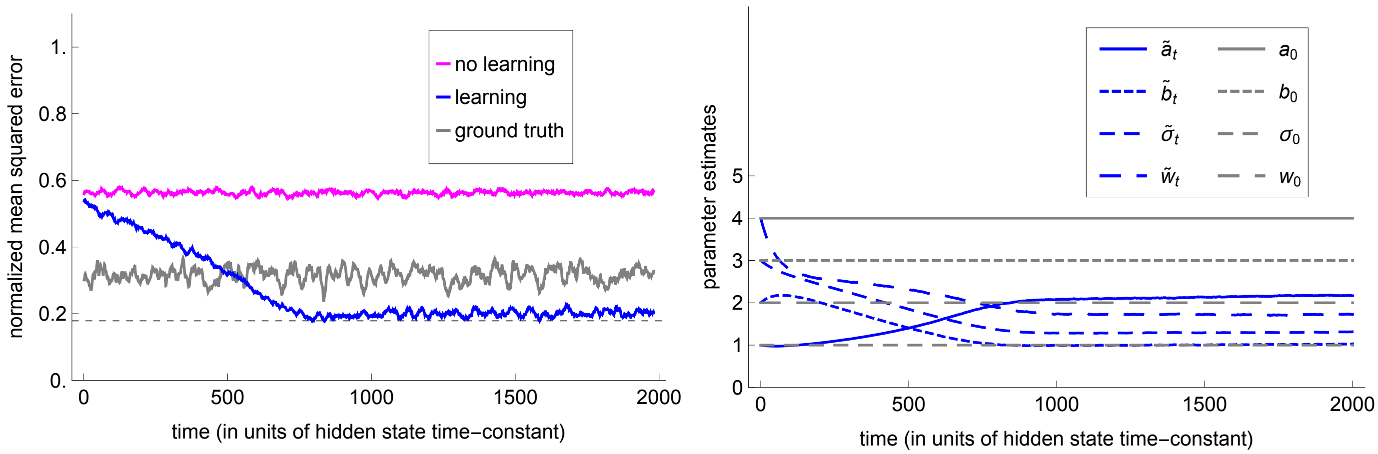

Fig. 5. **Online learning and filtering in the nonlinear model**. The time evolution of the MSE and parameter estimates are shown for the linear model of Section IV-A (see Fig. 1 caption for details). Left: the moving average of the normalized MSE (time window of 20 seconds) shows how the learning algorithm allows the filter performance to improve to a level that is better than that of a filter with fixed parameters set to the ground truth. However, it is still slightly worse than an optimal filter; the dashed black line shows the performance of a particle filter with 1000 particles with parameters set to the ground truth. Right: despite the low filter error, the parameter estimates do *not* converge to the ground truth. All curves are trial-averaged ($N = 100$ trials).

C. Online EM

Expectation maximization (EM) is a well-known method for offline parameter learning in partially observed stochastic systems [29], [30]. It is based on the following application of Jensen's inequality:

$$\begin{aligned}\mathcal{L}_t(\theta) - \mathcal{L}_t(\tilde{\theta}) &= \log \mathbb{E}_{\tilde{\theta}}\left[\frac{dP_\theta}{dP_{\tilde{\theta}}}\middle|\mathcal{F}_t^Y\right] \\ &\geq \mathbb{E}_{\tilde{\theta}}\left[\log\frac{dP_\theta}{dP_{\tilde{\theta}}}\middle|\mathcal{F}_t^Y\right] \doteq Q_t(\theta,\tilde{\theta}).\end{aligned} \tag{84}$$

Since $Q_t(\tilde{\theta},\tilde{\theta}) = 0$, by maximizing $Q_t(\theta,\tilde{\theta})$ with respect to θ (for fixed $\tilde{\theta}$), we obtain a non-negative change in the likelihood. EM thus produces a sequence of parameter estimates $\tilde{\theta}_k$, $k = 0, 1, 2, ...$ with non-decreasing likelihood by iterating the following procedure: compute the quantity $Q_t(\theta,\tilde{\theta}_k)$ (the 'expectation' or 'E' step in EM), then set $\tilde{\theta}_{k+1} = \mathrm{argmax}_\theta Q_t(\theta,\tilde{\theta}_k)$ (the 'maximization' or 'M' step in EM).

If a parametrization is chosen such that the complete-data log-likelihood[2] takes the form of an exponential family, i.e. $\Psi(\theta)\cdot S_t$, where Ψ is a vector-valued function of the parameters and S_t is a vector of functionals of the hidden state and observation trajectories, then $Q_t(\theta,\tilde{\theta}) = \Psi(\theta)\cdot\hat{S}_t(\tilde{\theta}) + R(\tilde{\theta})$, where

$$\hat{S}_t(\tilde{\theta}) = \mathbb{E}_{\tilde{\theta}}\left[S_t\middle|\mathcal{F}_t^Y\right], \tag{85}$$

and $R(\tilde{\theta})$ is some term that is independent of θ. The 'M' step can be done explicitly if the equation $\partial_\theta\Psi(\theta)\cdot\hat{S}_t(\tilde{\theta}) = 0$ has a unique closed-form solution. Meanwhile, the 'E' step consists of computing $\hat{S}_t(\tilde{\theta})$, which involves certain nonlinear smoothed functionals of the forms $\mathbb{E}_{\tilde{\theta}}\left[\int_0^t \varphi_1(X_s)dX_s\middle|\mathcal{F}_t^Y\right]$, $\mathbb{E}_{\tilde{\theta}}\left[\int_0^t \varphi_2(X_s)dY_s\middle|\mathcal{F}_t^Y\right]$, and $\mathbb{E}_{\tilde{\theta}}\left[\int_0^t \varphi_3(X_s)ds\middle|\mathcal{F}_t^Y\right]$, with possibly distinct functions $\varphi_1,\varphi_2,\varphi_3$. In general, these smoothed functionals are computed using a forward-backward smoothing algorithm, which is not suitable for online learning. In a few select cases, the smoothed functionals admit a finite-dimensional solution (see [32] and the remarks on p.99 of [30]), or even a finite-dimensional *recursive* solution [31], [33].

In [31], the smoothed functionals of the linear-Gaussian model are expressed (using the Fisher identity) in terms of derivatives of the incomplete-data log-likelihood, or a generalization thereof. This enables a recursive computation of the smoothed quantities of interest, and the auxiliary variables that need to be integrated (called sensitivity equations) are very similar to Eqs. (51)-(56). The relation between smoothed functionals and the sensitivity equations has been known for a long time (see [34] and Section 10.2 in [23]).

Several authors [35]- [38] have introduced the idea of a fully recursive form of EM, called *online EM*. In the references above, online EM has been explicitly formulated for HMMs and State Space Models (SSMs) by integrating the recursive smoothing algorithm using the online parameter estimate. This stochastic approximation approach to EM is thus very similar to the gradient-based approach used here and in the references discussed in Section V-A. Although we are not aware of a similar formulation for continuous-time models, the same idea could be applied to the recursions found by [31], [33] for the linear case. In nonlinear models, online EM could be formulated by making use of recursive particle approximations of the smoothing functionals (e.g. by applying the methods in [39], [40] to a suitable time discretization of the SDEs). As an alternative, assumed-density or projection filters could be used to approximate the recursive smoothed functionals.

