An Ensemble Kalman–Bucy Filter for correlated observation noise

Sebastian Ertel, Wilhelm Stannat

February 16, 2022

1 Introduction

The aim of filtering algorithms is to approximate the state of a signal from noisy and potentially incomplete or indirect observations. Ensemble Kalman–Bucy filters[1] (EnKBFs) accomplish this by employing a system of interacting particles, with the same initial value as the signal and which are propagated according to the signal dynamics with an added interaction term that nudges them in the direction of the observations.

EnKBFs are widely used in many scientific fields such as meteorology and the geosciences. In this work we derive an EnKBF for correlated observation noise and analyse its mean-field limit.

More precisely for an arbitrary timeframe $T > 0$ we consider a $\mathbb{R}^{d_x}$-valued signal process $(X_t)_{t\in[0,T]}$ determined by the stochastic differential equation (SDE)

$$\mathrm{d}X_t = B_t(X_t)\mathrm{d}t + C_t(X_t)\mathrm{d}W_t + \tilde{C}_t(X_t)\mathrm{d}V_t. \tag{1}$$

The drift function $B : [0,T] \times \mathbb{R}^{d_x} \to \mathbb{R}^{d_x}$ and the diffusion functions $C : [0,T] \times \mathbb{R}^{d_x} \to \mathbb{R}^{d_x \times d_w}, \tilde{C} : [0,T] \times \mathbb{R}^{d_x} \to \mathbb{R}^{d_x \times d_v}$ are Borel-measurable and satisfy the usual linear growth and global Lipschitz conditions with Lipschitz constants $\mathrm{Lip}(B), \mathrm{Lip}(C), \mathrm{Lip}\left(\tilde{C}\right)$. Furthermore W and V shall be two independent $\mathbb{R}^{d_w}$- and $\mathbb{R}^{d_v}$-dimensional Brownian motions.

The R^{d_y}-valued observation process $(Y_t)_{t\in[0,T]}$ is then given by

$$\mathrm{d}Y_t = H_t(X_t)\mathrm{d}t + \Gamma_t \mathrm{d}V_t,$$

where both $H : [0,T] \times \mathbb{R}^{d_x} \to \mathbb{R}^{d_y}$ and $\Gamma : [0,T] \to \mathbb{R}^{d_y \times d_v}$ are continuous. As is usual we assume that $Y_0 = 0$.

[1] The time continuous version of classical Ensemble Kalman filters.

The goal of stochastic filtering is to compute or at least approximate the conditional distribution of X_t given all past observations $(Y_s)_{s\in[0,t]}$, which we denote by

$$\eta_t := \mathbb{P}\left(\ X_t \in \cdot \mid Y_{0:t}\ \right),$$

for all times $t \in [0,T]$. η is called the posterior distribution. We shall denote the integral of a testfunction ϕ with respect to η_t by $\eta_t(\phi) := \int_{\mathbb{R}^{d_x}} \phi(x)\ \eta_t(\mathrm{d}x)$.

For any sufficiently regular and integrable testfunction ϕ the weak form of the Kushner–Stratonovich equation [BaCr], [NueReiRoz]

$$\begin{aligned}\mathrm{d}\eta_t(\phi) = \eta_t(L_t\phi)\mathrm{d}t + &\left(\eta_t\left(H_t^{\mathrm{T}}\phi\right) - \eta_t\left(H_t^{\mathrm{T}}\right)\eta_t\left(\phi\right) + \eta_t\left((\nabla\phi)^{\mathrm{T}}\tilde{C}_t\right)\right) \\ &\cdots R_t^{-1}\left(\mathrm{d}Y_t - \eta_t(H_t)\mathrm{d}t\right)\end{aligned} \tag{2}$$

is satsfied. Hereby L_t is the usual Fokker–Planck operator associated to (1) and $R_t := \Gamma_t\Gamma_t^{\mathrm{T}}$. In many applications the dimension of the signal d_x is too large so that an approximation of (2) using standard numerical PDE solvers becomes unfeasible. Instead practitioners rely on Monte-Carlo methods, among which the EnKBF has proven successful, to approximate η.

This paper is organized as follows:

- In the next section we formally derive a McKean–Vlasov equation that is consistent with (2), meaning that if a solution exists, its law will evolve according to (2). For the uncorrelated case $\tilde{C} = 0$ this has already been done in [PaStRe]. We extend their result to arbitrary $\tilde{C}$ and also give a fully vectorized description of all coefficients of the resulting McKean–Vlasov SDE.

- In section 3 we use this consistent representation to derive an EnKBF. We also discuss its well posedness and its mean-field limit.

- Section 4 is devoted to proving the well posedness of the McKean–Vlasov equation governing the mean-field limit of the previously derived EnKBF. For linear signals with Gaussian initial conditions and linear observations this result is well known. A proof for nonlinear signals that requires the observation function H to be bounded and employs a different version of the EnKBF can be found in [CoNiNuRe]. We extend the previous results to linear observations, which is a case that often occurs in practice, by using a mixture between a fixed point and a deterministic localization argument.

- In section 5 we derive a propagation of chaos result and discuss the inconsistency of the mean-field limit.

- Finally in section 6 we discuss how our well posedness result can be extended to other popular EnKBFs.

2 Mean-Field representation of the posterior

Just as in [PaStRe] we aim to represent the posterior through a McKean–Vlasov equation, i.e. we want to find a diffusion process $\left(\bar{X}_t\right)_{t\geq 0}$ such that for all times $t > 0$, its marginal law $\bar{\eta}_t$ is given by the posterior η_t. To this end we assume throughout this section that all functions appearing are sufficiently regular and integrable, so that we can always differentiate and integrate whenever necessary.

Let $\bar{W}$ be an independent copy of the Brownian motion W and $\bar{V}$ be an independent copy of V. To determine $\bar{X}$ we make the Ansatz

$$\mathrm{d}\bar{X}_t = B_t(\bar{X}_t)\mathrm{d}t + C_t(\bar{X}_t)\mathrm{d}\bar{W}_t + \tilde{C}_t(\bar{X}_t)\mathrm{d}\bar{V}_t + K_t\left(\bar{X}_t, \bar{\eta}_t\right)\mathrm{d}Y_t + a_t\left(\bar{X}_t, \bar{\eta}_t\right)\mathrm{d}t,$$

for two functions K and a depending on both the state and the (marginal) law of the process $\bar{X}$. Furthermore let us make the following notational convention.

