

# *Long-time stability and accuracy of the ensemble Kalman--Bucy filter for fully observed processes and small measurement noise*

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

de Wiljes, J., Reich, S. and Stannat, W. (2018) Long-time stability and accuracy of the ensemble Kalman--Bucy filter for fully observed processes and small measurement noise. SIAM Journal on Applied Dynamical Systems, 17 (2). pp. 1152-1181. ISSN 1536-0040 doi: 10.1137/17M1119056 Available at <https://centaur.reading.ac.uk/78674/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

Published version at: <http://dx.doi.org/10.1137/17M1119056>

To link to this article DOI: <http://dx.doi.org/10.1137/17M1119056>

Publisher: Society for Industrial and Applied Mathematics

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

## **CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

## Long-Time Stability and Accuracy of the Ensemble Kalman–Bucy Filter for Fully Observed Processes and Small Measurement Noise\*

Jana de Wiljes<sup>†</sup>, Sebastian Reich<sup>‡</sup>, and Wilhelm Stannat<sup>§</sup>

**Abstract.** The ensemble Kalman filter has become a popular data assimilation technique in the geosciences. However, little is known theoretically about its long term stability and accuracy. In this paper, we investigate the behavior of an ensemble Kalman–Bucy filter applied to continuous-time filtering problems. We derive mean field limiting equations as the ensemble size goes to infinity as well as uniform-in-time accuracy and stability results for finite ensemble sizes. The later results require that the process is fully observed and that the measurement noise is small. We also demonstrate that our ensemble Kalman–Bucy filter is consistent with the classic Kalman–Bucy filter for linear systems and Gaussian processes. We finally verify our theoretical findings for the Lorenz-63 system.

**Key words.** data assimilation, Kalman–Bucy filter, ensemble Kalman filter, stability, accuracy, asymptotic behavior

**AMS subject classifications.** 65C05, 62M20, 93E11, 62F15, 86A22

**DOI.** 10.1137/17M1119056

**1. Introduction.** In this paper, we consider the continuous-time filtering problem [Jaz70, BC08] for diffusion processes of type

$$(1) \quad dX_t = f(X_t)dt + \sqrt{2}CdB_t$$

and observations,  $Y_t$ , given by

$$(2) \quad dY_t = h(X_t)dt + R^{1/2}dB_t.$$

Here  $X_t$  denotes the state variable of the  $N_x$ -dimensional diffusion process with Lipschitz-continuous drift  $f : \mathbb{R}^{N_x} \rightarrow \mathbb{R}^{N_x}$  and constant diffusion tensors  $D = CC^T$  and  $C \in \mathbb{R}^{N_x \times N_w}$ . The observations  $Y_t$  are  $N_y$ -dimensional with forward map  $h : \mathbb{R}^{N_x} \rightarrow \mathbb{R}^{N_y}$  and measurement error covariance matrix  $R \in \mathbb{R}^{N_y \times N_y}$ . Finally,  $W_t \in \mathbb{R}^{N_w}$  and  $B_t \in \mathbb{R}^{N_y}$  denote independent

\*Received by the editors March 1, 2017; accepted for publication (in revised form) by E. Kostelich February 20, 2018; published electronically April 17, 2018.

<http://www.siam.org/journals/siads/17-2/M111905.html>

**Funding:** This research has been partially funded by Deutsche Forschungsgemeinschaft (DFG) through grant CRC 1294 “Data Assimilation”, Project (A02) “Long-time stability and accuracy of ensemble transform filter algorithms”.

<sup>†</sup>Universität Potsdam, Institut für Mathematik, Karl-Liebknecht-Str. 24/25, D-14476 Potsdam, Germany ([wiljes@uni-potsdam.de](mailto:wiljes@uni-potsdam.de)).

<sup>‡</sup>Universität Potsdam, Institut für Mathematik, Karl-Liebknecht-Str. 24/25, D-14476 Potsdam, Germany, and University of Reading, Department of Mathematics and Statistics, Whiteknights, PO Box 220, Reading RG6 6AX, UK ([sreich@math.uni-potsdam.de](mailto:sreich@math.uni-potsdam.de)).

<sup>§</sup>TU Berlin, Institut für Mathematik, Str. des 17. Juni 136, D-10623 Berlin, Germany ([stannat@math.tu-berlin.de](mailto:stannat@math.tu-berlin.de)).

Brownian motion of dimensions  $N_w$  and  $N_y$ , respectively. It is well known that the filtering distribution  $\pi_t$ , i.e., the conditional distribution in  $X_t$  for given observations  $Y_s$ ,  $s \in [0, t]$ , satisfies the Kushner–Zakai equation [Jaz70, BC08], which we write as an evolution equation in the expectation values

$$(3) \quad \pi_t[g] = \int_{\mathbb{R}^{N_x}} g(x) \pi_t(x) dx$$

of smooth and bounded functions  $g : \mathbb{R}^{N_x} \rightarrow \mathbb{R}$ , i.e.,

$$(4) \quad d\pi_t[g] = \pi_t[f \cdot \nabla g]dt + \pi_t[\nabla \cdot D\nabla g]dt + (\pi_t[gh] - \pi_t[g]\pi_t[h])^T R^{-1} (dY_t - \pi_t[h]dt).$$

In order to have a properly formulated filtering problem, we also have to specify the distribution at initial time  $t = 0$ .

Popular numerical methods for approximating solutions to (4) include direct finite-difference or finite-element discretizations of (4) and sequential Monte Carlo methods, also called particle filters [BC08, DdFe01]. These methods lead to consistent approximations but are typically restricted to low-dimensional problems. In recent years, particle filter methods have become popular, which are applicable to higher-dimensional problems but are no longer consistent. These include the ensemble Kalman filter (EnKF) [Eve06, LSZ15, RC15], which is now widely used in the geosciences.

Abstractly spoken, particle filters are defined as follows. First, one defines  $M$  weighted random variables  $X_t^i$ , called particles, which are i.i.d. at initial time  $t = 0$  with distribution  $\pi_0$ , and weights  $w_t^i \geq 0$  with  $w_0^i = 1/M$  at initial time. A particle filter is then characterized by appropriate evolution laws for the particles and the weights. Most known particle filters lead to particles which remain identically distributed while no longer being independent, so called interacting particle systems [Mor13]. If the weights are furthermore kept uniform either through resampling or appropriately defined evolution equations, then expectation can be taken with respect to the law  $\pi_t^M$  of the  $M$ th particle and consistency of a particle filter means that  $\lim_{M \rightarrow \infty} \pi_t^M[g] \rightarrow \pi_t[g]$ .

The classic bootstrap filter [AMGC02] uses (1) for the evolution of the particles and (2) for the update of the weights in combination with an appropriate resampling strategy in order to avoid the weights to degenerate. The EnKF, on the contrary, introduces modified evolution equations for the particles which include the observations and keep the weights uniform instead. Most available EnKF formulations are stated for discrete-in-time observations [Eve06]. While the robust behavior of EnKFs has been demonstrated for many applications primarily arising from the geosciences, our theoretical understanding of their long-time stability and accuracy is still rather limited. Large sample size limits have been, for example, investigated in [GMT11, KM15] and it has been demonstrated that the EnKF converges to the classic Kalman filter for linear systems (1), linear observations (2), and Gaussian initial conditions. Using concepts from shadowing, [GTH13] showed that the EnKF is stable and accurate uniformly in time for hyperbolic dynamical systems provided the ensemble size is larger than the dimension of the chaotic attractor. Stability and ergodicity of EnKFs have also been studied in [TMK16]. The authors demonstrate that the extended system consisting of (1), (2), and the filter algorithm possesses a unique ergodic invariant measure provided the existence of an

appropriate Lyapunov function can be guaranteed. While such ergodicity results of [MH12] are important, they do not imply the accuracy of a filter. In fact, it is well known that EnKF can diverge and techniques, such as ensemble inflation [Eve06], have been developed in order to stabilize a filter. Furthermore, it has been rigorously demonstrated, for example, in [KLS14] that a judicious choice of inflation can lead to uniform-in-time accurate state estimates. At the same time, [KMT15] provides an example of catastrophic filter divergence, i.e., an exponential blow-up of the ensemble systems, for a linear forward map  $h(x) = Hx$  with strongly nonnormal operator  $H$ .

In this paper, we investigate a time-continuous EnKF formulation which is consistent with the classic Kalman filter in the linear case and which is also stable and accurate uniformly in time without additional ensemble inflation. In this first study, we will assume for simplicity that the system is fully observable, i.e.,  $h(x) = x$  in (2), and that the measurement errors are small. These assumptions can be relaxed under appropriate assumptions on the stochastic process (1) and the observation process (2), well known from the theory of classic Kalman filter theory (i.e., observability and controllability) [Jaz70]. We will also investigate in future work whether the proposed filter formulations can prevent catastrophic filter divergence for strongly nonlinear and partially observed systems.

The specific ensemble Kalman–Bucy filter (EnKBF) formulation, which we will investigate in this paper, is given by drawing  $M$  independent realizations (called particles or ensemble members)  $X_0^i \sim \pi_0$ , which then follow the system of differential equations

$$(5) \quad dX_t^i = f(X_t^i)dt + D(P_t^M)^{-1}(X_t^i - \bar{x}_t^M)dt - \frac{1}{2}Q_t^M R^{-1} (h(X_t^i)dt + \bar{h}_t^M dt - 2dY_t)$$

for  $t \geq 0$ . These equations of motion for the particles are closed through the empirical estimates

$$(6) \quad \bar{x}_t^M = \frac{1}{M} \sum_{i=1}^M X_t^i, \quad P_t^M = \frac{1}{M-1} \sum_{i=1}^M (X_t^i - \bar{x}_t^M)(X_t^i - \bar{x}_t^M)^T$$

and

$$(7) \quad \bar{h}_t^M = \frac{1}{M} \sum_{i=1}^M h(X_t^i), \quad Q_t^M = \frac{1}{M-1} \sum_{i=1}^M (X_t^i - \bar{x}_t^M)(h(X_t^i) - \bar{h}_t^M)^T.$$

Finally, given a solution of (5), we define the empirical expectation values of a function  $g$  and the empirical distribution  $\hat{\pi}_t^M$  by

$$(8) \quad \bar{g}_t^M := \hat{\pi}_t^M[g], \quad \hat{\pi}_t^M(x) := \frac{1}{M} \sum_{i=1}^M \delta(x - X_t^i),$$

respectively. Here  $\delta(\cdot)$  denotes the standard Dirac delta function. The formulation (5) has been stated first in [BR10, BR12]. Alternative ensemble Kalman–Bucy formulations include stochastically perturbed formulations [Rei11, LSZ15, RC15] and the extended ensemble Kalman–Bucy filter, whose exponential stability and propagation of chaos properties have been studied in [DMKT16].

In case  $P_t^M$  is not invertible, which is surely the case for  $M \leq N_x$ , the inverse of  $P_t^M$  is replaced by its generalized inverse  $(P_t^M)^+$ . This generalization is unproblematic from a mathematical perspective since  $(P_t^M)^+$  gets multiplied by a vector which is in the range of  $P_t^M$ , and we show that the equations are well-posed in section 2. At the same time it is known that  $M \ll N_x$  often requires application of localization [Eve06, RC15] in order to obtain a full rank approximation of the covariance matrix and to prevent filter divergence. The impact of localization has been studied in [Ton17] from a rigorous mathematical perspective for high-dimensional linear systems.

Given the evolution equations (5), one can derive associated evolution equations for the ensemble mean,  $\bar{x}_t^M$ , and the ensemble covariance matrix,  $P_t^M$ . These are given by

$$(9) \quad d\bar{x}_t^M = \bar{f}_t^M dt - Q_t^M R^{-1}(\bar{h}_t^M dt - dY_t)$$

with  $\bar{f}_t^M = \hat{\pi}_t^M[f]$  and

$$(10) \quad \begin{aligned} \frac{d}{dt} P_t^M &= \frac{1}{M-1} \sum_{i=1}^M \{ (f(X_t^i) - \bar{f}_t)(X_t^i - \bar{x}_t)^T + (X_t^i - \bar{x}_t)(f(X_t^i) - \bar{f}_t)^T \} \\ &\quad + \{ D(P_t^M)^+ P_t^M + P_t^M (P_t^M)^+ D \} - Q_t^M R^{-1} (Q_t^M)^T. \end{aligned}$$

We will study the behavior of the EnKBF for fully observed processes, i.e.,  $h(x) = x$  and regular measurement error covariance matrix  $R$  in sections 2 and 3. More specifically, it is shown in section 2 that strong solutions of (5) exist for all times and are unique. This result implies that catastrophic filter divergence [KMT15] cannot arise under the setting considered in this paper. Next, uniform-in-time stability and accuracy of (5) are proven in section 3 under the additional assumption that  $R = \epsilon I$ ,  $\epsilon > 0$ , sufficiently small, and that  $M > N_x$ , i.e., the empirical covariance matrix  $P_t^M$  is invertible. In sections 4 and 5, we return to the filtering problem for general observation operator,  $h$ , and measurement error covariance matrix  $R$ . It is demonstrated in section 4 that in the case of linear systems, (9) and (10) are consistent with the classic Kalman–Bucy filtering equations [Jaz70].

Note that this does not imply that the empirical distribution of the extended EnKBF is asymptotically normal. In fact, we will identify in section 5 its asymptotic distribution for  $M \rightarrow \infty$ . To this end we will prove in Theorem 16 that the ensemble  $X_t^i$ ,  $1 \leq i \leq M$ , converges as  $M \rightarrow \infty$  to independent solutions  $\hat{X}_t^i$ ,  $i = 1, 2, 3, \dots$ , of the following McKean–Vlasov equation:

$$(11) \quad d\hat{X}_t = f(\hat{X}_t)dt + D(\mathcal{P}_t)^{-1}(\hat{X}_t - \bar{x}_t)dt - \frac{1}{2}Q_t R^{-1} \left( h(\hat{X}_t)dt + \bar{h}_t dt - 2dY_t \right),$$

with  $\bar{x}_t = \hat{\pi}_t[x]$ ,  $\bar{h}_t = \hat{\pi}_t[h]$ ,

$$(12) \quad \mathcal{P}_t = \text{Cov}(\hat{X}_t, \hat{X}_t), \quad Q_t = \text{Cov}(\hat{X}_t, h(\hat{X}_t)).$$

Here  $\hat{\pi}_t$  denotes the distribution of  $\hat{X}_t$ .