[2]We note that a limitation of EM in the continuous-time model is that the identification of parameters of the diffusion term g_θ has to be treated differently from that of drift parameters in f_θ and h_θ. This is due to the fact that there is no reference measure for the complete model that is independent of the diffusion parameters. The parameters of the diffusion term are therefore not included in θ, but are estimated separately from the quadratic variations of hidden state and observation. This issue is discussed in more detail in [31], Section IV-B. This issue is avoided in the gradient-based method here because the reference measure restricted to the observations is independent of all parameters, including the ones of the diffusion term.

D. State augmentation algorithms

The idea is to treat the unknown parameter as a random variable that is either static ($d\theta_t = 0$) or has dynamics that are coupled to the hidden state. In both cases the parameter may be estimated online by solving the filtering problem for the augmented state (X_t, θ_t). While this presents clear advantages for known dynamics of the hidden parameter, it introduces a new parameter estimation problem for the parameters of the dynamics of θ_t, called *hyperparameters*. A static prior for θ_t is problematic because the resulting filter will usually not be stable, with negative implications (see [41]) on the behavior of particle filters that are needed to solve the augmented filtering problem (but see [42], where stability conditions are discussed for the discrete-time case). In addition, for many interesting models, the parameter space may be of much higher dimension than the state space, introducing high computational costs for filtering of the augmented state.

E. Maximum-likelihood filtering and identification

The opposite of state augmentation was explored in [43], where the hidden state is also estimated via maximum likelihood, instead of the usual filtering paradigm using minimum mean-squared error. Equations for the maximum-likelihood state and parameter estimates are then derived. Although these equations are not directly suitable for recursive identification, they are very similar to the ones obtained by us in Section II. It remains a curiosity that the approach of [43] has rarely been cited and has not been further developed.

VI. Conclusions

The problem of estimating parameters in partially observed systems is very old and relevant to many applications. However, the majority of the literature on this subject is written for discrete-time processes and for offline learning, while despite of its enormous importance for filtering and control theory, the continuous-time case has received little attention. Online gradient ascent in continuous time has only recently been studied in [17]. The use of a change of measure in order to express the likelihood function in terms of the filter is not new, but it seems to be underexploited. To the best of our knowledge, the only appearance is in the technical report by [43]. We found it appropriate to revisit this approach and to extend the work of [17] to the partially observable case.

The main difficulty and open problem is to find conditions on the generative model that are easy to verify, sufficient for the convergence of the algorithm, and not too restrictive. Ideally, we would be able to obtain necessary conditions, but to our knowledge the theory that would allow this has not yet been developed. Even though the currently provided sufficient conditions turn out to be too strong for many applications, the convergence proofs hold as long as the conclusions of the two propositions hold. Therefore, statements about the convergence can be deduced by checking the conclusions of those propositions as soon as the complete filter is designed. We hope that this open question will be resolved in the future, as we do not find the work-around to be very satisfactory.

Let us briefly comment on the numerical examples that we provided. As we showed numerically, the algorithm is capable of improving filter performance even if the models are unidentifiable and the learning rate constant, even though this cannot be expected. In addition, the second numerical example showed that the performance of the filter can be improved even beyond what is possible with fixed parameters. This result could lead to new ways of improving the performance of approximate filters by using the additional degrees of freedom given by the online parameter estimates for both adaptation (learning) and reduced filter error. It remains to be explored whether this feature applies to a large enough class of approximate filters to be useful for practical applications.

Appendix A
Proof of proposition 1

(i) Existence of μ_θ follows from condition 1(a), and uniqueness of μ_θ follows from condition 1(b) (see [18]).

(ii) We have

$$\frac{1}{t}\mathcal{L}_t(\theta) = \frac{1}{t}\int_0^t l(\mathcal{X}_s(\theta), \theta) ds + \frac{1}{t}\int_0^t \psi_h(\theta, M_s(\theta)) \cdot dV_s. \quad (86)$$

By (i) and Birkhoff's ergodic theorem, the first term on the RHS converges to $\int_{\mathbb{R}^D} l(x,\theta)\mu_\theta(dx) = \tilde{\mathcal{L}}(\theta)$ almost surely as $t \to \infty$. This implies convergence in probability of the first term to $\tilde{\mathcal{L}}(\theta)$. It remains to be shown that the second term converges to zero in probability as $t \to \infty$. Consider the martingale $M_t = \int_0^t \psi_h(\theta, M_s(\theta)) \cdot dV_s$. From Itô isometry, condition 1(e), and (iv), it follows that for t large enough,

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_0^t \psi_h(\theta, M_s(\theta)) \cdot dV_s\right)^2\right] \\
&= \mathbb{E}\left[\int_0^t \|\psi_h(\theta, M_s(\theta))\|^2 ds\right] \\
&\leq \mathbb{E}\left[\int_0^t C(1 + ||M_s(\theta)||^q) ds\right] \\
&\leq \mathbb{E}\left[\int_0^t C(1 + ||\mathcal{X}_s(\theta)||^q) ds\right] \\
&\leq Ct\left(1 + \mathbb{E}[\sup_{s\leq t} ||\mathcal{X}_s(\theta)||^q]\right) \\
&\leq Ct(1 + C'\sqrt{t}).
\end{aligned} \quad (87)$$

In short, for t large enough, we have $\mathrm{Var}[M_t] \leq Kt^{3/2}$ for some $K > 0$. Therefore,

$$\mathrm{Var}\left[\tfrac{1}{t}M_t\right] \leq Kt^{-1/2} \to 0, \quad t \to \infty, \quad (88)$$

which means that the second term on the RHS of Eq. (86) converges to zero in mean square. Since convergence in mean square implies convergence in probability, the proof of (ii) is completed.

(iii) Under conditions 1(a,b,c), this follows from Proposition 1 and Theorem 3 in [20] by noting that under the conditions we impose, the constants $C(y)$ in [20] can be chosen to be independent of the parameter (which is denoted y in that paper). Note that condition 1(a) implies condition (H_b) in [20]. A version with global constants is found as Theorem A.1 in [17], but no proof is given. A proof of Proposition 1 of [20] with global constants is found as a special case of Theorem 2 in [19]. The global constants for Theorem 3 in [20] follow from the global constants in Theorems 1 and 2 in [20].

(iv) Using conditions 1 (a,b,d), this follows from Propositions 1 and 2 in [19]. ■

APPENDIX B
PROOF OF PROPOSITION 2

(i) We have that $\partial_\theta \tilde{\mathcal{L}}(\theta) = \lim_{t\to\infty} \frac{1}{t}\partial_\theta \mathcal{L}_t(\theta)$, if the derivative exists and the limit exists in probability. Due to Condition 1 (e), the derivative

$$\begin{aligned}\frac{1}{t}\partial_\theta \mathcal{L}_t(\theta) &= \frac{1}{t}\int_0^t \partial_\theta l(\mathcal{X}_s(\theta),\theta)ds \\ &\quad + \frac{1}{t}\int_0^t \partial_\theta \psi_h(\theta, M_s(\theta))dV_s \\ &= \frac{1}{t}\int_0^t F(\mathcal{X}_s(\theta),\theta)ds \\ &\quad + \frac{1}{t}\int_0^t H(\theta, M_s(\theta))dV_s \end{aligned} \tag{89}$$

exists. This converges to $\int_{\mathbb{R}^D} F(x,\theta)\mu_\theta(dx)$ by an argument analogous to the one in the proof of Proposition 1 (ii).

