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

Simone Carlo Surace and **Jean-Pascal Pfister**
Institute of Neuroinformatics,
University of Zürich and ETH Zürich,
Zürich, Switzerland.

Editors: Oren Anava, Marco Cuturi, Azadeh Khalghi, Vitaly Kuznetsov and Alexander Rakhlin

Keywords: Online Learning, Parameter Estimation, Maximum likelihood, Diffusion Process, Change of Measure

Abstract

We revisit the problem of estimating the parameters of a partially observed stochastic process (X_t, Y_t) with a continuous time parameter, where X_t is the hidden state process and Y_t is the observed process. The estimation is to be done online, i.e. the parameter estimate should be updated recursively based on the observation filtration $\sigma\{Y_s, s \leq t\}$. Online parameter estimation is a challenging problem that needs to be solved for designing adaptive filters and for stochastic control in all cases where the system is unknown or changing over time, with applications in robotics, neuroscience, or finance. Here, we use the representation of the log-likelihood function in terms of the Radon-Nikodym derivative of the probability measure restricted to the observation (the observation likelihood) with respect to a reference measure under which Y_t is a Wiener process. This log-likelihood can be computed by using the stochastic filter. Using stochastic gradient ascent on the likelihood function, we obtain an algorithm for the time evolution of the parameter estimate. Although this approach is based on theoretical results that have been known for several decades, this explicit method of recursive parameter estimation has remained unpublished.

1. Introduction

We consider a partially observed diffusion process under the probability measure P

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

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

where X_t is called the hidden state or signal of the system, and Y_t is called the observation, and W_t, V_t are independent Brownian motions (signal and observation noise respectively). The system also depends on a static parameter θ. The classical filtering problem is to find the conditional distribution of X_t conditioned on the history of observations $F_t^Y = \sigma\{Y_s, s \leq t\}$.

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

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

is a (P, F_t^Y)-Brownian motion. By applying Girsanov's theorem, we can change to a measure $\tilde{P}$ such that Y_t is a $(\tilde{P}, F_t^Y)$-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

$$E_P\left[\frac{dP}{d\tilde{P}}\Big|F_t^Y\right] = \exp\left[\int_0^t \hat{h}_s dY_s - \frac{1}{2}\int_0^t \hat{h}_s^2 ds\right], \tag{4}$$

In this paper, we consider the problem of finding an estimator $\tilde{\theta}_t$ that is F_t^Y-measurable, such as to estimate θ online from the stream of observations.[1] 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 observation as a function of the model parameters. Since the reference measure $\tilde{P}$ does not depend on θ, we can express the marginal log-likelihood function of the observations in terms of the optimal filter as

$$\mathcal{L}_t(\theta) = \log E_P\left[\frac{dP}{d\tilde{P}}\Big|F_t^Y\right] = \int_0^t \hat{h}_s dY_s - \frac{1}{2}\int_0^t \hat{h}_s^2 ds. \tag{5}$$

For an in-depth discussion 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. Bain and Crişan (2009).

2. Methods

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

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

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

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

where the first factor of the integrand, denoted by

$$\hat{h}_s^\theta \doteq \frac{d}{d\theta}\hat{h}_s, \tag{8}$$

is called the *filter derivative* of h with respect to θ. It captures both the explicit parameter dependence of h and the implicit dependence due to the dependence of the conditional expectation E_P on θ. It may be computed by using the filtering equation applied to the function h,

$$d\hat{h}_s = (\widehat{\mathcal{A}h})_s ds + \left((\widehat{h^2})_s - \hat{h}_s^2\right)(dY_s - \hat{h}_s ds), \tag{9}$$

where $\mathcal{A} = f(x,\theta)\partial_x + \frac{1}{2}g^2(x,\theta)\partial_x^2$ is the infinitesimal generator of the process X. After differentiating by θ, we obtain the following stochastic differential equation (SDE) for the filter derivative:

$$d\hat{h}_s^\theta = (\widehat{\mathcal{A}h})_s^\theta ds + \left((\widehat{h^2})_s^\theta - 2\hat{h}_s\hat{h}_s^\theta\right)(dY_s - \hat{h}_s ds) - \left((\widehat{h^2})_s - \hat{h}_s^2\right)\hat{h}_s^\theta ds. \tag{10}$$

1. We will denote all parameter estimates that are estimated using the maximum likelihood method by '˜' (such as $\tilde{\theta}$), whereas filter estimates (conditional expectations with respect to the observation filtration) are denoted by 'ˆ' (such as $\hat{h}_s$)

In general, both Eq. (9) and (10) couple to an infinite set of moment equations that completely describes the filter. In very rare cases, the filter has a finite-dimensional representation that allows the filter to be described by a finite number of statistics. In all other cases, the filtering problem can only be solved approximately, and the type of approximation scheme dictates the method for computing the filter derivatives. We shall not go into the details of this question in this paper, but we will present several examples of both exact and approximate filters below for which the calculation of filter derivatives will be made explicit.