Using Itô's formula, it is then easy to derive from (11) the weak formulation of the nonlinear Fokker–Planck equation driving the distribution  $\hat{\pi}_t$  of  $\hat{X}_t$

$$(13) \quad \begin{aligned} d\hat{\pi}_t[g] = \hat{\pi}_t \left[ \nabla g \cdot \left\{ f dt + DP_t^{-1}(x - \hat{\pi}_t[x])dt - \frac{1}{2} Q_t R^{-1}(h(x)dt + \hat{\pi}_t[h]dt - 2dY_t) \right\} \right] \\ + \hat{\pi}_t \left[ \frac{1}{2} \nabla \cdot Q_t R^{-1} Q_t^T \nabla g dt \right]. \end{aligned}$$

Note the difference between (13) and the Kushner–Zakai equation (4).

Some numerical results, supporting our theoretical estimates, will be presented in section 6 using a stochastically perturbed Lorenz-63 system [Lor63, LSZ15].

## 2. Well-posedness of the ensemble Kalman–Bucy filter for fully observed processes.

In this section, we specify the problem setting which is investigated in detail in this paper. We will also derive a first well-posedness result for the system (5)–(7) implying that the filter is not subject to catastrophic filter divergence. More specifically, we assume that the process is fully observed, i.e.,  $h(x) = x$ , that the diffusion tensor  $D$  has full rank, and that the drift function  $f$  is globally Lipschitz continuous. Since the ensemble size,  $M$ , will be fixed in this section, we also drop the superscript  $M$  in (5). Hence (5) is replaced by

$$(14) \quad dX_t^i = f(X_t^i)dt + DP_t^+(X_t^i - \bar{x}_t)dt - \frac{1}{2} P_t R^{-1} (X_t^i dt + \bar{x}_t dt - 2dY_t),$$

$i = 1, \dots, M$ . The standard inner product in  $R^{N_x}$  will be denoted by  $\langle \cdot, \cdot \rangle$  and we recall that

$$(15) \quad \langle a, b \rangle = \text{tr}(ba^T).$$

Hence we quickly verify that

$$(16) \quad \frac{1}{M-1} \sum_{i=1}^M \langle X_t^i - \bar{x}_t, DP_t^+(X_t^i - \bar{x}_t) \rangle = \text{tr}(DP_t^+ P_t)$$

and

$$(17) \quad \frac{1}{M-1} \sum_{i=1}^M \langle X_t^i - \bar{x}_t, P_t R^{-1} (X_t^i - \bar{x}_t) \rangle = \text{tr}(P_t R^{-1} P_t) = \|R^{-\frac{1}{2}} P_t\|_F^2.$$

Here  $\|A\|_F$  denotes the Frobenius norm of a matrix  $A$ . We also introduce the notation  $\langle A, B \rangle = \text{tr}(BA^T)$ , i.e.,  $\|A\|_F^2 = \langle A, A \rangle$ .

We now investigate the  $l_2$ -norm of the ensemble deviations from the mean, i.e.,

$$(18) \quad V_t = \frac{1}{M-1} \sum_{i=1}^M \|X_t^i - \bar{x}_t\|^2,$$

which satisfies the evolution equation

$$\begin{aligned}
 \frac{1}{2} \frac{dV_t}{dt} &= \frac{1}{M-1} \sum_{i=1}^M \langle X_t^i - \bar{x}_t, f(X_t^i) - \bar{f}_t \rangle + \frac{1}{M-1} \sum_{i=1}^M \langle X_t^i - \bar{x}_t, DP_t^+(X_t^i - \bar{x}_t) \rangle \\
 &\quad - \frac{1}{2\varepsilon} \frac{1}{M-1} \sum_{i=1}^M \langle X_t^i - \bar{x}_t, P_t R^{-1} (X_t^i - \bar{x}_t) \rangle \\
 (19) \quad &= \frac{1}{M-1} \sum_{i=1}^M \langle X_t^i - \bar{x}_t, f(X_t^i) - \bar{f}_t \rangle + \text{tr}(DP_t^+ P_t) - \frac{1}{2} \|R^{-\frac{1}{2}} P_t\|_F^2 \\
 &= \frac{1}{M-1} \sum_{i=1}^M \langle X_t^i - \bar{x}_t, f(X_t^i) - f(\bar{x}_t) \rangle + \text{tr}(DP_t^+ P_t) - \frac{1}{2} \|R^{-\frac{1}{2}} P_t\|_F^2.
 \end{aligned}$$

Here we have used

$$(20) \quad \sum_i \langle X_t^i - \bar{x}_t, f(\bar{x}_t) - \bar{f}_t \rangle = 0$$

and that the evolution equation (9) for the mean,  $\bar{x}_t$ , reduces to

$$(21) \quad d\bar{x}_t = \bar{f}_t dt - P_t R^{-1} (\bar{x}_t dt - dY_t)$$

in our setting.

**Lemma 1.** *The Frobenius norm of  $P_t$  satisfies*

$$(22) \quad \frac{1}{\sqrt{M}} V_t \leq \|P_t\|_F \leq V_t.$$

*Proof.* We first note the following identity:

$$(23) \quad \|P_t\|_F^2 = \frac{1}{(M-1)^2} \sum_{i,j} \langle X_t^i - \bar{x}_t, X_t^j - \bar{x}_t \rangle^2.$$

For the proof of the upper bound it is now sufficient to observe that

$$\begin{aligned}
 \frac{1}{(M-1)^2} \sum_{i,j} \langle X_t^i - \bar{x}_t, X_t^j - \bar{x}_t \rangle^2 &\leq \frac{1}{(M-1)^2} \sum_{i,j} \|X_t^i - \bar{x}_t\|^2 \|X_t^j - \bar{x}_t\|^2 \\
 (24) \quad &= \left( \frac{1}{M-1} \sum_i \|X_t^i - \bar{x}_t\|^2 \right)^2.
 \end{aligned}$$

For the proof of the lower bound observe that

$$\begin{aligned}
 \frac{1}{(M-1)^2} \sum_{i,j} \langle X_t^i - \bar{x}_t, X_t^j - \bar{x}_t \rangle^2 &\geq \frac{1}{(M-1)^2} \sum_i \|X_t^i - \bar{x}_t\|^4 \\
 (25) \quad &\geq \frac{1}{M} \left( \frac{1}{M-1} \sum_i \|X_t^i - \bar{x}_t\|^2 \right)^2.
 \end{aligned}$$

This completes the proof. ■



**Remark 2.** We recall the standard relations between the Frobenius and the spectral norm of a matrix  $A$ , i.e.,

$$(26) \quad \|A\| \leq \|A\|_F$$

and

$$(27) \quad \|A\|_F \leq \sqrt{N_x} \|A\|.$$

We are now ready to obtain uniform-in-time upper and lower bounds on  $V_t$ . First, we can estimate the first term of (19) from above and from below as follows:

$$(28) \quad \frac{1}{M-1} \sum_i \langle X_t^i - \bar{x}_t, f(X_t^i) - f(\bar{x}_t) \rangle \leq L_+ \frac{1}{M-1} \sum_i \|X_t^i - \bar{x}_t\|^2 = L_+ V_t$$

and

$$(29) \quad \frac{1}{M-1} \sum_i \langle X_t^i - \bar{x}_t, f(X_t^i) - f(\bar{x}_t) \rangle \geq L_- \frac{1}{M-1} \sum_i \|X_t^i - \bar{x}_t\|^2 = L_- V_t,$$

respectively, where

$$(30) \quad L_+ := \sup_{x \neq y} \frac{\langle f(x) - f(y), x - y \rangle}{\|x - y\|^2} \quad L_- := \inf_{x \neq y} \frac{\langle f(x) - f(y), x - y \rangle}{\|x - y\|^2}$$

are the upper and lower control on the “dissipativity” constant of  $f$ . We clearly have  $L_+ \leq \|f\|_{\text{Lip}}$  and  $L_- \geq -\|f\|_{\text{Lip}}$  for globally Lipschitz continuous  $f$  with Lipschitz constant  $\|f\|_{\text{Lip}}$ . Provided  $V_t \neq 0$ , we also find that

$$(31) \quad \lambda^{\min}(D) \leq \text{tr}(DP_t^+ P_t) \leq \text{tr}(D).$$

Here,  $\lambda^{\min}(A)$  and  $\lambda^{\max}(A)$  denote the smallest and largest singular values of a matrix  $A$ , respectively.

Finally, the third term in (19) can be estimated from above and from below using

$$(32) \quad \lambda^{\min}(R^{-1}) \|P_t\|_F^2 = \text{tr}(P_t \lambda^{\min}(R^{-1}) P_t) \leq \text{tr}(P_t R^{-1} P_t) = \|R^{-\frac{1}{2}} P_t\|_F^2$$

and

$$(33) \quad \|R^{-\frac{1}{2}} P_t\|_F^2 = \text{tr}(P_t R^{-1} P_t) \leq \text{tr}(P_t \lambda^{\max}(R^{-1}) P_t) = \lambda^{\max}(R^{-1}) \text{tr}(P_t^2) = (\lambda^{\min}(R))^{-1} \|P_t\|_F^2,$$

which follow from the inequalities  $P_t \lambda^{\min}(R^{-1}) P_t \leq P_t R^{-1} P_t \leq P_t \lambda^{\max}(R^{-1}) P_t$ , where  $\leq$  is meant in the sense of (symmetric) positive (semi-) definite matrices.

Inserting these estimates and the previous two identities into (19), we first obtain the upper bound

$$(34) \quad \frac{1}{2} \frac{dV_t}{dt} \leq L_+ V_t + \text{tr}(D) - \frac{\lambda^{\min}(R^{-1})}{2M} V_t^2.$$

This implies that  $V_t \leq \max\{V_0, \lambda^{\max}(R)L_+M + \sqrt{(\lambda^{\max}(R)ML_+)^2 + 2\lambda^{\max}(R)M \operatorname{tr}(D)}\}$  uniformly in  $t$ . Similarly, we obtain the lower bound

$$(35) \quad \frac{1}{2} \frac{dV_t}{dt} \geq L_- V_t + \lambda^{\min}(D) - \frac{\lambda^{\max}(R^{-1})}{2} V_t^2,$$

which implies that  $V_t \geq \min\{V_0, \lambda^{\min}(R)L_- + \sqrt{(\lambda^{\min}(R)L_-)^2 + 2\lambda^{\min}(R)\lambda^{\min}(D)}\}$  uniformly in  $t$  and  $V_t > 0$  provided  $V_0 > 0$ .

**Theorem 3.** Assume that the drift term  $f$  in (1) is globally Lipschitz continuous and satisfies a linear growth condition

$$(36) \quad \|f(x)\| \leq \tilde{c}_1(1 + \|x\|)$$

for an appropriate  $\tilde{c}_1 > 0$ . Then the system (14) together with (6)–(7) possesses strong solutions for all times  $t \geq 0$ .

*Proof.* We can decompose the equations (14) into ordinary differential equations in  $X_t^i - \bar{x}_t$ ,  $i = 1, \dots, M$  and (21) for the mean,  $\bar{x}_t$ . Since the  $l_2$ -norm,  $V_t$ , remains bounded, the equations in  $X_t^i - \bar{x}_t$  are well-posed. Furthermore, since  $\|P_t\|$  remains bounded as well, the combined drift term in (21), written as

$$(37) \quad d\bar{x}_t = f(\bar{x}_t)dt + b_t dt - P_t R^{-1}(\bar{x}_t dt - dY_t),$$

with  $b_t = \bar{f}_t - f(\bar{x}_t)$ , is Lipschitz continuous in  $\bar{x}_t$  and, hence, satisfies a linear growth condition, i.e.,

$$(38) \quad \|f(\bar{x}_t) + b_t - P_t R^{-1} \bar{x}_t\| \leq \|f(\bar{x}_t) - P_t R^{-1} \bar{x}_t\| + \|\bar{f}_t - f(\bar{x}_t)\| \leq \tilde{c}_2(1 + \|\bar{x}_t\|)$$

for an appropriate  $\tilde{c}_2 > 0$ , and, hence, strong solutions to (21) exist for all times [Øks00]. ■

**Remark 4.** For the analysis of the asymptotic behavior of  $M \rightarrow \infty$  the upper bound on  $V_t$  is not sufficient, because it diverges as  $M \rightarrow \infty$ . However, since we need a control only locally in time, we can use (34) to estimate  $\frac{1}{2} \frac{d}{dt} V_t \leq L_+ V_t + \operatorname{tr}(D)$  which implies the upper bound

$$(39) \quad V_t \leq e^{2L_+ t} \left( V_0 + \frac{\operatorname{tr}(D)}{L_+} \right),$$

which becomes uniform in  $M$  (but, of course, not in  $t$ ) if the particles at time  $t = 0$  are chosen with uniform upper bound on  $V_0 = V_0^M$ .

**3. Accuracy of the ensemble Kalman–Bucy filter for finite ensemble sizes and small measurement noise.** The goal of this section is to derive bounds on the estimation error

$$(40) \quad e_t = X_t^{\text{ref}} - \bar{x}_t,$$

where  $X_t^{\text{ref}}$  denotes the reference trajectory of (1) which generated the data. We again restrict the discussion to fully observed processes and globally Lipschitz-continuous drift functions  $f$ .

In addition, we assume the error covariance to be of the type  $R = \varepsilon I$  with sufficiently small  $\varepsilon > 0$ , implying

$$(41) \quad dY_t = X_t^{\text{ref}} dt + \sqrt{\varepsilon} dB_t,$$

and that  $P_t^M$  is invertible which necessitates that  $M > N_x$ . We drop the superscript  $M$  from all relevant quantities throughout this section, as we are interested in the accuracy of the filter for fixed ensemble size,  $M$ .