The representation of $\mathcal{H}$ in terms of the invariant measure and its derivative follows from Condition 1 (e) and Proposition 2 (iv).

(ii) This follows from (i) and the fact that F is in $\bar{\mathbb{G}}$ (condition 2).

(iii) Let $q > 0$ be given and choose $r > 0$ such that $\int_{\mathbb{R}^D} \frac{1+||x||^q}{1+||x||^r}dx = K_r < \infty$. By condition 1 and Theorem 1 in [20], the density $\rho(x,\theta)$ of μ_θ with respect to the Lebesgue measure dx on $\mathbb{R}^D$ admits a constant C_r such that $\rho(x,\theta) \le \frac{C_r}{1+||x||^r}$. Then, we have

$$\begin{aligned}&\int_{\mathbb{R}^D} (1+||x||^q)\,\mu_\theta(dx) \\ &\qquad \le C_r \int_{\mathbb{R}^D} \frac{1+||x||^q}{1+||x||^r}dx = C_r K_r \doteq K_q,\end{aligned} \tag{90}$$

which proves inequality (32).

(iv) Using condition 1 and Theorem 1 in [20] once more, we find that the vector $\rho'(x,\theta)$ of densities of $\nu_{\theta,i} = \partial_{\theta_i}\mu_\theta$ with respect to the Lebesgue measure,

$$\rho_i'(x,\theta) \doteq \frac{d\nu_{\theta,i}}{dx}(x,\theta) = \partial_{\theta_i}\rho(x,\theta),$$

admits a constant C_r' such that $||\rho'(x,\theta)|| \le \frac{C_r'}{1+||x||^r}$. Therefore,

$$\begin{aligned}&\int_{\mathbb{R}^D} (1+||x||^q)\,|\nu_{\theta,i}(dx)| \\ &\qquad = \int_{\mathbb{R}^D} (1+||x||^q)\,|\rho_i'(x,\theta)|dx \\ &\qquad \le \int_{\mathbb{R}^D} (1+||x||^q)\,||\rho'(x,\theta)||dx \\ &\qquad \le C_r' \int_{\mathbb{R}^D} \frac{1+||x||^q}{1+||x||^r}dx = C_r' K_r \doteq K_q',\end{aligned} \tag{91}$$

which proves inequality (33).

(v) By condition 2, $q, K > 0$ can be chosen such that the functions $l, F, \partial_\theta F, H$ grow at most as $K(1+||x||^q)$ for all $\theta \in \Theta$. From this and the first part of the present Lemma, it follows that

$$\begin{aligned}\tilde{\mathcal{L}}(\theta) &= \int_{\mathbb{R}^D} l(x,\theta)\mu_\theta(dx) \\ &\le K\int_{\mathbb{R}^D} (1+||x||^q)\mu_\theta(dx) \le KK_q.\end{aligned} \tag{92}$$

By a similar calculation, we have

$$\|g(\theta)\| \le KK_q. \tag{93}$$

For $\|\mathcal{H}(\theta)\|$, observe that

$$\begin{aligned}\|\mathcal{H}(\theta)\| &\le \left\|\int_{\mathbb{R}^D} \partial_\theta F(x,\theta)\mu_\theta(dx)\right\| \\ &\quad + \left\|\int_{\mathbb{R}^D} F(x,\theta)\nu_\theta(dx)\right\| \\ &\le KK_q + \left\|\int_{\mathbb{R}^D} F(x,\theta)\nu_\theta(dx)\right\|,\end{aligned} \tag{94}$$

where the first term on the RHS was treated in the same way as in the bound for $\tilde{\mathcal{L}}(\theta)$ and $\|g(\theta)\|$. For the second term, we observe that

$$\begin{aligned}\left\|\int_{\mathbb{R}^D} F(x,\theta)\nu_\theta(dx)\right\|^2 &= \sum_{i,j=1}^p \left(\int_{\mathbb{R}^D} F_i(x,\theta)\nu_{\theta,j}(dx)\right)^2 \\ &\le \sum_{i,j=1}^p \left(\int_{\mathbb{R}^D} |F_i(x,\theta)|\,|\nu_{\theta,j}(dx)|\right)^2 \\ &\le \sum_{i,j=1}^p \left(\int_{\mathbb{R}^D} ||F(x,\theta)||\,|\nu_{\theta,j}(dx)|\right)^2 \le p^2K^2K_q'^2\end{aligned} \tag{95}$$

Inequality (34) then follows by setting $C = 3KK_q + pKK_q'$. ■

APPENDIX C
LEMMAS

Here, we adapt the Lemmas of [17] to fit the present setting. As in [17], the proofs of the lemmas require results from [20], but in a slightly more general form than what was needed in [17]. Despite the strong similarities between our proofs and the proofs in [17], for the convenience of the reader we shall write them out in full detail and in the appropriate notation.

Note that Lemma 3 in [17] was included in Proposition 2 (iii)-(v) in this paper.

For the Lemmas 1-4 below, we assume that conditions 1-3 hold and that the first exit time from Θ is infinite (see the proof of Theorem 1). In addition, we define the following. Let $\kappa, \lambda > 0$ and define the $(P_{\theta_0}, \mathcal{F}_t)$-stopping times $\sigma_0 = 0$ and $\sigma_k, \tau_k,\ k \in \mathbb{N}$ as

$$\tau_k \doteq \inf\left\{t > \sigma_{k-1} : \|g(\tilde{\theta}_t)\| \geq \kappa\right\}, \tag{96}$$

$$\begin{aligned}\sigma_k \doteq \sup\Big\{t > \tau_k : \tfrac{1}{2}\|g(\tilde{\theta}_{\tau_k})\| \leq \|g(\tilde{\theta}_s)\| \\ \leq 2\|g(\tilde{\theta}_{\tau_k})\|, s \in [\tau_k, t] \text{ and } \int_{\tau_k}^{t} \gamma_s ds \leq \lambda\Big\}\end{aligned} \tag{97}$$

Lemma 1: Let $\eta > 0$ and define

$$\Gamma_{k,\eta} \doteq \int_{\tau_k}^{\sigma_k+\eta} \gamma_s \left(F(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) - g(\tilde{\theta}_s)\right) ds \tag{98}$$

Then, with probability one,

$$\lim_{k\to\infty} \|\Gamma_{k,\eta}\| = 0. \tag{99}$$

Proof: Consider the function $G(x,\theta) = F(x,\theta) - g(\theta)$. By definition, we have

$$\int_{\mathbb{R}^D} G(x,\theta)\mu_\theta(dx) = 0, \tag{100}$$

and by Proposition 2 (ii) we have that the components of $G(x,\cdot)$ are in $\bar{\mathbb{G}}$. Therefore, by Proposition 1 (iii), the Poisson equation

$$\mathcal{A}_{\mathcal{X}} v(x,\theta) = G(x,\theta), \qquad \int_{\mathbb{R}^D} v(x,\theta)\mu_\theta(dx) = 0 \tag{101}$$

has a unique twice differentiable solution with

$$||v(x,\theta)|| + ||\partial_\theta v(x,\theta)|| + ||\partial_\theta^2 v(x,\theta)|| \leq K'(1 + ||x||^{q'}). \tag{102}$$