Notation 1. *For any function f, integrable with respect to the joint law of $\bar{W}$ and $\bar{V}$, its expectation with respect to said law shall be denoted by*

$$\mathbb{E}_{\cancel{Y}}[f] := \int f(\bar{w}, \bar{v})\ \mathbb{P}^{\bar{W}}(\mathrm{d}\bar{w})\ \mathbb{P}^{\bar{V}}(\mathrm{d}\bar{v}).$$

Since $\bar{W}$ and $\bar{V}$ are jointly independent of Y this means that the expectation of a function f with respect to the joint law of Y, $\bar{V}$ and $\bar{W}$ is given by

$$\mathbb{E}[f] = \int \mathbb{E}_{\cancel{Y}}[f(y,\cdot,\cdot)]\ \mathbb{P}^{Y}(\mathrm{d}y).$$

Let $\phi : \mathbb{R}^{d_x} \to \mathbb{R}$ be some testfunction. Using Itô's formula and taking the expectation with respect to the joint law of $\bar{W}$ and $\bar{V}$, we derive for $\bar{\eta}_t(\phi) := \mathbb{E}_{\cancel{Y}}\left[\phi\left(\bar{X}_t\right)\right]$ the equation

$$\begin{aligned}\mathrm{d}\bar{\eta}_t(\phi) = \bar{\eta}_t(L_t\phi)\mathrm{d}t + \bar{\eta}_t\left(\nabla\phi\cdot K_t\left(\cdot,\bar{\eta}_t\right)\right)\mathrm{d}Y_t + \bar{\eta}_t\left(\nabla\phi\cdot a_t\left(\cdot,\bar{\eta}_t\right)\right)\mathrm{d}t + \cdots \\ \cdots + \frac{1}{2}\bar{\eta}_t\left(\mathrm{tr}\left[\phi'' K_t\left(\cdot,\bar{\eta}_t\right)R_t K_t\left(\cdot,\bar{\eta}_t\right)^{\mathrm{T}}\right]\right)\mathrm{d}t.\end{aligned} \tag{3}$$

Hereby ϕ'' denotes the Hessian of ϕ.

Since $\bar{\eta}_t$ shall coincide with the posterior η_t, it must also adhere to the Kushner–Stratonovich equation (2). By comparing the terms on the right-hand side of both equations (2) and (3), we get the two consistency conditions

$$\bar{\eta}_t\left(\nabla\phi\cdot K_t\left(\cdot,\bar{\eta}_t\right)\right) = \left(\bar{\eta}_t\left(H_t^{\mathrm{T}}\phi\right) - \bar{\eta}_t\left(H_t^{\mathrm{T}}\right)\bar{\eta}_t\left(\phi\right) + \bar{\eta}_t\left(\left(\nabla\phi\right)^{\mathrm{T}}\tilde{C}_t\right)\right)R_t^{-1} \tag{4}$$

and

$$\begin{aligned}\bar{\eta}_t\left(\nabla\phi\cdot a_t\left(\cdot,\bar{\eta}_t\right)\right) + \frac{1}{2}\bar{\eta}_t\left(\mathrm{tr}\left[\phi'' K_t\left(\cdot,\bar{\eta}_t\right)R_t K_t\left(\cdot,\bar{\eta}_t\right)^{\mathrm{T}}\right]\right) \\ = -\left(\bar{\eta}_t\left(H_t^{\mathrm{T}}\phi\right) - \bar{\eta}_t\left(H_t^{\mathrm{T}}\right)\bar{\eta}_t\left(\phi\right) + \bar{\eta}_t\left(\left(\nabla\phi\right)^{\mathrm{T}}\tilde{C}_t\right)\right)R_t^{-1}\bar{\eta}_t\left(H_t\right)\end{aligned} \tag{5}$$

Assuming that $\bar{\eta}_t$ admits a sufficiently regular density, which shall also be denoted by $\bar{\eta}_t$, we derive more direct characterizations of the two terms K and a in the following.

2.1 Consistency of the Kalman-gain

First we investigate the consistency condition (4) for the Kalman gain term K. To this end we make the following notational conventions that will be used throughout this section.

Notation 2. *Throughout this paper the divergence of a Matrix-valued function A shall be interpreted columnwise, i.e. if $A(x) \in \mathbb{R}^{m\times k}$ we have*

$$\mathrm{div}(A) := (\mathrm{div}(A_{\cdot,1}), \cdots, \mathrm{div}(A_{\cdot,k})) := \left(\sum_{j=1}^{m} \partial_{x_j} A_{j,1}, \cdots, \sum_{j=1}^{m} \partial_{x_j} A_{j,k}\right).$$

Furthermore we also interpret the scalar product between a matrix and a vector columnwise. Thus the scalar product between A and the gradient ∇ gives to following row-vector-valued differential operator

$$A \cdot \nabla := (A_{\cdot,1} \cdot \nabla, \cdots, A_{\cdot,k} \cdot \nabla) := \left(\sum_{j=1}^{m} A_{j,1}\partial_{x_j}, \cdots, \sum_{j=1}^{m} A_{j,k}\partial_{x_j}\right).$$

Employing integration by parts and the Fundamental Theorem of Calculus of Variations, we see that (4) can be interpreted as the weak form of the partial differential equation

$$-\mathrm{div}\left(\bar{\eta}_t K_t(\cdot, \bar{\eta}_t)\right) + \mathrm{div}\left(\bar{\eta}\tilde{C}_t\right) R_t^{-1} = \left(H^{\mathrm{T}} - \bar{\eta}_t(H^{\mathrm{T}})\right) R_t^{-1}\bar{\eta}_t \tag{6}$$

If we denote the dependence of K on $\tilde{C}$ by $K^{\tilde{C}}$, then we clearly have

$$K_t^{\tilde{C}} = K_t^0 + \tilde{C}_t R_t^{-1}, \tag{7}$$

with

$$-\mathrm{div}\left(\bar{\eta}_t K_t^0(\cdot, \bar{\eta}_t)\right) = \left(H_t^{\mathrm{T}} - \bar{\eta}_t(H_t^{\mathrm{T}})\right) R_t^{-1}\bar{\eta}_t. \tag{8}$$

Thus $K^{\tilde{C}}$ is just a translation of K^0, the Kalman gain for the uncorrelated case, which in turn is defined uniquely up to $\bar{\eta}_t$-harmonic functions.

2.2 Consistency of the correctional transport term

Next we investigate the consistency equation (5) for the correctional transport term a. Note that while K also shows up in this equation, it is already fully determined by (4).