We find that the estimation error satisfies the evolution equation

$$(42) \quad de_t = f(X_t^{\text{ref}})dt + \sqrt{2}C dW_t - \bar{f}_t dt - \frac{1}{\varepsilon} P_t (e_t dt + \varepsilon^{1/2} dB_t).$$

We introduce the squared estimation error norm  $E_t = \|e_t\|^2/2 = \langle e_t, e_t \rangle/2$ . Then Ito's formula implies that

$$(43) \quad \begin{aligned} dE_t = & \langle f(X_t^{\text{ref}}) - \bar{f}_t, X_t^{\text{ref}} - \bar{x}_t \rangle dt - \frac{1}{\varepsilon} \langle e_t, P_t e_t \rangle dt \\ & + \langle e_t, \sqrt{2}C dW_t \rangle - \frac{1}{\sqrt{\varepsilon}} \langle e_t, P_t dB_t \rangle + \text{tr}(D) dt + \frac{1}{2\varepsilon} \text{tr}(P_t^2) dt, \end{aligned}$$

which can be rewritten as

$$(44) \quad dE_t = \mathcal{E}_t dt + dM_t$$

with

$$(45) \quad \mathcal{E}_t = \langle X_t^{\text{ref}} - \bar{x}_t, f(X_t^{\text{ref}}) - \bar{f}_t \rangle - \frac{1}{\varepsilon} \langle e_t, P_t e_t \rangle + \text{tr}(D) + \frac{1}{2\varepsilon} \|P_t\|_F^2$$

and the martingale

$$M_t = \int_0^t \langle e_s, -\varepsilon^{-1/2} P_s dB_s + \sqrt{2}C dW_s \rangle, \quad t \geq 0.$$

To make further progress we need bounds for the smallest and largest singular values  $\lambda_t^{\min} = \lambda^{\min}(P_t)$  and  $\lambda_t^{\max} = \lambda^{\max}(P_t)$  of  $P_t$ , respectively. An upper bound for the largest singular value has already been derived in section 2, since  $\lambda_t^{\max} = \|P_t\| \leq \|P_t\|_F \leq V_t$ . Since  $P_t$  is assumed to be invertible, the explicit evolution equation for  $P_t$  reduces to

$$(46) \quad \frac{d}{dt} P_t = \frac{1}{M-1} \sum_i (f(X_t^i) - \bar{f}(t))(X_t^i - \bar{x}_t)^T + (X_t^i - \bar{x}_t)(f(X_t^i) - \bar{f}(t))^T + 2D - \frac{1}{\varepsilon} P_t^2.$$

Next, we make use of the fact that  $P_t$  can be diagonalized, i.e., there are orthogonal matrices  $Q_t$  and diagonal matrices  $\Lambda_t$  such that

$$(47) \quad P_t = Q_t^T \Lambda_t Q_t.$$

While the orthogonal matrices  $Q_t$  are, in general, only continuous in  $t$ , the diagonal matrix of singular values can be chosen to be differentiable in  $t$  [Rel69]. As shown in [DE99], the evolution equation for diagonal matrix of eigenvalues,  $\Lambda_t$ , is of the form

$$(48) \quad \frac{d\Lambda_t}{dt} = \text{diag}(Q_t U_t Q_t^T) + 2\text{diag}(Q_t D Q_t^T) - \frac{1}{\varepsilon} \Lambda_t^2$$

with

$$(49) \quad U_t := \frac{1}{M-1} \sum_i \{f(X_t^i) - f(\bar{x}_t)\} \{X_t^i - \bar{x}_t\}^T + \{X_t^i - \bar{x}_t\} \{f(X_t^i) - f(\bar{x}_t)\}^T.$$

Here  $\text{diag}(A)$  denotes a diagonal matrix with diagonal entries equal to the diagonal of  $A$ . More specifically, the diagonal entries of  $\text{diag}(Q_t U_t Q_t^T)$  are given by

$$(50) \quad (\text{diag}(Q_t U_t Q_t^T))_{ii} = e_i^T Q_t U_t Q_t^T e_i,$$

where  $e_i \in \mathbb{R}^{N_x}$  denotes the  $i$ th basis vector in  $\mathbb{R}^{N_x}$ .

Next, we derive the following estimate using the fact that  $f$  is globally Lipschitz continuous. Then, given any unit vector  $v$ , it holds that

$$(51) \quad \begin{aligned} & \left| \frac{1}{M-1} \sum_i \langle f(X_t^i) - f(\bar{x}_t), v \rangle \langle X_t^i - \bar{x}_t, v \rangle \right| \\ & \leq \left( \frac{1}{M-1} \sum_i \langle f(X_t^i) - f(\bar{x}_t), v \rangle^2 \right)^{\frac{1}{2}} \left( \frac{1}{M-1} \sum_i \langle X_t^i - \bar{x}_t, v \rangle^2 \right)^{\frac{1}{2}} \\ & \leq \|f\|_{\text{Lip}} V_t \leq \|f\|_{\text{Lip}} \sqrt{N_x M} \|P_t\|, \end{aligned}$$

where we have used  $V_t \leq \sqrt{N_x M} \|P_t\|$ .

Hence setting  $v = Q_t^T e_i$ , we obtain

$$(52) \quad |(\text{diag}(Q_t U_t Q_t^T))_{ii}| \leq 2\|f\|_{\text{Lip}} \sqrt{N_x M} \|P_t\| = 2\|f\|_{\text{Lip}} \sqrt{N_x M} \lambda_t^{\max}.$$

Since  $\lambda_t^{\max} = (\Lambda_t)_{ii}$  for some index  $i$ , we hence deduce that

$$(53) \quad \frac{d\lambda_t^{\max}}{dt} \leq 2\|f\|_{\text{Lip}} \sqrt{N_x M} \lambda_t^{\max} + 2\lambda^{\max}(D) - \frac{(\lambda_t^{\max})^2}{\varepsilon}.$$

This implies that

$$(54) \quad \lambda_t^{\max} \leq \max \left\{ \lambda_0^{\max}, \varepsilon \|f\|_{\text{Lip}} \sqrt{N_x M} + \sqrt{\varepsilon^2 \|f\|_{\text{Lip}}^2 N_x M + 2\varepsilon \lambda^{\max}(D)} \right\}.$$

Hence we have shown the following lemma.

**Lemma 5 (upper bound on spectral radius of  $P_t$ ).** *There is a constant*

$$(55) \quad C_1 = C_1(\|f\|_{\text{Lip}}, M, N_x, D, \varepsilon_0)$$

such that  $\lambda_0^{\max} \leq C_1 \varepsilon^{1/2}$  at initial time  $t = 0$  implies  $\lambda_t^{\max} \leq C_1 \varepsilon^{1/2}$  for all times and all  $\varepsilon \leq \varepsilon_0$ .

We now use our upper bound on  $\lambda_t^{\max} = \|P_t\|_2$  from Lemma 5 in order to get the estimate

$$(56) \quad |(\text{diag}(Q_t U_t Q_t^T))_{ii}| \leq 2L\sqrt{N_x M} C_1 \varepsilon^{1/2}.$$

Hence, we deduce that

$$(57) \quad \frac{d\lambda_t^{\min}}{dt} \geq -2\|f\|_{\text{Lip}}\sqrt{N_x M} C_1 \varepsilon^{1/2} + 2\lambda^{\min}(D) - \frac{(\lambda_t^{\min})^2}{\varepsilon}$$

and

$$(58) \quad \lambda_t^{\min} \geq \min\{\lambda_0^{\min}, \sqrt{2\varepsilon\lambda^{\min}(D) - 2\varepsilon^{\frac{3}{2}}\|f\|_{\text{Lip}}\sqrt{N_x M} C_1}\},$$

which implies the desired lower bound on  $\lambda_t^{\min}$ . Here  $\lambda^{\min}(D)$  denotes the smallest eigenvalue of  $D$ . We now fix  $\varepsilon_0 > 0$  such that

$$(59) \quad \lambda^{\min}(D) - \sqrt{\varepsilon}\|f\|_{\text{Lip}}\sqrt{N_x M} C_1 > 0.$$

**Lemma 6 (lower bound on smallest singular value of  $P_t$ ).** *There is a constant*

$$(60) \quad C_2 = C_2(\|f\|_{\text{Lip}}, M, N_x, D, \varepsilon_0)$$

such that  $\lambda_0^{\min} \geq C_2 \varepsilon^{1/2}$  at initial time  $t = 0$  implies  $\lambda_t^{\min} \geq C_2 \varepsilon^{1/2}$  for all  $t > 0$  and all  $\varepsilon \leq \varepsilon_0$ .

**Remark 7.** *The upper and lower bounds for the largest and smallest, respectively, eigenvalue of  $P_t$  depend on the ensemble size,  $M$ . This dependence can be eliminated for the price of the estimates no longer being valid uniformly in time. We now derive such  $M$ -independent upper and lower bounds. Let us assume that*

$$(61) \quad \|f\|_{\text{Lip}} V_s \leq \lambda^{\max}(D)$$

for all  $s \in [0, t]$ . Such a bound can be found because of (39) and for  $\varepsilon$  sufficiently small, i.e.,  $\varepsilon \leq \varepsilon_t$ . Then (53) implies that

$$(62) \quad \lambda_s^{\max} \leq 2(\lambda^{\max}(D)\varepsilon)^{1/2}$$

for all  $s \in [0, t]$  and all  $\varepsilon \leq \varepsilon_t$ . Similarly, (57) implies that

$$(63) \quad \lambda_s^{\min} \geq (2\lambda^{\min}(D)\varepsilon)^{1/2}.$$

Hence we have traded the  $M$ -dependent constants  $C_1$  and  $C_2$  in the previous two lemmas by  $M$ -independent constants  $\tilde{C}_1 = 2\lambda^{\max}(D)^{1/2}$  and  $\tilde{C}_2 = \lambda^{\min}(D)^{1/2}$ , respectively. However, the estimates hold for  $\varepsilon \leq \varepsilon_t$  only, where the upper bound  $\varepsilon_t = \varepsilon_t(\|f\|_{\text{Lip}}, D)$  decreases in time.

The upper and lower bounds of the eigenvalues of  $P_t$  obtained in the previous two lemmas hold with constants  $C_1$  and  $C_2$  independent of the driving Wiener processes. They only depend on the initial conditions (which might be random), but we can impose deterministic

bounds on the spectral radius of the covariance matrix. Hence we can take expectations on both sides of (44) in order to obtain the following integral inequality:

$$(64) \quad \mathbb{E}[E_t] \leq \mathbb{E}[E_0] + \int_0^t \mathbb{E}[\mathcal{E}_s] \, ds \leq \mathbb{E}[E_0] + \int_0^t \mathbb{E} \left[ \operatorname{tr}(D) + N_x C_1^2 + 2\varepsilon^{1/4} \|f\|_{\text{Lip}} C_1^{1/2} (N_x M)^{1/4} E_s^{1/2} - 2 \frac{C_2 - \varepsilon^{1/2} \|f\|_{\text{Lip}}}{\varepsilon^{1/2}} E_s \right] ds,$$

where we used

$$(65) \quad \begin{aligned} \langle f(X_t^{\text{ref}}) - \bar{f}_t, X_t^{\text{ref}} - \bar{x}_t \rangle &= \langle f(X_t^{\text{ref}}) - f(\bar{x}_t), X_t^{\text{ref}} - \bar{x}_t \rangle + \langle f(\bar{x}_t) - \bar{f}_t, X_t^{\text{ref}} - \bar{x}_t \rangle \\ &\leq 2L_+ E_t + \sqrt{2} \|f\|_{\text{Lip}} V_t^{1/2} E_t^{1/2} \\ &\leq 2\|f\|_{\text{Lip}} \left( E_t + \varepsilon^{1/4} C_1^{1/2} (MN_x)^{1/4} E_t^{1/2} \right). \end{aligned}$$

The next step is to close the right-hand side in  $\mathbb{E}[E_s]$ . To this end, we first derive the following  $\omega$ -wise estimate:

$$(66) \quad \begin{aligned} \mathcal{E}_s &\leq \left( \operatorname{tr}(D) + N_x C_1^2 + 2\varepsilon^{1/4} \|f\|_{\text{Lip}} C_1^{1/2} (N_x M)^{1/4} E_s^{1/2} - 2 \frac{C_2 - \varepsilon^{1/2} \|f\|_{\text{Lip}}}{\varepsilon^{1/2}} E_s \right) \\ &\leq C_3 + \varepsilon^{1/4} C_4 E_s^{1/2} - 2 \frac{C_2 - \varepsilon^{1/2} \|f\|_{\text{Lip}}}{\varepsilon^{1/2}} E_s \\ &\leq \left( C_3 + \varepsilon \frac{C_4^2}{C_2} \right) - \frac{C_2 - 2\varepsilon^{1/2} \|f\|_{\text{Lip}}}{\varepsilon^{1/2}} E_s \\ &=: \Phi(E_s) \end{aligned}$$

for  $C_3 = \operatorname{tr}(D) + N_x C_1^2$ ,  $C_4 = 2\|f\|_{\text{Lip}} C_1^{1/2} (N_x M)^{1/4}$ , and a linear function  $\Phi(E_s)$ . Taking expectations and using  $\mathbb{E}[\Phi(E_s)] = \Phi(\mathbb{E}[E_s])$  we arrive at the integral inequality

$$(67) \quad \mathbb{E}[E_t] \leq \mathbb{E}[E_0] + \int_0^t \Phi(\mathbb{E}[E_s]) \, ds,$$

and we can now apply the Gronwall lemma or comparison techniques for integral inequalities. More precisely, let  $\alpha = \varepsilon^{-1/2}(C_2 - 2\varepsilon^{1/2}\|f\|_{\text{Lip}}) > 0$ , then the time-dependent Ito's-formula implies that

$$(68) \quad e^{\alpha t} \mathbb{E}[E_t] \leq \mathbb{E}[E_0] + \int_0^t e^{\alpha s} \left( C_3 + \varepsilon \frac{C_4^2}{C_2} \right) ds$$

and, hence,

$$(69) \quad \mathbb{E}[E_t] \leq e^{-\alpha t} \mathbb{E}[E_0] + \alpha^{-1} K$$

with  $K := C_3 + \varepsilon \frac{C_4^2}{C_2}$ . Note that  $\alpha^{-1} = \mathcal{O}(\varepsilon^{\frac{1}{2}})$ . Hence we have shown the following theorem.