Let $u(t,x,\theta) = \gamma_t v(x,\theta)$, and apply Itô's lemma to each component of u:

$$\begin{aligned} u_i(\sigma, \tilde{\mathcal{X}}_\sigma, \tilde{\theta}_\sigma) - u_i(\tau, \tilde{\mathcal{X}}_\tau, \tilde{\theta}_\tau) = \int_\tau^\sigma \partial_s u_i(s, \tilde{\mathcal{X}}_s, \tilde{\theta}_s) ds \\ + \int_\tau^\sigma \mathcal{A}_{\mathcal{X}} u_i(s, \tilde{\mathcal{X}}_s, \tilde{\theta}_s) ds + \int_\tau^\sigma \mathcal{A}_\theta u_i(s, \tilde{\mathcal{X}}_s, \tilde{\theta}_s) ds \\ + \int_\tau^\sigma \gamma_s \mathrm{tr}\left[\hat{\Sigma}(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s)^\top \partial_\theta^\top \partial_x u_i(s, \tilde{\mathcal{X}}_s, \tilde{\theta}_s)\right] ds \\ + \int_\tau^\sigma \partial_x u_i(s, \tilde{\mathcal{X}}_s, \tilde{\theta}_s) \Sigma(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) d\mathcal{B}_s \\ + \int_\tau^\sigma \gamma_s \partial_\theta u_i(s, \tilde{\mathcal{X}}_s, \tilde{\theta}_s) H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) dV_s, \end{aligned} \tag{103}$$

where $\mathcal{A}_{\mathcal{X}}$ and $\mathcal{A}_\theta$ are the infinitesimal generators of the processes $\mathcal{X}_t$ and $\tilde{\theta}_t$, respectively, $\hat{\Sigma}(x,\theta)$ denotes the $(D \times n_y)$-matrix consisting of the rows $n'+1, n'+2, ..., n'+n_y$ of the matrix $\Sigma(x,\theta)$, and $\partial_\theta^\top \partial_x u_k(s,x,\theta)_{ij} = \partial_{\theta_i} \partial_{x_j} u_k(s,x,\theta)$. Using the Poisson equation and the previous identity, we obtain

$$\begin{aligned} \Gamma_{k,\eta} &= \int_{\tau_k}^{\sigma_k+\eta} \gamma_s \left(F(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) - g(\tilde{\theta}_s)\right) ds \\ &= \int_{\tau_k}^{\sigma_k+\eta} \gamma_s G(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) ds = \int_{\tau_k}^{\sigma_k+\eta} \gamma_s \mathcal{A}_{\mathcal{X}} v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) ds \\ &= \int_{\tau_k}^{\sigma_k+\eta} \mathcal{A}_{\mathcal{X}} u(s, \tilde{\mathcal{X}}_s, \tilde{\theta}_s) ds \\ &= \gamma_{\sigma_k+\eta} v(\tilde{\mathcal{X}}_{\sigma_k+\eta}, \tilde{\theta}_{\sigma_k+\eta}) - \gamma_{\tau_k} v(\tilde{\mathcal{X}}_{\tau_k}, \tilde{\theta}_{\tau_k}) \\ &\quad - \int_{\tau_k}^{\sigma_k+\eta} \dot{\gamma}_s v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) ds - \int_{\tau_k}^{\sigma_k+\eta} \gamma_s \mathcal{A}_\theta v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) ds \\ &\quad - \int_{\tau_k}^{\sigma_k+\eta} \gamma_s^2 \mathrm{tr}\left[\hat{\Sigma}(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s)^\top \partial_\theta^\top \partial_x v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s)\right] ds \\ &\quad - \int_{\tau_k}^{\sigma_k+\eta} \gamma_s \partial_x v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) \Sigma(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) d\mathcal{B}_s \\ &\quad - \int_{\tau_k}^{\sigma_k+\eta} \gamma_s^2 \partial_\theta v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) dV_s. \end{aligned} \tag{104}$$

Define

$$J_t^{(1)} \doteq \gamma_t \sup_{s \leq t} ||v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s)||. \tag{105}$$

By using Proposition 1, we have

$$\begin{aligned} \mathbb{E}\left[\left(J_t^{(1)}\right)^2\right] &= \mathbb{E}\left[\gamma_t^2 \sup_{s\leq t} ||v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s)||^2\right] \\ &\leq K\gamma_t^2 \mathbb{E}\left[1 + \sup_{s\leq t} ||\tilde{\mathcal{X}}_s||^q\right] \\ &= K\gamma_t^2 \left(1 + \mathbb{E}\left[\sup_{s\leq t} ||\tilde{\mathcal{X}}_s||^q\right]\right) \\ &\leq KK'\gamma_t^2(1 + \sqrt{t}) \\ &\leq K''\gamma_t^2\sqrt{t}, \end{aligned} \tag{106}$$

where the first two inequalities use Proposition 1 (iii) and (iv), respectively. We choose a $r > 0$ such that $\gamma_t^2 t^{1/2+2r} \to 0$ for $t \to \infty$ (this is possible due to Condition 3), and we pick $T > 0$ large enough such that $\gamma_t^2 t^{1/2+2r} \leq 1$ for $t \geq T$. In addition, for each $0 < \delta < r$ we define the event $A_{t,\delta} \doteq \{J_t^{(1)} t^{r-\delta} \geq 1\}$. For $t \geq T$,

$$\begin{aligned} \mathbb{P}(A_{t,\delta}) \leq \mathbb{E}\left[J_t^{(1)} t^{r-\delta}\right] \leq \mathbb{E}\left[\left(J_t^{(1)}\right)^2\right] t^{2r-2\delta} \\ \leq K''\gamma_t^2 t^{1/2+2r-2\delta} \leq K'' t^{-2\delta}, \end{aligned} \tag{107}$$

where (106) was used in the second inequality.[3] We therefore have that

$$\sum_{n=1}^\infty \mathbb{P}(A_{2^n,\delta}) < \infty. \tag{108}$$

By the Borel-Cantelli Lemma, only finitely many events $A_{2^n,\delta}$ can occur. Therefore, there is a random index n_0 such that

[3] The first inequality in (107) is elementary: For a nonnegative random variable Y with law p, we have

$$\mathbb{P}(Y \geq 1) = \int_1^\infty p(dy) \leq \int_1^\infty y p(dy) \leq \int_0^\infty y p(dy) = \mathbb{E}(Y).$$

$J_{2^n}^{(1)} 2^{n(r-\delta)} \leq 1$ for all $n \geq n_0$. Alternatively, we can say that there is a finite positive random variable ξ and a deterministic $n_1 \in \mathbb{N}$ such that

$$J_{2^n}^{(1)} 2^{n(r-\delta)} \leq \xi, \quad n \geq n_1 \tag{109}$$

(e.g. choose $\xi = \max\{\max_{1 \leq n' \leq n_0} J_{2^{n'}}^{(1)} 2^{n'(r-\delta)}, 1\}$). For $t \in [2^n, 2^{n+1}]$ and $n \geq n_1$, we therefore have

$$\begin{aligned} J_t^{(1)} &= \gamma_t \sup_{s \leq t} ||v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s)|| \\ &\leq \gamma_{2^n} \sup_{s \leq t} ||v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s)|| \\ &\leq \gamma_{2^n} \sup_{s \leq 2^{n+1}} ||v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s)|| \\ &\leq K \gamma_{2^{n+1}} \sup_{s \leq 2^{n+1}} ||v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s)|| \\ &= K J_{2^{n+1}}^{(1)} \\ &\leq K \frac{\xi}{2^{(n+1)(r-\delta)}} \\ &\leq K \frac{\xi}{t^{r-\delta}}, \end{aligned} \tag{110}$$

and as a consequence, $J_t^{(1)} \to 0$ with probability one as $t \to \infty$.

