ON THE CONTINUOUS TIME LIMIT OF THE ENSEMBLE KALMAN FILTER

THERESA LANGE, WILHELM STANNAT

ABSTRACT. We present recent results on the existence of a continuous time limit for Ensemble Kalman Filter algorithms. In the setting of continuous signal and observation processes, we apply the original Ensemble Kalman Filter algorithm proposed by [2] as well as a recent variant [6] to the respective discretizations and show that in the limit of decreasing stepsize the filter equations converge to an ensemble of interacting (stochastic) differential equations in the ensemble-mean-square sense. Our analysis also allows for the derivation of convergence rates with respect to the stepsize.

An application of our analysis is the rigorous derivation of continuous ensemble filtering algorithms consistent with discrete approximation schemes. Conversely, the continuous time limit allows for a better qualitative and quantitative analysis of the time-discrete counterparts using the rich theory of dynamical systems in continuous time.

1. Introduction

In this paper, we aim to give a rigorous derivation of a continuous time limit of the Ensemble Kalman Filter. The Ensemble Kalman Filter (EnKF), or Ensemble Kalman-Bucy Filter (EnKBF) when considered in continuous time, is a data assimilation technique which since its invention in the 1990s gained wide popularity in many scientific fields such as oceanography or meteorology. The general idea is the following: a $d$-dimensional Markovian signal $X$ described by

\begin{equation}
\mathrm{d}X_t = f(X_t)\,\mathrm{d}t + \mathbf{Q}\,\mathrm{d}W_t,
\end{equation}

is unknown but can be observed through a $p$-dimensional process $Y$

\begin{equation}
\mathrm{d}Y_t = g(X_t)\,\mathrm{d}t + \mathbf{C}\,\mathrm{d}V_t
\end{equation}

where $W$ and $V$ are independent Brownian motions. Based on the model assumptions and the real-time measurements, a filter calculates an estimate of the signal at the current time $t$. In case of the EnK(B)F, the signal is estimated in a Monte-Carlo fashion, i.e. an ensemble of initial conditions $X_0^{(i)}$, $i = 1, \ldots, M$, is propagated according to (1) and

2010 Mathematics Subject Classification. 60H35, 93E11, 60F99.

Key words and phrases. Ensemble Kalman Filter, Continuous time limit.
such that its ensemble mean

$$\bar{x}_t := \frac{1}{M} \sum_{i=1}^{M} X_t^{(i)}$$

yields an estimate for the true signal at time $t$.

In [1], the authors give a continuous-time formulation of these equations in the case of linear observations which we generalize as follows:

$$dX_t^{(i)} = f\left(X_t^{(i)}\right) dt + Q^\frac{1}{2} dW_t^{(i)}$$

$$+ \frac{1}{M - 1} E_t G_t^T C^{-1} \left(dY_t + C^\frac{1}{2} dV_t^{(i)} - g\left(X_t^{(i)}\right) dt\right),$$

where

$$E_t := \left[X_t^{(i)} - \bar{x}_t\right]_{i=1,...,M} \in \mathbb{R}^{d \times M},$$

$$G_t := \left[g\left(X_t^{(i)}\right) - \bar{g}_t\right]_{i=1,...,M} \in \mathbb{R}^{p \times M}$$

with

$$\bar{g}_t := \frac{1}{M} \sum_{i=1}^{M} g\left(X_t^{(i)}\right).$$

The continuous formulation is of great use when investigating the mathematical properties of the filter by means of the rich theory of stochastic differential equations and continuous time dynamical systems. In the numerical application, however, only the discrete algorithm can be implemented. To relate both formulations to each other one analyzes the continuous time limit whose existence implies that the continuous formulation is an intrinsic result of the discrete filter. This means that each property we derive for the continuous formulation is also a property of the discrete scheme independent of the discretization step.

The above continuous formulation was achieved by considering the classical EnKF algorithm (cf. [2]) and rearranging terms such that the result resembled an Euler-Maruyama scheme of a stochastic differential equation yielding (4). A similar solution can be found in [3] in the context of inverse problems where the authors derive a tamed Euler-Maruyama type discretization using an EnKF algorithm given in [4].

As known from classical results, the Euler-Maruyama scheme strongly converges to the solution of the underlying stochastic differential equation under some regularity assumptions on the coefficients (cf. [5]). Nevertheless, a rigorous analysis of the convergence of the Euler-Maruyama discretization is still required to verify the aforementioned consistency.

A first attempt to rigorously show a continuous time limit for the EnKF in the context of inverse problems is given in [7] but only for a simplified example. In this work, we investigate the limit for a more general setting and explicitly show convergence to (4) with effective rates in terms of the stepsize in Theorem 4.1 and Corollary 4.2 for explicit rates.
It turns out that assuring the additional integrability condition for the explicit rates is not that easy. The reason for this originates in the stochastic perturbation of the observation in the update step of the filter. In [2], the authors showed that omitting this perturbation yields an underestimation of the resulting covariance of the ensemble. Since this additional source of noise complicates the attempts of achieving better rates, one may consider different filtering approaches avoiding this particular step.

For instance consider the so called deterministic filtering algorithms (such as the Ensemble Square Root Filters (EnSRF), see e.g. [11]) that aim at transforming the ensemble without additional noise and in such a way that the correct ensemble statistics are obtained. Recently in [6], a modified version of (4) was analyzed:

\[ dX_t^{(i)} = f(X_t^{(i)})dt + \frac{1}{2}QP_t^{-1}(X_t^{(i)} - \bar{x}_t)dt \]

(8) \[ + \frac{1}{M-1}E_tG_t^TC_t^{-1}(dY_t + \frac{1}{2}(g(X_t^{(i)}) - \bar{g}_t)dt - g(X_t^{(i)})dt) \]

with the ensemble covariance matrix

\[ P_t := \frac{1}{M-1}\sum_{i=1}^{M} (X_t^{(i)} - \bar{x}_t)(X_t^{(i)} - \bar{x}_t)^T. \]

In this formulation which we will call modified formulation as opposed to the classical formulation (4), the noise terms are replaced by

\[ Q^{1/2}dW_t^{(i)} \sim \frac{1}{2}QP_t^{-1}(X_t^{(i)} - \bar{x}_t)dt, \]

(10) \[ C^{1/2}dV_t^{(i)} \sim \frac{1}{2}(g(X_t^{(i)}) - \bar{g}_t)dt. \]

Note that with this choice the evolution equations for the ensemble mean \( \bar{x}_t \) and the covariance matrix \( P_t \) read

\[ d\bar{x}_t = \bar{f}_t dt + \frac{1}{M-1}E_tG_t^TC_t^{-1}(dY_t - \bar{g}_t dt) \]

and

\[ \frac{d}{dt}P_t = \frac{1}{M-1}\sum_{i=1}^{M} \left( (f(X_t^{(i)}) - \bar{f}_t)(X_t^{(i)} - \bar{x}_t)^T + \right. \]

\[ \left. (X_t^{(i)} - \bar{x}_t)(f(X_t^{(i)}) - \bar{f}_t)^T \right) \]

\[ + Q - \frac{1}{(M-1)^2}E_tG_t^TC_t^{-1}G_tE_t^T, \]

which in the linear case coincide with the Kalman-Bucy equations explaining the particular choice of perturbation.

For [6], the authors of [6] were able to prove long-time stability and accuracy properties of the EnKBF. We therefore also give a continuous
time limit analysis for (25) which might be a first contribution especially to future discussions on how to treat the noise numerically in context of achieving better approximation results.

The paper is organized as follows: after specifying in Section 2 the setting we consider throughout this paper, we shortly review in Section 3 the EnKF algorithm for the above mentioned versions. In Section 4 we turn to the continuous time limit analysis and discuss our results both on the classical and the modified formulation as given in Theorem 4.1 and Theorem 4.4, respectively. The proofs will then be given in Section 5 where we use various properties of the continuous and the discrete time processes involved which we give in Appendix A and Appendix B, respectively.

2. Setting

Consider the $d$-dimensional signal process modeled by

$\text{(13)} \quad dX_t = f(X_t)dt + Q^{\frac{1}{2}}dW_t, \quad X_t \in \mathbb{R}^d$

with drift $f : \mathbb{R}^d \to \mathbb{R}^d$, $Q \in \mathbb{R}^{d \times d}$ symmetric, positive definite, and $W = (W_t)_{t \geq 0}$ a $d$-dimensional standard Brownian motion.

The signal is observed via a $p$-dimensional process $Y \subset \mathbb{R}^p$ modeled by

$\text{(14)} \quad dY_t = g(X_t)dt + C^{\frac{1}{2}}dV_t, \quad Y_t \in \mathbb{R}^p$

with forward map $g : \mathbb{R}^d \to \mathbb{R}^p$, $C \in \mathbb{R}^{p \times p}$ symmetric, positive definite, and $V = (V_t)_{t \geq 0}$ a $p$-dimensional standard Brownian motion.

Further $X_0$, $W$ and $V$ are mutually independent. Throughout this paper we assume that both $f$ and $g$ are Lipschitz-continuous with Lipschitz-constants $\|f\|_{\text{Lip}}$ and $\|g\|_{\text{Lip}}$, and bounded by $\|f\|_{\infty}$ and $\|g\|_{\infty}$, respectively, with $\| \cdot \|_{\infty}$ the usual supremum norm, as well as that all random entities are defined on a given probability space $(\Omega, \mathcal{F}, P)$.

In some time horizon $T$ we consider the partition

$0 = t_0 < t_1 < \cdots < t_L = T$

with step-size $h > 0$, i.e. $t_{k+1} = t_k + h = (k + 1)h$.

The Euler-Maruyama scheme yields a time-discretization of the signal and observation process as follows:

$\text{(15)} \quad X_{t_k} = X_{t_{k-1}} + hf(X_{t_{k-1}}) + Q^{\frac{1}{2}}(W_{t_k} - W_{t_{k-1}})$

$\text{(16)} \quad Y_{t_k} = Y_{t_{k-1}} + hg(X_{t_{k-1}}) + C^{\frac{1}{2}}(V_{t_k} - V_{t_{k-1}})$

$\text{(17)} \quad \Rightarrow \Delta Y_k := Y_k - Y_{t_{k-1}} = hg(X_{t_{k-1}}) + C^{\frac{1}{2}}(V_{t_k} - V_{t_{k-1}})$

Thus $\Delta Y$ is the discrete-time observation process for the above time-discretization of $X$. 

3. Ensemble Kalman Filter algorithms

3.1. The classical formulation. We adopt the formulation in [3] as follows: assume that at time \( t_{k-1} \), we have an ensemble of \( M \) analyzed ensemble members \( \left( X_{t_k}^{(i),a} \right)_{i=1,...,M} \). The forecast step yields a new ensemble consisting of the forecasted members, i.e. for each \( i = 1, ..., M \) obtain

\[
X_{t_k}^{(i),f} = X_{t_{k-1}}^{(i),a} + hf \left( X_{t_{k-1}}^{(i),a} \right) + Q_0 \tilde{W}_k^{(i)}
\]

where \( \left( \tilde{W}_k^{(i)} \right)_{i=1,...,M} \) is an i.i.d. sequence of samples of \( \mathcal{N}(0, h \text{Id}) \). The ensemble mean is then given by

\[
\bar{x}_k = \frac{1}{M} \sum_{i=1}^{M} X_{t_k}^{(i),f}.
\]

In the update step we stochastically perturb the new observation, i.e. define

\[
\Delta Y_k^{(i)} := \Delta Y_k + C_x \tilde{V}_k^{(i)},
\]

where \( \left( \tilde{V}_k^{(i)} \right)_{i=1,...,M} \) is an i.i.d. sequence of samples of \( \mathcal{N}(0, h \text{Id}) \), and formulate an analyzed ensemble

\[
X_{t_k}^{(i),a} = X_{t_k}^{(i),f} + K_k \left( \Delta Y_k^{(i)} - h g \left( X_{t_k}^{(i),f} \right) \right)
\]

with Kalman gain matrix

\[
K_k = \frac{1}{M - 1} E_k^f \left( G_k^f \right)^T \left( C + \frac{h}{M - 1} G_k^f \left( G_k^f \right)^T \right)^{-1}
\]

where

\[
E_k^f = \left[ \bar{x}_k^f - X_{t_k}^{(i),f} \right]_{i=1,..,M},
\]

\[
G_k^f = \left[ g \left( X_{t_k}^{(i),f} \right) - \bar{g}_k^f \right]_{i=1,...,M}, \quad \bar{g}_k^f = \frac{1}{M} \sum_{i=1}^{M} g \left( X_{t_k}^{(i),f} \right).
\]

In total, the algorithm for the ensemble members as well as centered ensemble members reads:

**Algorithm 3.1.** Forecast step:

\[
X_{t_k}^{(i),f} = X_{t_{k-1}}^{(i),a} + hf \left( X_{t_{k-1}}^{(i),a} \right) + Q_0 \tilde{W}_k^{(i)},
\]

\[
X_{t_k}^{(i),f} - \bar{x}_k^f = X_{t_{k-1}}^{(i),a} - \bar{x}_k^a + h \left( f \left( X_{t_{k-1}}^{(i),a} \right) - f_{k-1}^a \right) + Q_0 \tilde{W}_k^{(i)} - \bar{W}_k^{(i)}
\]

Update step:

\[
X_{t_k}^{(i),a} = X_{t_k}^{(i),f} + K_k \left( \Delta Y_k + C_x \tilde{V}_k^{(i)} - h g \left( X_{t_k}^{(i),f} \right) \right),
\]

\[
X_{t_k}^{(i),a} - \bar{x}_k^a = X_{t_k}^{(i),f} - \bar{x}_k^f + K_k \left( C_x \tilde{V}_k^{(i)} - \bar{V}_k^{(i)} \right) - h \left( g \left( X_{t_k}^{(i),f} \right) - \bar{g}_k^f \right)
\]
Remark 3.1. In the case of a linear observation operator, i.e. \( g(x) = Gx \), an alternative interpretation of the Kalman gain matrix is obtained from the following optimization problem (see Proposition 3.2 in [1]):

minimizing the mean-square distance of the ensemble members to the forecast as well as the information provided by the innovation, i.e. minimizing the functional

\[
J_k^{(i)}(X) := \frac{1}{2} \left\| \Delta Y_k^{(i)} - hGX \right\|_C^2 + \frac{1}{2} \left\| X - X_{tk}^{(i),f} \right\|_{P_k^f}^2
\]

where for a positive definite matrix \( A \) we use the Mahalanobis norm

\[
\| x \|^2_A := \langle A^{-1}x, x \rangle = x^T A^{-1} x,
\]

leads to the minimizer given by (21) for each \( i = 1, ..., M \) with gain

\[
K_k = P_k^f G^T \left( C + hGP_k^f G^T \right)^{-1}, \quad P_k^f = \frac{1}{M-1} E_k \left( G_k^f \right)^T.
\]

In the general nonlinear case, however, we cannot employ such a derivation since the above minimization problem cannot be solved analytically for general \( g \). Thus we choose a similar structure of the gain matrix to obtain Equation (22) by replacing

\[
P_k^f G^T \sim \frac{1}{M-1} E_k \left( G_k^f \right)^T,
\]

\[
GP_k^f G^T \sim \frac{1}{M-1} G_k \left( G_k^f \right)^T.
\]

3.2. The modified formulation. In discrete time, the choice [10] translates into the addition of similarly motivated terms of the form

\[
Q_k^{1/2} W_k^{(i)} := \frac{h}{2} Q \left( P_{k-1}^a \right)^{-1} \left( X_{tk-1}^{(i),a} - \bar{x}_{k-1}^{a} \right),
\]

\[
C_k^{1/2} \Gamma_k^{(i)} := \frac{h}{2} \left( g \left( X_{tk}^{(i),f} \right) - \bar{g}_k \right).
\]

This in total yields a discrete-time algorithm for the ensemble members as well as the centered ensemble members:

Algorithm 3.2. Forecast step:

\[
X_{tk}^{(i),f} = X_{tk-1}^{(i),a} + hf \left( X_{tk-1}^{(i),a} \right) + \frac{h}{2} Q \left( P_{k-1}^a \right)^{-1} \left( X_{tk-1}^{(i),a} - \bar{x}_{k-1}^{a} \right),
\]

\[
X_{tk}^{(i),f} - \bar{x}_k^{f} = X_{tk-1}^{(i),a} - \bar{x}_{k-1}^{a} + h \left( f \left( X_{tk-1}^{(i),a} \right) - \bar{f}_k^{a} \right) + \frac{h}{2} Q \left( P_{k-1}^a \right)^{-1} \left( X_{tk-1}^{(i),a} - \bar{x}_{k-1}^{a} \right)
\]

Update step:

\[
X_{tk}^{(i),a} = X_{tk}^{(i),f} + K_k \left( \Delta Y_k - \frac{h}{2} \left( g \left( X_{tk}^{(i),f} \right) + \bar{g}_k \right) \right),
\]

\[
X_{tk}^{(i),a} - \bar{x}_k = X_{tk}^{(i),f} - \bar{x}_k - \frac{h}{2} K_k \left( g \left( X_{tk}^{(i),f} \right) - \bar{g}_k \right).
\]
4. The continuous time limit

Similar to [7], introduce for \( t \in [t_{k-1}, t_k] \) the notation

\[
\eta(t) = t_{k-1}, \quad \eta_+(t) = t_k, \quad \nu(t) = k - 1, \quad \nu_+(t) = k.
\]

For the classical formulation we obtain the following continuous time limit result:

**Theorem 4.1.** Consider Algorithm 3.1. If the initial ensemble satisfies the following properties

\[
\sum_{i=1}^{M} E \left[ \left\| X_0^{(i),a} - X_0^{(i)} \right\|^2 \right] \in O \left( h^{2\gamma} \right),
\]

\[
\sum_{i=1}^{M} E \left[ \left\| X_0^{(i),a} - \bar{x}_0^{a} \right\|^2 \right] < \infty, \quad E \left[ \left( \sum_{i=1}^{M} \left\| X_0^{(i),a} - \bar{x}_0^{a} \right\|^2 \right)^{2} \right] < \infty
\]

where \( \gamma < \frac{1}{2} \) is the Hölder-coefficient coming from \( W, V, \) and \( Y \), then there exists a continuous process \( (L(s))_{s \geq 0} \) such that

\[
\sup_{t \in [0,T]} E \left[ e^{-\int_{0}^{t} L(s) ds} \left( \sum_{i=1}^{M} \left\| X_0^{(i),a} - X_0^{(i)} \right\|^2 \right) \right] \in O \left( h^{2\gamma} \right).
\]

Assuming further an integrability assumption on \( L \) we also obtain an explicit convergence rate:

**Corollary 4.2.** If, moreover, there exists a \( \delta > 0 \) such that

\[
\sup_{t \in [0,T]} E \left[ e^{2\delta \int_{0}^{t} L(s) ds} \left( \sum_{i=1}^{M} \left\| X_0^{(i),a} - X_0^{(i)} \right\|^2 \right) \right] < \infty
\]

then

\[
\sup_{t \in [0,T]} E \left[ \left( \sum_{i=1}^{M} \left\| X_0^{(i),a} - \bar{x}_0^{a} \right\|^2 \right)^{\frac{4}{2+\delta}} \right] \in O \left( h^{2\gamma + \frac{4}{2+\delta}} \right).
\]

To establish the exponential moment estimate (38) turns out to be rather difficult. The reason for that lies in the following fact: \( L(t) \) is an affine linear functional of

\[
\mathcal{V}_t := \frac{1}{M-1} \sum_{i=1}^{M} \left\| X_0^{(i)} - \bar{x}_t \right\|^2
\]

which reduces the problem of finding exponential moment estimates of \( \int_{0}^{t} L(s) ds \) to the same problem for \( \int_{0}^{t} \mathcal{V}_s ds \). However, \( \mathcal{V} \) satisfies the following differential inequality (see Lemma A.1)

\[
d\mathcal{V}_t \leq 2(Lf)_+ \mathcal{V}_t dt + \text{tr}(Q) dt + dN_t
\]

where

\[
(Lf)_+ := \sup_{x \neq y} \frac{\langle f(x) - f(y), x - y \rangle}{\| x - y \|^2}
\]
and
\[ \frac{dN_t}{N_t} = 2 \sum_{i=1}^{M} \left( X_t^{(i)} - \bar{x}_t, Q_{t}^{(i)2} dW_t^{(i)} + \frac{1}{M-1} E_t G_t C^{-\frac{1}{2}} dV_t^{(i)} \right) \]

hence the above problem requires the control of the stochastic integral with respect to \( N \). But note that the quadratic variation of \( N \) satisfies
\[ \langle N \rangle_t \leq 4 t (M - 1)^2 \left( \| Q_{t}^{\frac{1}{2}} \|_{\infty}^2 + \frac{4M \| g \|_{\infty}^2 \| C^{-\frac{1}{2}} \|_{\infty}^2}{M - 1} \right) \mathcal{V}_t \]
which gives no effective control on the growth of \( N \) and hence on \( \mathcal{V} \).

**Remark 4.3.** Without the use of a stochastic perturbation of the observation in the classical EnKF formulation (4), however, the strategy of the proof of Lemma 21 in [6] is indeed applicable to our setting hence we can find a \( \delta > 0 \) such that for the corresponding process \( L \) it holds
\[ \sup_{t \in [0,T]} E \left[ e^{2\delta \int_0^t L(s) \, ds} \right] < \infty. \]

As already mentioned in the Introduction we obtain better error estimates for the modified formulation:

**Theorem 4.4.** Consider Algorithm 3.2. If the initial ensemble satisfies
\[ \sum_{i=1}^{M} E \left[ \| X_0^{(i),a} - X_0^{(i)} \|_2^2 \right] \in O(h) \]
and has bounded inverse covariance matrix, then
\[ E \left[ \sup_{t \in [0,T]} \sum_{i=1}^{M} \| X_t^{(i),a} - X_t^{(i)} \|_2^2 \right] \in O(h). \]

Note that this statement is stronger than the result in Theorem 4.1 which is due to the fact that the coefficients of [8] are bounded as will be shown in Section 5. In particular, this yields that for all \( t \in [0, T] \)
\[ \sum_{i=1}^{M} \| X_t^{(i)} - \tilde{X}_t^{(i)} \| \leq \tilde{C} |t - \eta(t)|^\rho \]
for some \( \rho < \frac{1}{2} \), thus we may concentrate our analysis on the time points \( t_k \) of the time interval partition.

5. **Proofs of Theorem 4.1 and Theorem 4.4**

5.1. **Preliminaries.** For the classical formulation we introduce the 'continuous embedding' of the update step
\[ \tilde{X}_t^{(i)} := X_0^{(i),a} + \int_{\eta(t)}^{t} f \left( X_s^{(i),a} \right) ds + \int_{\eta(t)}^{t} Q_{s}^{2} dW_s^{(i)} + \int_{\eta(t)}^{t} K_{\nu(s)} \nu(s) ds + \int_{\eta(t)}^{t} K_{\nu(s)} C^{2} dV_s^{(i)} \]
and consider the decomposition

\begin{equation}
\|X_{\eta(t)}^{(i),a} - X_t^{(i)}\|^2 \leq 2 \left( \|X_{\eta(t)}^{(i),a} - \tilde{X}_{\eta(t)}^{(i)}\|^2 + 2 \left\|\tilde{X}_{\eta(t)}^{(i)} - X_t^{(i)}\right\|^2 \right)
\end{equation}

where

\begin{equation}
\tilde{X}_t^{(i)} - X_t^{(i)} = X_{\eta(t)}^{(i),a} - X_{\eta(t)}^{(i)}
\end{equation}

\begin{align*}
+ \int_{\eta(t)}^t f \left( X_{\eta(s)}^{(i),a} \right) - f \left( \tilde{X}_s^{(i)} \right) \, ds \\
+ \int_{\eta(t)}^t (K_{\nu+}(s) - \frac{1}{M-1} E_s G_s^T C^{-1}) \, dY_s \\
+ \int_{\eta(t)}^t (K_{\nu+}(s) - \frac{1}{M-1} E_s G_s^T C^{-1}) C \, dV_s^{(i)} \\
- \int_{\eta(t)}^t (K_{\nu+}(s)\, g \left( X_{\eta(s)}^{(i),a} \right) - \frac{1}{M-1} E_s G_s^T C^{-1} g \left( \tilde{X}_s^{(i)} \right) \right) \, ds \\
- \int_{\eta(t)}^t \frac{1}{M-1} E_s G_s^T C^{-1} \left( g \left( \tilde{X}_s^{(i)} \right) - g \left( X_s^{(i)} \right) \right) \, ds.
\end{align*}

In the modified formulation consider the decomposition

\begin{equation}
\|X_{\eta(t)}^{(i),a} - X_t^{(i)}\|^2 \leq 2 \left( \|X_{\eta(t)}^{(i),a} - X_{\eta(t)}^{(i)}\|^2 + 2 \left\|X_{\eta(t)}^{(i)} - X_t^{(i)}\right\|^2 \right)
\end{equation}

where

\begin{equation}
X_{\eta(t)}^{(i),a} - X_{\eta(t)}^{(i)}
\end{equation}

\begin{align*}
= \tilde{X}_0^{(i)} - X_0^{(i)} \\
+ \int_0^{\eta(t)} f \left( X_{\eta(s)}^{(i),a} \right) - f \left( X_s^{(i)} \right) \, ds \\
+ \int_0^{\eta(t)} \frac{1}{2} Q \left( \left( P_{\nu(s)}^{a} \right)^{-1} \left( X_{\eta(s)}^{(i),a} - \tilde{x}_{\nu(s)}^{a} \right) - P_{-1}^{-1} \left( X_s^{(i)} - \tilde{x}_s \right) \right) \, ds \\
+ \int_0^{\eta(t)} \left( K_{\nu+}(s) - \frac{1}{M-1} E_s G_s^T C^{-1} \right) \, dY_s \\
- \int_0^{\eta(t)} \frac{1}{2} \left( K_{\nu+}(s) \, g \left( X_{\eta(s)}^{(i),a} \right) + \tilde{g}_{\nu+}(s) \right) - \frac{1}{M-1} E_s G_s^T C^{-1} \left( g \left( X_s^{(i)} \right) + \tilde{g}_s \right) \right) \, ds.
\end{align*}