By again using integration by parts and the Fundamental Theorem of Calculus of Variations, we derive the differential form of consistency condition (4)

$$
\begin{aligned}
&-\operatorname{div}\left(\bar{\eta}_t a_t\left(\cdot, \bar{\eta}_t\right)\right)+\frac{1}{2} \sum_{i, j=1}^{d_x} \partial_{x_i x_j}^2\left(\bar{\eta}_t K_t\left(\cdot, \bar{\eta}_t\right) R_t K_t\left(\cdot, \bar{\eta}_t\right)^{\mathrm{T}}\right)_{j i} \\
&\cdots-\operatorname{div}\left(\bar{\eta}_t \tilde{C}_t\right) R_t^{-1} \bar{\eta}_t\left(H_t\right)=-\left(H_t^{\mathrm{T}}-\bar{\eta}_t(H_t^{\mathrm{T}})\right) R_t^{-1} \bar{\eta}_t(H_t) \bar{\eta}_t .
\end{aligned} \tag{9}
$$

We aim to simplify this expression. Therefore we first simply use the product rule to derive

$$
\begin{aligned}
&\sum_{i, j=1}^{d_x} \partial_{x_i x_j}^2\left(\bar{\eta}_t K_t\left(\cdot, \bar{\eta}_t\right) R_t K_t\left(\cdot, \bar{\eta}_t\right)^{\mathrm{T}}\right)_{j i} \\
&=\sum_{j, l} \partial_{x_j}\left(\operatorname{div}\left(\bar{\eta}_t K_t\left(\cdot, \bar{\eta}_t\right)\right)_l\left(R_t K_t^{\mathrm{T}}\left(\cdot, \bar{\eta}_t\right)\right)_{l j}\right)+\cdots \\
&\cdots+\sum_j \partial_{x_j} \bar{\eta}_t \sum_l\left(\sum_i K_t\left(\cdot, \bar{\eta}_t\right)_{i l} \partial_{x_i}\left(R_t K_t^{\mathrm{T}}\left(\cdot, \bar{\eta}_t\right)\right)_{l j}\right) .
\end{aligned}
$$

Note that since

$$
\sum_{l=1}^{d_y}\left(\sum_{i=1}^{d_x} K_t\left(\cdot, \bar{\eta}_t\right)_{i l} \partial_{x_i}\left(R_t K_t^{\mathrm{T}}\left(\cdot, \bar{\eta}_t\right)\right)_{l j}\right)=\left(\left(K_t\left(\cdot, \bar{\eta}_t\right) \cdot \nabla\right) R_t K_t^{\mathrm{T}}\left(\cdot, \bar{\eta}_t\right)\right)_j
$$

we can rewrite this equality more compactly as

$$
\begin{aligned}
&\sum_{i, j=1}^{d_x} \partial_{x_i x_j}^2\left(\bar{\eta}_t K_t\left(\cdot, \bar{\eta}_t\right) R_t K_t\left(\cdot, \bar{\eta}_t\right)^{\mathrm{T}}\right)_{j i} \\
&=\operatorname{div}\left(K_t\left(\cdot, \bar{\eta}_t\right)\ R_t \operatorname{div}\left(\bar{\eta}_t K_t\left(\cdot, \bar{\eta}_t\right)\right)^{\mathrm{T}}\right)+\operatorname{div}\left(\bar{\eta}_t\left(\left(K_t\left(\cdot, \bar{\eta}_t\right) \cdot \nabla\right) R_t K_t^{\mathrm{T}}\left(\cdot, \bar{\eta}_t\right)\right)^{\mathrm{T}}\right) .
\end{aligned}
$$

Using the consistency equation of the Kalman gain term (6) gives us the identity

$$
\begin{aligned}
&\sum_{i, j=1}^{d_x} \partial_{x_i x_j}^2\left(\bar{\eta}_t K_t\left(\cdot, \bar{\eta}_t\right) R_t K_t\left(\cdot, \bar{\eta}_t\right)^{\mathrm{T}}\right)_{j i} \\
&=-\operatorname{div}\left(\bar{\eta}_t K_t\left(\cdot, \bar{\eta}_t\right)\left(H_t-\bar{\eta}_t(H_t)\right)\right)+\operatorname{div}\left(K_t\left(\cdot, \bar{\eta}_t\right) \operatorname{div}\left(\bar{\eta}_t \tilde{C}_t\right)^{\mathrm{T}}\right) \\
&\cdots+\operatorname{div}\left(\bar{\eta}_t\left(\left(K_t\left(\cdot, \bar{\eta}_t\right) \cdot \nabla\right) R_t K_t^{\mathrm{T}}\left(\cdot, \bar{\eta}_t\right)\right)^{\mathrm{T}}\right) .
\end{aligned}
$$

Thus we can rewrite (9) into

$$
\begin{aligned}
&-\operatorname{div}\left(\bar{\eta}_t a_t\left(\cdot,\bar{\eta}_t\right)\right)-\frac{\operatorname{div}\left(\bar{\eta}_t K_t\left(\cdot,\bar{\eta}_t\right)\left(H_t-\bar{\eta}_t(H_t)\right)\right)}{2}-\operatorname{div}\left(\bar{\eta}_t\tilde{C}_t\right)R_t^{-1}\bar{\eta}_t\left(H_t\right)\\
&\cdots+\frac{\operatorname{div}\left(K_t\left(\cdot,\bar{\eta}_t\right)\operatorname{div}\left(\bar{\eta}_t\tilde{C}_t\right)^{\mathrm{T}}\right)}{2}+\frac{\operatorname{div}\left(\bar{\eta}_t\left(\left(K_t\left(\cdot,\bar{\eta}_t\right)\cdot\nabla\right)R_t K_t^{\mathrm{T}}\left(\cdot,\bar{\eta}_t\right)\right)^{\mathrm{T}}\right)}{2}\\
&=-\left(H_t^{\mathrm{T}}-\bar{\eta}_t(H_t^{\mathrm{T}})\right)R_t^{-1}\bar{\eta}_t(H_t)\bar{\eta}_t.
\end{aligned}
$$

Just as in [PaStRe] one sees immediatley that

$$
\begin{aligned}
a_t\left(\cdot,\bar{\eta}_t\right)=&-\frac{K_t\left(\cdot,\bar{\eta}_t\right)\left(H_t+\bar{\eta}_t(H_t)\right)}{2}+\frac{\left(\left(K_t\left(\cdot,\bar{\eta}_t\right)\cdot\nabla\right)R_t K_t^{\mathrm{T}}\left(\cdot,\bar{\eta}_t\right)\right)^{\mathrm{T}}}{2}\\
&\cdots+\frac{K_t\left(\cdot,\bar{\eta}_t\right)\operatorname{div}\left(\bar{\eta}_t\tilde{C}_t\right)^{\mathrm{T}}}{2\ \bar{\eta}_t}+\Omega_t^0,
\end{aligned}
\tag{10}
$$

where Ω_t^0 is an arbitrary $\bar{\eta}_t$-harmonic field.