Next, define

$$J_t^{(2)} = \int_0^t \Big|\Big| \dot{\gamma}_s v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) ds + \gamma_s \mathcal{A}_\theta v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) ds + \gamma_s^2 \text{tr}\left[\hat{\Sigma}(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s)^\top \partial_\theta^\top \partial_x v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s)\right] \Big|\Big| ds \tag{111}$$

Due to the PGP of H (condition 2), the boundedness of $\hat{\Sigma}$ (condition 1 (b)) and the PGP of v (Proposition 1 (iii)), we have

$$\begin{aligned} \sup_{t>0} \mathbb{E}\left[J_t^{(2)}\right] &\leq K \int_0^\infty \left(\dot{\gamma}_s + \gamma_s^2\right) \left(1 + \mathbb{E}[||\tilde{\mathcal{X}}_s||^q]\right) ds \\ &\leq KC \int_0^\infty \left(\dot{\gamma}_s + \gamma_s^2\right) ds < \infty. \end{aligned} \tag{112}$$

In the first inequality we additionally used the fact that $\mathcal{A}_\theta$ contains at least a factor of γ_t, in the second one we relied on Proposition 1 (iv) and in the third inequality we used Condition 3 . Thus $J_t^{(2)}$ converges to a finite random variable with probability one.

Lastly, we have the term

$$\begin{aligned} J_t^{(3)} = &\int_0^t \gamma_s \partial_x v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) \Sigma(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) d\mathcal{B}_s \\ &+ \int_0^t \gamma_s^2 \partial_\theta v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) dV_s. \end{aligned} \tag{113}$$

By using Itô isometry, the boundedness of Σ, the PGP of H and v, and Proposition 1 (iv), we obtain

$$\begin{aligned} \sup_{t>0} \mathbb{E}\left[||J_t^{(3)}||^2\right] &= \int_0^\infty \gamma_s^2 \mathbb{E}\left[\|\partial_x v \Sigma\|^2\right] ds \\ &\quad + \int_0^\infty \gamma_s^4 \mathbb{E}\left[\|\partial_\theta v H\|^2\right] ds \\ &\quad + 2\int_0^\infty \gamma_s^3 \text{tr}\mathbb{E}\left[\hat{\Sigma} H^\top \partial_\theta^\top \partial_x v\right] ds \\ &\leq CK \int_0^\infty \left(\gamma_s^2 + \gamma_s^3 + \gamma_s^4\right) \left(1 + \mathbb{E}[||\tilde{\mathcal{X}}_s||^q]\right) ds \\ &\leq CKC' \int_0^\infty \left(\gamma_s^2 + \gamma_s^3 + \gamma_s^4\right) ds < \infty. \end{aligned} \tag{114}$$

Thus, by Doob's martingale convergence theorem, $J_t^{(3)}$ converges to a square integrable random variable with probability one.

Finally, we note that

$$\begin{aligned} ||\Gamma_{k,\eta}|| \leq\ & J_{\sigma_k+\eta}^{(1)} + J_{\tau_k}^{(1)} + J_{\sigma_k+\eta}^{(2)} - J_{\tau_k}^{(2)} \\ &+ ||J_{\sigma_k+\eta}^{(3)} - J_{\tau_k}^{(3)}|| \to 0, \quad k \to \infty, \end{aligned} \tag{115}$$

which concludes the proof. ■

Lemma 2: Let L be the Lipschitz constant of g. Choose $\lambda > 0$ such that for a given $\kappa > 0$ (this is the parameter of the stopping times τ_k) we have $3\lambda + \frac{\lambda}{4\kappa} = \frac{1}{2L}$. For k large enough and $\eta > 0$ small enough, $\int_{\tau_k}^{\sigma_k+\eta} \gamma_s ds > \lambda$. In addition, with probability one, $\frac{\lambda}{2} \leq \int_{\tau_k}^{\sigma_k} \gamma_s ds \leq \lambda$.

Proof: This proof goes through exactly like the proof of Lemma 3.2 in [17], with the only modification that the martingale in that proof takes the form

$$\int_0^t \gamma_s \frac{g(\tilde{\theta}_s)}{R_s} H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) dV_s.$$

■

Lemma 3: Suppose that $\tilde{\theta}_t \in \Theta$ for $t \geq 0$ and that there is an infinite number of intervals $[\tau_k, \sigma_k)$. There is a $\beta > 0$ such that for $k > k_0$,

$$\tilde{\mathcal{L}}(\tilde{\theta}_{\sigma_k}) - \tilde{\mathcal{L}}(\tilde{\theta}_{\tau_k}) \geq \beta \tag{116}$$

almost surely.

Proof: By using Itô's lemma and the parameter update SDE (26), we obtain four terms:

$$\begin{aligned} \tilde{\mathcal{L}}(\tilde{\theta}_{\sigma_k}) - \tilde{\mathcal{L}}(\tilde{\theta}_{\tau_k}) &= \int_{\tau_k}^{\sigma_k} \gamma_s \left\| g(\tilde{\theta}_s) \right\|^2 ds \\ &\quad + \int_{\tau_k}^{\sigma_k} \gamma_s g(\tilde{\theta}_s) H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) dV_s \\ &\quad + \int_{\tau_k}^{\sigma_k} \frac{\gamma_s^2}{2} \text{tr}\left[H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s)^\top \mathcal{H}(\tilde{\theta}_s) H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s)\right] ds \\ &\quad + \int_{\tau_k}^{\sigma_k} \gamma_s g(\tilde{\theta}_s) \cdot \left[F(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) - g(\tilde{\theta}_s)\right] ds \\ &= \Omega_{1,k} + \Omega_{2,k} + \Omega_{3,k} + \Omega_{4,k}, \end{aligned} \tag{117}$$

where $\mathcal{H}$ is used to denote the Hessian of $\tilde{\mathcal{L}}$. By virtue of the definition of the stopping times and Lemmas **??** and 2,

$$\begin{aligned} \Omega_{1,k} &= \int_{\tau_k}^{\sigma_k} \gamma_s \left\| g(\tilde{\theta}_s) \right\|^2 ds \\ &\geq \frac{\left\| g(\tilde{\theta}_{\tau_k}) \right\|^2}{4} \int_{\tau_k}^{\sigma_k} \gamma_s ds \geq \frac{\left\| g(\tilde{\theta}_{\tau_k}) \right\|^2}{8} \lambda(\kappa). \end{aligned} \tag{118}$$