Let us now turn to the 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 likelihood to update the parameter estimate online as new data is reaching the observer. The time-dependent parameter estimate $\tilde{\theta}_t$ evolves as

$$d\tilde{\theta}_t = \eta_t \frac{d}{d\theta} d\mathcal{L}_t(\theta)\Big|_{\theta=\tilde{\theta}_t} = \eta_t \hat{h}_t^\theta \left(dY_t - \hat{h}_t dt \right) \Big|_{\theta=\tilde{\theta}_t}, \tag{11}$$

where $\eta_t > 0$ is a time-dependent learning rate. The processes $\hat{h}_s$ and $\hat{h}_s^\theta$ have to be integrated along with all other SDEs, making use of the instantaneous parameter estimate $\tilde{\theta}_t$. We call Eq. (11) a *parameter learning rule* for θ. The generalization to multiple parameters is straightforward, since the gradient ascent does not mix parameters. Thus, the same equations hold for each parameter θ in a parameter vector Θ independently.

In Eq. (11), we introduced a time-dependent learning rate η_t for three reasons. The first is that gradient ascent is by definition dimensionally ill-defined, since it identifies a co-vector (the gradient of the cost function) with a vector (the tangent vector of the learning trajectory) in a possibly non-Euclidean parameter space. This identification is not preserved under a change of coordinates, and therefore the learning trajectories depend on the chosen parametrization. A time-dependent learning rate partially solves this problem by admitting a diagonal Riemannian metric, which makes this identification in a consistent way. The second reason, which is a special case of the first, but the only one we shall use in this paper, is the possibility to safe-guard against sign-changes in parameter estimates that have to be either strictly non-negative or non-positive for stability reasons. Making the learning rate proportional to the parameter estimate (i.e. $\eta_t = \eta\tilde{\theta}_t$ with η some positive constant) will prevent that parameter from changing sign, effectively introducing a boundary. The third reason is to make the learning rate explicitly dependent on time for non-stationary models. However, all models considered in this paper will be stationary.

3. Examples and numerical validation

Here, we consider two different example filtering problems and show explicitly how the parameter learning rules are derived. We also numerically validate the learning method. 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.

3.1 One-dimensional Kalman-Bucy filter (linear filtering problem)

We shall first consider the simple case of the linear filtering problem, for which the exact solution is available. Here, we have three parameters, i.e. $\Theta = (a, \sigma, w)$, where $a, \sigma > 0$ and $w \in \mathbb{R}$, and we have $f(x, a, \sigma, w) = -\theta x$, $g(x, a, \sigma, w) = \sigma$ and $h(x, a, \sigma, w) = wx$, such that the filtering problem reads

$$dX_t = -aX_t dt + \sigma dW_t, \qquad dY_t = wX_t dt + dV_t. \tag{12}$$

We make the observation that this model is non-identifiable as a model for the observation process Y_t, since any rescaling of the parameters that preserves $\frac{w^2\sigma^2}{2a}$ also preserves the statistics of Y_t. The model becomes identifiable by treating either σ or w as known. Alternatively, one may choose a parametrization for which X_t has unit variance (i.e. $\sigma = \sqrt{2a}$).

The online parameter update equations read

$$d\tilde{a}_t = \eta_a \tilde{a}_t \tilde{w}_t \mu_t^a \left(dY_t - \tilde{w}_t \mu_t dt\right), \tag{13}$$
$$d\tilde{\sigma}_t = \eta_\sigma \tilde{\sigma}_t \tilde{w}_t \mu_t^\sigma \left(dY_t - \tilde{w}_t \mu_t dt\right), \tag{14}$$
$$d\tilde{w}_t = \eta_w \tilde{w}_t \left(\mu_t + \tilde{w}_t \mu_t^w\right) \left(dY_t - \tilde{w}_t \mu_t dt\right). \tag{15}$$