Remark 3. *We note that*

$$
\frac{\operatorname{div}\left(\bar{\eta}_t\tilde{C}_t\right)}{\bar{\eta}_t}=\left(\nabla\log\bar{\eta}_t\right)^{\mathrm{T}}\tilde{C}_t+\operatorname{div}\left(\tilde{C}_t\right)
$$

Therefore we can rewrite (10) *into*

$$
\begin{aligned}
a_t\left(\cdot,\bar{\eta}_t\right)=&-\frac{K_t\left(\cdot,\bar{\eta}_t\right)\left(H_t+\bar{\eta}_t(H_t)-\left(\tilde{C}_t^{\mathrm{T}}\nabla\log\bar{\eta}_t+\operatorname{div}\left(\tilde{C}_t\right)^{\mathrm{T}}\right)\right)}{2}+\cdots\\
&\cdots+\frac{\left(K_t\left(\cdot,\bar{\eta}_t\right)\cdot\nabla\left(R_t K_t^{\mathrm{T}}\left(\cdot,\bar{\eta}_t\right)\right)\right)^{\mathrm{T}}}{2}+\Omega_t^0.
\end{aligned}
$$

This also shows that, just as stated in [NueReiRoz], for constant $\tilde{C}_t(x)=\tilde{C}_t$, the correlated observation case can be intepreted in terms of the uncorrelated case with a modfied observation map $\tilde{H}_t:=H_t-\tilde{C}_t\nabla\log\eta_t$.

Remark 4. *Note that in the case of one-dimensional observations $d_y=1$, the term R_t is scalar and $K_t\left(\cdot,\bar{\eta}_t\right)$ is a $\mathbb{R}^{d_x}$ valued function, one sees immediatley that*

$$
\left(\left(K_t\left(\cdot,\bar{\eta}_t\right)\cdot\nabla\right)R_t K_t^{\mathrm{T}}\left(\cdot,\bar{\eta}_t\right)\right)^{\mathrm{T}}=R_t\left(K_t\left(\cdot,\bar{\eta}_t\right)\cdot\nabla\right)K_t\left(\cdot,\bar{\eta}_t\right).
$$

This is the convective change of the field K under a flow with velocity K.

3 The ensemble Kalman-Bucy filter for correlated observations

When B and H are linear and $C, \tilde{C}$ are constant, it is well known that if η_0 is Gaussian, then the posterior η_t will also be Gaussian for all times $t \geq 0$. We denote its mean by $\bar{m}_t \in \mathbb{R}^{d_x}$ and its covariance by $\bar{P}_t \in \mathbb{R}^{d_x \times d_x}$, i.e. $\bar{\eta}_t = \eta_t = \mathcal{N}\left(\bar{m}_t, \bar{P}_t\right)$.

Note that in this case one can easily verify that for any matrix $A \in \mathbb{R}^{d_x \times k}$ with $k \in \mathbb{N}$ arbitrary, it holds that

$$\operatorname{div}\left(A\bar{\eta}_t\right) = -\left(x - \bar{m}_t\right)^{\mathrm{T}} \bar{P}_t^{-1}\, A\bar{\eta}_t. \tag{11}$$

This motivates the Ansatz $K_t^0 = \bar{P}_t H_t^{\mathrm{T}} \bar{\eta}_t$, which in turn results in the Kalman gain

$$K_t = \left(\bar{P}_t H_t^{\mathrm{T}} + \tilde{C}_t\right) R_t^{-1}.$$

Since K_t only depends on the distribution $\bar{\eta}_t$, but not on the state variable $\bar{X}_t$, it follows that

$$\left(\left(K_t\left(\cdot, \bar{\eta}_t\right) \cdot \nabla\right) R_t K_t^{\mathrm{T}}\left(\cdot, \bar{\eta}_t\right)\right)^{\mathrm{T}} = 0$$

and therefore we obtain by using (11) and by setting $\Omega^0 = 0$

$$a_t\left(x, \bar{\eta}_t\right) = -\left(\bar{P}_t H_t^{\mathrm{T}} + \tilde{C}_t\right) R_t^{-1} \left(\frac{H_t\left(x + \bar{m}_t\right)}{2} + \tilde{C}_t^{\mathrm{T}} \bar{P}_t^{-1} \frac{x - \bar{m}_t}{2}\right).$$

Thus in the linear Gaussian case the following corollary holds.

Corollary 5. *Assume that X_0 is Gaussian, B_t, H_t are linear and $C_t, \tilde{C}_t$ are constant. The solution $\bar{X}$ to*

$$\mathrm{d}\bar{X}_t = B_t \bar{X}_t \mathrm{d}t + C_t \mathrm{d}\bar{W}_t + \tilde{C}_t \mathrm{d}\bar{V}_t + \left(\bar{P}_t H_t^{\mathrm{T}} + \tilde{C}_t\right) R_t^{-1} \left(\mathrm{d}Y_t - \frac{H_t\left(\bar{X}_t + \bar{m}_t\right)}{2}\mathrm{d}t\right)$$
$$\cdots - \left(\bar{P}_t H_t^{\mathrm{T}} + \tilde{C}_t\right) R_t^{-1} \tilde{C}_t^{\mathrm{T}} \bar{P}_t^{-1} \frac{\bar{X}_t - \bar{m}_t}{2} \mathrm{d}t. \tag{12}$$

satisfying the initial condition $\mathrm{Law}\left(\bar{X}_0\right) = \eta_0 = \mathrm{Law}\left(X_0\right)$*, satisfies* $\mathrm{Law}\left(\bar{X}_t\right) = \eta_t$ *for all time $t \geq 0$ and is thus a consistent mean-field representation of the posterior η.*

□

From this corollary one can deduce that the mean $\bar{m}$ and the covariance $\bar{P}$ satisfy the Kalman–Bucy equations [Jaz, chapter 7, page 228]

$$\mathrm{d}\bar{m}_t = B_t \bar{m}_t \mathrm{d}t + \left(\bar{P}_t H_t^{\mathrm{T}} + \tilde{C}_t\right) R_t^{-1} \left(\mathrm{d}Y_t - H_t \bar{m}_t \mathrm{d}t\right) \tag{13a}$$

$$\frac{\mathrm{d}\bar{P}_t}{\mathrm{d}t} = B_t\bar{P}_t + \bar{P}_t B_t^{\mathrm{T}} + C_t C_t^{\mathrm{T}} + \tilde{C}_t \tilde{C}_t^{\mathrm{T}} - \left(\bar{P}_t H_t^{\mathrm{T}} + \tilde{C}_t\right) R_t^{-1} \left(H_t \bar{P}_t + \tilde{C}_t^{\mathrm{T}}\right), \quad (13b)$$

which again shows the consistency of $\bar{X}$ in the linear Gaussian case.

Of course one can consider system (13) detached from (12) and easily show the global existence and uniqueness of solutions. This can then be used to prove the global existence and uniqueness of solutions to (12) as the following remark shows.

Remark 6. *To prove global existence and uniqueness of solutions to (12), one can argue similar to [CrDMJaRu, Remark 2.1] and make a fixpoint Ansatz that linearizes the problem by replacing the covariance matrix with a symmetric positive definite (spd) matrix P_t. This results in a linear McKean-Vlasov equation, which admits a unique global solution $\bar{X}^P$.*
Since solutions to (13b), starting from a regular matrix $\bar{P}_0$, are always spd, it is possible to choose P to be the solution of (13b). One can then verify easily that $\bar{X}^P$ indeed satisfies (12).
Uniqueness can easily be derived from the uniqueness of solutions to (13b) and the uniqueness of solutions to linear McKean–Vlasov equations.