We define

$$R_t = \begin{cases} ||g(\tilde{\theta}_{\tau_k})||, & t \in [\tau_k, \sigma_k) \text{ for some k} \geq 1, \\ \kappa, & \text{else} \end{cases}, \tag{119}$$

such that we can write

$$\begin{aligned}\Omega_{2,k} &= \int_{\tau_k}^{\sigma_k} \gamma_s g(\tilde{\theta}_s) H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) dV_s \\ &= \left\| g(\tilde{\theta}_{\tau_k}) \right\| \int_{\tau_k}^{\sigma_k} \gamma_s \frac{g(\tilde{\theta}_s)}{R_s} H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) dV_s.\end{aligned} \tag{120}$$

Since $||g(\tilde{\theta}_s)||/R_s \leq 2$, it follows from the Itô isometry, condition 2, and Proposition 1 (iv) that

$$\begin{aligned}&\sup_{t\geq 0} \mathbb{E}\left[\left(\int_0^t \gamma_s \frac{g(\tilde{\theta}_s)}{R_s} H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) dV_s \right)^2 \right] \\ &\quad \leq \sup_{t\geq 0} \int_0^t \mathbb{E}\left[\gamma_s^2 \frac{||g(\tilde{\theta}_s)||^2}{R_s^2} ||H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s)||^2 \right] ds \\ &\quad \leq 4 \int_0^\infty \gamma_s^2 \mathbb{E}\left[\left\| H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) \right\|^2 \right] ds \\ &\quad \leq 4K \int_0^\infty \gamma_s^2 \left(1 + \mathbb{E}\left[||\tilde{\mathcal{X}}_s||^q \right] \right) ds < \infty.\end{aligned} \tag{121}$$

By Doob's martingale convergence theorem, the martingale $M_t = \int_0^t \gamma_s \frac{g(\tilde{\theta}_s)}{R_s} H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) dV_s$ converges to a finite random variable M as $t \to \infty$. Thus for any $\epsilon > 0$, there is a k_0 such that almost surely we have $\Omega_{2,k} \leq ||g(\tilde{\theta}_{\tau_k})||\epsilon$ for all $k \geq k_0$.

Next, we consider $\Omega_{3,k}$. Using condition 2 and Propositions 1 and 2, we obtain

$$\begin{aligned}&\sup_{t\geq 0} \mathbb{E}\left[\left| \int_0^t \frac{\gamma_s^2}{2} \mathrm{tr}\left[H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s)^\top \mathcal{H}(\tilde{\theta}_s) H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) \right] ds \right| \right] \\ &\leq \sup_{t\geq 0} \mathbb{E}\left[\int_0^t \frac{\gamma_s^2}{2} \left| \mathrm{tr}\left[H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s)^\top \mathcal{H}(\tilde{\theta}_s) H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) \right] \right| ds \right] \\ &\quad \leq \int_0^\infty \frac{\gamma_s^2}{2} \mathbb{E}\left[||H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s)||^2 ||\mathcal{H}(\tilde{\theta}_s)|| \right] ds \\ &\quad \leq K \int_0^\infty \frac{\gamma_s^2}{2} \left(1 + \mathbb{E}\left[||\tilde{\mathcal{X}}_s||^q \right] \right) ds < \infty,\end{aligned} \tag{122}$$

from which it follows that

$$\int_0^t \frac{\gamma_s^2}{2} \mathrm{tr}\left[H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s)^\top \mathcal{H}(\tilde{\theta}_s) H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) \right] ds$$

converges to a finite random variable as $t \to \infty$. Thus, with probability one, $\Omega_{3,k}$ must converge to zero as $k \to \infty$.

Finally, we consider the term $\Omega_{4,k}$ and define $G(x, \theta) = g(\theta) \cdot [F(x, \theta) - g(\theta)]$. The function G satisfies $\int_{\mathbb{R}^D} G(x, \theta) \mu_\theta(dx) = 0$. By Proposition 1 (iii), for each $\theta \in \Theta$ the Poisson equation $\mathcal{A}_\mathcal{X} v(x, \theta) = G(x, \theta)$ (where $\mathcal{A}_\mathcal{X}$ is the infinitesimal generator of the process $\mathcal{X}_t$) has a unique solution v with $\int_{\mathbb{R}^D} v(x, \theta) \mu_\theta(dx) = 0$. Let $u(t, x, \theta) \doteq \gamma_t v(x, \theta)$ and apply Itô's lemma

$$\begin{aligned}u(\sigma, \tilde{\mathcal{X}}_\sigma, \tilde{\theta}_\sigma) - u(\tau, \tilde{\mathcal{X}}_\tau, \tilde{\theta}_\tau) &= \int_\tau^\sigma \partial_s u(s, \tilde{\mathcal{X}}_s, \tilde{\theta}_s) ds \\ &+ \int_\tau^\sigma \mathcal{A}_\mathcal{X} u(s, \tilde{\mathcal{X}}_s, \tilde{\theta}_s) ds + \int_\tau^\sigma \mathcal{A}_\theta u(s, \tilde{\mathcal{X}}_s, \tilde{\theta}_s) ds \\ &+ \int_\tau^\sigma \gamma_s \mathrm{tr}\left[\hat{\Sigma}(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s)^\top \partial_x \partial_\theta u(s, \tilde{\mathcal{X}}_s, \tilde{\theta}_s) \right] ds \\ &+ \int_\tau^\sigma \partial_x u(s, \tilde{\mathcal{X}}_s, \tilde{\theta}_s) \Sigma(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) d\mathcal{B}_s \\ &+ \int_\tau^\sigma \gamma_s \partial_\theta u(s, \tilde{\mathcal{X}}_s, \tilde{\theta}_s) H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) dV_s,\end{aligned} \tag{123}$$

where $\hat{\Sigma}(x, \theta)$ denotes the $(D \times n_y)$-matrix consisting of the rows $n' + 1, n' + 2, ..., n' + n_y$ of the matrix $\Sigma(x, \theta)$, and $\partial_x \partial_\theta u(s, x, \theta)_{ij} = \partial_{\theta_i} \partial_{x_j} u(s, x, \theta)$. Using the Poisson equation and the previous identity, we obtain

$$\begin{aligned}\Omega_{4,k} &= \int_{\tau_k}^{\sigma_k} \gamma_s g(\tilde{\theta}_s) \cdot \left[F(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) - g(\tilde{\theta}_s) \right] ds \\ &= \int_{\tau_k}^{\sigma_k} \gamma_s G(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) ds = \int_{\tau_k}^{\sigma_k} \gamma_s \mathcal{A}_\mathcal{X} v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) ds \\ &= \int_{\tau_k}^{\sigma_k} \mathcal{A}_\mathcal{X} u(s, \tilde{\mathcal{X}}_s, \tilde{\theta}_s) ds \\ &= \gamma_{\sigma_k} v(\tilde{\mathcal{X}}_{\sigma_k}, \tilde{\theta}_{\sigma_k}) - \gamma_{\tau_k} v(\tilde{\mathcal{X}}_{\tau_k}, \tilde{\theta}_{\tau_k}) \\ &\quad - \int_{\tau_k}^{\sigma_k} \partial_s \gamma_s v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) ds - \int_{\tau_k}^{\sigma_k} \gamma_s \mathcal{A}_\theta v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) ds \\ &\quad - \int_{\tau_k}^{\sigma_k} \gamma_s^2 \mathrm{tr}\left[\hat{\Sigma}(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s)^\top \partial_\theta^\top \partial_x v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) \right] ds \\ &\quad - \int_{\tau_k}^{\sigma_k} \gamma_s \partial_x v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) \Sigma(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) d\mathcal{B}_s \\ &\quad - \int_{\tau_k}^{\sigma_k} \gamma_s^2 \partial_\theta v(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) dV_s\end{aligned} \tag{124}$$

Note that this has the same structure as Eq. (104). By following the steps in the proof of Lemma 1, we find that $\Omega_{4,k} \to 0$ as $k \to \infty$ with probability one.