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 F_t^Y). We introduced the conditional mean $\mu_t = \hat{X}_t$ and its filter derivatives μ_t^a, μ_t^σ and μ_t^w. In fact, μ_t and the variance P_t of the Kalman-Bucy filter (Kalman and Bucy, 1961) evolve as

$$d\mu_t = -\tilde{a}_t \mu_t dt + \tilde{w}_t P_t (dY_t - \tilde{w}_t \mu_t dt), \quad \mu_0 = 0, \tag{16}$$
$$dP_t = \left(\tilde{\sigma}_t^2 - 2\tilde{a}_t P_t - \tilde{w}_t^2 P_t^2\right) dt, \quad P_0 = \frac{\tilde{\sigma}_0^2}{2\tilde{a}_0}, \tag{17}$$

where the initialization of P_0 reflects the prior belief of the variance of X_0 based on the initial parameter estimates, and the filter derivatives of the mean and variance satisfy

$$d\mu_t^a = -\left[\mu_t + \left(\tilde{a}_t + \tilde{w}_t^2 P_t\right) \mu_t^a + \tilde{w}_t^2 \mu_t P_t^a\right] dt + \tilde{w}_t P_t^a dY_t, \tag{18}$$
$$dP_t^a = -\left[2P_t + 2\left(\tilde{a}_t + \tilde{w}_t^2 P_t\right) P_t^a\right] dt, \tag{19}$$
$$d\mu_t^\sigma = -\left[\left(\tilde{a}_t + \tilde{w}_t^2 P_t\right) \mu_t^\sigma + \tilde{w}_t^2 \mu_t P_t^\sigma\right] dt + \tilde{w}_t P_t^\sigma dY_t, \tag{20}$$
$$dP_t^\sigma = \left[2\tilde{\sigma}_t - 2\left(\tilde{a}_t + \tilde{w}_t^2 P_t\right) P_t^\sigma\right] dt, \tag{21}$$
$$d\mu_t^w = -\left[2\tilde{w}_t \mu_t P_t + \left(\tilde{a}_t + \tilde{w}_t^2 P_t\right) \mu_t^w + \tilde{w}_t^2 \mu_t P_t^w\right] dt + \left[P_t + \tilde{w}_t P_t^w\right] dY_t, \tag{22}$$
$$dP_t^w = -\left[2\tilde{w}_t P_t^2 + 2\left(\tilde{a}_t + \tilde{w}_t^2 P_t\right) P_t^w\right] dt, \tag{23}$$
$$\mu_0^a = \mu_0^\sigma = \mu_0^w = 0, \quad P_0^a = -\frac{\tilde{\sigma}_0^2}{2\tilde{a}_0^2}, \quad P_0^\sigma = \frac{\tilde{\sigma}_0}{\tilde{a}_0}, \quad P_0^w = 0, \tag{24}$$

where the right-hand sides of the equations are evaluated with parameters set to the current estimated values.

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 $\tilde{w}_0 = w = 3$ and $\eta_w = 0$. The performance of the algorithm is

normalized MSE	without learning	with 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 1: 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.

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 $a = 1$, $\sigma = 2$, 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 positive learning rates $\eta_a = \eta_\sigma = 0.03$, 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.

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

Consider the following system with four positive parameters $\Theta = (a, b, \sigma, w)$:

$$dX_t = X_t \left(a - bX_t^2\right) dt + \sigma dW_t, \qquad dY_t = wX_t dt + dV_t. \tag{25}$$

In this problem the hidden state X_t has a bimodal stationary distribution with modes at $\pm\sqrt{a/b}$. Since the observation model is linear like in Section 3.1, the parameter learning rules are expressed in terms of the posterior mean $\mu_t = \hat{X}_t$ as

$$d\tilde{a}_t = \eta_a \tilde{a}_t \tilde{w}_t \mu_t^a \left(dY_t - \tilde{w}_t \mu_t dt\right), \tag{26}$$

$$d\tilde{b}_t = \eta_b \tilde{b}_t \tilde{w}_t \mu_t^b \left(dY_t - \tilde{w}_t \mu_t dt\right), \tag{27}$$

$$d\tilde{\sigma}_t = \eta_\sigma \tilde{\sigma}_t \tilde{w}_t \mu_t^\sigma \left(dY_t - \tilde{w}_t \mu_t dt\right), \tag{28}$$

$$d\tilde{w}_t = \eta_w \tilde{w}_t \left(\mu_t + \tilde{w}_t \mu_t^w\right) \left(dY_t - \tilde{w}_t \mu_t dt\right), \tag{29}$$