Further note that

\begin{equation}
K_{\nu+}(s)\, g(x) - \frac{1}{M-1} E_s G_s^T C^{-1} g(y)
\end{equation}

\begin{align*}
= K_{\nu+}(s) \left( g(x) - g(y) \right) + \left( K_{\nu+}(s) - \frac{1}{M-1} E_s G_s^T C^{-1} \right) g(y) \\
= \left( K_{\nu+}(s) - \frac{1}{M-1} E_s G_s^T C^{-1} \right) g(x) + \frac{1}{M-1} E_s G_s^T C^{-1} \left( g(x) - g(y) \right).
\end{align*}
An essential tool in the analysis is the uniform control of the Lipschitz constants of the coefficients. The coefficients that both the classical and the modified formulation have in common are

\[
\begin{align*}
(55) & \quad \left\| K_{\nu_+}(t) - \frac{1}{M-1} E_t G_t^T C^{-1} \right\|^2, \\
(56) & \quad \left\| K_{\nu_+}(t) \right\|^2, \\
(57) & \quad \left\| \frac{1}{M-1} E_t G_t^T C^{-1} \right\|^2,
\end{align*}
\]

thus we provide the following preliminary estimates:

**Lemma 5.1.** There exist constants \(C_1, C_2, C_3\) independent of \(h\) such that for both Algorithms 3.1 and 3.2 it holds

\[
(58) \quad \left\| K_{\nu_+}(t) - \frac{1}{M-1} E_t G_t^T C^{-1} \right\|^2 \leq C_1 \left( h^2 \left( \nu_{\nu_+}^f(t) \right)^2 + (1 + \nu_t) \left( \frac{1}{M-1} \sum_{i=1}^M \left\| X_{\eta_+}^{(i),f} - X_t^{(i)} \right\|^2 \right) \right).
\]

\[
(59) \quad \left\| K_{\nu_+}(t) \right\|^2 \leq C_2 \nu_{\nu_+}^f(t),
\]

\[
(60) \quad \left\| \frac{1}{M-1} E_t G_t^T C^{-1} \right\|^2 \leq C_3 \nu_t
\]

where

\[
\begin{align*}
\nu_{\nu_+}^f(t) & := \frac{1}{M-1} \sum_{i=1}^M \left\| X_{\eta_+}^{(i),f} - \bar{x}_{\nu_+}^f(t) \right\|^2, \\
\nu_t & := \frac{1}{M-1} \sum_{i=1}^M \left\| X_t^{(i)} - \bar{x}_t \right\|^2.
\end{align*}
\]

In both the classical and the modified formulation, we can control \(\nu_t\) as well as \(\nu_{\nu_+}^f(t)\) in a sense made precise in Appendix A and Appendix B respectively.

**Proof.** On (58):

\[
(62) \quad \begin{align*}
(55) & = \left\| \frac{1}{M-1} E_{\nu_+}^f(t) \left( G_{\nu_+}^f(t) \right)^T \left( C + \frac{h}{M-1} G_{\nu_+}^f(t) \left( G_{\nu_+}^f(t) \right)^T \right)^{-1} - C^{-1} \right\| \\
& \quad + \left\| \frac{1}{M-1} \left( E_{\nu_+}^f(t) \left( G_{\nu_+}^f(t) \right)^T - E_t G_t^T \right) C^{-1} \right\|^2 \\
& \leq 2 \left\| \frac{1}{M-1} E_{\nu_+}^f(t) \left( G_{\nu_+}^f(t) \right)^T \right\|^2 \left\| C + \frac{h}{M-1} G_{\nu_+}^f(t) \left( G_{\nu_+}^f(t) \right)^T \right\|^{-1} - C^{-1} \right\|^2 \\
& \quad + 2 \left\| \frac{1}{M-1} \left( E_{\nu_+}^f(t) \left( G_{\nu_+}^f(t) \right)^T - E_t G_t^T \right) \right\|^2 \left\| C^{-1} \right\|^2.
\end{align*}
\]
For this we estimate by boundedness of $g$
\[(63)\]
\[
\left\| \frac{1}{M-1} E_{\nu+}^f \left( G_{\nu+}^f \right)^T \right\| = \left\| \frac{1}{M-1} \sum_{i=1}^{M} \left( X^{(i),f}_{\nu+} - \tilde{x}^{f}_{\nu+} \right) \left( g \left( X^{(i),f}_{\nu+} \right) - g^{f}_{\nu+} \right)^T \right\| 
\leq 2 \|g\|_{\infty} \sqrt{\frac{M}{M-1}} \left( \mathcal{V}^f_{\nu+}(t) \right)^{\frac{1}{2}}.
\]

Further
\[(64)\]
\[
\left\| \left( \mathcal{C} + \frac{h}{M-1} G_{\nu+}^f \left( G_{\nu+}^f \right)^T \right)^{-1} C^{-1} \right\| = \left\| \left( \mathcal{C} + \frac{h}{M-1} G_{\nu+}^f \left( G_{\nu+}^f \right)^T \right)^{-1} \left( \text{Id} - \left( \mathcal{C} + \frac{h}{M-1} G_{\nu+}^f \left( G_{\nu+}^f \right)^T \right) C^{-1} \right) \right\|
\leq h \left\| C^{-1} \right\|^2 \left\| \frac{1}{M-1} G_{\nu+}^f \left( G_{\nu+}^f \right)^T \right\| \leq h \frac{4 \|g\|_{\infty}^2 \|C^{-1}\|^2 M}{M-1}. \]

Finally
\[(65)\]
\[
\left\| \frac{1}{M-1} \left( E_{\nu+}^f \left( G_{\nu+}^f \right)^T - E_t G_t^f \right) \right\| = \left\| \frac{1}{M-1} \sum_{i=1}^{M} \left( \left( X^{(i),f}_{\nu+} - \tilde{x}^{f}_{\nu+} \right) - \left( X_t^{(i)} - \tilde{x}_t \right) \right) \left( g \left( X^{(i),f}_{\nu+} \right) - g \left( \tilde{x}^{f}_{\nu+} \right) \right)^T 
+ \left( X_t^{(i)} - \tilde{x}_t \right) \left( g \left( X_t^{(i)} \right) - g \left( \tilde{x}_t \right) \right) \left( g \left( X_t^{(i)} \right) - g \left( \tilde{x}_t \right) \right)^T \right\|
\leq \left( \frac{1}{M-1} \sum_{i=1}^{M} \left\| \left( X^{(i),f}_{\nu+} - \tilde{x}^{f}_{\nu+} \right) - \left( X_t^{(i)} - \tilde{x}_t \right) \right\|^2 \right)^{\frac{1}{2}} \left( \frac{1}{M-1} \sum_{i=1}^{M} \left\| g \left( X^{(i),f}_{\nu+} \right) - g \left( \tilde{x}^{f}_{\nu+} \right) \right\|^2 \right)^{\frac{1}{2}}
+ \left( \frac{1}{M-1} \sum_{i=1}^{M} \left\| X_t^{(i)} - \tilde{x}_t \right\|^2 \right)^{\frac{1}{2}}
\left( \frac{1}{M-1} \sum_{i=1}^{M} \left\| g \left( X^{(i),f}_{\nu+} \right) - g \left( \tilde{x}^{f}_{\nu+} \right) \right\|^2 - \left( g \left( X_t^{(i)} \right) - g \left( \tilde{x}_t \right) \right)^2 \right)^{\frac{1}{2}}
\leq 2 \left( \left\| g \right\|_{\infty} \sqrt{\frac{M}{M-1}} + \left\| g \right\|_{\text{Lip}} \left( \mathcal{V}_t^{(i)} \right)^{\frac{1}{2}} \right) \left( \frac{1}{M-1} \sum_{i=1}^{M} \left\| X^{(i),f}_{\nu+} - X_t^{(i)} \right\|^2 \right)^{\frac{1}{2}}. \]
In total we obtain
\begin{align}
&\left\| K_{t+} - \frac{1}{M-1} E_t G^T C^{-1} \right\|^2 \\
&\leq 16 \left\| C^{-1} \right\|^2 \left( \frac{h^2 \| g \|_\infty^4 \| g \|_{L^p}^2 \| C^{-1} \|^2 M^2}{(M-1)^2} \left( \mathcal{V}_{t+}^{f} \right)^2 \right) \\
&+ \left( \frac{4 \| g \|_{L^\infty}^2 M}{M-1} + \| g \|_{L^p}^2 \mathcal{V}_t \right) \left( \frac{1}{M-1} \sum_{i=1}^M \left\| X_{\eta_i(t)}^{(i),f} - X_t^{(i)} \right\|^2 \right).
\end{align}

On (59): using (63) we obtain
\begin{align}
&\left\| K_{t+} \right\| \leq \left\| \frac{1}{M-1} E_t G^T C^{-1} \right\| C^{-1} \\
&\leq 2 \| g \|_\infty \left\| C^{-1} \right\| \sqrt{\frac{M}{M-1}} \left( \mathcal{V}_{t+}^{f} \right)^{1/2},
\end{align}
i.e.
\begin{align}
&\left\| K_{t+} \right\|^2 \leq \frac{4 \| g \|_{L^\infty}^2 \left\| C^{-1} \right\|^2 M \mathcal{V}_t}{M-1}.
\end{align}

On (60): similar to (63) we obtain
\begin{align}
&\left\| \frac{1}{M-1} E_t G^T C^{-1} \right\|^2 \leq \frac{4 \| g \|_{L^\infty}^2 \left\| C^{-1} \right\|^2 M}{M-1} \mathcal{V}_t.
\end{align}

5.2. **Proof of Theorem 4.1** Note that pathwise it holds by Hölder-continuity of \( W^{(i)}, V^{(i)} \), and \( Y \):
\begin{align}
&\left\| X_{\eta_t(a)}^{(i),a} - \hat{X}_t^{(i)} \right\| \leq h \left( \| f \|_\infty + \| g \|_\infty \left\| K_{t+} \right\| \right) \\
&+ h^\gamma \left( \left\| Q - \beta \right\| \left\| W^{(i)} \right\|_{H^\delta} + \| Y \|_{H^\delta} \left\| K_{t+} \right\| + \left\| C^\frac{1}{2} \right\| \left\| V^{(i)} \right\|_{H^\delta} \left\| K_{t+} \right\| \right)
\end{align}
where \( \gamma < \frac{1}{2} \), and thus
\begin{align}
&\left\| X_{\eta_t(a)}^{(i),a} - \hat{X}_t^{(i)} \right\|^2 \leq 5 \left( h^2 \left( \| f \|_\infty^2 + \| g \|_\infty^2 \left\| K_{t+} \right\|^2 \right) \\
&+ h^{2\gamma} \left( \left\| Q - \beta \right\|^2 \left\| W^{(i)} \right\|^2_{H^\delta} + \| Y \|^2_{H^\delta} \left\| K_{t+} \right\|^2 + \left\| C^\frac{1}{2} \right|^2 \left\| V^{(i)} \right\|^2_{H^\delta} \left\| K_{t+} \right\|^2 \right) \right).
\end{align}