**Theorem 8 (estimation error).** *If the measurement error variance  $\varepsilon$  is chosen sufficiently small, the initial ensemble is chosen such that  $P_0$  is invertible and the bounds of Lemmas 5 and 6 are satisfied at initial time, then the mean squared estimation error is of order  $\varepsilon^{1/2}$  asymptotically in time.*

Using Markov's inequality the above estimate on the measurement error now yields for fixed  $t$  the following estimate:

$$(70) \quad \mathbb{P}[E_t \geq c\varepsilon^q] \leq \frac{1}{c\varepsilon^q} \mathbb{E}[E_t] = \mathcal{O}(\varepsilon^{1/2-q}).$$

In particular, for any  $q \in (0, 1/2)$  the estimation error  $E_t = \|e_t\|^2/2$  is of order  $\mathcal{O}(\varepsilon^q)$  with probability close to one. Note that this does not imply that for a given realization of the EnKBF, the estimation error  $E_t$  will be small all the time, i.e., that  $\sup_{t \geq 0} E_t$  (or  $\max_{t \in [0, T]} E_t$ ) is of order  $\mathcal{O}(\varepsilon^q)$  with probability close to one. This latter statement requires a pathwise control, i.e., a (locally) uniform in time control of  $E_t$ , which we will derive in the next step. To this end note that (44) together with the inequality (66) imply the pathwise estimate

$$(71) \quad \begin{aligned} E_t &\leq e^{-\alpha t} E_0 + \frac{K}{\alpha} (1 - e^{-\alpha t}) + \int_0^t e^{-\alpha(t-s)} dM_s \\ &= e^{-\alpha t} E_0 + \frac{K}{\alpha} (1 - e^{-\alpha t}) + e^{-\alpha t} M_t + \alpha \int_0^t e^{-\alpha(t-s)} (M_t - M_s) ds, \end{aligned}$$

hence

$$(72) \quad \sup_{t \leq T} E_t \leq \left( E_0 + \frac{K}{\alpha} \right) + \sup_{t \leq T} \left( e^{-\alpha t} |M_t| + \alpha \int_0^t e^{-\alpha(t-s)} |M_t - M_s| ds \right).$$

In order to control the third term, first note that the quadratic variation of the martingale is given as

$$(73) \quad \langle M \rangle_t = \int_0^t \varepsilon^{-1} \|P_s e_s\|^2 + 2 \|C e_s\|^2 dr,$$

so that

$$(74) \quad \langle M \rangle_t - \langle M \rangle_s = \int_s^t \varepsilon^{-1} \|P_r e_r\|^2 + 2 \|C e_r\|^2 dr \leq (C_1 + 2) \int_s^t E_r dr.$$

In the following let  $L_{T,\delta} := \sup_{0 \leq s < t \leq T} |M_t - M_s| / (\langle M \rangle_t - \langle M \rangle_s)^{\frac{1}{2}-\delta}$  for  $\delta \in (0, \frac{1}{2})$ . Theorem 5.1 in [BY82] now implies for any  $\gamma \geq 1$  that there exists a finite constant  $C_{\delta,\gamma}$  such that

$$(75) \quad \mathbb{E}[(L_{T,\delta})^\gamma]^{\frac{1}{\gamma}} \leq C_{\delta,\gamma} \mathbb{E}[\langle M \rangle_T^{\delta\gamma}]^{\frac{1}{\gamma}}.$$

Combining the last estimate with Theorem 8 we obtain for  $\gamma\delta \leq 1$  that

$$(76) \quad \mathbb{E}[(L_{T,\delta})^\gamma]^{\frac{1}{\gamma}} \leq C_{\delta,\gamma} \mathbb{E}[\langle M \rangle_T^{\delta\gamma}]^{\frac{1}{\gamma}} \leq C_{\delta,\gamma} \mathbb{E}[\langle M \rangle_T]^\delta \leq C_{\delta,\gamma} (C_1 + 2)^\delta \mathbb{E} \left[ \int_0^T E_t dt \right]^\delta \leq C \varepsilon^{\frac{\delta}{2}}$$

for some constant  $C$ , depending on  $\gamma, \delta, T, C_1$  and on the bound on the mean squared error obtained in Theorem 8. We can therefore estimate

$$\begin{aligned}
 (77) \quad & \mathbb{E} \left[ \sup_{t \leq T} e^{-\alpha t} |M_t| + \alpha \int_0^t e^{-\alpha(t-s)} |M_t - M_s| \, ds \right] \\
 & \leq \mathbb{E} \left[ \sup_{t \leq T} \left( e^{-\alpha t} \langle M \rangle_t^{\frac{1}{2}-\delta} + \alpha \int_0^t e^{-\alpha(t-s)} (\langle M \rangle_t - \langle M \rangle_s)^{\frac{1}{2}-\delta} \, ds \right) L_{T,\delta} \right] \\
 & \leq (C_1 + 2) \mathbb{E} \left[ \sup_{t \leq T} \left( e^{-\alpha t} t^{\frac{1}{2}-\delta} + \alpha \int_0^t e^{-\alpha(t-s)} (t-s)^{\frac{1}{2}-\delta} \, ds \right) \sup_{t \leq T} E_t^{\frac{1}{2}-\delta} L_{T,\delta} \right] \\
 & \leq (C_1 + 2) \frac{\Gamma(\frac{3}{2}-\delta)}{\alpha^{\frac{1}{2}-\delta}} \mathbb{E} \left[ \sup_{t \leq T} E_t^{\frac{1}{2}-\delta} L_{T,\delta} \right].
 \end{aligned}$$

Applying Young's inequality with  $p = \frac{1}{\frac{1}{2}-\delta}$  and  $q = \frac{1}{\frac{1}{2}+\delta}$  we can further estimate the right-hand side from above by

$$(78) \quad (C_1 + 2) \frac{\Gamma(\frac{3}{2}-\delta)}{\alpha^{\frac{1}{2}-\delta}} \mathbb{E} \left[ \sup_{t \leq T} E_t^{\frac{1}{2}-\delta} L_{T,\delta} \right] \leq \left( \frac{1}{2} - \delta \right) \mathbb{E} \left[ \sup_{t \leq T} E_t \right] + \frac{C}{\alpha^{\frac{1-2\delta}{1+2\delta}}} \mathbb{E} \left[ L_{T,\delta}^{\frac{1}{\frac{1}{2}+\delta}} \right],$$

for some finite constant  $C$  depending on  $C_2$  and  $\delta$ . Taking expectation in (72) and using (76) to estimate the third term gives

$$\begin{aligned}
 (79) \quad & \mathbb{E} \left[ \sup_{t \leq T} E_t \right] \leq \left( \mathbb{E}[E_0] + \frac{K}{\alpha} \right) + \left( \frac{1}{2} - \delta \right) \mathbb{E} \left[ \sup_{t \leq T} E_t \right] + \frac{C}{\alpha^{\frac{1-2\delta}{1+2\delta}}} \mathbb{E} \left[ L_{T,\delta}^{\frac{1}{\frac{1}{2}+\delta}} \right] \\
 & \leq \left( \mathbb{E}[E_0] + \frac{K}{\alpha} \right) + \left( \frac{1}{2} - \delta \right) \mathbb{E} \left[ \sup_{t \leq T} E_t \right] + \frac{C}{\alpha^{\frac{1-2\delta}{1+2\delta}}} \varepsilon^{\frac{\delta}{1+2\delta}}
 \end{aligned}$$

with some different constant  $C$ . Under the assumptions of Theorem 8, in particular  $\mathbb{E}[E_0] = \mathcal{O}(\varepsilon^{\frac{1}{2}})$ , and thus  $\alpha^{-1} = \mathcal{O}(\varepsilon^{\frac{1}{2}})$  for  $\varepsilon \leq \varepsilon_0$ ,  $\varepsilon_0$  sufficiently small, we can now find for any  $\eta \in (0, \frac{1}{4})$  a finite constant  $C$  such that

$$(80) \quad \mathbb{E} \left[ \sup_{t \leq T} E_t \right] \leq C \varepsilon^{\frac{1}{2}-\eta}.$$

In particular,

$$(81) \quad \mathbb{P} \left[ \sup_{t \leq T} E_t \geq c \varepsilon^q \right] \leq \frac{1}{c \varepsilon^q} \mathbb{E} \left[ \sup_{t \leq T} E_t \right] = \mathcal{O} \left( \varepsilon^{1/2-\eta-q} \right),$$

which implies that for any  $q \in (0, 1/2)$  the estimation error  $E_t = \|e_t\|^2/2$  is of order  $\mathcal{O}(\varepsilon^q)$  uniformly on  $[0, T]$  with probability close to one.



**4. Consistency of the ensemble Kalman–Bucy filter for linear systems.** In this section, we provide a detailed analysis of the EnKBF in the case of linear model dynamics, i.e.,  $f(x) = Ax + b$ , linear forward map, i.e.  $h(x) = Hx$ , full rank diffusion tensor,  $D$ , and initial ensemble,  $X_0^i$ , chosen such that  $P_0^M$  is invertible. Then the EnKBF (5) reduces to

$$(82) \quad dX_t^i = (AX_t^i + b)dt + D(P_t^M)^{-1}(X_t^i - \bar{x}_t^M)dt - \frac{1}{2}P_t^M H^T R^{-1} (HX_t^i dt + H\bar{x}_t^M dt - 2dY_t),$$

$i = 1, \dots, M$ , from which we can extract the equation for the empirical mean,  $\bar{x}_t$ ,

$$(83) \quad d\bar{x}_t^M = A\bar{x}_t^M dt + bdt - P_t^M H^T R^{-1} (H\bar{x}_t^M dt - dY_t)$$

and the equation for the empirical covariance matrix, as defined in (6),

$$(84) \quad \frac{d}{dt}P_t^M = AP_t^M + P_t^M A^T + 2D - P_t^M H^T R^{-1} H P_t^M$$

provided  $P_t^M$  has full rank. These equations correspond exactly to the classic Kalman–Bucy filter formulas for the mean and the covariance matrix [Jaz70]. However, while one would set  $P_0^M$  and  $\bar{x}_0^M$  equal to the mean and the covariance matrix, respectively, of the given initial Gaussian distribution  $N(\bar{x}_0, P_0)$  in the classic Kalman–Bucy filter formulation, the  $P_t^M$  and  $\bar{x}_t^M$  arise in our context from sampling from the initial distribution, i.e.,  $X_0^i \sim N(\bar{x}_0, P_0)$ .

**Remark 9.** *It is well known that solutions to (84) have full rank for all  $t > 0$  even if the initial  $P_0^M$  is singular. However, note that (84) holds true only if  $P_0^M$  is nonsingular and that the diffusion induced contribution in (84) needs to be replaced by  $D(P_t^M)^+ P_t^M$  otherwise. This discrepancy between the Riccati equation for the classic Kalman–Bucy filter and the EnKBF is caused by our interacting particle approximation to the diffusion term in (1).*

We will now investigate the asymptotic behavior of the EnKBF in the large ensemble size limit. More specifically, we will show that the empirical distribution of the EnKBF converges under appropriate conditions towards a distribution with mean and covariance determined by the Kalman–Bucy filtering equations. Note that this does not imply that the empirical distribution of the EnKBF converges to the conditional distribution  $\pi_t$  given by the solution of the Kushner–Zakai equation (4), but by the nonlinear Fokker–Planck equation (13) instead as we will show in section 5.

Let us first state the following a.s. result on the asymptotic behavior of  $P_t^M$ .

**Proposition 10.** *Let  $\pi_0$  be the initial distribution on  $\mathbb{R}^{N_x}$  with finite second moments and invertible covariance matrix with entries*

$$(85) \quad \bar{P}_0(k, l) = \pi_0[x_k x_l] - \pi_0[x_k] \pi_0[x_l],$$

*$1 \leq k, l \leq N_x$ . Let  $X_0^i$ ,  $i = 1, 2, \dots$ , be i.i.d.  $(\pi_0)$ , and let  $\bar{P}_t$  be the solution of the Kalman–Bucy filtering equation (99) with initial condition  $\bar{P}_0$ . Then there exists a constant*

$$(86) \quad \tilde{C} = \tilde{C}\left(t, A, D, H^T R^{-1} H, \max_{0 \leq s \leq t} \|\bar{P}_s\|_F, \sup_{M \geq 2} V_0^M\right)$$

such that

$$(87) \quad \|P_t^M - \bar{P}_t\|_F^2 \leq e^{t\tilde{C}} \|P_0^M - \bar{P}_0\|_F^2,$$

where  $V_0^M$  is defined by (18) with  $t = 0$ .

Note that the strong law of large numbers implies that  $\sup_{M \geq 2} V_0^M < \infty$   $\pi_0$ -a.s.

*Proof.* Using the dynamical equations (84) for  $P_t^M$  and (99) for  $\bar{P}_t$  (which, of course, coincides with (84)), we immediately obtain that

$$(88) \quad \frac{1}{2} \frac{d}{dt} \|P_t^M - \bar{P}_t\|_F^2 \leq \langle A(P_t^M - \bar{P}_t), P_t^M - \bar{P}_t \rangle + \langle (P_t^M - \bar{P}_t) A^T, P_t^M - \bar{P}_t \rangle \\ - \langle P_t^M H^T R^{-1} H P_t^M - \bar{P}_t H^T R^{-1} H \bar{P}_t, P_t^M - \bar{P}_t \rangle.$$

Using

$$(89) \quad \langle P_t^M H^T R^{-1} H P_t^M - \bar{P}_t H^T R^{-1} H \bar{P}_t, P_t^M - \bar{P}_t \rangle \\ = \langle P_t^M H^T R^{-1} H (P_t^M - \bar{P}_t), P_t^M - \bar{P}_t \rangle + \langle (P_t^M - \bar{P}_t) H^T R^{-1} H \bar{P}_t, P_t^M - \bar{P}_t \rangle \\ \leq \|H^T R^{-1} H\|_F (\|P_t^M\|_F + \|\bar{P}_t\|_F) \|P_t^M - \bar{P}_t\|_F^2,$$

we arrive at the following differential inequality:

$$(90) \quad \frac{1}{2} \frac{d}{dt} \|P_t^M - \bar{P}_t\|_F^2 \leq (2\|A\|_F + \|H^T R^{-1} H\|_F (\|P_t^M\|_F + \|\bar{P}_t\|_F)) \|P_t^M - \bar{P}_t\|_F^2.$$

Integrating up the last inequality w.r.t. time  $t$  yields

$$(91) \quad \|P_t^M - \bar{P}_t\|_F^2 \leq \exp \left( 4t\|A\|_F + \|H^T R^{-1} H\|_F \int_0^t (\|P_s^M\|_F + \|\bar{P}_s\|_F) ds \right) \|P_0^M - \bar{P}_0\|_F^2.$$

In the next step we will need a uniform in  $M$  upper bound on  $\|P_t^M\|_F$  that holds (locally) uniform w.r.t. time  $t$ . To this end first note that (39) implies

$$(92) \quad \|P_t^M\|_F \leq V_t^M \leq e^{t\|A\|_F} \left( V_0^M + \frac{\text{tr}(D)}{\|A\|_F} \right),$$

thereby using  $L_+ \leq \|A\|_F$ . Since the solution  $\bar{P}_t$  of (99) is continuous, hence, also locally bounded, we can estimate the exponential in (91) from above by

$$2t \left( 2\|A\|_F + \|R^{-1}\|_F \|H\|_F^2 \left( e^{t\|A\|_F} \left( V_0^M + \frac{\text{tr}(D)}{\|A\|_F} \right) + \max_{0 \leq s \leq t} \|\bar{P}_s\|_F \right) \right)$$

which implies the assertion. ■

We can now state our main result on the asymptotic consistency of the ensemble Kalman filter.