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 3.1, the filtering problem is not exactly solvable. We use the projection filter on the manifold of Gaussian densities introduced by Brigo et al. (1999), 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

$$d\mu_t = \left[\tilde{a}_t \mu_t - \tilde{b}_t \mu_t^3 - \left(3\tilde{b}_t + \tilde{w}_t^2\right) \mu_t P_t\right] dt + \tilde{w}_t P_t dY_t, \quad \mu_0 = 0, \tag{30}$$

$$dP_t = \left[\tilde{\sigma}_t^2 + \left(2\tilde{a}_t - \tilde{w}_t^2 P_t^2 - 6\tilde{b}_t(\mu_t^2 + P_t)\right) P_t\right] dt, \tag{31}$$

where the initial variance as a function of the initial parameter estimates is given by

$$P_0 = \Gamma\left(\tilde{a}_0, \tilde{b}_0, \tilde{\sigma}_0\right) = \frac{\int_{-\infty}^{\infty} x^2 e^{\tilde{\sigma}_0^{-2}\left(\tilde{a}_0 x^2 - \frac{1}{2}\tilde{b}_0 x^4\right)} dx}{\int_{-\infty}^{\infty} e^{\tilde{\sigma}_0^{-2}\left(\tilde{a}_0 x^2 - \frac{1}{2}\tilde{b}_0 x^4\right)} dx}. \tag{32}$$

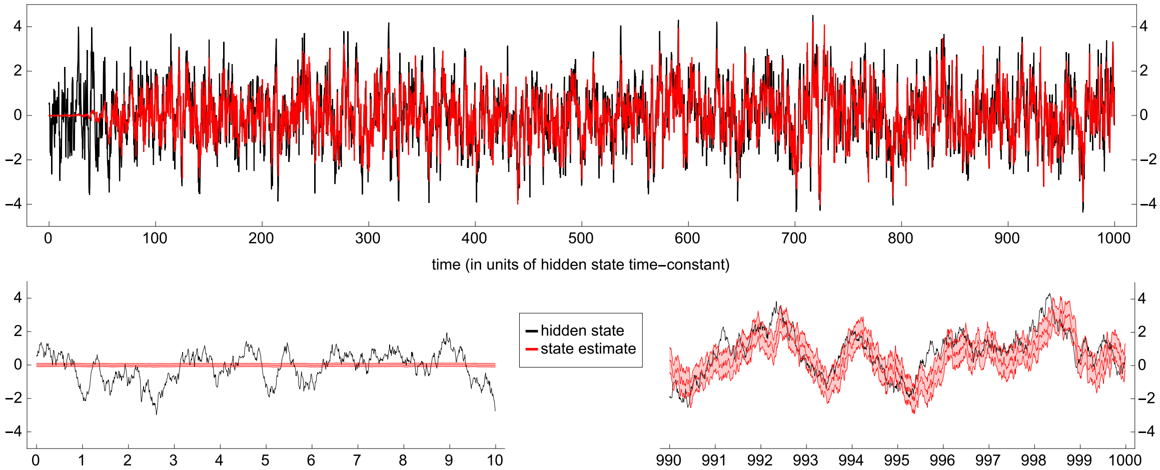

Figure 1: **Online learning and filtering in the linear model**. The hidden state X_t (black) and Kalman-Bucy state estimate μ_t (red, shaded region shows μ_t $\pm$ one standard deviation $\sqrt{P_t}$, c.f. Eqs. (16,17)) are shown for the linear model of Section 3.1 with parameters $a = 1$, $\sigma = 2$, $w = 3$. 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 is known), and the learning rates are $\eta_a = \eta_\sigma = 0.03$ and $\eta_w = 0$. 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.