By Lemma 5.1 and Appendix \( B \) \( \| K_k \|^2 \) is bounded in expectation for any \( k \) and we get
\begin{align}
&\mathbb{E} \left[ \left\| X_{\eta_t(a)}^{(i),a} - \hat{X}_t^{(i)} \right\|^2 \right] \in \mathcal{O} \left( h^{2\gamma} \right).
\end{align}
Further observe that the observation process $Y$ can be written as

$$dY_t = g \left( X^{\text{ref}}_t \right) dt + C^\perp dV_t$$

where $X^{\text{ref}}$ is the reference trajectory that generates the observations. Using this in (73) we obtain by the Itô formula

$$\frac{1}{2} d \left\| \tilde{X}_t^{(i)} - X_t^{(i)} \right\|^2$$

$$= \left\langle \tilde{X}_t^{(i)} - X_t^{(i)}, f \left( \tilde{X}_t^{(i)} \right) - f \left( X_t^{(i)} \right) \right\rangle dt$$

$$- \left\langle \tilde{X}_t^{(i)} - X_t^{(i)}, \frac{1}{M-1} E_t G_t^T C^{-1} \left( g \left( \tilde{X}_t^{(i)} \right) - g \left( X_t^{(i)} \right) \right) \right\rangle dt$$

$$+ \left\langle \tilde{X}_t^{(i)} - X_t^{(i)}, f \left( X^{(i),a}_{\eta(t)} \right) - f \left( \tilde{X}_t^{(i)} \right) + \left( K_{\nu^+(t)} - \frac{1}{M-1} E_t G_t^T C^{-1} \right) g \left( X^{\text{ref}}_t \right) \right.\right.$$

$$\left. - \left( K_{\nu^+(t)} g \left( X^{(i),f}_{\eta^+(t)} \right) - \frac{1}{M-1} E_t G_t^T C^{-1} g \left( \tilde{X}_t^{(i)} \right) \right) \right\rangle dt$$

$$+ \text{tr} \left( \left( K_{\nu^+(t)} - \frac{1}{M-1} E_t G_t^T C^{-1} \right) C \left( K_{\nu^+(t)} - \frac{1}{M-1} E_t G_t^T C^{-1} \right)^T \right) dt$$

$$+ \left\langle \tilde{X}_t^{(i)} - X_t^{(i)}, \left( K_{\nu^+(t)} - \frac{1}{M-1} E_t G_t^T C^{-1} \right) C^\perp \left( dV_t + dV_t^{(i)} \right) \right\rangle$$

$$\leq \frac{1}{2} + \| f \|_{\text{Lip}} + \| g \|_{\text{Lip}} \| C^{-1} \| \| V_t \|_{\text{Lip}} \left\| \tilde{X}_t^{(i)} - X_t^{(i)} \right\|^2 dt$$

$$+ \left( 2 \| f \|_{\text{Lip}} \| X^{(i),a}_{\eta(t)} - \tilde{X}_t^{(i)} \| + 2 \| g \|_{\text{Lip}} \| K_{\nu^+(t)} \| \| X^{(i),f}_{\eta^+(t)} - \tilde{X}_t^{(i)} \| \right.$$\n
$$\left. + \left( 4 \| g \|_{\infty} + \| C \| \right) \| K_{\nu^+(t)} - \frac{1}{M-1} E_t G_t^T C^{-1} \|_{F} \right) dt$$

$$+ \left\langle \tilde{X}_t^{(i)} - X_t^{(i)}, \left( K_{\nu^+(t)} - \frac{1}{M-1} E_t G_t^T C^{-1} \right) C^\perp \left( dV_t + dV_t^{(i)} \right) \right\rangle.$$

Observe that

$$\sum_{i=1}^{M} \left\| X^{(i),f}_{\eta^+(t)} - \tilde{X}_t^{(i)} \right\|^2$$

(75) $$= \sum_{i=1}^{M} \left\| X^{(i),a}_{\eta(t)} - \tilde{X}_t^{(i)} + hf \left( X^{(i),a}_{\eta(t)} \right) + Q^\perp \hat{W}^{(i)}_{\nu^+(t)} \right\|^2$$

$$\leq 3 \left( \sum_{i=1}^{M} \left\| X^{(i),a}_{\eta(t)} - \tilde{X}_t^{(i)} \right\|^2 + h^2 \| f \|_{\infty}^2 M + \| Q^\perp \|^2 \sum_{i=1}^{M} \| \hat{W}^{(i)}_{\nu^+(t)} \|^2 \right)$$

by boundedness of $f$. With this and Lemma 5.1 we obtain in total

(76) $$\frac{1}{2} \sum_{i=1}^{M} \left\| \tilde{X}_t^{(i)} - X_t^{(i)} \right\|^2 \leq L(t) \sum_{i=1}^{M} \left\| \tilde{X}_t^{(i)} - X_t^{(i)} \right\|^2 dt + R(t) dt + dN_t$$
where

\[
L(t) := \frac{1}{2} + \|f\|_{\text{Lip}} + \|g\|_{\text{Lip}}^2 \|C^{-1}\|_{\mathbb{V}_t} + \kappa \left( \frac{4\|g\|_{\text{Lip}}^2 M}{M - 1} + \|g\|_{\text{Lip}}^2 \|V_t\| \right),
\]

\[
R(t) := \left(2\|f\|_{\text{Lip}}^2 + 6\|g\|_{\text{Lip}}^2 \|K_{\nu_t}^{(i)}(t)\|^2 + \kappa \left( \frac{4\|g\|_{\text{Lip}}^2 M}{M - 1} + \|g\|_{\text{Lip}}^2 \|V_t\| \right) \right) \sum_{i=1}^{M} \|X_{\eta_t}^{(i)\ast} - \bar{X}_t^{(i)}\|^2
\]

\[
+ 6h^2 \|f\|_{\infty}^2 M + 6\|Q^T\|_2^2 \sum_{i=1}^{M} \|\bar{W}_{\nu_t}^{(i)}\|^2
\]

\[
+ \kappa \left( \frac{h}{M - 1} + \|g\|_{\text{Lip}}^2 \|V_t\| \right) \left( h^2 \|f\|_{\infty}^2 M + \|Q^T\|_2^2 \sum_{i=1}^{M} \|\bar{W}_{\nu_t}^{(i)}\|^2 \right),
\]

\[
d\mathcal{N}_t := \sum_{i=1}^{M} \left( \bar{X}_t^{(i)} - X_t^{(i)} \right), \left( K_{\nu_t}^{(i)} \right) \left( \frac{1}{M - 1} - E g G^T C^{-1} \right) C^{1/2} \left( dV_t + dV_t^{(i)} \right)
\]

with \(\kappa := 64\|C^{-1}\|^2 \left(4\|g\|_{\infty}^2 + \|C\| \right) \frac{M}{M - 1}\). By the Itô product rule

\[
d\mathbb{E} \left[ e^{-\int_0^t 2L(s)ds} \left( \sum_{i=1}^{M} \|\bar{X}_t^{(i)} - X_t^{(i)}\|^2 \right) \right] \leq 2e^{-\int_0^t 2L(s)ds} \left( R(t) \right) dt + d\mathbb{N}_t
\]

\[
\leq \mathbb{E} \left[ \sum_{i=1}^{M} \|\bar{X}_t^{(i)} - X_t^{(i)}\|^2 \right] + 2 \mathbb{E} \left[ \int_0^t e^{-\int_0^r 2L(s)ds} dr \mathbb{E} \left[ R(s) \right] ds \right]
\]

Now use (71) as well as Appendix B.1 and Remark A.2 for

\[
\mathbb{E} \left[ \mathcal{V}_t \left\| K_{\nu_t}^{(i)} \right\| \right] \leq \frac{1}{2} \left( \mathcal{V}_t \right)^2 + \frac{1}{2} \left\| K_{\nu_t}^{(i)} \right\|^4
\]

to deduce that it holds

\[
\sup_{t \in [0,T]} \mathbb{E} \left[ R(t) \right] \in O \left( h^{2\gamma} \right).
\]

By the assumption on the initial conditions this then yields

\[
\sup_{t \in [0,T]} \mathbb{E} \left[ e^{-\int_0^t 2L(s)ds} \left( \sum_{i=1}^{M} \|\bar{X}_t^{(i)} - X_t^{(i)}\|^2 \right) \right] \in O \left( h^{2\gamma} \right)
\]

and hence the claim

\[
\sup_{t \in [0,T]} \mathbb{E} \left[ e^{-\int_0^t 2L(s)ds} \left( \sum_{i=1}^{M} \|X_{\eta_t}^{(i)\ast} - X_t^{(i)}\|^2 \right) \right] \in O \left( h^{2\gamma} \right).
5.3. Proof of Theorem 4.4. First of all observe from (53)

\[ X^{(i)}_{\eta(t)} - X^{(i)}_t = -\int_{\eta(t)}^t f(X^{(i)}_s) + \frac{1}{2} QP^{-1}(X^{(i)}_s - \bar{x})_s - \frac{1}{2} M^{-1} E_s G_s^T C^{-1}(g(X^{(i)}_s) + \bar{y}_s) \, ds \]

\[ - \int_{\eta(t)}^t \frac{1}{M-1} E_s G_s^T C^{-1} dY_s \]

\[ = -\int_{\eta(t)}^t f(X^{(i)}_s) + \frac{1}{2} QP^{-1}(X^{(i)}_s - \bar{x})_s + \frac{1}{M-1} E_s G_s^T C^{-1} g(X^{(i)}_s) \]

\[ - \frac{1}{2} M^{-1} E_s G_s^T C^{-1}(g(X^{(i)}_s) + \bar{y}_s) \, ds \]

\[ - \int_{\eta(t)}^t \frac{1}{M-1} E_s G_s^T C^{-\frac{1}{2}} dV_s. \]

With this we obtain

\[ \frac{1}{M-1} \sum_{i=1}^M \left\| X^{(i)}_{\eta(t)} - X^{(i)}_t \right\|^2 \]

\[ \leq \frac{2}{M-1} \sum_{i=1}^M \left\| \int_{\eta(t)}^t f(X^{(i)}_s) + \frac{1}{2} QP^{-1}(X^{(i)}_s - \bar{x})_s + \frac{1}{M-1} E_s G_s^T C^{-1} g(X^{(i)}_s) \right\|^2 \]

\[ - \frac{1}{2} M^{-1} E_s G_s^T C^{-\frac{1}{2}} dV_s \right\|^2 \]

\[ =: (I) + (II). \]

Using the estimate (60) in combination with Appendix A.2 as well as boundedness of \( f \) and \( g \) we deduce

\[ (85) \quad (I) \in O(h^2). \]

Further we obtain that the process

\[ \int_0^1 \frac{1}{M-1} E_s G_s^T C^{\frac{1}{2}} dV_s \]

is Hölder-continuous with coefficient \( \rho < \frac{1}{2} \) (cf. [32]). In total, this yields

\[ (86) \quad \frac{1}{M-1} \sum_{i=1}^M \left\| X^{(i)}_{\eta(t)} - X^{(i)}_t \right\|^2 \in O(h^{2\rho}). \]

Now let \( 1 \leq k \leq L \), then by using the semimartingale representation of \( Y \) and applying the Cauchy-Schwarz-inequality, we obtain
\[ \frac{1}{M-1} \sum_{i=1}^{M} \left\| X_{tk}^{(i),a} - X_{tk}^{(i)} \right\|^2 \]

\[ = \frac{1}{M-1} \sum_{i=1}^{M} \left\| X_0^{(i),a} - X_0^{(i)} + \int_{0}^{tk} f \left( X_{\eta(s)}^{(i),a} \right) - f \left( X_s^{(i)} \right) ds \right\|^2 \]

\[ + \int_{0}^{tk} \frac{1}{2} Q \left( (P_{\nu(s)}^{a})^{-1} \left( X_{\eta(s)}^{(i),a} - \bar{x}_{\nu(s)}^a \right) - P_s^{-1} \left( X_s^{(i)} - \bar{x}_s \right) \right) ds \]

\[ + \int_{0}^{tk} \left( K_{\nu(s)}(s) - \frac{1}{M-1} E_s G_s^T C^{-1} \right) g \left( X_s^{(i)} \right) ds \]

\[ + \int_{0}^{tk} \left( K_{\nu(s)}(s) - \frac{1}{M-1} E_s G_s^T C^{-1} \right) C_2^T dV_s \]

\[ - \frac{1}{2} \left( K_{\nu(s)}(s) g \left( X_{\eta(s)}^{(i),a} \right) + G_{\nu(s)} \right) \]

\[ - \frac{1}{M-1} E_s G_s^T C^{-1} \left( g \left( X_s^{(i)} \right) + \bar{g}_s \right) \right) ds \right\|^2 \]

\[ \leq 6 \left( \frac{1}{M-1} \sum_{i=1}^{M} \left\| X_0^{(i),a} - X_0^{(i)} \right\|^2 + (1) + (2) + (3) + (4) + (5) \right). \]

Term (1) can be estimated by using the Lipschitz-continuity of \( f \).