For all $\epsilon > 0$ with probability one, we have for k large enough

$$\begin{aligned}\tilde{\mathcal{L}}(\tilde{\theta}_{\sigma_k}) - \tilde{\mathcal{L}}(\tilde{\theta}_{\tau_k}) &= \Omega_{1,k} + \Omega_{2,k} + \Omega_{3,k} + \Omega_{4,k} \\ &\geq \Omega_{1,k} - ||\Omega_{2,k}|| - ||\Omega_{3,k}|| - ||\Omega_{4,k}|| \\ &\geq \frac{1}{8} \lambda(\kappa) ||g(\tilde{\theta}_{\tau_k}||^2 - \epsilon ||g(\tilde{\theta}_{\tau_k}|| - 2\epsilon.\end{aligned} \tag{125}$$

The Lemma then follows by choosing $\epsilon = \min\{\frac{\lambda(\kappa)\kappa^2}{32}, \frac{\lambda(\kappa)}{32}\}$ and $\beta = \frac{\lambda(\kappa)\kappa^2}{32}$. ∎

Lemma 4: Under the conditions of Lemma 3 there is a $0 < \beta_1 < \beta$ such that for $k > k_0$

$$\tilde{\mathcal{L}}(\tilde{\theta}_{\tau_k}) - \tilde{\mathcal{L}}(\tilde{\theta}_{\sigma_{k-1}}) \geq -\beta_1 \tag{126}$$

almost surely.

Proof: As in Lemma 3, we obtain

$$\begin{aligned}\tilde{\mathcal{L}}(\tilde{\theta}_{\tau_k}) - \tilde{\mathcal{L}}(\tilde{\theta}_{\sigma_{k-1}}) &\geq \int_{\sigma_{k-1}}^{\tau_k} \gamma_s g(\tilde{\theta}_s) H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) dV_s \\ &+ \int_{\sigma_{k-1}}^{\tau_k} \frac{\gamma_s^2}{2} \mathrm{tr}\left[H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s)^\top \mathcal{H}(\tilde{\theta}_s) H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) \right] ds \\ &+ \int_{\sigma_{k-1}}^{\tau_k} \gamma_s g(\tilde{\theta}_s) \cdot \left[F(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) - g(\tilde{\theta}_s) \right] ds.\end{aligned} \tag{127}$$

It is sufficient to show that the RHS converges to zero almost surely. Due to Eq. (119), the first term can be rewritten as

$$\kappa \int_{\sigma_{k-1}}^{\tau_k} \gamma_s \frac{g(\tilde{\theta}_s)}{R_s} H(\tilde{\mathcal{X}}_s, \tilde{\theta}_s) dV_s. \tag{128}$$

Using the argument from the proof of Lemma 3, this converges to zero almost surely as $k \to \infty$. The treatment of the second and third terms is identical to the treatment of the terms $\Omega_{3,k}$ and $\Omega_{4,k}$ in the proof of Lemma 3. ■

Appendix D
Proof of Theorem 1

First, define the first exit time from Θ

$$\tau = \inf\left\{t \geq 0 : \; \tilde{\theta}_t \notin \Theta\right\}. \tag{129}$$

If $\tau < \infty$, since the paths of $\tilde{\theta}_t$ are continuous, we have $\tilde{\theta}_\tau \in \partial\Theta$. Furthermore, since $d\tilde{\theta}_t = 0$ on $\partial\Theta$, we have $\tilde{\theta}_t \in \partial\Theta$ for all $t \geq \tau$.

Next, consider the case when $\tau = \infty$, which implies that $\tilde{\theta}_t \in \Theta$ for all $t \geq 0$. Consider the case when there is a finite number of stopping times τ_k. Then, there is a finite T such that $||g(\tilde{\theta}_t)|| < \kappa$ for $t \geq T$. Therefore, since κ can be chosen arbitrarily small, $\lim_{t\to\infty} ||g(\tilde{\theta}_t)|| = 0$. Next, suppose that the number of stopping times τ_k is infinite. By Lemmas 3 and 4 there is a k_0 and constants $\beta > \beta_1 > 0$ such that for all $k \geq k_0$ with probability one

$$\tilde{\mathcal{L}}(\tilde{\theta}_{\sigma_k}) - \tilde{\mathcal{L}}(\tilde{\theta}_{\tau_k}) \geq \beta \tag{130}$$

$$\tilde{\mathcal{L}}(\tilde{\theta}_{\tau_k}) - \tilde{\mathcal{L}}(\tilde{\theta}_{\sigma_{k-1}}) \geq -\beta_1 > -\beta. \tag{131}$$

Thus, we have

$$\begin{aligned}
&\tilde{\mathcal{L}}(\tilde{\theta}_{\tau_{n+1}}) - \tilde{\mathcal{L}}(\tilde{\theta}_{\tau_{k_0}}) \\
&\quad = \sum_{k=k_0}^{n} \left[\tilde{\mathcal{L}}(\tilde{\theta}_{\sigma_k}) - \tilde{\mathcal{L}}(\tilde{\theta}_{\tau_k}) + \tilde{\mathcal{L}}(\tilde{\theta}_{\tau_{k+1}}) - \tilde{\mathcal{L}}(\tilde{\theta}_{\sigma_k})\right] \\
&\quad \geq (n+1-k_0)(\beta - \beta_1).
\end{aligned} \tag{132}$$

Since $\beta - \beta_1 > 0$, when $n \to \infty$, $\tilde{\mathcal{L}}(\tilde{\theta}_{\tau_{n+1}}) \to \infty$ almost surely, and therefore $\tilde{\mathcal{L}}(\tilde{\theta}_t) \to \infty$ almost surely. This is a contradiction to Proposition 2 (v), which states that $\tilde{\mathcal{L}}$ is bounded from above. Therefore, there are almost surely only a finite number of stopping times τ_k. ■

Acknowledgment

The authors would like to thank Philipp Harms and Ofer Zeitouni for helpful discussions.

References

[1] A. Bain and D. Crişan, *Fundamentals of stochastic filtering*, ser. Stochastic Modelling and Applied Probability. Springer, New York, 2009, vol. 60.

[2] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Trans. ASME Ser. D. J. Basic Engrg.*, vol. 83, pp. 95–108, 1961.

[3] V. E. Beneš, "Exact finite-dimensional filters for certain diffusions with nonlinear drift," *Stochastics*, vol. 5, no. 1-2, pp. 65–92, 1981.

[4] ——, "New exact nonlinear filters with large Lie algebras," *Systems & Control Letters*, vol. 5, no. 4, pp. 217–221, 1985.

[5] F. E. Daum, "Exact finite-dimensional nonlinear filters," *IEEE Transactions on Automatic Control*, vol. 31, no. 7, pp. 616–622, 1986.