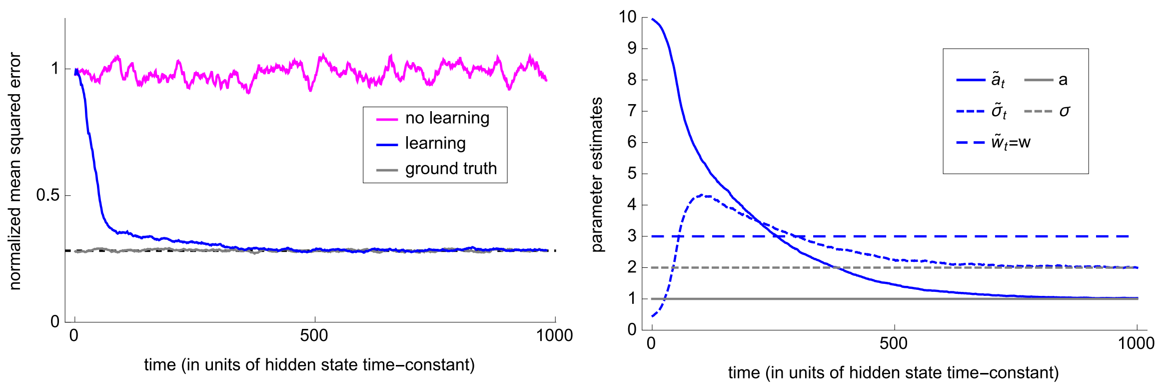

Figure 2: **Online learning and filtering in the linear model**. The time evolution of the MSE and parameter estimates are shown for the linear model of Section 3.1 (see Fig. 1 caption for details). Left: the moving average of the normalized MSE (time window of 20 seconds) shows how the learning algorithm 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).

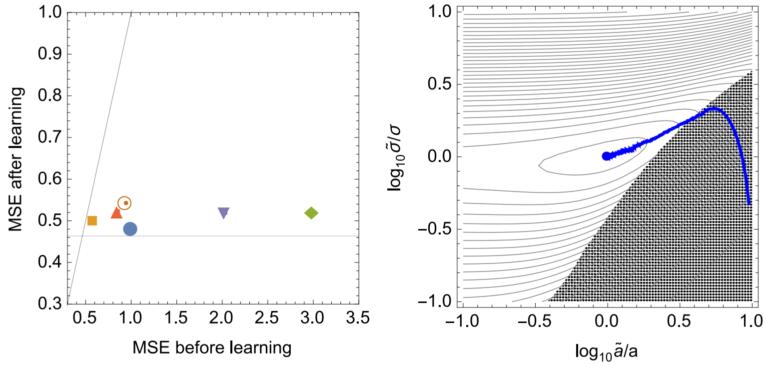

Figure 3: **Online learning and filtering in the linear model**. Left: the filter error before and after learning is shown for different initial conditions of the parameter estimates. For all initial conditions, the learning algorithm achieves a dramatical improvement of the filter error. The horizontal 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 log likelihood function from Eq. (5) for fixed $t = 1000$ in the parameter subspace spanned by $\tilde{a}$ and $\tilde{\sigma}$ for $\tilde{w} = w = 3$ has a single global maximum near $\tilde{a} = a$ and $\tilde{\sigma} = \sigma$. The shading shows the region where the function is non-concave, and the blue line is the trial-averaged learning trajectory from Fig. 2.

By differentiating Eqs. (30,31) with respect to the parameters, we obtain the following equations for the filter derivatives:

$$d\mu_t^a = \left[\mu_t + \alpha_t \mu_t^a + \beta_t P_t^a\right] dt + \tilde{w}_t P_t^a dY_t, \tag{33}$$

$$dP_t^a = \left[2P_t + A_t \mu_t^a + B_t P_t^a\right] dt, \tag{34}$$

$$d\mu_t^b = \left[-\mu_t \left(\mu_t^2 + 3P_t\right) + \alpha_t \mu_t^b + \beta_t P_t^b\right] dt + \tilde{w}_t P_t^a dY_t, \tag{35}$$

$$dP_t^b = \left[-6P_t \left(\mu_t^2 + P_t\right) + A_t \mu_t^b + B_t P_t^b\right] dt, \tag{36}$$

$$d\mu_t^\sigma = \left[\alpha_t \mu_t^\sigma + \beta_t P_t^\sigma\right] dt + \tilde{w}_t P_t^\sigma dY_t, \tag{37}$$

$$dP_t^\sigma = \left[2\tilde{\sigma}_t + A_t \mu_t^\sigma + B_t P_t^\sigma\right] dt, \tag{38}$$

$$d\mu_t^w = \left[-2\tilde{w}_t \mu_t P_t + \alpha_t \mu_t^w + \beta_t P_t^w\right] dt + \left[P_t + \tilde{w}_t P_t^w\right] dY_t, \tag{39}$$