In (2), the integrand can be decomposed in the following way:

\[ \left( P_{\nu(s)}^{a} \right)^{-1} \left( X_{\eta(s)}^{(i),a} - \bar{x}_{\nu(s)}^a \right) - P_s^{-1} \left( X_s^{(i)} - \bar{x}_s \right) \]

\[ = \left( \left( P_{\nu(s)}^{a} \right)^{-1} - P_s^{-1} \right) \left( X_{\eta(s)}^{(i),a} - \bar{x}_{\nu(s)}^a \right) + P_s^{-1} \left( X_{\eta(s)}^{(i),a} - \bar{x}_{\nu(s)}^a \right) - \left( X_s^{(i)} - \bar{x}_s \right) \]

\[ = - \left( P_{\nu(s)}^{a} \right)^{-1} \left( P_{\nu(s)}^{a} - P_s \right) P_s^{-1} \left( X_{\eta(s)}^{(i),a} - \bar{x}_{\nu(s)}^a \right) + P_s^{-1} \left( X_{\eta(s)}^{(i),a} - \bar{x}_{\nu(s)}^a \right) - \left( X_s^{(i)} - \bar{x}_s \right). \]

It holds with \( V^a \) as defined in (127)

\[ \left\| P_{\nu(s)}^{a} - P_s \right\|^2 \leq 8 \left( \left( V_{\nu(s)}^{a} \right) + \mathcal{V}_s \right) \left( \frac{1}{M-1} \sum_{i=1}^{M} \left\| X_{\nu(s)}^{(i),a} - X_s^{(i)} \right\|^2 \right) \]

which yields by Appendix A and Appendix B.2

\[ (2) \leq \frac{2tk \| Q \|^2}{(\lambda_T)^2} \left( \frac{v_{\eta(s)}^{a,a} (v_{\eta(s)}^{a,a} + v_{\eta}^{a})}{(p_T^a)^2} + 1 \right) \int_{0}^{tk} \frac{1}{M-1} \sum_{i=1}^{M} \left\| X_{\eta(s)}^{(i),a} - X_s^{(i)} \right\|^2 ds. \]

Term (3), (4) and (5) can be estimated via Lemma 5.1 where in (4) we apply the Burkholder-Davis-Gundy inequality to obtain for any \( t \in [0, T] \)
(91) \[ E \left[ \sup_{r \in [0,T]} \int_0^t \left( K_{r+}(r) - \frac{1}{M-1} E_r G_r^T C^{-1} \right) C^\frac{1}{2} dV_r \right]^2 \] 
\leq C_{BDG} \left( C^\frac{1}{2} \right)^2 \int_0^t E \left[ \left\| K_{r+}(s) - \frac{1}{M-1} E_s G_s^T C^{-1} \right\|^2 \right] ds 
\leq 16 C_{BDG} \left( C^{-1} \right)^2 \left( C^\frac{1}{2} \right)^2 
\left( h^2 t^8 \|g\|_\infty^6 \|C^{-1}\|^2 \int_0^t M^3 \right) + \|g\|_{\text{Lip}}^2 \left( f_T^* + v_T^* \right) \int_0^t E \left[ \frac{1}{M-1} \sum_{i=1}^M \left\| X_{\eta_{i+}(s)}^{(i),f} - X_{\eta_{i+}(s)}^{(i),f} \right\|^2 \right] ds \right). 

Further it holds by the forecast step (32) 
(92) 
\frac{1}{M-1} \sum_{i=1}^M \left\| X_{\eta_{i+}(s)}^{(i),f} - X_{\eta_{i+}(s)}^{(i),f} \right\|^2 
\leq 3 \left( \frac{1}{M-1} \sum_{i=1}^M \left\| X_{\eta_{i+}(s)}^{(i),a} - X_{\eta_{i+}(s)}^{(i),a} \right\|^2 \right) + 3 h^2 \left( \frac{\|f\|_\infty^2 M}{M-1} + \frac{\|Q\|^2}{4 (p_T^*)^2 (p_T^*)} \right) 
\leq 3 \left( \frac{1}{M-1} \sum_{i=1}^M \left\| X_{\eta_{i+}(s)}^{(i),a} - X_{\eta_{i+}(s)}^{(i),a} \right\|^2 \right) + 3 h^2 r^*. 

In total we obtain 
(93) 
E \left[ \sup_{r \in [0,T]} \frac{1}{M-1} \sum_{i=1}^M \left\| X_{\eta_{i+}(s)}^{(i),a} - X_{\eta_{i+}(s)}^{(i),a} \right\|^2 \right] 
\leq 6 E \left[ \frac{1}{M-1} \sum_{i=1}^M \left\| X_{\eta_{i+}(s)}^{(i),a} - X_{\eta_{i+}(s)}^{(i),a} \right\|^2 \right] 
+ L_T \int_0^T E \left[ \sup_{r \in [0,T]} \frac{1}{M-1} \sum_{i=1}^M \left\| X_{\eta_{i+}(s)}^{(i),a} - X_{\eta_{i+}(s)}^{(i),a} \right\|^2 \right] ds + R_T h^2 
where 
(94) 
L_T := 18 \|g\|_{\text{Lip}}^2 \left( v_T^* + v_T^* \right) + 6 T \left( \|f\|_{\text{Lip}}^2 + \frac{2 \|Q\|^2}{(\lambda_T^*)^2} \left( \frac{v_T^* + v_T^*}{(p_T^*)^2} + 1 \right) + 24 \|g\|_{\text{Lip}}^2 \left( C^{-1} \right)^2 M \|g\|_{\text{Lip}}^2 \left( 6 v_T^* + 7 v_T^* \right) \right), 

(95) 
R_T := 48 h^2 T \|C^{-1}\|^2 M 
\left( 3 T \|g\|_{\infty}^2 + C_{BDG} \left( C^\frac{1}{2} \right)^2 \right) \left( 16 \|g\|_{\infty}^6 \|C^{-1}\|^2 M^3 \int_0^t M^3 \right) + 6 \|g\|_{\text{Lip}}^2 \left( v_T^* + v_T^* \right) r^* 
\leq 3 T \|g\|_{\infty}^2 \|g\|_{\text{Lip}}^2 v_T^* r^*. 

Now since by the initial discussion of this subsection and the Burkholder-
Davis-Gundy inequality it holds

\[ E \left[ \sup_{t \in [0,T]} \frac{1}{M-1} \sum_{i=1}^{M} \left\| X^{(i)}_{\eta(t)} - X^{(i)}_{t} \right\|^2 \right] \in O(h), \]

we obtain by assumption on the initial conditions and by a Gronwall
argument a constant \( \tilde{C}_T > 0 \) such that

\[ E \left[ \sup_{t \in [a,T]} \frac{1}{M-1} \sum_{i=1}^{M} \left\| X^{(i),a}_{\eta(t)} - X^{(i)}_{t} \right\|^2 \right] \leq \tilde{C}_T h \]

which concludes the proof.

6. Conclusion and Outlook

The Ensemble Kalman Filter is a powerful tool in the field of data
assimilation, and its numerics are widely explored when applied to a
variety of high-dimensional models as such arising in the geosciences.
Our understanding of its theoretical properties, however, is still rather
limited. Some investigations in this direction used the continuous ver-
sion of the Ensemble Kalman Filter in context of observations arriving
with high-frequency, and present first results on stability and accuracy.
In this work, we investigated how the discrete filtering scheme used in
the numerics and the continuous filtering scheme used for the mathem-
atical analysis relate to each other in that we conducted a continuous
time limit argument. In the literature, first results on this can be found
in [1], as well as [3] and [7] in the context of inverse problems, but these
lack a rigorous proof and effective rates for the limit. Similar to [1],
we first considered the classical Ensemble Kalman Filter formulation
which uses a stochastically perturbed observation ensemble, and our
result reads as follows:

Choosing the Euler-Maruyama approximation of the signal and the ob-
servation as their respective time-discretization, the expected ensemble-
mean-square error of the corresponding Ensemble Kalman Filter formu-
lation and the classical continuous version converges to zero uniformly
in time with convergence rate given in Theorem 4.1.

In the attempt of proving better rates, the stochastic perturbation of
the observations posed a serious problem. When omitting this pertur-
bation step the convergence speed can be improved; according to [2],
however, such a perturbation is necessary for the algorithm to give the
correct statistics of the ensemble.

These findings led us to consider alternative versions of the Ensemble
Kalman Filter that do not rely on the stochastic perturbation step and
still produce the correct statistics. Recently in [6], the authors analyzed
a modified continuous Ensemble Kalman Filter scheme originating from
deterministic transformations of the ensemble. The very promising
stability and accuracy properties shown in [6] proposed this formulation as an appropriate candidate for our investigations. We summarize our result to:

Choosing the Euler-Maruyama approximation of the signal and the observation as their respective time-discretization, the ensemble-mean-square error of the corresponding modified Ensemble Kalman Filter formulation and the modified continuous version converges to zero locally uniformly in time in expectation with convergence rate given in Theorem 4.4.

One may further consider investigating the convergence of a mixture of both formulations as mentioned in [10] where the classical forecast step is combined with the modified update step.

The extension to unbounded observation maps $g$ remains an open problem at present. Work is in progress though on the linear case.

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

REFERENCES

[1] Kelly, David T. B., Law, Kody J., Stuart, Andrew M., Well-posedness and accuracy of the ensemble Kalman filter in discrete and continuous time. Nonlinearity, Vol. 27, 2579–2603, 2014
[2] Burgers, Gerrit, van Leeuwen, Peter J., Evensen, Geir, Analysis Scheme in the Ensemble Kalman Filter. Monthly Weather Review, Vol. 126, 1719–1724, 1998
[3] Schillings, Claudia, Stuart, Andrew M., Analysis of the Ensemble Kalman Filter for Inverse Problems. SIAM J. Numerical Analysis, Vol. 55, 1264–1290, 2017
[4] Iglesias, Marco A., Law, Kody J., Stuart, Andrew M., Ensemble Kalman methods for inverse problems. Inverse Problems, Vol. 29, No. 4, 045001, 2013
[5] Kloeden, Peter E., Platen, Eckhard, Numerical Solution of Stochastic Differential Equations. Applications of Mathematics, Springer-Verlag, Berlin, Vol. 23, 1992
[6] de Wiljes, Jana, Reich, Sebastian, Stannat, Wilhelm, Long-time stability and accuracy of the ensemble Kalman-Bucy filter for fully observed processes and small measurement noise. SIAM J. Appl. Dyn. Syst., Vol. 17, No. 2, 1152–1181, 2018
[7] Blömker, Dirk, Schillings, Claudia, Wacker, Phillip, A strongly convergent numerical scheme from Ensemble Kalman inversion. arXiv preprint arXiv:1703.06767v3, 2018
[8] Law, Kody J., Stuart, Andrew M., Zygalakis, Konstantinos, Data assimilation: a mathematical introduction. Springer-Verlag, Berlin, 2015
[9] Roynette, Bernard, Mouvement brownien et espaces de besov. Stochastics and Stochastic Reports, Vol. 43, No. 3-4, 221–260, 1993
[10] Bergemann, Kay, Reich, Sebastian, An ensemble Kalman-Bucy filter for continuous data assimilation. Meteorologische Zeitschrift, Vol. 21, No. 3, 213–219, 2012


APPENDIX A. CONTROL OF THE CONTINUOUS-TIME PROCESSES

A.1. Classical formulation.

**Lemma A.1.** $(\mathcal{V}_t)$ can be controlled $\omega$-wise locally in time $t$. Precisely, it holds

\[
\sup_{t \in [0,T]} \mathcal{V}_t \leq e^{AT} \mathcal{V}_0 + C + e^{AT} \sup_{t \in [0,T]} \int_0^t e^{-As} dN_s
\]

for constants $A = A(f)$ and $C = C(Q, f, T)$, and a martingale $N$.

**Proof.** First observe the following: since $(W^{(i)})_{i=1, \ldots, M}$ and $(V^{(i)})_{i=1, \ldots, M}$ are independent standard Brownian motions, the quadratic variation for each $i = 1, \ldots, M$ reads

\[
\langle W^{(i)} - \bar{w} \rangle_t = \frac{M-1}{M} t = \langle V^{(i)} - \bar{v} \rangle_t.
\]

By the Itô formula we thus obtain

\[
d\mathcal{V}_t = \frac{1}{M-1} \sum_{i=1}^M d\left\| X^{(i)}_t - \bar{x}_t \right\|^2
\]

\[
= \frac{1}{M-1} \sum_{i=1}^M 2 \langle X^{(i)}_t - \bar{x}_t, f \left( X^{(i)}_t \right) - \bar{f}_t \rangle dt
\]

\[
+ 2 \langle X^{(i)}_t - \bar{x}_t, Q^{\frac{1}{2}} \left( dW^{(i)}_t - d\bar{w}_t \right) \rangle
\]

\[
+ 2 \langle X^{(i)}_t - \bar{x}_t, \frac{1}{M-1} E_t G^T_t C^{-\frac{1}{2}} \left( dV^{(i)}_t - d\bar{v}_t \right) \rangle
\]

\[
- 2 \langle X^{(i)}_t - \bar{x}_t, \frac{1}{M-1} E_t G^T_t C^{-1} \left( g \left( X^{(i)}_t \right) - \bar{g}_t \right) \rangle dt
\]

\[
+ \text{tr}(Q) dt + \frac{1}{(M-1)^2} \text{tr} \left( E_t G^T_t C^{-1} G_t E^T_t \right) dt.
\]

Observe now that

\[
\sum_{i=1}^M \langle X^{(i)}_t - \bar{x}_t, f(\bar{x}_t) - \bar{f}_t \rangle = 0
\]

and thus

\[
\sum_{i=1}^M \langle X^{(i)}_t - \bar{x}_t, f \left( X^{(i)}_t \right) - \bar{f}_t \rangle = \sum_{i=1}^M \langle X^{(i)}_t - \bar{x}_t, f \left( X^{(i)}_t \right) - f(\bar{x}_t) \rangle.
\]
Thus with the martingale $N$ 

Applying the Itô product rule yields 
we estimate

and therefore it holds

Note that

Further it holds

Thus

Therefore with the martingale $N$ with

we estimate

Applying the Itô product rule yields

and therefore it holds

Note that

is a continuous martingale and thus locally bounded in time $t$. Hence we can control $\mathcal{V}_t$ $\omega$-wise locally in $t$, since for any $T > 0$ we get

\[ \sup_{t \in [0,T]} \mathcal{V}_t \leq e^{2(Lf)_+ T} \mathcal{V}_0 + \frac{\text{tr}(Q)}{2(Lf)_+} (e^{2(Lf)_+ t} - 1) + \int_0^t e^{2(Lf)_+ (t-s)} d\mathcal{N}_s. \]
Remark A.2. Since one can estimate by boundedness of $g$

$$\frac{d}{dt} \langle \mathcal{N} \rangle_t \leq \frac{4}{(M - 1)^2} \left( \|Q^{\frac{1}{2}}\|^2 + \frac{4M\|g\|_{\infty}^2}{M - 1}\|C^{-\frac{1}{2}}\|^2 \right) \mathcal{N}_t,$$

one can further show by using Equation (108) and a Gronwall argument that

$$\sup_{t \in [0, T]} \mathbb{E} [\mathcal{N}_t]^2 < \infty.$$

A.2. Modified formulation.