[6] ——, "Solution of the Zakai equation by separation of variables," *IEEE Transactions on Automatic Control*, vol. 32, no. 10, pp. 941–943, 1987.

[7] C. D. Charalambous and R. J. Elliott, "Classes of nonlinear partially observable stochastic optimal control problems with explicit optimal control laws," *SIAM Journal on Control and Optimization*, vol. 36, no. 2, pp. 542–578, 1998.

[8] F. Gustafsson and G. Hendeby, "Some Relations Between Extended and Unscented Kalman Filters," *IEEE Transactions on Signal Processing*, vol. 60, no. 2, pp. 545–555, Mar. 2012.

[9] E. A. Wan and R. Van Der Merwe, "The unscented Kalman filter for nonlinear estimation," in *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373*. IEEE, 2000, pp. 153–158.

[10] D. Brigo, B. Hanzon, and F. LeGland, "A differential geometric approach to nonlinear filtering: the projection filter," *IEEE Transactions on Automatic Control*, vol. 43, no. 2, pp. 247–252, 1998.

[11] D. Brigo, B. Hanzon, and F. Le Gland, "Approximate nonlinear filtering by projection on exponential manifolds of densities," *Bernoulli*, vol. 5, no. 3, pp. 495–534, 1999.

[12] A. Doucet and A. M. Johansen, "A tutorial on particle filtering and smoothing: fifteen years later," in *The Oxford handbook of nonlinear filtering*. Oxford Univ. Press, Oxford, 2011, pp. 656–704.

[13] G. Evensen, "The Ensemble Kalman Filter: theoretical formulation and practical implementation," *Ocean Dynamics*, vol. 53, no. 4, pp. 343–367, Nov. 2003.

[14] D. Crişan and J. Xiong, "Approximate McKean–Vlasov representations for a class of SPDEs," *Stochastics An International Journal of Probability and Stochastic Processes*, vol. 82, no. 1, pp. 53–68, Feb. 2010.

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

[16] F. LeGland and L. Mevel, "Recursive estimation in hidden Markov models," in *Decision and Control, 1997., Proceedings of the 36th IEEE Conference on*, 1997, pp. 3468–3473.

[17] J. Sirignano and K. Spiliopoulos, "Stochastic Gradient Descent in Continuous Time," *arXiv.org*, Nov. 2016.

[18] A. Y. Veretennikov, "On polynomial mixing bounds for stochastic differential equations," *Stochastic Processes and their Applications*, vol. 70, no. 1, pp. 115–127, Oct. 1997.

[19] E. Pardoux and A. Y. Veretennikov, "On the Poisson equation and diffusion approximation 1," *The Annals of Probability*, vol. 29, no. 3, pp. 1061–1085, Dec. 2001.

[20] ——, "On the Poisson equation and diffusion approximation 2," *The Annals of Probability*, vol. 31, no. 3, pp. 1166–1192, May 2003.

[21] ——, "On the Poisson equation and diffusion approximation 3," *The Annals of Probability*, vol. 33, no. 3, pp. 1111–1133, May 2005.

[22] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, Jul. 2000.

[23] O. Cappé, E. Moulines, and T. Ryden, *Inference in Hidden Markov Models*. Springer Science & Business Media, Apr. 2006.

[24] R. van Handel, *Hidden markov models*. Unpublished notes, Jul. 2008.

[25] W. Khreich, E. Granger, A. Miri, and R. Sabourin, "A survey of techniques for incremental learning of HMM parameters," *Information Sciences*, vol. 197, no. C, pp. 105–130, Aug. 2012.

[26] F. LeGland and L. Mevel, "Recursive identification of HMMs with observations in a finite set," in *Proceedings of 1997 36th IEEE Conference on Decision and Control*. IEEE, 1995, pp. 216–221 vol.1.

[27] V. B. Tadic, "Analyticity, Convergence, and Convergence Rate of Recursive Maximum-Likelihood Estimation in Hidden Markov Models," *IEEE Transactions on Information Theory*, vol. 56, no. 12, pp. 6406–6432, Nov. 2010.

[28] I. B. Collings, V. Krishnamurthy, and J. B. Moore, "On-line identification of hidden Markov models via recursive prediction error techniques," *IEEE Transactions on Signal Processing*, vol. 42, no. 12, pp. 3535–3539, 1994.

[29] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society ...*, 1977.

[30] A. Dembo and O. Zeitouni, "Parameter estimation of partially observed continuous time stochastic processes via the EM algorithm," *Stochastic Processes and their Applications*, vol. 23, no. 1, pp. 91–113, 1986.

[31] C. D. Charalambous and A. Logothetis, "Maximum likelihood parameter estimation from incomplete data via the sensitivity equations: The continuous-time case," *IEEE Transactions on Automatic Control*, vol. 45, no. 5, pp. 928–934, 2000.

[32] O. Zeitouni and B. Z. Bobrovsky, "On the joint nonlinear filtering-smoothing of diffusion processes," *System Control Letters*, vol. 7, pp. 317–321, 1986.

[33] R. J. Elliott and V. Krishnamurthy, "Exact finite-dimensional filters for maximum likelihood parameter estimation of continuous-time linear Gaussian systems," *SIAM Journal on Control and Optimization*, vol. 35, no. 6, pp. 1908–1923, 1997.

[34] N. Gupta and R. Mehra, “Computational aspects of maximum likelihood estimation and reduction in sensitivity function calculations,” *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 774–783, 1974.

[35] V. Krishnamurthy and J. B. Moore, “On-Line estimation of hidden Markov model parameters based on the Kullback-Leibler information measure,” *IEEE Transactions on Signal Processing*, vol. 41, no. 8, pp. 2557–2573, 1993.

[36] G. Mongillo and S. Deneve, “Online Learning with Hidden Markov Models,” *Neural Computation*, vol. 20, no. 7, pp. 1706–1716, Jul. 2008.

[37] O. Cappé, “Online sequential monte carlo em algorithm,” *Statistical Signal Processing*, 2009.

[38] ——, “Online EM Algorithm for Hidden Markov Models,” *Journal of Computational and Graphical Statistics*, vol. 20, no. 3, pp. 728–749, Jan. 2011.

[39] ——, “Recursive computation of smoothed functionals of hidden Markovian processes using a particle approximation,” *Monte Carlo Methods and Applications*, vol. 7, no. 1-2, pp. 81–92, 2001.

[40] P. Del Moral, A. Doucet, and S. Singh, “Forward Smoothing using Sequential Monte Carlo,” *arXiv.org*, Dec. 2010.

[41] P. Rebeschini and R. van Handel, “Can local particle filters beat the curse of dimensionality?” *Annals of Applied Probability*, vol. 25, no. 5, pp. 2809–2866, Oct. 2015.

[42] A. Papavasiliou, “Parameter estimation and asymptotic stability in stochastic filtering,” *Stochastic Processes and their Applications*, vol. 116, no. 7, pp. 1048–1065, Jul. 2006.

[43] J. M. F. Moura and S. K. Mitter, “Identification and Filtering–optimal Recursive Maximum Likelihood Approach,” Massachusetts Institute of Technology and Laboratory for Information and Decision Systems, Tech. Rep., 1986.