$$dP_t^w = \left[-2\tilde{w}_t P_t^2 + A_t \mu_t^w + B_t P_t^w\right] dt, \tag{40}$$

$$\mu_0^a = \mu_0^b = \mu_0^\sigma = \mu_0^w = 0, \tag{41}$$

$$P_0^a = \frac{\partial}{\partial \tilde{a}_0} \Gamma\left(\tilde{a}_0, \tilde{b}_0, \tilde{\sigma}_0\right), \quad P_0^b = \frac{\partial}{\partial \tilde{b}_0} \Gamma\left(\tilde{a}_0, \tilde{b}_0, \tilde{\sigma}_0\right), \tag{42}$$

$$P_0^\sigma = \frac{\partial}{\partial \tilde{\sigma}_0} \Gamma\left(\tilde{a}_0, \tilde{b}_0, \tilde{\sigma}_0\right), \quad P_0^w = 0, \tag{43}$$

where we introduced the following auxiliary processes

$$\alpha_t = \tilde{a}_t - \tilde{w}_t^2 P_t - 3\tilde{b}_t \left(\mu_t^2 + P_t\right), \qquad \beta_t = -\left(\tilde{w}_t^2 + 3\tilde{b}_t\right) \mu_t, \tag{44}$$

$$A_t = -12\tilde{b}_t \mu_t P_t, \qquad B_t = 2\tilde{a}_t - 2\tilde{w}_t^2 P_t - 6\tilde{b}_t \left(\mu_t^2 + 2P_t\right). \tag{45}$$

We numerically tested the learning algorithm for this nonlinear model by simulating a system with $a = 4$, $b = 3$, $\sigma = 1$ and $w = 2$, 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 (Doucet et al., 2000).

4. Discussion

We revisited the problem of online parameter estimation in a partially observed diffusion process. Using a change of measure, we were able to express the log likelihood function of the observed data in terms of the filter for the hidden state. Then, using a stochastic gradient ascent on the log likelihood function, we derived learning rules for the parameters that we then tested numerically.

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. In an offline setting, the method of choice for most problems is the Expectation Maximization (EM) algorithm. The expectation step in EM is performed over the conditional distribution of all latent states given the observations, and therefore requires smoothing or path estimation. This dependence on smoothing seems to be common to many other approaches, see also Sutter et al. (2015), Movellan et al. (2002), Part II of Cappé et al. (2006). Since smoothing, in contrast to filtering, requires information from the future, offline algorithms that rely on smoothing usually cannot be easily transferred to an online setting.

We note that 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 less widely known than it deserves to be. To the best of our knowledge, the only appearance is in the technical report by Moura and Mitter (1986). We found it appropriate to revisit this approach and to bring it to the attention of the community as a fruitful way of deriving workable online learning rules.

Our derivation of parameter learning rules is tractable in all cases where h is differentiable with respect to the parameters and where the filter is either exactly solvable or a finite-dimensional approximation is known. Moreover, as we showed in two simple examples, the algorithm is capable of improving filter performance when the system parameters are unknown. This holds even if the models are unidentifiable (i.e. have parameter redundancies that make different parameters produce the same statistics of the observed process).

Even in the cases where the model is identifiable, the log likelihood function need not be globally concave, as we saw in the linear model (Fig. 3). It is therefore not easy to provide general convergence results using tools from convex optimization, and it is an outstanding problem to find conditions for global concavity of the log likelihood function and more generally, for convergence of the parameter estimates.

The second numerical example showed that the parameter estimates do not need to converge to the ground truth, and 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.