Lemma A.3. $\mathcal{V}_t$ is bounded for all $t \in [0, T]$, i.e.

$$(111) \quad \mathcal{V}_t \leq v^*_T := e^{2(Lf)T} \mathcal{V}_0 + \frac{\text{tr}(Q)}{2(Lf)_+} \left( e^{2(Lf)T} - 1 \right).$$

Proof. It holds

$$(112) \quad \frac{1}{2} \frac{d}{dt} \mathcal{V}_t = \frac{1}{M - 1} \sum_{i=1}^{M} \left\langle X_t^{(i)} - \bar{x}_t, f(X_t^{(i)}) - \bar{f}_t \right\rangle + \left\langle X_t^{(i)} - \bar{x}_t, \frac{1}{2}QP_t^{-1} \left( X_t^{(i)} - \bar{x}_t \right) \right\rangle - \left\langle X_t^{(i)} - \bar{x}_t, \frac{1}{2}M - 1 E_tG_t^T C^{-1} \left( g(X_t^{(i)}) - \bar{g}_t \right) \right\rangle$$

$$= \frac{1}{M - 1} \sum_{i=1}^{M} \left\langle X_t^{(i)} - \bar{x}_t, f(X_t^{(i)}) - f(\bar{x}_t) \right\rangle + \frac{\text{tr}(Q)}{2} - \frac{1}{2(M - 1)^2} \text{tr} \left( E_tG_t^T C^{-1}G_tE_t^T \right)$$

$$= \frac{1}{M - 1} \sum_{i=1}^{M} \left\langle X_t^{(i)} - \bar{x}_t, f(X_t^{(i)}) - f(\bar{x}_t) \right\rangle + \frac{\text{tr}(Q)}{2} - \frac{1}{2} \| \frac{1}{M - 1} E_tG_t^T C^{-\frac{1}{2}} \|^2_F$$

which yields the claim after applying a Gronwall argument. $\Box$

Recall the evolution equation of $P_t$

$$(113) \quad \frac{d}{dt} P_t = \frac{1}{M - 1} \sum_{i=1}^{M} \left( \left( f(X_t^{(i)}) - \bar{f}_t \right) \left( X_t^{(i)} - \bar{x}_t \right)^T + \left( X_t^{(i)} - \bar{x}_t \right) \left( f(X_t^{(i)}) - \bar{f}_t \right)^T \right)$$

$$+ Q - \frac{1}{(M - 1)^2} E_tG_t^T C^{-1}G_tE_t^T.$$
Lemma A.4. If $\lambda_{\min}^T$ is bounded away from 0, then there exists a constant $\lambda^*_T > 0$ bounded away from 0 such that

$$(114) \quad \lambda_{\min}^T \geq \lambda^*_T \quad \forall t \in [0, T].$$

Note that, consequently, $\|P_t^{-1}\|$ is bounded from above by $(\lambda^*_T)^{-1}$.

**Proof.** First of all observe that for any $v \in \mathbb{R}^d$ it holds

$$(115) \quad \langle P_t v, v \rangle = \frac{1}{M - 1} \sum_{i=1}^{M} \left\langle X_t^{(i)} - \bar{x}_t, v \right\rangle^2.$$

We proceed as in the proof of Lemma 6 in [6]. Consider a diagonalization of $P_t$, i.e. orthogonal matrices $U_t$ and diagonal matrices $\Lambda_t$ such that

$$(116) \quad P_t = U_t^T \Lambda_t U_t.$$

Then by Equation (113)

$$(117) \quad \frac{d}{dt} \Lambda_t = \text{diag} \left( U_t \left( \frac{1}{M - 1} \sum_{i=1}^{M} \left( f \left( X_t^{(i)} \right) - \bar{f}_t \right) \left( X_t^{(i)} - \bar{x}_t \right)^T + \left( X_t^{(i)} - \bar{x}_t \right)^T \left( f \left( X_t^{(i)} \right) - \bar{f}_t \right) U_t^T \right) + \text{diag} \left( U_t Q U_t^T \right) \right) - \text{diag} \left( U_t \left( \frac{1}{(M - 1)^2} E_t G_t^T C^{-1} G_t E_t^T \right) U_t^T \right).$$

Note that for any matrix $A \in \mathbb{R}^{d \times d}$

$$(118) \quad \left\langle \text{diag} \left( U_t A U_t^T \right) \right\rangle_{ll} = e_t^T U_t A U_t^T e_t$$

and $\|U_t e_t\| = 1$ with $(e_t)$ the standard orthonormal basis in $\mathbb{R}^d$.

Let $v_t = U_t^T e_t$, then $\|v_t\| = 1$ and we estimate

$$(119) \quad \left\| \left( \text{diag} \left( U_t \left( \frac{1}{M - 1} \sum_{i=1}^{M} \left( f \left( X_t^{(i)} \right) - \bar{f}_t \right) \left( X_t^{(i)} - \bar{x}_t \right)^T U_t^T \right) \right) \right)_{ll} \right\| = \frac{1}{M - 1} \sum_{i=1}^{M} \left\langle f \left( X_t^{(i)} \right) - f \left( \bar{x}_t \right), v_t \right\rangle \left\langle X_t^{(i)} - \bar{x}_t, v_t \right\rangle \\
= \left( \frac{1}{M - 1} \sum_{i=1}^{M} \left\langle f \left( X_t^{(i)} \right) - f \left( \bar{x}_t \right), v_t \right\rangle \left\langle X_t^{(i)} - \bar{x}_t, v_t \right\rangle \right)^{\frac{1}{2}} \left( \frac{1}{M - 1} \sum_{i=1}^{M} \left\langle X_t^{(i)} - \bar{x}_t, v_t \right\rangle^2 \right)^{\frac{1}{2}} \\
\leq \|f\|_{\text{Lip}(\mathcal{V}_t)} \left( P_t v_t, v_t \right)^{\frac{1}{2}}.$$
Further (120)
\[
\left| \left( \text{diag} \left( U_t \frac{1}{(M-1)^2} E_t G_t^T C^{-1} G_t E_t^T U_t^T \right) \right) \right| = \frac{1}{(M-1)^2} \langle C^{-1} G_t E_t^T v_l, G_t E_t^T v_l \rangle
\]
\[
\leq \| C^{-1} \| \left| \frac{1}{M-1} G_t E_t^T v_l \right|^2
\]
\[
\leq \| C^{-1} \| \left( \frac{1}{M-1} \sum_{i=1}^M \| g \left( X_i^{(i)} \right) - g \left( \bar{x}_t \right) \| \right)^2 \left( \frac{1}{M-1} \sum_{i=1}^M \langle X_i^{(i)} - \bar{x}_t, v_l \rangle^2 \right)
\]
\[
\leq \| C^{-1} \| \| g \|_{Lip}^2 \langle P_t v_l, v_l \rangle.
\]

Thus by Young’s inequality we obtain for some \( \epsilon > 0 \)
(121)
\[
\frac{d}{dt} \langle \Lambda_t \rangle_{ll} \geq -2 \| f \|_{Lip} \| v^*_T \langle P_t v_l, v_l \rangle^2 + \lambda_{\text{min}}(Q) - \| g \|_{Lip}^2 \left\| C^{-1} \right\| v^*_T \langle P_t v_l, v_l \rangle
\]
\[
\geq \lambda_{\text{min}}(Q) - 2 \| f \|_{Lip}^2 v^*_T \epsilon - \left( \frac{2}{\epsilon} + \| g \|_{Lip}^2 \left\| C^{-1} \right\| v^*_T \right) \langle P_t v_l, v_l \rangle.
\]

Choose \( \epsilon > 0 \) small enough such that
(122)
\[
q_\epsilon := \lambda_{\text{min}}(Q) - 2 \| f \|_{Lip}^2 v^*_T \epsilon > 0.
\]

Now since \( \langle P_t v_l, v_l \rangle = \langle \Lambda_t \rangle_{ll} \), it holds for \( l \) such that \( \lambda_{\text{lmin}}^l = \langle \Lambda_t \rangle_{ll} \) and with
(123)
\[
\alpha_\epsilon := \frac{2}{\epsilon} + \| g \|_{Lip}^2 \left\| C^{-1} \right\| v^*_T
\]
that
(124)
\[
\frac{d}{dt} \lambda_{\text{lmin}}^l \geq q_\epsilon - \alpha_\epsilon \lambda_{\text{lmin}}^l,
\]
thus by a Gronwall argument there exists \( \lambda_{l, \epsilon}^* > 0 \) such that if \( \lambda_{0}^{\text{lmin}} > 0 \) it holds
(125)
\[
\lambda_{l}^{\text{lmin}} \geq \lambda_{l, \epsilon}^*.
\]

\[\square\]

**Appendix B. Control of the discrete-time processes**

**B.1. Classical formulation.** By Algorithm 3.1 we obtain for any \( k \)
Thus for

\begin{equation}
\mathcal{V}_{k+1}^{(i),f} = \frac{1}{M - 1} \sum_{i=1}^{M} \left\| X_{t_k}^{(i),f} - \bar{x}_{k+1} \right\|^2
\end{equation}

and

\begin{equation}
\mathcal{V}_{k}^{a} := \frac{1}{M - 1} \sum_{i=1}^{M} \left\| X_{t_k}^{(i),a} - \bar{x}_{k} \right\|^2, \quad W_{k} := \frac{1}{M - 1} \sum_{i=1}^{M} \left\| W_{t_k}^{(i)} - \bar{w}_{k} \right\|^2
\end{equation}

this gives by the Cauchy-Schwarz-inequality

\begin{equation}
\mathcal{V}_{k+1}^{f} \leq \left(2 + 2h(Lf)_{+} + 5h^2 \left\| f \right\|_{Lip}^2\right) \mathcal{V}_{k}^{a} + 3 \left\| Q_{+}^{f} \right\|^2 \mathcal{W}_{k+1}
\end{equation}

which yields with the estimate (68)

\begin{equation}
\|K_{k+1}\|^2 \leq \frac{4\|g\|^2 M \| C^{-1} \|^2 M}{M - 1} \left(2 + 2h(Lf)_{+} + 5h^2 \left\| f \right\|_{Lip}^2\right) \mathcal{V}_{k}^{a} + 3 \left\| Q_{+}^{f} \right\|^2 \mathcal{W}_{k+1}
\end{equation}

By taking expectation in Equation (126), we furthermore estimate

\begin{equation}
\mathbb{E} \left[ \mathcal{V}_{k+1}^{f} \right] \leq \left(1 + 2h(Lf)_{+} + 4h^2 \left\| f \right\|_{Lip}^2\right) \mathbb{E} \left[ \mathcal{V}_{k}^{a} \right] + \left\| Q_{+}^{f} \right\|^2 \mathbb{E} \left[ \mathcal{W}_{k+1} \right]
\end{equation}

by independence of \(X_{t_k}^{(i),a}\) and \(W_{t_k}^{(i)}\). Similar to the continuous case it holds

\begin{equation}
W_{k+1}^{(i)} - \bar{w}_{k+1} \sim \mathcal{N} \left(0, \frac{M - 1}{M} h \text{Id} \right)
\end{equation}

and thus \(\mathbb{E} \left[ \mathcal{W}_{k+1} \right] = h\), i.e.