To approximate (12) one uses an interacting particle system. For $M \in \mathbb{N}$ let $W^i,\ i = 1, \cdots, M$ and $V^i,\ i = 1, \cdots, M$ be independent copies of W and V. Let $X^i,\ i = 1, \cdots, M$ be the solution of

$$\begin{aligned} \mathrm{d}X_t^i &= B_t X_t^i \mathrm{d}t + C_t \mathrm{d}W_t^i + \tilde{C}_t \mathrm{d}V_t^i \\ &\cdots + \left(P_t^M H_t^{\mathrm{T}} + \tilde{C}_t\right) R_t^{-1} \left(\mathrm{d}Y_t - \frac{H_t\left(X_t^i + x_t^M\right)}{2}\mathrm{d}t\right) \\ &\cdots - \left(P_t^M H_t^{\mathrm{T}} + \tilde{C}_t\right) R_t^{-1} \tilde{C}_t^{\mathrm{T}} \left(P_t^M\right)^+ \frac{X_t^i - x_t^m}{2}\mathrm{d}t,\ i = 1, \cdots, M \end{aligned} \quad (14)$$

where

$$x_t^M := \frac{1}{M}\sum_{i=1}^M X_t^i,\ P_t^M := \frac{1}{M-1}\sum_{i=1}^M \left(X_t^i - x_t^M\right)\left(X_t^i - x_t^M\right)^{\mathrm{T}}$$

denote the ensemble average and the ensemble covariance matrix. $\left(P_t^M\right)^+$ denotes the Moore–Penrose pseudoinverse of P_t^M.

System (14) is called the Ensemble Kalman–Bucy filter (EnKBF). It can easily be generalized to nonlinear signals

$$\begin{aligned} \mathrm{d}X_t^i &= B_t\left(X_t^i\right)\mathrm{d}t + C_t\left(X_t^i\right)\mathrm{d}W_t^i + \tilde{C}_t\left(X_t^i\right)\mathrm{d}V_t^i \\ &\cdots + \left(P_t^M H_t^{\mathrm{T}} + \tilde{C}_t\left(X_t^i\right)\right) R_t^{-1}\left(\mathrm{d}Y_t - \frac{H_t\left(X_t^i + x_t^M\right)}{2}\mathrm{d}t\right) \\ &\cdots - \left(P_t^M H_t^{\mathrm{T}} + \tilde{C}_t\left(X_t^i\right)\right)\ R_t^{-1}\ \tilde{C}_t^{\mathrm{T}}\left(X_t^i\right)\left(P_t^M\right)^+ \frac{X_t^i - x_t^M}{2}\mathrm{d}t \end{aligned} \tag{15}$$

for $i = 1, \cdots, M$.

While in this case the EnKBF does not approximate a consistent mean-field representation of the posterior, it is still widely used in practice.

Equation (15) is a system of nonlinear SDEs, in which the coefficents do not satisfy linear growth properties and which therefore falls outside the standard existence theory for SDEs. Nevertheless in [LaSt] the existence and uniqueness of solutions was proven for uncorrelated observation noise. More precisely it was assumed that C is constant and $\tilde{C} = 0$.

However, generalizing this result to the correlated case (15) is not straightforward, as the pseudoinverse $\left(P_t^M\right)^+$ does not depend continuously on the ensemble members X^i, $i = 1, \cdots, M$ and actually has singularities where P_t^M changes its rank. Indeed for $M = 2$ it is easy to see that

$$\left(P_t^M\right)^+\left(X_t^i - x_t^M\right) = \mathbb{1}_{[X_t^1 \neq X_t^2]}\frac{\left(X_t^i - x_t^M\right)}{2\left|X_t^i - x_t^M\right|^2},\ i = 1, 2,$$

which becomes singular when the two particles collide.

To avoid this issue we use the following regularized matrix inversion.

Notation 7. *Let P be an arbitrary symmetric positive semidefinite (spsd) matrix. Then we define for any $\epsilon > 0$ and $n \in \mathbb{N}$*

$$P^{+\epsilon,n} := \left(P^n + \epsilon I\right)^{-1} P^{n-1}.$$

Note that for $n = 2$ this is indeed an approximation of the pseudoinverse as $\lim_{\epsilon\to 0} P^{+\epsilon,2} = P^+$.

Furthermore let us make the following notational and algebraic remark.

Remark 8. *Throughout this paper $|\cdot|$ shall denote the standard Euclidian norm. If the input is a matrix $A \in \mathbb{R}^{m\times k}$ the result is the Frobeniusnorm $|A| = \sqrt{\sum_{i=1}^m \sum_{j=1}^k A_{ij}^2} = \sqrt{\mathrm{tr}\left[AA^{\mathrm{T}}\right]}$. It is easy to see that for any quadratic matrix $A \in \mathbb{R}^{k\times k}$, the trace can be bounded by the Frobeniusnorm in the following way $\mathrm{tr}A \leq \sqrt{k}|A|$. If, on the other hand, $A \in \mathbb{R}^{k\times k}$ is spsd, then, by using the singular value decomposition, one sees immediatley that $|A| \leq \mathrm{tr}A$.*

As $P^{+\epsilon,n}$ depends Lipschitz-continuously on P, the following theorem is easy to prove.

Theorem 9. *Beside the standard Lipschitz assumptions on the coefficients, let us assume that C is bounded and that $\tilde{C}$ is constant. Then for every $\epsilon > 0$ and $n \in \mathbb{N}$ there exists a unique solution to*

$$\begin{aligned}
\mathrm{d}X_t^i &= B_t\left(X_t^i\right)\mathrm{d}t + C_t\left(X_t^i\right)\mathrm{d}W_t^i + \tilde{C}_t \mathrm{d}V_t^i \\
&\cdots + \left(P_t^M H_t^{\mathrm{T}} + \tilde{C}_t\left(X_t^i\right)\right) R_t^{-1}\left(\mathrm{d}Y_t - \frac{H_t\left(X_t^i + x_t^M\right)}{2}\mathrm{d}t\right) \\
&\cdots - \left(P_t^M H_t^{\mathrm{T}} + \tilde{C}_t\right)\ R_t^{-1}\ \tilde{C}_t^{\mathrm{T}}\left(P_t^M\right)^{+\epsilon,n}\frac{X_t^i - x_t^M}{2}\mathrm{d}t
\end{aligned} \tag{16}$$

on every fixed time intervall $[0,T]$.