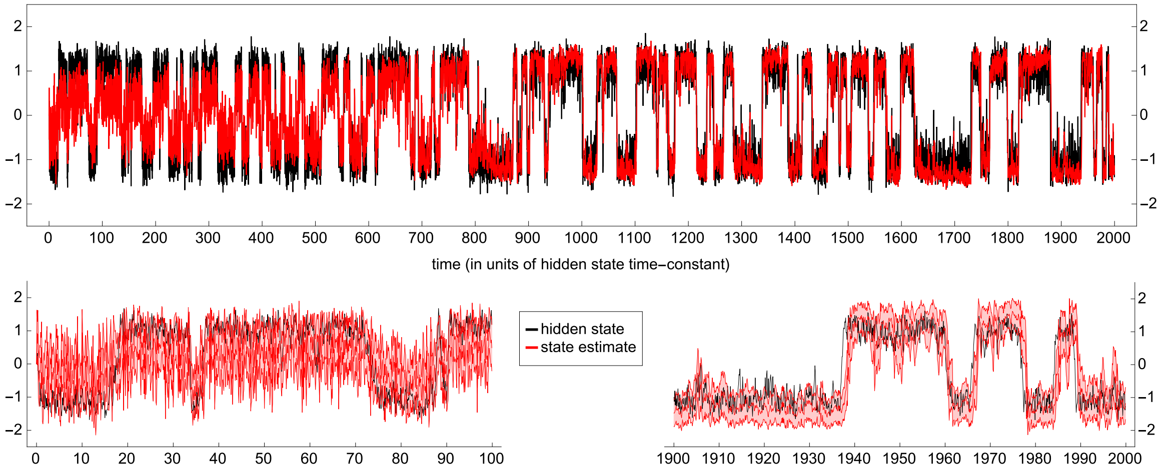

Figure 4: **Online learning and filtering in the nonlinear model**. The hidden state X_t (black) and mean μ_t of the projection filter are shown for the bimodal model of Section 3.2 with parameters $a = 4$, $b = 3$, $\sigma = 1$ and $w = 2$, $\tilde{a}_0 = 1$, $\tilde{b}_0 = 2$, $\tilde{\sigma}_0 = 3$, $\tilde{w}_0 = 4$, $\eta_a = \eta_b = \eta_w = 10^{-1}$ and $\eta_\sigma = 0.04$. 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.

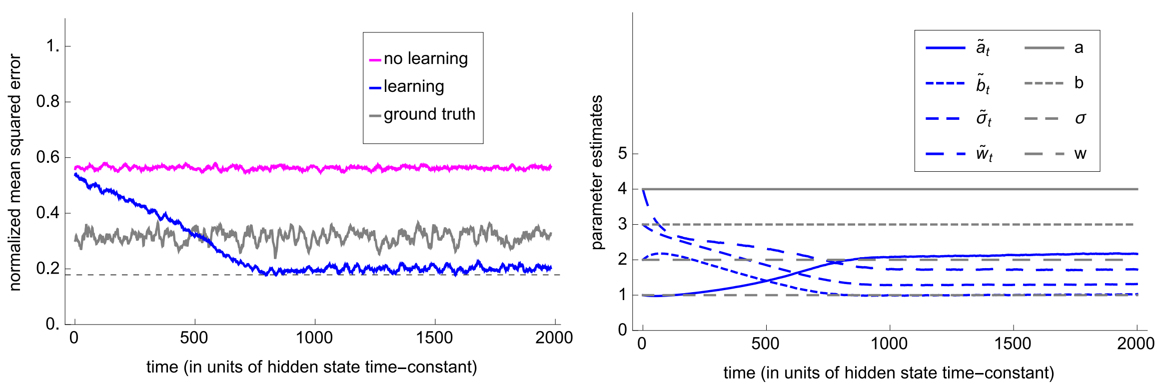

Figure 5: **Online learning and filtering in the nonlinear model**. The time evolution of the MSE and parameter estimates are shown for the linear model of Section 3.1 (see Fig. 1 caption for details). Left: the moving average of the normalized MSE (time window of 20 seconds) shows how the learning algorithm 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).

References

Alan Bain and Dan Crişan. *Fundamentals of stochastic filtering*, volume 60 of *Stochastic Modelling and Applied Probability*. Springer, New York, 2009.

Damiano Brigo, Bernard Hanzon, and François Le Gland. Approximate nonlinear filtering by projection on exponential manifolds of densities. *Bernoulli*, 5(3):495–534, 1999.

Olivier Cappé, Eric Moulines, and Tobias Ryden. *Inference in Hidden Markov Models*. Springer Science & Business Media, April 2006.

Arnaud Doucet, S Godsill, and Christophe Andrieu. On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing*, 10(3):197–208, July 2000.

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

José M F Moura and Sanjoy K Mitter. Identification and Filtering–optimal Recursive Maximum Likelihood Approach. Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, 1986.

Javier R Movellan, P Mineiro, and R J Williams. A Monte Carlo EM approach for partially observable diffusion processes: Theory and applications to neural networks. *Neural Computation*, 14(7): 1507–1544, July 2002.

T Sutter, Arnab Ganguly, and H Koeppl. A variational approach to path estimation and parameter inference of hidden diffusion processes. *arXiv.org*, 2015.