\begin{equation}
\mathbb{E} \left[ \mathcal{V}_{k+1}^{f} \right] \leq \left(1 + 2h(Lf)_{+} + 4h^2 \left\| f \right\|_{Lip}^2\right) \mathbb{E} \left[ \mathcal{V}_{k}^{a} \right] + \left\| Q_{+}^{f} \right\|^2 h
\end{equation}

and

\begin{equation}
\mathbb{E} \left[ \|K_{k+1}\|^2 \right] \leq \frac{4\|g\|^2 M \| C^{-1} \|^2 M}{M - 1} \left(1 + 2h(Lf)_{+} + 4h^2 \left\| f \right\|_{Lip}^2\right) \mathbb{E} \left[ \mathcal{V}_{k}^{a} \right] + \left\| Q_{+}^{f} \right\|^2 h
\end{equation}

\begin{equation}
=: K_3 \left( \mathbb{K}_4(h) \mathbb{E} \left[ \mathcal{V}_{k}^{a} \right] + \left\| Q_{+}^{f} \right\|^2 h \right).
\end{equation}
Thus \( E \left[ \| K_{k+1} \|^2 \right] \) as well as \( E \left[ \| K_{k+1} \|^4 \right] \) are locally bounded in \( k \) according to the following lemma:

**Lemma B.1.** If \( E [V_0^a] < \infty \) and \( E \left[ (V_0^a)^2 \right] < \infty \), then first and second moment of \( V_k^a \) are locally bounded in \( k \), i.e. for \( h \ll 1 \) there exist constants \( C_1, C_2, C_3, \) and \( C_4 \), all independent of \( h \), such that

\[
E [V_k^a] \leq e^{C_T} (E [V_0^a] + TC_2)
\]

as well as

\[
E \left[ (V_k^a)^2 \right] \leq C_3 \exp (TC_4).
\]

**Proof.** Using the Cauchy-Schwarz inequality as well as (138) yields with Algorithm 3.1 the following estimate

\[
V_{k+1}^a - V_k^a = \frac{1}{M-1} \sum_{i=1}^{M} \left| X_{t_k}^{(i),a} - x_{k+1}^a \right|^2 - \left| X_{t_k}^{(i),a} - x_k^a \right|^2
\]

\[
= \frac{1}{M-1} \sum_{i=1}^{M} \left( (X_{t_k}^{(i),a} - x_{k+1}^a) + (X_{t_k}^{(i),a} - x_k^a) \right) \left( (X_{t_k}^{(i),a} - x_{k+1}^a) - (X_{t_k}^{(i),a} - x_k^a) \right)
\]

\[
\leq \left( h + 2h(Lf)_+ + 5h^2 \| f \|_{Lip}^2 + \frac{8h(1 + h)M}{M - 1} \| g \|_{\infty}^2 \tilde{K}_1 \tilde{K}_2(h) \right) V_k^a
\]

\[
+ \left( 2 + h + \frac{8h(1 + h)M}{M - 1} \| g \|_{\infty}^2 \tilde{K}_1 \right) \| Q_{\tilde{X}}^2 \| W_{k+1} + 2 \| K_{k+1} \|^2 \| C_4 \|^2 \tilde{V}_{k+1} + 2 N_{k+1}
\]

with

\[
V_k := \frac{1}{M-1} \sum_{i=1}^{M} \left| \tilde{V}_k^{(i)} - \tilde{v}_k \right|^2
\]

and the martingale

\[
N_{k+1}
\]

\[
:= \frac{1}{M-1} \sum_{i=1}^{M} \left( X_{t_k}^{(i),a} - x_k^a + h f (X_{t_k}^{(i),a}) - \tilde{f}_k^a, \tilde{W}_k^{(i)} (\tilde{W}_{k+1}^{(i)} - \tilde{w}_{k+1}) \right)
\]

\[
+ \left( X_{t_k}^{(i),a} - x_k^a + h f (X_{t_k}^{(i),a}) - \tilde{f}_k^a - hK_{k+1} (g (X_{t_k+1}^{(i)}) - \tilde{g}_k^{f}) , \tilde{W}_k^{(i)} (\tilde{W}_{k+1}^{(i)} - \tilde{w}_{k+1}) \right)
\]

\[
+ \frac{d}{M-1} \sum_{i=1}^{M} \int_{t_k}^{t_{k+1}} \left( X_{t_k}^{(i),a} - x_k^a + h f (X_{t_k}^{(i),a}) - \tilde{f}_k^a , Q_{\tilde{X}}^2 dW_s^{(i)} \right)
\]

\[
\left( X_{t_k}^{(i),a} - x_k^a + h f (X_{t_k}^{(i),a}) - \tilde{f}_k^a , hK_{k+1} (g (X_{t_k+1}^{(i)}) - \tilde{g}_k^{f}) , K_{k+1} C_4^2 dV_s^{(i)} \right).
\]
Thus by independence of $V_{k+1}$, $W_{k+1}$ and $V_k^a$ for each $k$ and since $E[W_k] = h = E[V_k]$ for each $k$ we obtain

\begin{equation}
E\left[V_{k+1}^a\right] \leq (1 + hC_1(h)) E[V_k^a] + hC_2(h)
\end{equation}

where

\begin{align}
C_1(h) &:= 1 + 2(Lf)_+ + 5h\|f\|_{lip}^2 + \frac{8(1 + h)M}{M - 1} \|g\|_{\infty}^2 \tilde{K}_1(h) \tilde{K}_2(h) + 2\left\|\mathcal{C}_1^2\right\|^2 \tilde{K}_3 \tilde{K}_4(h), \\
C_2(h) &:= \left\|Q_1^\perp\right\|^2 \left(2 + h + \frac{8h(1 + h)M}{M - 1} \|g\|_{\infty}^2 \tilde{K}_1 + 2\left\|\mathcal{C}_1^2\right\|^2 \tilde{K}_3(h)\right).
\end{align}

Therefore by a Gronwall argument

\begin{equation}
E[V_k^a] \leq (1 + hC_1(h))^k E[V_0^a] + \sum_{j=0}^{k-1} (1 + hC_1(h))^j hC_2(h)
\end{equation}

and since $h = \frac{T}{L}$, we achieve for $h \ll 1$ the final estimate

\begin{equation}
E[V_k^a] \leq e^{TC_1(1)} (E[V_0^a] + TC_2(1))
\end{equation}

for all $1 \leq k \leq L$. Thus $E[V_k^a]$ is bounded for all $1 \leq k \leq L$ since $E[V_0^a] < \infty$.

Now consider the result of applying a Gronwall argument to Equation (136): introduce

\begin{align}
\bar{C}_1(h) &:= 1 + 2(Lf)_+ + 5h\|f\|_{lip}^2 + \frac{8(1 + h)M}{M - 1} \|g\|_{\infty}^2 \tilde{K}_1 \tilde{K}_2(h), \\
\bar{C}_2(h) &:= \left\|Q^\perp_1\right\|^2 \left(2 + h + \frac{8h(1 + h)M}{M - 1} \|g\|_{\infty}^2 \tilde{K}_1\right),
\end{align}

then

\begin{equation}
V_k^a \leq \left(1 + h\bar{C}_1(h)\right)^k V_0^a + \sum_{j=1}^{k} \left(1 + h\bar{C}_1(h)\right)^{k-j} \left(\bar{C}_2(h)W_j + 2\|K_j\|^2 \left\|\mathcal{C}^2_1\right\|^2 V_j + 2N_j\right),
\end{equation}

By the Cauchy-Schwarz inequality we obtain

\begin{equation}
(V_k^a)^2 \leq 4 \left(1 + h\bar{C}_1(h)\right)^{2k} (V_0^a)^2 + 4\left(\sum_{j=1}^{k} \left(1 + h\bar{C}_1(h)\right)^{k-j} \bar{C}_2(h)W_j\right)^2
+ 16 \left(\sum_{j=1}^{k} \left(1 + h\bar{C}_1(h)\right)^{k-j} \|K_j\|^2 \left\|\mathcal{C}^2_1\right\|^2 V_j\right)^2 + 16 \left(\sum_{j=1}^{k} \left(1 + h\bar{C}_1(h)\right)^{k-j} N_j\right)^2.
\end{equation}

\begin{align}
&=: (I) + (II) + (III) + (IV).
\end{align}
Again we estimate for $h \ll 1$

\begin{equation}
(1 + h\tilde{C}_1(h))^k \leq e^{T\tilde{C}_1(1)}
\end{equation}

which gives

\begin{equation}
E[(I)] \leq e^{2T\tilde{C}_1(1)}E\left[ (\mathcal{V}_0^a)^2 \right].
\end{equation}

Since $\tilde{W}_k^{(i)} \sim \mathcal{N}(0, h\text{Id})$ this yields by (99)

\begin{equation}
E\left[ \left\| \tilde{W}_k^{(i)} - \bar{w}_k \right\|^4 \right] = 3 \left( \frac{M - 1}{M} h \right)^2
\end{equation}

and thus $E[(W_k)^2] \leq 3h^2$, therefore we obtain

\begin{equation}
E[(II)] \leq 12Te^{2T\tilde{C}_1(1)}\tilde{C}_2(1)^2.
\end{equation}

For (III) we use that also $E\left[ (V_k)^2 \right] \leq 3h^2$ and by (129)

\begin{equation}
\|K_j\|^4 \leq 2\tilde{K}_1^2 \left( \tilde{K}_2(h)^2 \left( \mathcal{V}_{j-1}^a \right)^2 + \|Q^a_1\|^4 (W_j)^2 \right).
\end{equation}

Thus by independence of $K_j$ and $V_j$ for any $j$ this yields

\begin{equation}
E[(III)] \leq 96Th \|C^a_1\|^4 e^{2T\tilde{C}_1(1)}\tilde{K}_2^2 \left( \tilde{K}_2(h)^2 \sum_{j=1}^k E\left[ (\mathcal{V}_{j-1}^a)^2 \right] + 3Th \|Q^a_1\|^4 \right).
\end{equation}

Now on (IV): recall that $\left( \tilde{W}_k^{(i)} \right)$ and $\left( \tilde{V}_k^{(i)} \right)$ are i.i.d. sequences independent of each other. Then by independence of Brownian increments, the Itô isometry as well as (129) it holds

\begin{equation}
E\left[ \left( \sum_{j=1}^k N_j \right)^2 \right] = \sum_{j=1}^k E\left[ (N_j)^2 \right]
\end{equation}

\begin{align}
= \sum_{j=1}^k \frac{1}{(M - 1)^2} \\
+ \sum_{i=1}^M E\left[ \int_0^{t_j} \left( \int_0^{t_j} \langle C_{k_j}^T K_j (X_{i_{j-1}}^{(i),a} - \bar{x}_{j-1}^a) + h (f (X_{i_{j-1}}^{(i),a}) - \bar{f}_{j-1}^a) \rangle 1_{\{s \in [t_{j-1}, t_j)\}}, dW_s^{(i)} \rangle \right) \right] \right]
\end{align}
\[
\leq \sum_{j=1}^{k} \frac{h}{M-1} \left( \left\| Q_k^j \right\|^2 \left( 1 + 2h(L_f)_+ + 4h^2 \left\| f \right\|_{\text{Lip}}^2 \right) E \left[ V_j^{n-1} \right] \\
+ 2 \left\| C_k^j \right\|^2 \left( \left( 1 + 2h(L_f)_+ + 4h^2 \left\| f \right\|_{\text{Lip}}^2 \right) E \left[ \left\| K_j \right\|^2 V_j^{n-1} \right] \\
+ 4h^2 \left\| g \right\|_{\infty}^2 \left\| f \right\|_{\text{Lip}}^2 \right) E \left[ \left\| K_j \right\|^4 \right] \right) \right)
\]
\[
\leq \sum_{j=0}^{k-1} \frac{h}{M-1} \left( 2 \left\| C_k^j \right\|^2 \left( 1 + 2h(L_f)_+ + 4h^2 \left\| f \right\|_{\text{Lip}}^2 \right) \left( 1 + 2 \left\| C_k^j \right\|^2 \right) \right)
\]
\[
+ \left\| Q_k^j \right\|^2 \left( 1 + 2h(L_f)_+ + 4h^2 \left\| f \right\|_{\text{Lip}}^2 \right) \left( 1 + 2 \left\| C_k^j \right\|^2 \right) \left\| g \right\|_{\infty}^2 \left\| f \right\|_{\text{Lip}}^2 \frac{M}{M-1} \left\| f \right\|_{\text{Lip}}^2 \right).
\]

Thus by the Cauchy-Schwarz inequality we obtain an estimate of the form
\[
(155) \quad E \left[ \left( V_k^n \right)^2 \right] \leq \tilde{C}_1(h) + h \tilde{C}_2(h) \sum_{j=0}^{k-1} E \left[ \left( V_j^n \right)^2 \right]
\]
which for \( h \ll 1 \) and by a Gronwall argument leads to
\[
(156) \quad E \left[ \left( V_k^n \right)^2 \right] \leq \tilde{C}_1(1) \exp \left( T \tilde{C}_2(1) \right)
\]
and this concludes the proof. \( \Box \)

**B.2. Modified formulation.**

**Lemma B.2.** If for the smallest eigenvalue of \( P_0^a \) it holds
\[
\lambda_{\text{min}} \left( P_0^a \right) > 0,
\]
then there exists a constant \( p^*_T > 0 \) such that
\[
(158) \quad \left\| (P_k^a)^{-1} \right\| \leq \frac{1}{p^*_T}
\]
and constants \( v_{T^*}^a, v_{T^*}^f \) (depending on \( \lambda_{\text{min}} \left( P_0^a \right) \) but independent of \( h \)) such that
\[
(159) \quad V_k^n \leq v_{T^*}^a \text{ and } V_k^f \leq v_{T^*}^f \quad \text{(see (173) and (174))}
\]
for all \( 1 \leq k \leq L \).