**Theorem 11.** Suppose that  $X_0^i$ ,  $i = 1, 2, 3, \dots$ , are i.i.d.  $(\pi_0)$  where the initial distribution  $\pi_0$  has finite second-order moments and invertible covariance matrix (85). Let  $\bar{P}_t$  be the solution of the Kalman–Bucy filtering equation (99) with initial condition  $\bar{P}_0$ , and let  $\bar{x}_t$  be the unique solution of

$$(93) \quad d\bar{x}_t = A\bar{x}_t dt + b dt - \bar{P}_t H^T R^{-1} (H\bar{x}_t dt - dY_t)$$

with initial condition  $\bar{x}_0 := \pi_0[x]$ . Then  $\lim_{M \rightarrow \infty} \bar{x}_t^M = \bar{x}_t$  in  $L^2$ , in particular in probability, for all  $t \geq 0$ .

**Proof.** Since  $X_0^i$  are i.i.d., the strong law of large numbers implies that  $\lim_{M \rightarrow \infty} P_0^M = \bar{P}_0$   $\pi_0$ -a.s. and in  $L^2$ , since  $\pi_0$  has finite second moments, thus  $\lim_{M \rightarrow \infty} P_t^M = \bar{P}_t$  a.s. and in  $L^2$  for  $t \geq 0$  due to Proposition 10.

To see that  $\bar{x}_t^M$  converges towards the unique solution  $\bar{x}_t$  of (93) note that

$$(94) \quad d(\bar{x}_t^M - \bar{x}_t) = A(\bar{x}_t^M - \bar{x}_t) dt - (P_t^M H^T R^{-1} H \bar{x}_t^M - \bar{P}_t H^T R^{-1} H \bar{x}_t) dt \\ + (P_t^M - \bar{P}_t) H^T R^{-1} dY_t$$

and, consequently,

$$(95) \quad \|\bar{x}_t^M - \bar{x}_t\| \leq \|\bar{x}_0^M - \bar{x}_0\| + \int_0^t (\|A\|_F + \|H^T R^{-1} H\|_F \|\bar{P}_s\|_F) \|\bar{x}_s^M - \bar{x}_s\| ds \\ + \int_0^t \|H^T R^{-1} H\|_F \|P_s^M - \bar{P}_s\|_F \|\bar{x}_s^M\| ds + \left\| \int_0^t (P_s^M - \bar{P}_s) H^T R^{-1} dY_s \right\|.$$

Taking expectations we arrive at

$$(96) \quad \mathbb{E} [\|\bar{x}_t^M - \bar{x}_t\|] \leq \mathbb{E} [\|\bar{x}_0^M - \bar{x}_0\|] + \int_0^t (\|A\|_F + \|H\|_F^2 \|R^{-1}\|_F \|\bar{P}_s\|_F) \mathbb{E} [\|\bar{x}_s^M - \bar{x}_s\|] ds \\ + \int_0^t \|H\|_F^2 \|R^{-1}\|_F \mathbb{E} [\|P_s^M - \bar{P}_s\|_F] \|\bar{x}_s^M\| ds \\ + \mathbb{E} \left[ \left\| \int_0^t (P_s^M - \bar{P}_s) H^T R^{-1} dY_s \right\| \right].$$

Using  $\lim_{M \rightarrow \infty} \mathbb{E} [\|P_t^M - \bar{P}_t\|_F^2] = 0$  it follows that

$$(97) \quad \lim_{M \rightarrow \infty} \mathbb{E} \left[ \left\| \int_0^t (P_s^M - \bar{P}_s) H^T R^{-1} dY_s \right\| \right] = 0$$

by dominated convergence, and then Gronwall's lemma implies that  $\lim_{M \rightarrow \infty} \mathbb{E} [\|\bar{x}_t^M - \bar{x}_t\|] = 0$ . ■

**Remark 12.** It is well known that if  $(A, H)$  is observable, i.e.,  $\text{rank} [H^T, (HA)^T, \dots, (HA^{N_x-1})^T] = N_x$ , and  $(A, C)$  is controllable, i.e.,  $\text{rank} [C, AC, \dots, A^{N_x-1}C] = N_x$ , then there exists a unique positive definite solution  $P_\infty$  of the matrix Riccati equation

$$(98) \quad 0 = AP_\infty + P_\infty A^T + 2D - P_\infty H^T R^{-1} H P_\infty,$$

and the solution  $P_t$  of the matrix Riccati equation

$$(99) \quad \frac{d}{dt}P_t = AP_t + P_tA^T + 2D - P_tH^TR^{-1}HP_t,$$

converges for any initial condition  $P_0$  towards  $P_\infty$  as  $t \rightarrow \infty$  with exponential rate  $\lambda < \lambda_*$ , where

$$(100) \quad \lambda_* := \inf\{-\operatorname{Re}(\lambda) \mid \lambda \text{ eigenvalue of } A - P_\infty H^TR^{-1}H\}.$$

(See [KS72, Theorem 4.11], and [OP96, Lemma 2.2].)

Now recall that we have assumed in sections 2 and 3 that  $h(x) = x$ , i.e.,  $H = I$ , and that  $D = CC^T$  has full rank. In other words, we have assumed a restricted case of (nonlinear) controllability and observability. It would be of interest to explore, in as far the conditions of sections 2 and 3 can be relaxed while maintaining the well-posedness, stability, and accuracy of the associated EnKBF.

**5. Asymptotic limiting equations for the extended EnKBF.** In this section, we will derive the non-Markovian stochastic differential equation (11) with (12) of McKean–Vlasov type. We first have to show now that (11) is well-posed. To this end we assume that  $f, h$  are globally Lipschitz continuous and that the initial condition  $\hat{X}_0$  has finite second moments with invertible covariance matrix  $\mathcal{P}_0$ . Recall that—given  $X_t = X_t^{\text{ref}}$ —the observation process  $Y_t$  can be interpreted as Brownian motion with covariance operator  $R$  and drift term  $h(X_t^{\text{ref}})$ , so that we can solve (11) uniquely up to the first time  $\tau$  where  $\mathcal{P}_\tau$  becomes singular. Clearly,  $\tau > 0$  a.s. (w.r.t. the distribution of  $\{Y_s\}$ ). Using Itô's formula, it is then straightforward to see that the distribution  $\hat{\pi}_t$  of  $\hat{X}_t$  indeed satisfies the nonlinear Fokker–Planck equation (13) (up to time  $\tau$ ).

**5.1. Lower bounds on  $\lambda^{\min}(\mathcal{P}_t)$  and well-posedness of (11).** We will prove in Lemma 15 a strictly positive lower bound on the smallest eigenvalue  $\lambda^{\min}(\mathcal{P}_t)$  of  $\mathcal{P}_t$  locally uniformly w.r.t.  $t$ , a.s. w.r.t. the distribution of  $\{Y_s\}$ , under appropriate assumptions on the coefficients  $f, h, D$ , and  $R$ . This implies, in particular, that  $\mathcal{P}_t$  will stay invertible for all  $t$ , a.s. and yields existence and uniqueness of a strong solution of (11) for all times  $t$  (for typical observation  $\{Y_s\}$ ). On the other hand, using the algebraic identity

$$(101) \quad (P_s^M)^{-1} - \mathcal{P}_s^{-1} = (P_s^M)^{-1} (\mathcal{P}_s - P_s^M) \mathcal{P}_s^{-1},$$

we also obtain the following control:

$$(102) \quad \|(P_s^M)^{-1} - \mathcal{P}_s^{-1}\|_2 \leq C(t)^2 \|\mathcal{P}_s - P_s^M\|_2, \quad s \leq t,$$

for the distance between the inverse covariance matrix of the EnKBF and  $\mathcal{P}_t$ . Here,  $C(t)$  is a joint upper bound of  $\|\mathcal{P}_s^{-1}\|_2$  and  $\|(P_s^M)^{-1}\|_2$  (uniform in  $M$ ) for  $s \leq t$ .

To this end let us first state the dynamical equations for the mean  $\bar{x}_t$  and the covariance matrix  $\mathcal{P}_t$  (analogous to (9) and (10) for the EnKBF):

$$(103) \quad d\bar{x}_t = \bar{f}_t dt - \mathcal{Q}_t R^{-1} (\bar{h}_t dt - dY_t), \quad t < \tau,$$

with  $\bar{f}_t = \mathbb{E}[f(\hat{X}_t)]$  and

$$(104) \quad \frac{d}{dt} \mathcal{P}_t = \mathbb{E} \left[ (f(\hat{X}_t) - \bar{f}_t)(\hat{X}_t - \bar{x}_t)^T + (\hat{X}_t - \bar{x}_t)(f(\hat{X}_t) - \bar{f}_t)^T \right] + 2D - \mathcal{Q}_t R^{-1} \mathcal{Q}_t^T, \quad t < \tau.$$

**Lemma 13.** *We have that*

$$(105) \quad \frac{1}{\sqrt{N_x}} \mathbb{E} \left[ \|\hat{X}_t - \bar{x}_t\|^2 \right] \leq \|\mathcal{P}_t\|_F \leq \mathbb{E} \left[ \|\hat{X}_t - \bar{x}_t\|^2 \right], \quad t \leq \tau.$$

*Proof.* Similar to the proof of Lemma 1, we have the following.

Upper bound:

$$(106) \quad \begin{aligned} \|\mathcal{P}_t\|_F^2 &= \sum_{k,l} \mathbb{E} \left[ \left( \hat{X}_t - \bar{x}_t \right)(k) \left( \hat{X}_t - \bar{x}_t \right)(l) \right]^2 \\ &\leq \sum_{k,l} \mathbb{E} \left[ \left( \hat{X}_t - \bar{x}_t \right)^2(k) \right] \mathbb{E} \left[ \left( \hat{X}_t - \bar{x}_t \right)^2(l) \right] = \mathbb{E} \left[ \|\hat{X}_t - \bar{x}_t\|^2 \right]^2. \end{aligned}$$

Lower bound:

$$(107) \quad \|\mathcal{P}_t\|_F^2 = \sum_{k,l} \mathbb{E} \left[ \left( \hat{X}_t - \bar{x}_t \right)(k) \left( \hat{X}_t - \bar{x}_t \right)(l) \right]^2 \geq \sum_k \mathbb{E} \left[ \left( \hat{X}_t - \bar{x}_t \right)^2(k) \right]^2. \quad \blacksquare$$

**Lemma 14.** *For all  $t < \tau$  there exists some finite constant  $C_4(t)$ —independent of  $\{Y_s\}$ —such that*

$$(108) \quad \sup_{0 \leq s \leq t} \mathbb{E} \left[ \|\hat{X}_s - \bar{x}_s\|^2 \right] \leq C_4(t).$$

*Proof.* The difference  $\hat{X}_t - \bar{x}_t$  satisfies the ordinary differential equation

$$(109) \quad \frac{d}{dt} (\hat{X}_t - \bar{x}_t) = (f(\hat{X}_t) - \bar{f}_t) + D\mathcal{P}_t^{-1} (\hat{X}_t - \bar{x}_t) - \frac{1}{2} \mathcal{Q}_t R^{-1} (h(\hat{X}_t) - \bar{h}_t)$$

up to time  $\tau$  so that for  $t < \tau$

$$(110) \quad \begin{aligned} \frac{d}{dt} \mathbb{E} \left[ \|\hat{X}_t - \bar{x}_t\|^2 \right] &= 2\mathbb{E} \left[ \langle f(\hat{X}_t) - \bar{f}_t, \hat{X}_t - \bar{x}_t \rangle \right] + 2\mathbb{E} \left[ \langle D\mathcal{P}_t^{-1} (\hat{X}_t - \bar{x}_t), \hat{X}_t - \bar{x}_t \rangle \right] \\ &\quad - \mathbb{E} \left[ \langle \mathcal{Q}_t R^{-1} (h(\hat{X}_t) - \bar{h}_t), \hat{X}_t - \bar{x}_t \rangle \right] \\ &\leq 2L_+ \mathbb{E} \left[ \|\hat{X}_t - \bar{x}_t\|^2 \right] + 2 \operatorname{tr}(D), \end{aligned}$$

thereby using

$$(111) \quad \mathbb{E} \left[ \langle \mathcal{Q}_t R^{-1} (h(\hat{X}_t) - \bar{h}_t), \hat{X}_t - \bar{x}_t \rangle \right] = \|R^{-1/2} \mathcal{Q}_t^T\|_F^2 \geq 0.$$

This implies the same bound

$$(112) \quad \text{Var}(\hat{X}_t) := \mathbb{E} \left[ \|\hat{X}_t - \bar{x}_t\|^2 \right] \leq e^{2L_+ t} \left( \mathbb{E} \left[ \|\hat{X}_0 - \bar{x}_0\|^2 \right] + \frac{\text{tr}(D)}{L_+} \right)$$

as stated in Remark 4 for the EnKBF for  $h(x) = x$ ; therefore,

$$(113) \quad \sup_{0 \leq s \leq t} \text{Var}(\hat{X}_s) = \sup_{0 \leq s \leq t} \mathbb{E} \left[ \|\hat{X}_s - \bar{x}_s\|^2 \right] \leq C_4(t)$$

for some finite constant  $C_4(t)$  depending on  $t$ . Note that  $C_4(t)$  clearly is independent of  $\{Y_s\}$ . ■

**Lemma 15.** Let  $\|f\|_{\text{Lip}}^2 < 2\lambda^{\min}(D)\|R^{-1}\|_{\text{F}}\|h\|_{\text{Lip}}^2$ . If

$$(114) \quad \lambda^{\min}(\mathcal{P}_0) \geq \kappa_- := \frac{2\lambda^{\min}(D)\|R^{-1}\|_{\text{F}}\|h\|_{\text{Lip}}^2 - \|f\|_{\text{Lip}}^2}{2\|R^{-1}\|_{\text{F}}^2\|h\|_{\text{Lip}}^4 C_4(t)},$$

where  $C_4(t)$  is the upper bound (108) obtained in Lemma 14, then  $\lambda^{\min}(\mathcal{P}_s) \geq \kappa_-$  for all  $s < \tau \wedge t$ . In particular,  $\tau > t$ .