Proof. Denote by ξ the explosion time of the ensemble. First we note that by the linearity of stochastic differentials the following equation for the mean holds up to the explosion time ξ

$$\begin{aligned}
\mathrm{d}x_t^M &= \underbrace{\frac{1}{M}\sum_{i=1}^M B_t(X_t^i)\,\mathrm{d}t}_{:=b_t^M} + \frac{1}{M}\sum_{i=1}^M C_t(X_t^i)\mathrm{d}W_t^i + \tilde{C}_t\mathrm{d}V_t^i + \cdots \\
&\cdots + \left(P_t^M H_t^{\mathrm{T}} + \tilde{C}_t\right) R_t^{-1}\left(\mathrm{d}Y_t - H_t x_t^M \mathrm{d}t\right),
\end{aligned}$$

which gives us

$$\begin{aligned}
\mathrm{d}\left(X_t^i - x_t^M\right) &= \left(B_t(X_t^i) - b_t^M\right)\mathrm{d}t + \frac{1}{M}\sum_{j=1}^M\left(C_t(X^i)\mathrm{d}W_t^i - C_t(X^j)\mathrm{d}W_t^j\right) \\
\cdots + \frac{1}{M}\sum_{j=1}^M\left(\tilde{C}\mathrm{d}V_t^i - \tilde{C}\mathrm{d}V_t^j\right) &- \left(P_t^M H_t^{\mathrm{T}} + \tilde{C}_t\right) R_t^{-1}\left(H_t + \tilde{C}_t^{\mathrm{T}}\left(P_t^M\right)^{+\epsilon,n}\right)\frac{X_t^i - x_t^M}{2}\mathrm{d}t.
\end{aligned}$$

By employing Itô's formula we derive the following evolution equation for the empirical covariance matrix up to explosion time ξ

$$\begin{aligned}
\mathrm{d}P_t^M &= \ll B_t, X_t \gg_N \mathrm{d}t - \frac{1}{2}\left(P_t^M H_t^{\mathrm{T}} + \tilde{C}_t\right) R_t^{-1}\left(H_t + \tilde{C}_t^{\mathrm{T}}\left(P_t^M\right)^{+\epsilon,n}\right) P_t^M \mathrm{d}t \\
\cdots &- \frac{1}{2}P_t^M\left(H_t^{\mathrm{T}} + \left(P_t^M\right)^{+\epsilon,n}\tilde{C}_t\right) R_t^{-1}\left(H_t P_t^M + \tilde{C}_t^{\mathrm{T}}\right)\mathrm{d}t + \frac{M}{M-1}\tilde{C}_t\tilde{C}_t^{\mathrm{T}}\mathrm{d}t \\
\cdots &+ \frac{1}{M^2}\sum_{i,j=1}^M\left(C_t(X_t^i)C_t(X_t^i)^{\mathrm{T}} + \frac{C_t(X_t^j)C_t(X_t^j)^{\mathrm{T}}}{M-1}\right)\mathrm{d}t + \mathrm{d}\mathfrak{m}_t,
\end{aligned}$$

where

$$\ll B_t, X_t \gg_N := \frac{1}{M-1}\sum_{i=1}^{M}\left(B_t(X_t^i) - b_t^M\right)\left(X_t^i - x_t^M\right)^{\mathrm{T}} \mathrm{d}t$$
$$\cdots + \frac{1}{M-1}\sum_{i=1}^{M}\left(X_t^i - x_t^M\right)\left(B_t(X_t^i) - b_t^M\right)^{\mathrm{T}} \mathrm{d}t.$$

and $\mathfrak{lm}$ denotes a local martingale whose concrete form does not matter for the following computations.

First we note that we can replace b_t^M by $B_t(x_t^M)$ in the definition of $\ll B_t, X_t \gg_N$ and thus obtain

$$\mathrm{tr} \ll B_t, X_t \gg_N \leq\ 2\mathrm{Lip}\left(B_t\right)\frac{1}{M-1}\sum_{i=1}^{M}\left|X_t^i - x_t^M\right|^2 = 2\mathrm{Lip}\left(B_t\right)\mathrm{tr}P_t^M.$$

Furthermore we note that since $P_t^M H_t^{\mathrm{T}} R_t^{-1} H_t P_t^M$ is spsd and since $\left|\left(P_t^M\right)^{+\epsilon,n} P_t^M\right| \leq 2\sqrt{d_x}$, we have

$$\begin{aligned}
&-\mathrm{tr}\left[\left(P_t^M H_t^{\mathrm{T}} + \tilde{C}_t\right) R_t^{-1}\left(H_t + \tilde{C}_t^{\mathrm{T}}\left(P_t^M\right)^{+}\right) P_t^M\right]\\
&\quad = -\mathrm{tr}\left[P_t^M H_t^{\mathrm{T}} R_t^{-1} H_t P_t^M\right] - \mathrm{tr}\left[P_t^M H_t^{\mathrm{T}} R_t^{-1}\tilde{C}_t^{\mathrm{T}}\left(P_t^M\right)^{+\epsilon,n} P_t^M\right]\\
&\quad \cdots - \mathrm{tr}\left[\tilde{C}_t R_t^{-1} H_t P_t^M\right] - \mathrm{tr}\left[\tilde{C}_t R_t^{-1}\tilde{C}_t^{\mathrm{T}}\left(P_t^M\right)^{+\epsilon,n} P_t^M\right]\\
&\quad \leq 4 d_x\left|\tilde{C}_t R_t^{-1} H_t\right|\ \mathrm{tr}\left[P_t^M\right] + 2 d_x\left|\tilde{C}_t R_t^{-1}\tilde{C}_t^{\mathrm{T}}\right|.
\end{aligned}$$

This gives us the stochastic differential inequality

$$\mathrm{dtr}P_t^M \leq 2\left(\mathrm{Lip}\left(B_t\right) + 4 d_x\left|\tilde{C}_t R_t^{-1} H_t\right|\right)\mathrm{tr}P_t^M \mathrm{d}t + 2 d_x\left|\tilde{C}_t R_t^{-1}\tilde{C}_t^{\mathrm{T}}\right|\mathrm{d}t + \mathrm{d}\mathfrak{lm}_t$$
$$\cdots + \sup_{s\leq t}\frac{M}{M-1}\mathrm{tr}\tilde{C}_t\tilde{C}_t^{\mathrm{T}}\mathrm{d}t + \sup_{s\leq t}\sup_{x\in\mathbb{R}^{d_x}}\frac{1}{M^2}\sum_{i,j=1}^{M}\left(\mathrm{tr}C_t(x)C_t(x)^{\mathrm{T}} + \frac{\mathrm{tr}C_t(x)C_t(x)^{\mathrm{T}}}{M-1}\right)\mathrm{d}t$$

Using the stochastic Gronwall Lemma [Scheu] and the fact that $\mathrm{tr}C_t(x)C_t(x)^{\mathrm{T}} = |C_t(x)|^2$ we can thus bound the covariance matrix by

$$\mathbb{E}\left[\sup_{t\leq T\wedge\xi}\sqrt{\mathrm{tr}P_t^M}\right] \leq \mathcal{C}\exp\left(\int_0^T \mathrm{Lip}\left(B_s\right) + 4 d_x\left|\tilde{C}_s R_s^{-1} H_s\right|\ \mathrm{d}s\right)\cdots$$
$$\cdots\sqrt{2 d_x T\ \sup_{t\leq T}\left|\tilde{C}_t R_t^{-1}\tilde{C}_t^{\mathrm{T}}\right| + \frac{TM}{M-1}\sup_{t\leq T}\left(\left|\tilde{C}_t\right|^2 + \sup_{x\in\mathbb{R}^{d_x}}|C_t(x)|^2\right)}$$

for some constant $\mathcal{C}$.