**Proof.** First of all note that \( \left\| (P_k^a)^{-1} \right\| = \left( \lambda_{\text{min}} \left( P_k^a \right) \right)^{-1} \) where \( \lambda_{\text{min}} \left( P_k^a \right) \), the smallest eigenvalue of \( P_k^a \), satisfies
\[
(160) \quad \lambda_{\text{min}} \left( P_k^a \right) = \inf_{v, \left\| v \right\| = 1} \left\langle P_k^a v, v \right\rangle.
\]
Thus let \( v \in \mathbb{R}^d \) with \( \| v \| = 1 \). By the following recursive formula (161)
\[
P_{k+1}^f = \frac{1}{M - 1} \sum_{i=1}^{M} \left( X_{i_{k+1}}^{(i),f} - \bar{x}_{k+1}^f \right) \left( X_{i_{k+1}}^{(i),f} - \bar{x}_{k+1}^f \right)^T
\]
\[
= P_k^a + h \left( \frac{1}{M - 1} \sum_{i=1}^{M} \left( X_{i_{k}}^{(i),a} - \bar{x}_{k}^a \right) \left( f \left( X_{i_{k}}^{(i),a} \right) - f \left( \bar{x}_{k}^a \right) \right)^T \right)
\]
\[
+ \left( f \left( X_{i_{k}}^{(i),a} \right) - f \left( \bar{x}_{k}^a \right) \right) \left( X_{i_{k}}^{(i),a} - \bar{x}_{k}^a \right)^T \right) + hQ
\]
\[
+ h^2 \frac{1}{M - 1} \sum_{i=1}^{M} \left( f \left( X_{i_{k}}^{(i),a} \right) - \bar{f}_k^a + \frac{1}{2} Q \left( P_k^{a} \right)^{-1} \left( X_{i_{k}}^{(i),a} - \bar{x}_{k}^a \right) \right)
\]
\[
\left( f \left( X_{i_{k}}^{(i),a} \right) - \bar{f}_k^a + \frac{1}{2} Q \left( P_k^{a} \right)^{-1} \left( X_{i_{k}}^{(i),a} - \bar{x}_{k}^a \right) \right)^T ,
\]
we obtain by the Young’s inequality for some \( \epsilon > 0 \) (162)
\[
\langle P_{k+1}^f, v \rangle \geq \langle P_k^a, v \rangle + 2h \frac{1}{M - 1} \sum_{i=1}^{M} \left( f \left( X_{i_{k}}^{(i),a} \right) - f \left( \bar{x}_{k}^a \right) , v \right) \langle X_{i_{k}}^{(i),a} - \bar{x}_{k}^a , v \rangle + h \langle Q, v \rangle
\]
\[
\geq \langle P_k^a, v \rangle - 2h \| f \|_{\text{Lip}} \left( \mathcal{V}_k^p \right)^{\frac{1}{2}} \langle P_k^a, v \rangle^{\frac{1}{2}} + h \langle Q, v \rangle
\]
\[
\geq \left( 1 - \frac{h}{\epsilon} \right) \langle P_k^a, v \rangle + h \left( \lambda_{\min} \left( Q \right) - \epsilon \| f \|_{\text{Lip}}^2 \mathcal{V}_k^p \right).
\]

By the specific structure of the Kalman gain matrix we obtain the following recursive form for the analyzed ensemble covariance matrix (163)
\[
P_{k+1}^a = \frac{1}{M - 1} \sum_{i=1}^{M} \left( X_{i_{k+1}}^{(i),a} - \bar{x}_{k+1}^a \right) \left( X_{i_{k+1}}^{(i),a} - \bar{x}_{k+1}^a \right)^T
\]
\[
= P_{k+1}^f - h \frac{1}{M - 1} \sum_{i=1}^{M} \left( X_{i_{k+1}}^{(i),f} - \bar{f}_{k+1}^f \right) \left( g \left( X_{i_{k+1}}^{(i),f} \right) - \bar{g}_{k+1}^f \right)^T K_{k+1}^T
\]
\[
+ \frac{h^2}{4} \frac{1}{M - 1} \sum_{i=1}^{M} K_{k+1} \left( g \left( X_{i_{k+1}}^{(i),f} \right) - \bar{g}_{k+1}^f \right) \left( g \left( X_{i_{k+1}}^{(i),f} \right) - \bar{g}_{k+1}^f \right)^T K_{k+1}^T
\]
thus
\begin{equation}
\langle P_{k+1}^a v, v \rangle \geq \langle P_{k+1}^f v, v \rangle
- \frac{h}{M - 1} \sum_{i=1}^{M} \langle X_{i,k+1}^{(i),f} - \bar{x}_{k+1}^f, v \rangle \langle K_{k+1} \left( g \left( X_{i,k+1}^{(i)} \right) - \bar{g}_{k+1}^f \right), v \rangle
\geq \langle P_{k+1}^f v, v \rangle
\end{equation}

Further we obtain with Algorithm 3.2
\begin{equation}
\mathcal{V}_{k+1}^f = \frac{1}{M - 1} \sum_{i=1}^{M} \| X_{i,k+1}^{(i),f} - \bar{x}_{k+1}^f \|^2
\leq \left( 1 + 2h(Lf)_+ + 4h^2 \| f \|_{\text{Lip}}^2 + h^2 \| f \|_{\text{Lip}} \| Q \| \left( P_k^a \right)^{-1} \right) \mathcal{V}_k^a
+ h \text{tr}(Q) + \frac{h^2}{4} \left( P_k^a \right)^{-1} \text{tr} \left( Q^2 \right)
\end{equation}

and
\begin{equation}
\mathcal{V}_{k+1}^a = \frac{1}{M - 1} \sum_{i=1}^{M} \| X_{i,k+1}^{(i),a} - \bar{x}_{k+1}^a \|^2
\leq \left( 1 + h^2 16 \| g \|_{\infty}^4 \| C^{-1} \|^2 M^2 \right) \mathcal{V}_{k+1}^f.
\end{equation}

Now assume that there exists a constant \( D > 0 \) such that
\begin{equation}
\lambda_{\text{min}} \left( P_k^a \right) \geq Dh.
\end{equation}

Then
\begin{equation}
\mathcal{V}_{k+1}^a \leq \exp \left( hC(1)(h, D) \right) \left( \mathcal{V}_k^a + hC(2)(D) \right)
\end{equation}

where
\begin{equation}
C(1)(h, D) := h \frac{16 \| g \|_{\infty}^4 \| C^{-1} \|^2 M^2}{(M - 1)^2} + 2(Lf)_+ + 4h \| f \|_{\text{Lip}}^2 + \frac{\| f \|_{\text{Lip}} \| Q \|}{D},
\end{equation}
\begin{equation}
C(2)(D) := \text{tr}(Q) + \frac{\text{tr}(Q^2)}{4D}.
\end{equation}

We make the ansatz
\begin{equation}
\mathcal{V}_k^a \leq \sum_{l=0}^{k-1} \exp \left( hC(1)(h, D)(k - l) \right) hC(2)(D).
\end{equation}
Then
\begin{equation}
\mathcal{V}_{k+1}^a \leq \sum_{l=0}^{k} \exp \left( hC^{(1)}(h, D)(((k + 1) - l) hC^{(2)}(D) \right)
\end{equation}
and for \( h \ll 1 \)
\begin{equation}
\mathcal{V}_{k+1}^a \leq C^{(2)}(D) \int_0^T e^{C^{(1)}(1, D)(T-s)} ds =: v_{a}^{*,D,T}.
\end{equation}
Using (165) in combination with Assumption (167) and the uniform bound on \( \mathcal{V}_k^a \), we also obtain the existence of a constant \( v_{f}^{*,D,T} \) such that
\begin{equation}
\mathcal{V}_{k+1}^f \leq v_{f}^{*,D,T}.
\end{equation}
This yields with Equations (162) and (164)
\begin{equation}
\langle P_{k+1}^a v, v \rangle \geq \left( 1 - \frac{h}{\epsilon} \right)^2 \langle P_k^a v, v \rangle + h \left( \frac{1 - h}{\epsilon} \right) \lambda_{\text{min}}(Q) - \epsilon \left( \|f\|_{\text{Lip}}^2 v_{f}^{*,D,T} + 16\|g\|_{\infty}^4 \|C^{-1}\|_2^2 M^2 v_{f}^{*,D,T} \right).
\end{equation}
Choose \( \epsilon > 0 \) such that for some \( \hat{h} > 0 \)
\begin{equation}
\frac{h}{\epsilon} \lambda_{\text{min}}(Q) + \epsilon \left( \|f\|_{\text{Lip}}^2 v_{f}^{*,D,T} + 16\|g\|_{\infty}^4 \|C^{-1}\|_2^2 M^2 v_{f}^{*,D,T} \right) \leq \frac{\lambda_{\text{min}}(Q)}{2}
\end{equation}
for all \( h < \hat{h} \), then
\begin{equation}
\langle P_{k+1}^a v, v \rangle \geq \left( 1 - \frac{h}{\epsilon} \right)^2 \langle P_k^a v, v \rangle + \frac{h}{\epsilon} \lambda_{\text{min}}(Q) \geq \left( \left( 1 - \frac{h}{\epsilon} \right)^2 + \frac{\lambda_{\text{min}}(Q)}{2D} \right) Dh.
\end{equation}
Observe that
\begin{equation}
\lim_{h \to 0} \left( 1 - \frac{h}{\epsilon} \right)^2 + \frac{\lambda_{\text{min}}(Q)}{2D} = 1 + \frac{\lambda_{\text{min}}(Q)}{2D} > 1,
\end{equation}
thus there exists an \( \tilde{h} > 0 \) such that for all \( h < \tilde{h} \) it holds
\begin{equation}
\langle P_{k+1}^a v, v \rangle \geq Dh. \quad \text{Since } v \text{ was chosen arbitrarily we may thus conclude}
\end{equation}
\begin{equation}
\lambda_{\text{min}} \left( P_{k+1}^a \right) \geq Dh.
\end{equation}
In total this yields that if
\begin{equation}
\lambda_{\text{min}} \left( P_0^a \right) \geq Dh,
\end{equation}
then
\begin{equation}
\lambda_{\text{min}} \left( P_k^a \right) \geq Dh.
for all $1 \leq k \leq L$ and $h < h^*$ with $h^* > 0$ small enough.

Further it holds
\begin{equation}
\langle P_{k+1}^a v, v \rangle \geq \left(1 - \frac{h}{\epsilon}\right)^{2(k+1)} \langle P_0^a v, v \rangle + \frac{\lambda_{\text{min}}(Q)}{2} \sum_{l=0}^{k} \left(1 - \frac{h}{\epsilon}\right)^{2(k-l)} h.
\end{equation}

Observe that for some $h_0 < \epsilon$ there exists a $\beta > 0$ such that
\begin{equation}
1 - \frac{h}{\epsilon} \geq \exp\left(-\left(\frac{1}{\epsilon} + \beta\right) h\right)
\end{equation}
for all $h < h_0$. Thus
\begin{equation}
\langle P_{k+1}^a v, v \rangle \geq \exp\left(-2\left(\frac{1}{\epsilon} + \beta\right)T\right) \langle P_0^a v, v \rangle + \frac{\lambda_{\text{min}}(Q)}{2} \int_0^{kh} \exp\left(-2\left(\frac{1}{\epsilon} + \beta\right)kh - s\right) ds
\end{equation}

Therefore if $h$ is small enough and $\lambda_{\text{min}}(P_0^a) > 0$, then
\begin{equation}
\lambda_{\text{min}}(P_k^a) \geq \exp\left(-2\left(\frac{1}{\epsilon} + \beta\right)T\right) \lambda_{\text{min}}(P_0^a) > 0
\end{equation}
and hence there exists a constant $p_T^* > 0$ such that
\begin{equation}
\| (P_k^a)^{-1} \| \leq \frac{1}{p_T^*}
\end{equation}
for all $1 \leq k \leq L$. \hfill \Box

Institut für Mathematik, Technische Universität Berlin, Straße des 17. Juni 136, D-10623 Berlin, and, Bernstein Center for Computational Neuroscience, Philippstr. 13, D-10115 Berlin, Germany

E-mail address: tlange@math.tu-berlin.de, stannat@math.tu-berlin.de