*Proof.* We will use the representation  $\lambda^{\min}(\mathcal{P}_t) = \inf_{\|v\|=1} \langle \mathcal{P}_t v, v \rangle$ . So fix  $v$  with  $\|v\| = 1$ . Then

$$(115) \quad \frac{d}{dt} \langle \mathcal{P}_t v, v \rangle = 2\mathbb{E} \left[ \langle f(\hat{X}_t) - \bar{f}_t, v \rangle \langle \hat{X}_t - \bar{x}_t, v \rangle \right] + 2\langle Dv, v \rangle - \langle R^{-1} \mathcal{Q}_t^T v, \mathcal{Q}_t^T v \rangle.$$

Using

$$(116) \quad \langle \mathcal{P}_t v, v \rangle = \mathbb{E} \left[ \langle \hat{X}_t - \bar{x}_t, v \rangle^2 \right]$$

and

$$(117) \quad \begin{aligned} \langle R^{-1} \mathcal{Q}_t^T v, \mathcal{Q}_t^T v \rangle &= \langle R^{-1} \mathbb{E}[(h(\hat{X}_t) - \bar{h}_t) \langle \hat{X}_t - \bar{x}_t, v \rangle], \mathbb{E}[(h(\hat{X}_t) - \bar{h}_t) \langle \hat{X}_t - \bar{x}_t, v \rangle] \rangle \\ &\leq \|R^{-1}\|_{\text{F}} \mathbb{E} \left[ \|h(\hat{X}_t) - \bar{h}_t\|^2 \right] \mathbb{E} \left[ \langle \hat{X}_t - \bar{x}_t, v \rangle^2 \right] \\ &\leq \|R^{-1}\|_{\text{F}} \|h\|_{\text{Lip}}^2 \text{Var}(\hat{X}_t) \langle \mathcal{P}_t v, v \rangle, \end{aligned}$$

we can estimate

$$(118) \quad \begin{aligned} \frac{d}{dt} \langle \mathcal{P}_t v, v \rangle &\geq -2\|f\|_{\text{Lip}} \text{Var}(\hat{X}_t)^{\frac{1}{2}} \|v\| \langle \mathcal{P}_t v, v \rangle^{\frac{1}{2}} + 2\langle Dv, v \rangle - \|h\|_{\text{Lip}}^2 \|R^{-1}\|_{\text{F}} \text{Var}(\hat{X}_t) \langle \mathcal{P}_t v, v \rangle \\ &\geq -2\|f\|_{\text{Lip}} C_4(t)^{1/2} \langle \mathcal{P}_t v, v \rangle^{\frac{1}{2}} + 2\langle Dv, v \rangle - \|h\|_{\text{Lip}}^2 \|R^{-1}\|_{\text{F}} C_4(t) \langle \mathcal{P}_t v, v \rangle \\ &\geq 2\lambda^{\min}(D) - \frac{\|f\|_{\text{Lip}}^2}{\|R^{-1}\|_{\text{F}} \|h\|_{\text{Lip}}^2} - 2\|h\|_{\text{Lip}}^2 \|R^{-1}\|_{\text{F}} C_4(t) \langle \mathcal{P}_t v, v \rangle. \end{aligned}$$

Now  $\lambda^{\min}(\mathcal{P}_0) \geq \kappa_-$  implies that  $\langle \mathcal{P}_0 v, v \rangle \geq \kappa_-$  and thus  $\langle \mathcal{P}_s v, v \rangle \geq \kappa_-$  for all  $s < \tau \wedge t$ . Hence  $\lambda^{\min}(\mathcal{P}_s) \geq \kappa_- > 0$  for all  $s < \tau \wedge t$  so that  $\tau > t$ , since otherwise  $\lim_{s \uparrow \tau} \lambda^{\min}(\mathcal{P}_s) = 0$ . ■

The lower bound on  $\lambda^{\min}(\mathcal{P}_t)$ , locally uniformly w.r.t.  $t$ , implies that the coefficients of (11) are globally Lipschitz on bounded time-intervals, which gives existence and uniqueness of strong solutions by standard results for all  $t$ , a.s. (w.r.t. the distribution of  $\{Y_s\}$ ).

**5.2. Convergence of the extended EnKBF to the solution of (11).** We are now ready to state our main result on the asymptotic behavior of the extended EnKBF.

**Theorem 16.** Assume that  $\|f\|_{\text{Lip}}^2 < 2\lambda^{\min}(D)\|R^{-1}\|_{\text{F}}\|h\|_{\text{Lip}}^2$ . Let  $\pi_0$  be a distribution on  $\mathbb{R}^{N_x}$  with finite support and invertible covariance matrix  $\mathcal{P}_0$  satisfying  $\lambda^{\min}(\mathcal{P}_0) \geq \kappa_-$ , where  $\kappa_-$  is as in Lemma 15. Let  $\hat{X}_t^i$  be solutions of the mean-field process (11) with initial conditions  $\hat{X}_0^i = X_0^i$  and  $X_0^i$  are i.i.d.  $(\pi_0)$ , so that the solutions  $\hat{X}_t^i$  to the mean field processes are i.i.d., too. Then

$$(119) \quad \lim_{M \rightarrow \infty} \mathbb{E} \left[ \frac{1}{M} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2 \right] = 0.$$

In particular,

$$(120) \quad \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{i=1}^M g(X_t^i) - \hat{\pi}_t[g] = 0$$

in  $L^2(\mathbb{P})$ , hence in probability, for any Lipschitz continuous function  $g$ . Here, the expectation is taken also w.r.t. the distribution of  $\{Y_s\}$ .

**Remark 17.** The last theorem implies by general theory that the empirical distribution  $\hat{\pi}_t^M$ , defined in (8), of the extended EnKBF with  $M$  ensemble members converges weakly towards the distribution  $\hat{\pi}_t$  of the mean field process (11) in probability w.r.t. the distribution of  $\{Y_s\}$ .

**Remark 18.** The conditions of Theorem 16 are satisfied for fully observed processes  $h(x) = x$ , measurement error covariance matrix  $R = \varepsilon I$ ,  $\varepsilon > 0$  sufficiently small, and full rank diffusion tensor  $D$ , i.e., for the filtering setting considered in sections 2 and 3.

*Proof of Theorem 16.* Itô's formula implies that

$$(121) \quad \begin{aligned} \mathbb{d} \left( \frac{1}{M} \sum_{i=1}^M \|\Delta X_t^i\|^2 \right) &= \frac{2}{M} \sum_{i=1}^M \langle f(X_t^i) - f(\hat{X}_t^i), \Delta X_t^i \rangle dt \\ &\quad + \frac{2}{M} \sum_{i=1}^M \langle D((P_t^M)^{-1}(X_t^i - \bar{x}_t^M) - \mathcal{P}_t^{-1}(\hat{X}_t^i - \bar{x}_t)), \Delta X_t^i \rangle dt \\ &\quad - \frac{1}{M} \sum_{i=1}^M \langle Q_t^M R^{-1}(h(X_t^i) + \bar{h}_t^M) - \mathcal{Q}_t R^{-1}(h(\hat{X}_t^i) + \bar{h}_t), \Delta X_t^i \rangle dt \\ &\quad + \frac{2}{M} \sum_{i=1}^M \langle (Q_t^M - \mathcal{Q}_t) R^{-1} dY_t, \Delta X_t^i \rangle \\ &\quad + \frac{1}{M} \text{tr}((Q_t^M - \mathcal{Q}_t) R^{-1} (Q_t^M - \mathcal{Q}_t)^{\text{T}}) dt \\ &= I + \dots + V, \end{aligned}$$

with the abbreviation  $\Delta X_t^i = X_t^i - \hat{X}_t^i$ . Our aim is to estimate the right-hand side of (121) in terms of  $\frac{1}{M} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2$  and then to apply the Gronwall inequality. This requires,

in particular, controlling the stochastic integral  $IV$  w.r.t. the observation  $\{Y_s\}$ . Using the decomposition  $dY_t = h(X_t^{\text{ref}}) dt + R^{1/2} dB_t$  we can split up the stochastic integral  $IV$  into

$$(122) \quad \begin{aligned} \frac{2}{M} \sum_{i=1}^M \langle (Q_t^M - \mathcal{Q}_t) R^{-1} dY_t, \Delta X_t^i \rangle &= \frac{2}{M} \sum_{i=1}^M \langle (Q_t^M - \mathcal{Q}_t) R^{-1} h(X_t^{\text{ref}}), \Delta X_t^i \rangle dt \\ &\quad + \frac{2}{M} \sum_{i=1}^M \langle (Q_t^M - \mathcal{Q}_t) R^{-1/2} dB_t, \Delta X_t^i \rangle \\ &= IVa + IVb. \end{aligned}$$

We can now estimate the right-hand side of the above equation for  $t \leq T$  from above as follows:

$$(123) \quad \begin{aligned} d \left( \frac{1}{M} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2 \right) &\leq U_M(t) \left( \frac{1}{M} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2 + R_M(t) \right) dt \\ &\quad + \frac{2}{M} \sum_{i=1}^M \langle (Q_t^M - \mathcal{Q}_t) R^{-1/2} dB_t, X_t^i - \hat{X}_t^i \rangle, \end{aligned}$$

thereby keeping the stochastic integral  $IVb$ . Here,

$$(124) \quad \begin{aligned} U_M(t) &= C_T \left( 1 + \left\| h(X_{0:T}^{\text{ref}}) \right\|_\infty^2 + \frac{1}{M} \sum_{i=1}^M \|\hat{X}_t^i\|^2 \right) \\ &\quad \times \left( 1 + \frac{1}{M} \sum_{i=1}^M \|X_t^i - \bar{x}_t^M\|^2 + \frac{1}{M} \sum_{i=1}^M \|\hat{X}_t^i - \bar{x}_t^M\|^2 \right), \end{aligned}$$

with some finite constant  $C_T$ , and a remainder  $R_M(t)$  that converges to zero in  $L^p(\mathbb{P})$  as  $M \rightarrow \infty$  for all finite  $p$ .

Indeed, this is obvious for term  $I$ , using that  $f$  is globally Lipschitz, for terms III, IVa, and V using (133) in Lemma 19 in the appendix and for term II it follows from (131) in Lemma 19 in the appendix in combination with (102).

Applying Itô's product formula to the process  $e^{-\int_0^t U_M(s) ds} \frac{1}{M} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2$  and taking expectations w.r.t. the distribution of  $\{Y_s\}$ , we arrive at the following estimate:

$$(125) \quad \mathbb{E} \left[ e^{-\int_0^t U_M(s) ds} \frac{1}{M} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2 \right] \leq C_T \mathbb{E} \left[ \int_0^t e^{-\int_0^s U_M(r) dr} U_M(s) R_M(s) ds \right]$$

for  $t \leq T$ . Since  $U_M R_M$  is bounded by some finite constant plus some power of  $\frac{1}{M} \sum_{i=1}^M \|\hat{X}_t^i\|^2$  and the latter one has some finite exponential moment by Lemma 21, it follows that

$$(126) \quad \lim_{M \rightarrow \infty} \mathbb{E} \left[ e^{-\alpha_T \int_0^t \frac{1}{M} \sum_{i=1}^M \|\hat{X}_s^i\|^2 ds} \frac{1}{M} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2 \right] = 0, \quad t \leq T,$$



for some  $\alpha_T > 0$ . Now, using Lemma 21 again, we also may now conclude that

$$(127) \quad \lim_{M \rightarrow \infty} \mathbb{E} \left[ \frac{1}{M} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\| \right]^2 \leq \sup_{M \geq 2} \mathbb{E} \left[ e^{\alpha_T \int_0^t \frac{1}{M} \sum_{i=1}^M \|\hat{X}_s^i\|^2 ds} \right] \\ \times \lim_{M \rightarrow \infty} \mathbb{E} \left[ e^{-\alpha_T \int_0^t \frac{1}{M} \sum_{i=1}^M \|\hat{X}_s^i\|^2 ds} \frac{1}{M} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2 \right] = 0$$

for all  $t \leq T$ . ■

**6. Numerical example.** We consider the stochastically perturbed Lorenz-63 system [Lor63, LSZ15], which leads to  $N_x = 3$ ,  $D = C = I_3$ , and drift term given by

$$(128) \quad f(x) = \begin{pmatrix} 10(x_2 - x_1) \\ (28 - x_3)x_1 - x_2 \\ x_1x_2 - \frac{8}{3}x_3 \end{pmatrix},$$

where  $x = (x_1, x_2, x_3)^T$ . Solutions of the Lorenz-63 system diverge exponentially fast and filtering is required in order to track a reference solution. Although (128) is only locally Lipschitz continuous, the results from this paper are likely to be applicable to the Lorenz-63 system due to the existence of a Lyapunov function.