Due to the assumptions the right-hand side of this inequality is finite. Similarly one can bound

$$.\mathbb{E}\left[\sup_{t\leq T\wedge\xi}\sqrt{\left|x_t^M\right|}\right]<+\infty,$$

which shows that $\mathbb{P}(\xi<T)=0$. Thus explosion can not occur in the time interval $[0,T]$, which yields the existence of a unique global solution to the EnKBF.

□

Throughout this paper we will make use of the following notation to make formulas less loaded.

Notation 10. *For any $\mathbb{R}^{d_x}$ random vector Z we denote*

$$\begin{aligned}\ll B_t, Z \gg := &\mathbb{E}_{\times}\left[\left(B_t(Z)-\mathbb{E}_{\times}[B_t(Z)]\right)\left(\bar{X}_t-\mathbb{E}_{\times}[Z]\right)^{\mathrm{T}}\right]\\ &\cdots+\mathbb{E}_{\times}\left[\left(\bar{X}_t-\mathbb{E}_{\times}[Z]\right)\left(B_t(Z)-\mathbb{E}_{\times}[B_t(Z)]\right)^{\mathrm{T}}\right].\end{aligned}$$

Letting the number of ensemble members M in the EnKBF (15) go to infinity gives us, at least formally, the mean-field equation

$$\begin{aligned}\mathrm{d}\bar{X}_t^\epsilon &= B_t\left(\bar{X}_t^\epsilon\right)\mathrm{d}t + C_t\left(\bar{X}_t^\epsilon\right)\mathrm{d}\bar{W}_t+\tilde{C}_t\left(\bar{X}_t^\epsilon\right)\mathrm{d}\bar{V}_t\\ &\cdots+\left(\bar{P}_t^\epsilon H_t^{\mathrm{T}}+\tilde{C}_t\left(\bar{X}_t^\epsilon\right)\right)R_t^{-1}\left(\mathrm{d}Y_t-\frac{H_t\left(\bar{X}_t^\epsilon+\bar{m}_t^\epsilon\right)}{2}\mathrm{d}t\right)\\ &\cdots-\left(\bar{P}_t^\epsilon H^{\mathrm{T}}+\tilde{C}_t\left(\bar{X}_t^\epsilon\right)\right)R_t^{-1}\tilde{C}_t^{\mathrm{T}}\left(\bar{X}_t^\epsilon\right)\left(\bar{P}^\epsilon\right)_t^{+\epsilon,n}\frac{\bar{X}_t^\epsilon-\bar{m}_t^\epsilon}{2}\mathrm{d}t.\end{aligned}\tag{17}$$

Just as in the case of the particle filter, it is easy to see by using Itô's formula and taking the expectation, that if $\tilde{C}_t$ is constant, then

$$\begin{aligned}\frac{\mathrm{d}\bar{P}_t^\epsilon}{\mathrm{d}t} =&\ll B_t,\bar{X}_t^\epsilon\gg+\mathbb{E}_{\times}\left[C_t\left(\bar{X}_t^\epsilon\right)C_t\left(\bar{X}_t^\epsilon\right)^{\mathrm{T}}\right]+\tilde{C}_t\tilde{C}_t^{\mathrm{T}}\\ &\cdots-\frac{1}{2}\left(\bar{P}_t^\epsilon H_t^{\mathrm{T}}+\tilde{C}_t\right)R_t^{-1}\left(H_t+\tilde{C}_t^{\mathrm{T}}\left(\bar{P}^\epsilon\right)_t^{+\epsilon,n}\right)\bar{P}_t^\epsilon\\ &\cdots-\frac{1}{2}\bar{P}_t^\epsilon\left(H_t^{\mathrm{T}}+\left(\bar{P}_t^\epsilon\right)^{+\epsilon,n}\tilde{C}_t\right)R_t^{-1}\left(H_t\bar{P}_t^\epsilon+\tilde{C}_t^{\mathrm{T}}\right)\end{aligned}\tag{18}$$

and therefore

$$\frac{\mathrm{dtr}\bar{P}_t^\epsilon}{\mathrm{d}t}\leq 2\left(\mathrm{Lip}\left(B_t\right)+2d_x\left|\tilde{C}_tR_t^{-1}H_t\right|\right)\ \mathrm{tr}\bar{P}_t^\epsilon+\sup_{x\in\mathbb{R}^{d_x}}\left|C_t(x)\right|^2+\left|\tilde{C}_t\right|^2+2d_x\left|\tilde{C}_tR_t^{-1}\tilde{C}_t^{\mathrm{T}}\right|$$

Using the deterministic Gronwall Lemma this differential inequality lets us derive

$$\sup_{t\le T} \operatorname{tr}\bar{P}_t^{\epsilon} \le \bar{\Psi}(T) := \exp\left(2\int_0^T \operatorname{Lip}(B_t) + d_x \left|\tilde{C}_t R_t^{-1} H_t\right| \,\mathrm{d}t\right)\cdots$$
$$\cdots\left(\operatorname{tr}\bar{P}_0 + \int_0^T \sup_{x\in\mathbb{R}^{d_x}} |C_t(x)|^2 + \left|\tilde{C}_t\right|^2 + d_x\left|\tilde{C}_t R_t^{-1}\tilde{C}_t^{\mathrm{T}}\right| \,\mathrm{d}t\right). \tag{19}$$

Note that (17) coincides with a generalization of (12) to nonlinear signal, with the exception that the precision matrix $\bar{P}_t^{-1}$ is replaced by the regularized pseudo-inverse $\bar{P}_t^{+\epsilon,n}$. The following theorem shows that the covariance matrix of solutions to (17) is always regular and the inverse is uniformely bounded in ϵ and t.