We apply the EnKBF with ensemble size  $M = 4$  for values of the measurement error variances  $\varepsilon \in \{10^{-1}, \dots, 10^{-4}, 10^{-5}\}$ . The stochastic evolution equations of the EnKBF are solved by the following modified Euler-Maruyama scheme

$$(129) \quad X_{n+1}^i = X_n^i + \Delta t f(X_n^i) + \Delta t (P_n^M)^{-1} (X_n^i - \bar{x}_n^M) - \frac{1}{2} P_n^M \left( P_n^M + \frac{\varepsilon}{\Delta t} I_3 \right)^{-1} \left( X_n^i + \bar{x}_n^i - 2 \frac{\Delta Y_n}{\Delta t} \right)$$

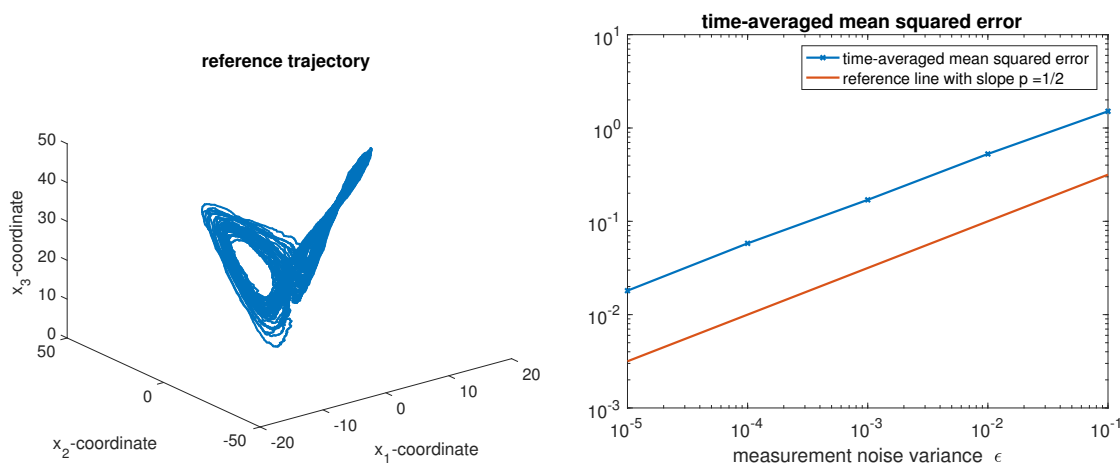
with step-size  $\Delta t = 0.00005$  over a total of  $10^7$  time-steps. Note that

$$(130) \quad \left( P_n^M + \frac{\varepsilon}{\Delta t} I_3 \right)^{-1} \approx \frac{\Delta t}{\varepsilon} I_3$$

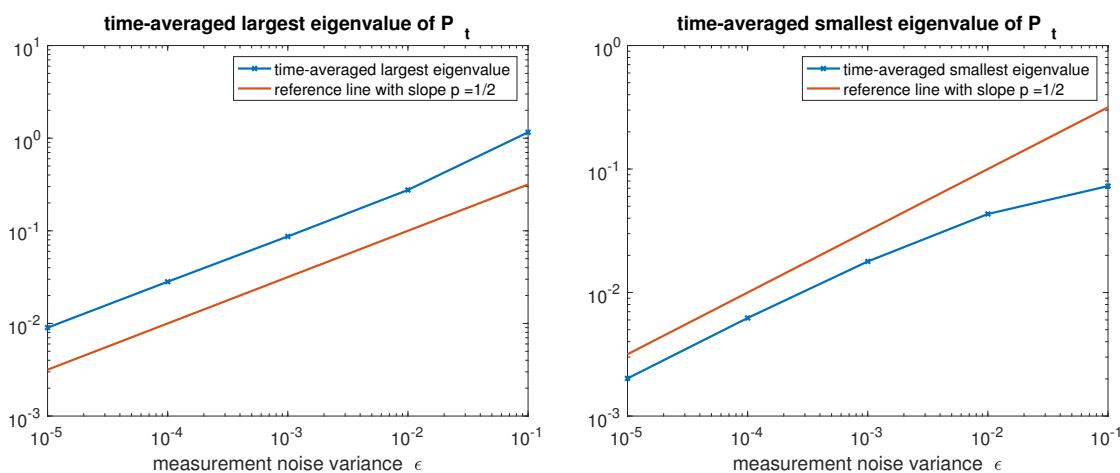
for  $\Delta t$  sufficiently small and the modification is introduced for numerical stability reasons. See [AKIR14] for more details.

The results can be found in Figures 1 and 2. The numerical results are in agreement with our theoretical findings, which predicted an  $\mathcal{O}(\varepsilon^{1/2})$  behavior of these quantities. While this scaling holds for the time-averaged mean squared error and the time-averaged largest eigenvalue of  $P_t^M$  for the whole range of considered values of  $\varepsilon$ , the time-averaged smallest eigenvalue truncates slightly off for the larger values of  $\varepsilon$ . We can also see that there is a gap between the smallest and largest eigenvalues of  $P_t^M$  on average.

We repeated the experiment for ensemble sizes of  $M = 2$  and  $M = 3$ , in which case  $P_t^M$  is singular. We still find that the time-averaged mean squared error is roughly of  $\mathcal{O}(\varepsilon^{1/2})$ . See Figure 3. The results are in line with those obtained in [GTH13] for hyperbolic dynamical systems. We will further investigate the theoretical properties of the EnKBF under singular  $P_t^M$  in a separate paper.



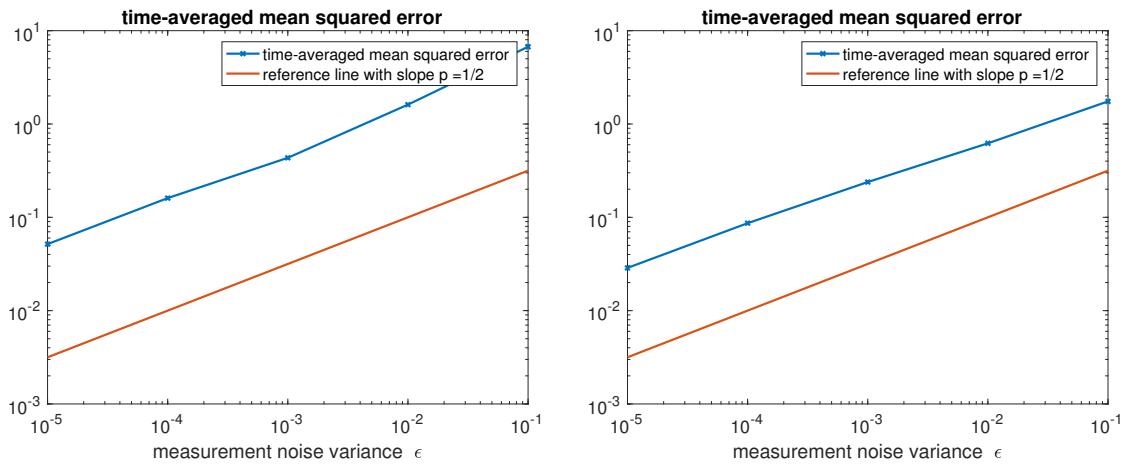
**Figure 1.** Reference trajectory (left panel) and time-averaged mean squared error as a function of the measurement error variance  $\epsilon$  (right panel).



**Figure 2.** Time-averaged largest (left panel) and smallest (right panel) eigenvalues of  $P_t$  as a function of the measurement error variance  $\epsilon$ .

**7. Conclusions.** In this paper, we have taken the first steps towards an understanding of the long-time behavior of the EnKBF and have derived limiting mean-field equations. Natural extensions include partially observed processes and configurations which lead to singular empirical covariance matrices  $P_t^M$ . We also plan to extend our analysis to other ensemble filter algorithms, such as the stochastically perturbed EnKBF and the ensemble transform particle filter. See, for example, [RC15] for more details.

**Appendix: Supplement to the proof of Theorem 16.** The purpose of this appendix is to provide two lemmata on the control of  $\|P_t^M - \mathcal{P}_t\|_F$  and on the existence of exponential moments of  $\int_0^t \frac{1}{M} \sum_{i=1}^M \|\hat{X}_s^i\|^2 ds$  used in the proof of Theorem 16.



**Figure 3.** Time-averaged mean squared error as a function of the measurement error variance  $\varepsilon$  for ensemble sizes  $M = 2$  (left panel) and  $M = 3$  (right panel).

**Lemma 19.** *We have that*

$$(131) \quad \|P_t^M - \mathcal{P}_t\|_F \leq 2\Sigma(t) \left( \frac{1}{M-1} \sum_i \|X_t^i - \hat{X}_t^i\|^2 \right)^{\frac{1}{2}} + R_M(t)$$

with  $\lim_{M \rightarrow \infty} R_M(t) = 0$  a.s. and in  $L^1(\mathbb{P})$ . Here

$$(132) \quad \Sigma(t) := \left( \frac{1}{M-1} \sum_i \|X_t^i - \bar{x}_t^M\|^2 \right)^{\frac{1}{2}} + \left( \frac{1}{M-1} \sum_i \|\hat{X}_t^i - \bar{x}_t\|^2 \right)^{\frac{1}{2}}.$$

Similarly,

$$(133) \quad \|Q_t^M - \mathcal{Q}_t\|_F \leq 2(1 + \|h\|_{\text{Lip}})\Sigma(t) \left( \frac{1}{M-1} \sum_i \|X_t^i - \hat{X}_t^i\|^2 \right)^{\frac{1}{2}} + S_M(t)$$

with  $\lim_{M \rightarrow \infty} S_M(t) = 0$  a.s. and in  $L^1(\mathbb{P})$ .

**Remark 20.** Note that the factor  $\Sigma(t)$  is locally bounded in  $t$  due to Lemma 14 and an appropriate generalization of Lemma 4.

*Proof of Lemma 19.* First, note that we can decompose

$$\begin{aligned}
 P_t^M - \mathcal{P}_t &= \frac{1}{M-1} \sum_{i=1}^M (X_t^i - \bar{x}_t^M) (X_t^i - \bar{x}_t^M)^T - \mathbb{E} \left[ (\hat{X}_t - \bar{x}_t) (\hat{X}_t - \bar{x}_t)^T \right] \\
 &= \frac{1}{M-1} \sum_{i=1}^M \left( X_t^i - \bar{x}_t^M - (\hat{X}_t^i - \bar{x}_t) \right) (X_t^i - \bar{x}_t^M)^T \\
 &\quad + \frac{1}{M-1} \sum_{i=1}^M (\hat{X}_t^i - \bar{x}_t) \left( X_t^i - \bar{x}_t^M - (\hat{X}_t^i - \bar{x}_t) \right)^T \\
 &\quad + \frac{1}{M-1} \sum_{i=1}^M (\hat{X}_t^i - \bar{x}_t) (\hat{X}_t^i - \bar{x}_t)^T - \mathbb{E} \left[ (\hat{X}_t - \bar{x}_t) (\hat{X}_t - \bar{x}_t)^T \right] \\
 &= \text{I} + \text{II} + \text{III}.
 \end{aligned} \tag{134}$$

In particular,  $\|P_t^M - \mathcal{P}_t\|_F \leq \|\text{I}\|_F + \|\text{II}\|_F + \|\text{III}\|_F$ . Term I can be estimated from above by

$$\begin{aligned}
 \|\text{I}\|_F &\leq \left( \frac{1}{M-1} \sum_{i=1}^M \|X_t^i - \bar{x}_t^M - (\hat{X}_t^i - \bar{x}_t)\|^2 \right)^{1/2} \left( \frac{1}{M-1} \sum_{i=1}^M \|X_t^i - \bar{x}_t^M\|^2 \right)^{1/2} \\
 &\leq \left( \left( \frac{1}{M-1} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2 \right)^{1/2} + \sqrt{\frac{M}{M-1}} \|x_t^M - \bar{x}_t\| \right) \left( \frac{1}{M-1} \sum_{i=1}^M \|X_t^i - \bar{x}_t^M\|^2 \right)^{1/2} \\
 &\leq \left( 2 \left( \frac{1}{M-1} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2 \right)^{1/2} + \sqrt{\frac{M}{M-1}} \left\| \frac{1}{M} \sum_{i=1}^M \hat{X}_t^i - \mathbb{E} [\hat{X}_t^i] \right\| \right) \\
 &\quad \times \left( \frac{1}{M-1} \sum_{i=1}^M \|X_t^i - \bar{x}_t^M\|^2 \right)^{1/2}.
 \end{aligned}$$

Similarly,

$$\begin{aligned}
 \|\text{II}\|_F &\leq \left( 2 \left( \frac{1}{M-1} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2 \right)^{1/2} + \sqrt{\frac{M}{M-1}} \left\| \frac{1}{M} \sum_{i=1}^M \hat{X}_t^i - \mathbb{E} [\hat{X}_t^i] \right\| \right) \\
 &\quad \times \left( \frac{1}{M-1} \sum_{i=1}^M \|\hat{X}_t^i - \bar{x}_t\|^2 \right)^{1/2}.
 \end{aligned} \tag{136}$$

Finally,

$$\begin{aligned}
 \|\text{III}\|_F &= \left\| \frac{1}{M} \sum_{i=1}^M \left( \hat{X}_t^i (\hat{X}_t^i)^T - \mathbb{E} [\hat{X}_t^i (\hat{X}_t^i)^T] \right) + \frac{1}{M(M-1)} (\hat{X}_t^i - \bar{x}_t) (\hat{X}_t^i - \bar{x}_t)^T \right\|_F \\
 &\leq \left\| \frac{1}{M} \sum_{i=1}^M \left( \hat{X}_t^i (\hat{X}_t^i)^T - \mathbb{E} [\hat{X}_t^i (\hat{X}_t^i)^T] \right) \right\|_F + \frac{1}{M} \left\| \frac{1}{M-1} \sum_{i=1}^M (\hat{X}_t^i - \bar{x}_t) (\hat{X}_t^i - \bar{x}_t)^T \right\|_F.
 \end{aligned} \tag{137}$$

Adding up all terms we arrive at the estimate

$$(138) \quad \|P_t^M - \mathcal{P}_t\|_F \leq 2\Sigma(t) \left( \frac{1}{M-1} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2 \right)^{1/2} + R_M(t)$$

with the remainder

$$(139) \quad R_M(t) = \Sigma(t) \sqrt{\frac{M}{M-1}} \left\| \frac{1}{M} \sum_{i=1}^M \hat{X}_t^i - \mathbb{E} [\hat{X}_t^i] \right\| + \|\text{III}\|_F.$$

The strong law of large numbers now implies that  $\lim_{M \rightarrow \infty} R_M(t) = 0$  in a.s. and in  $L^1(\mathbb{P})$ . The proof of the second estimate is done similarly.  $\blacksquare$

**Lemma 21.** *Let  $\hat{X}_t^i$ ,  $1 \leq i \leq M$ ,  $M \geq 2$ , be the solution of (11) with initial conditions i.i.d.  $(\pi_0)$  and suppose that  $\pi_0$  has bounded support contained in a ball with radius  $K$ . Then for all  $T > 0$  there exist  $\delta_0 > 0$  and  $\kappa_0 > 0$  depending on  $T$ , but independent of  $M$ , such that*

$$(140) \quad \mathbb{E} \left[ e^{\delta_0 \int_0^t \frac{1}{M} \sum_{i=1}^M \|\hat{X}_s^i\|^2 ds} \right] \leq e^{2\kappa_0 \left( \frac{K^2}{M} + \|h(X_{0:T}^{\text{ref}})\|_\infty^2 \right)} < +\infty \quad \forall t \leq T.$$

Here, the expectation is taken also w.r.t. the distribution of  $\{Y_s\}$ .

*Proof.* First, note that Itô's formula and (11) imply that

$$(141) \quad \begin{aligned} d \left( \frac{1}{M} \sum_{i=1}^M \|\hat{X}_t^i\|^2 \right) &= \frac{2}{M} \sum_{i=1}^M \langle f(\hat{X}_t^i), \hat{X}_t^i \rangle dt + \frac{2}{M} \sum_{i=1}^M \langle D\mathcal{P}_t^{-1}(\hat{X}_t^i - \bar{x}_t), \hat{X}_t^i \rangle dt \\ &\quad - \frac{1}{M} \sum_{i=1}^M \langle \mathcal{Q}_t R^{-1} h(\hat{X}_t^i), \hat{X}_t^i \rangle dt - \frac{1}{M} \sum_{i=1}^M \langle \mathcal{Q}_t R^{-1} \bar{h}_t, \hat{X}_t^i \rangle dt \\ &\quad + \frac{2}{M} \sum_{i=1}^M \langle \hat{X}_t^i, \mathcal{Q}_t R^{-1} dY_t \rangle + \frac{1}{M} \text{tr}(\mathcal{Q}_t R^{-1} \mathcal{Q}_t) dt. \end{aligned}$$

Using Lipschitz continuity of  $f$  and  $h$  and the Lemmata 14 and 15, the right-hand side can be estimated from above for  $t \leq T$  by

$$(142) \quad C(T) \left( 1 + \frac{1}{M} \sum_{i=1}^M \|\hat{X}_t^i\|^2 \right) + \frac{2}{M} \sum_{i=1}^M \langle \hat{X}_t^i, \mathcal{Q}_t R^{-1} dY_t \rangle$$

for some uniform constant  $C(T)$ . Since  $dY_t = h(X_t^{\text{ref}}) dt + R^{-1/2} dB_t$  we can further estimate from above for  $t \leq T$

$$(143) \quad C(T) \left( 1 + \|h(X_t^{\text{ref}})\|^2 + \frac{1}{M} \sum_{i=1}^M \|\hat{X}_t^i\|^2 \right) + \frac{2}{M} \sum_{i=1}^M \langle \hat{X}_t^i, \mathcal{Q}_t R^{-1/2} dB_t \rangle$$

for some possibly different constant  $C(T)$ . Itô's product rule now implies for  $\alpha := 1 + C(T)$  and  $t \leq T$

$$(144) \quad \begin{aligned} d \left( e^{-\alpha t} \frac{1}{M} \sum_{i=1}^M \|\hat{X}_t^i\|^2 \right) &\leq e^{-\alpha t} C(T) \left( 1 + \|h(X_t^{\text{ref}})\|^2 \right) dt - e^{-\alpha t} \left( \frac{1}{M} \sum_{i=1}^M \|\hat{X}_t^i\|^2 \right) dt \\ &\quad + e^{-\alpha t} \frac{2}{M} \sum_{i=1}^M \langle \hat{X}_t^i, \mathcal{Q}_t R^{-1/2} dB_t \rangle, \end{aligned}$$

which implies that

$$(145) \quad \begin{aligned} \int_0^t e^{-\alpha s} \frac{1}{M} \sum_{i=1}^M \|\hat{X}_s^i\|^2 ds &\leq \frac{1}{M} \sum_{i=1}^M \|\hat{X}_0^i\|^2 + C(T) \left( 1 + \|h(X_{0:T}^{\text{ref}})\|_\infty^2 \right) \\ &\quad + \int_0^t e^{-\alpha s} \frac{2}{M} \sum_{i=1}^M \langle \hat{X}_s^i, \mathcal{Q}_s R^{-1/2} dB_s \rangle. \end{aligned}$$

To simplify notation in the following, let

$$(146) \quad M_t := \int_0^t e^{-\alpha s} \frac{2}{M} \sum_{i=1}^M \langle \hat{X}_s^i, \mathcal{Q}_s R^{-1/2} dB_s \rangle$$

and observe that the quadratic variation  $\langle M \rangle_t$  can be estimated from above by

$$(147) \quad \begin{aligned} \langle M \rangle_t &= \frac{4}{M^2} \sum_{i=1}^M \int_0^t e^{-2\alpha s} \|\mathcal{Q}_s^T \hat{X}_s^i\|^2 ds \\ &\leq \frac{4 \|R^{-1/2}\|_{\text{F}}^2 \|h\|_{\text{Lip}}^2 C(T)^2}{M} \int_0^t e^{-\alpha s} \frac{1}{M} \sum_{i=1}^M \|\hat{X}_s^i\|^2 ds, \end{aligned}$$

using

$$(148) \quad \|\mathcal{Q}_s\|_{\text{F}}^2 \leq \|h\|_{\text{Lip}}^2 \mathbb{E} [\|\hat{X}_s - \bar{x}_s\|^2] \leq \|h\|_{\text{Lip}}^2 C(T)^2$$

and Lemma 14. The assumption on the initial condition now implies for  $\delta > 0$

$$(149) \quad \begin{aligned} \mathbb{E} \left[ e^{\delta \int_0^t e^{-\alpha s} \frac{1}{M} \sum_{i=1}^M \|\hat{X}_s^i\|^2 ds} \right] &\leq e^{\delta \left( \frac{K^2}{M} + C(T) \left( 1 + \|h(X_{0:T}^{\text{ref}})\|_\infty^2 \right) \right)} \mathbb{E} \left[ e^{\delta M_t} \right] \\ &\leq e^{\delta \left( \frac{K^2}{M} + C(T) \left( 1 + \|h(X_{0:T}^{\text{ref}})\|_\infty^2 \right) \right)} \mathbb{E} \left[ e^{2\delta^2 \langle M \rangle_t} \right]^{1/2} \\ &\leq e^{\delta \left( \frac{K^2}{M} + C(T) \left( 1 + \|h(X_{0:T}^{\text{ref}})\|_\infty^2 \right) \right)} \\ &\quad \times \mathbb{E} \left[ e^{2\delta^2 \frac{4 \|R^{-1/2}\|_{\text{F}}^2 \|h\|_{\text{Lip}}^2 C(T)^2}{M} \int_0^t e^{-\alpha s} \frac{1}{M} \sum_{i=1}^M \|\hat{X}_s^i\|^2 ds} \right]^{1/2}, \end{aligned}$$

thereby using the inequality

$$(150) \quad \begin{aligned} \mathbb{E} \left[ e^{\delta M_t} \right] &= \mathbb{E} \left[ e^{\frac{1}{2}(2\delta M_t - 2\delta^2 \langle M \rangle_t)} e^{\frac{1}{2}(2\delta^2 \langle M \rangle_t)} \right] \leq \mathbb{E} \left[ e^{2\delta M_t - 2\delta^2 \langle M \rangle_t} \right]^{1/2} \mathbb{E} \left[ e^{2\delta^2 \langle M \rangle_t} \right]^{1/2} \\ &= \mathbb{E} \left[ e^{2\delta^2 \langle M \rangle_t} \right]^{1/2}. \end{aligned}$$

Hence for  $\delta_0 > 0$  with

$$(151) \quad \delta_0 \frac{8 \|R^{-1/2}\|_F^2 \|h\|_{\text{Lip}}^2 C(T)^2}{M} < 1$$

it follows that

$$(152) \quad \mathbb{E} \left[ e^{\delta_0 \int_0^t e^{-\alpha s} \frac{1}{M} \sum_{i=1}^M \|\hat{X}_s^i\|^2 ds} \right] \leq e^{2\delta_0 \left( \frac{\kappa^2}{M} + C(T) \left( 1 + \|h(X_{0:T}^{\text{ref}})\|_\infty^2 \right) \right)} < e^{2\kappa_0 \left( \frac{\kappa^2}{M} + \|h(X_{0:T}^{\text{ref}})\|_\infty^2 \right)} < +\infty$$

for a suitable  $\kappa_0 > 0$ . ■

## REFERENCES

- [AKIR14] J. AMEZCUA, E. KALNAY, K. IDE, AND S. REICH, *Ensemble transform Kalman-Bucy filters*, Q. J. R. Meteorol. Soc., 140 (2014), pp. 995–1004.
- [AMGC02] M.S. ARULAMPALAM, S. MASKELL, N. GORDON, AND T. CLAPP, *A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking*, IEEE Trans. Signal Process., 50 (2002), pp. 174–188.
- [BC08] A. BAIN AND D. CRISAN, *Fundamentals of Stochastic Filtering*, Stoch. Model. Appl. Probab. 60, Springer-Verlag, New York, 2008.
- [BR10] K. BERGEMANN AND S. REICH, *A localization technique for ensemble Kalman filters*, Q. J. R. Meteorol. Soc., 136 (2010), pp. 701–707.
- [BR12] K. BERGEMANN AND S. REICH, *An ensemble Kalman-Bucy filter for continuous data assimilation*, Meteorolog. Zeitschrift, 21 (2012), pp. 213–219.
- [BY82] M. BARLOW AND M. YOR, *Semi-martingale inequalities via the Garcia-Rodemich-Rumsey lemma, and applications to local times*, J. Functional Anal., 49 (1982), pp. 198–229.
- [DdFe01] A. DOUCET, N. DE FREITAS, AND N. GORDON, EDS., *Sequential Monte Carlo Methods in Practice*, Springer-Verlag, New York, 2001.
- [DE99] L. DIECI AND T. EIROLA, *On smooth decompositions of matrices*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 800–819, <https://doi.org/10.1137/S0895479897330182>.
- [DMKT16] P. DEL MORAL, A. KURTZMANN, AND J. TUGAUT, *On the Stability and the Uniform Propagation of Chaos of a Class of Extended Ensemble Kalman-Bucy Filters*, Technical Report, INRIA Bordeaux Research Center, Talence, France, <https://arxiv.org/abs/1606.08256>, 2016.
- [Eve06] G. EVENSEN, *Data Assimilation. The Ensemble Kalman Filter*, Springer-Verlag, New York, 2006.
- [GMT11] F. LE GLAND, V. MONBET, AND V.D. TRAN, *Large sample asymptotics for the ensemble Kalman filter*, in The Oxford Handbook of Nonlinear Filtering, Oxford University Press, Oxford, 2011, pp. 598–631.
- [GTH13] C. GONZÁLEZ-TOKMAN AND B.R. HUNT, *Ensemble data assimilation for hyperbolic systems*, Phys. D, 243 (2013), pp. 128–142.
- [Jaz70] A.H. JAZWINSKI, *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.
- [KLS14] D.T. KELLY, K.J.H. LAW, AND A. STUART, *Well-posedness and accuracy of the ensemble Kalman filter in discrete and continuous time*, Nonlinearity, 27 (2014), pp. 2579–2604.
- [KM15] E. KWIATOWSKI AND J. MANDEL, *Convergence of the square root ensemble Kalman filter in the large ensemble limit*, SIAM/ASA J. Uncertain. Quantif., 3 (2015), pp. 1–17, <https://doi.org/10.1137/140965363>.

- [KMT15] D. KELLY, A.J. MAJDA, AND X.T. TONG, *Concrete ensemble Kalman filters with rigorous catastrophic filter divergence*, Proc. Natl. Acad. Sci. USA, 112 (2015), pp. 10589–10594.
- [KS72] H. KWAKERNAAK AND R. SIVAN, *Linear Optimal Control Systems*, Wiley Interscience, New York, 1972.
- [Lor63] E.N. LORENZ, *Deterministic non-periodic flows*, J. Atmos. Sci., 20 (1963), pp. 130–141.
- [LSZ15] K. LAW, A. STUART, AND K. ZYGALAKIS, *Data Assimilation: A Mathematical Introduction*, Springer-Verlag, Cham, 2015.
- [MH12] A. MAJDA AND J. HARLIM, *Filtering Complex Turbulent Systems*, Cambridge University Press, Cambridge, 2012.
- [Mor13] P. DEL MORAL, *Mean Field Simulation for Monte Carlo Integration*, CRC Press, Boca Raton, FL, 2013.
- [Øks00] B. ØKSENDAL, *Stochastic Differential Equations*, 5th ed., Springer-Verlag, Berlin, 1998.
- [OP96] D. OCONE AND E. PARDOUX, *Asymptotic stability of the optimal filter with respect to its initial condition*, SIAM J. Control Optim., 34 (1996), pp. 226–243, <https://doi.org/10.1137/S0363012993256617>.
- [RC15] S. REICH AND C.J. COTTER, *Probabilistic Forecasting and Bayesian Data Assimilation*, Cambridge University Press, New York, 2015.
- [Rei11] S. REICH, *A dynamical systems framework for intermittent data assimilation*, BIT, 51 (2011), pp. 235–249.
- [Rel69] F. RELICH, *Perturbation Theory for Linear Operators*, Gordon and Breach, New York, 1969.
- [TMK16] X.T. TONG, A.J. MAJDA, AND D. KELLY, *Nonlinear stability and ergodicity of ensemble based Kalman filters*, Nonlinearity, 29 (2016), pp. 657–691.
- [Ton17] X.T. TONG, *Performance Analysis of Local Ensemble Kalman Filter*, Technical Report, National University of Singapore, Singapore, <https://arxiv.org/abs/1705.10598>, 2017.