Lemma 11. *Let $T > 0$ be arbitrary but fixed and let $\bar{X}^\epsilon$ be a solution of (17) and denote its covariance matrix by $\bar{P}^\epsilon$. Assume that $\bar{P}_0$ is regular. If the coefficients of (17) are Lipschitz and if*

$$\inf_x \lambda_{\min}\left(C_t(x)C_t(x)^{\mathrm{T}}\right) + \lambda_{\min}\left(\tilde{C}_t\left(I - R_t^{-1}\right)\tilde{C}_t^{\mathrm{T}}\right) > 0 \text{ for every } t > 0, \tag{20}$$

then $\bar{P}_t^\epsilon$ is regular for all times $t \in [0,T]$. The norm of its inverse can be uniformely bounded by a constant $\underline{\Psi}(T)$, i.e. $\sup_{t\le T}\left|\left(\bar{P}_t^\epsilon\right)^{-1}\right| \le \underline{\Psi}(T)$, for all $\epsilon > 0$.

Proof. For the sake of convenience and notational brevity let us drop the dependence on ϵ in this proof, i.e. we write $\bar{P}_t$ instead of $\bar{P}_t^\epsilon$. Since $\bar{P}_t$ is a spsd matrix, there exists an orthogonal matrix Q_t and a diagonal matrix $\Lambda_t := \operatorname{diag}\left(\lambda_t^1, \cdots, \lambda_t^{d_x}\right)$ such that

$$\bar{P}_t = Q_t^{\mathrm{T}}\Lambda_t Q_t \text{ for all } t \ge 0. \tag{21}$$

It can be shown (see [DiEi]) that Q is continuous and while Λ_t is differentiable and satisfies

$$\frac{\mathrm{d}\Lambda_t}{\mathrm{d}t} = \operatorname{diag}\left(Q_t \frac{\mathrm{d}\bar{P}_t}{\mathrm{d}t} Q_t^{\mathrm{T}}\right). \tag{22}$$

Denote by e_i the i-th unit vector in $\mathbb{R}^{d_x}$, then this and (18) implies

$$\begin{aligned}\frac{\mathrm{d}\lambda_t^i}{\mathrm{d}t} &= e_i^{\mathrm{T}} Q_t \frac{\mathrm{d}\bar{P}_t}{\mathrm{d}t} Q_t^{\mathrm{T}} e_i \\ &= e_i^{\mathrm{T}} Q_t \ll B_t, \bar{X}_t \gg Q_t^{\mathrm{T}} e_i + \mathbb{E}_{\not X}\left[e_i^{\mathrm{T}} Q_t C_t\left(\bar{X}_t\right) C_t\left(\bar{X}_t\right)^{\mathrm{T}} Q_t^{\mathrm{T}} e_i\right] \\ &\cdots + e_i^{\mathrm{T}} Q_t \tilde{C}_t \tilde{C}_t^{\mathrm{T}} Q_t^{\mathrm{T}} e_i - e_i^{\mathrm{T}} Q_t\left(\bar{P}_t H_t^{\mathrm{T}} + \tilde{C}_t\right) R_t^{-1}\left(H_t + \tilde{C}_t^{\mathrm{T}} \bar{P}_t^{+\epsilon,n}\right)\bar{P}_t Q_t^{\mathrm{T}} e_i.\end{aligned}$$

We want to estimate λ^i from below. To this end we bound all three terms on the right-hand side from below to derive a differential inequality.

First we note that

$$
\begin{aligned}
& e_i^{\mathrm{T}} Q_t \ll B_t, \bar{X}_t \gg Q_t^{\mathrm{T}} e_i \\
& = 2\, \mathbb{E}_{\mathcal{X}} \left[\left(e_i^{\mathrm{T}} Q_t \left(B(\bar{X}_t) - B(\bar{m}_t) \right) \right) \left(e_i^{\mathrm{T}} Q_t \left(\bar{X}_t - \bar{m}_t \right) \right) \right] \\
& \geq -2 \sqrt{\mathbb{E}_{\mathcal{X}} \left[\left| e_i^{\mathrm{T}} Q_t^{\mathrm{T}} \left(B(\bar{X}_t) - B(\bar{m}_t) \right) \right|^2 \right]} \underbrace{\sqrt{\mathbb{E}_{\mathcal{X}} \left[\left| e_i^{\mathrm{T}} Q_t^{\mathrm{T}} \left(\bar{X}_t - \bar{m}_t \right) \right|^2 \right]}}_{= e_i^{\mathrm{T}} Q_t^{\mathrm{T}} \bar{P}_t Q_t e_i = \lambda_t^i} \\
& \geq -2 \mathrm{Lip}(B) \sqrt{\mathrm{tr} \bar{P}_t} \sqrt{\lambda_t^i}.
\end{aligned}
$$

By the variational definition of the minimal eigenvalue we also see that

$$
\mathbb{E}_{\mathcal{X}} \left[e_i^{\mathrm{T}} Q_t C_t \left(\bar{X}_t \right) C_t \left(\bar{X}_t \right)^{\mathrm{T}} Q_t^{\mathrm{T}} e_i \right] \geq \inf_x \lambda_{\min} \left(C(x) C(x)^{\mathrm{T}} \right).
$$

Finally it remains to bound

$$
\begin{aligned}
& e_i^{\mathrm{T}} Q_t \tilde{C}_t \tilde{C}_t^{\mathrm{T}} Q_t^{\mathrm{T}} e_i - e_i^{\mathrm{T}} Q_t \left(\bar{P}_t H_t^{\mathrm{T}} + \tilde{C}_t \right) R_t^{-1} \left(H_t + \tilde{C}_t^{\mathrm{T}} \bar{P}_t^{+} \right) \bar{P}_t Q_t^{\mathrm{T}} e_i \\
& = e_i^{\mathrm{T}} Q_t \left(\tilde{C}_t \tilde{C}_t^{\mathrm{T}} - \tilde{C}_t R_t^{-1} \tilde{C}_t^{\mathrm{T}} \bar{P}_t^{+\epsilon,n} \bar{P}_t \right) Q_t^{\mathrm{T}} e_i - e_i^{\mathrm{T}} Q_t \bar{P}_t H_t^{\mathrm{T}} R_t^{-1} H_t \bar{P}_t Q_t^{\mathrm{T}} e_i \\
& \cdots - e_i^{\mathrm{T}} Q_t \left(\tilde{C}_t R_t^{-1} H_t \bar{P}_t + \bar{P}_t H_t^{\mathrm{T}} R_t^{-1} \tilde{C}_t^{\mathrm{T}} \bar{P}_t^{+\epsilon,n} \bar{P}_t \right) Q_t^{\mathrm{T}} e_i
\end{aligned}
$$

from below.

By using the decompostion (21) of $\bar{P}_t$ into Q_t and Λ_t we see that

$$
\bar{P}_t^{+\epsilon,n} \bar{P}_t Q^{\mathrm{T}} e_i = \frac{\left(\lambda_t^i\right)^n}{\left(\lambda_t^i\right)^n + \epsilon} Q^T e_i.
$$

This, together with decomposition (21) of $\bar{P}_t$, lets us easily derive

$$
\begin{aligned}
e_i^{\mathrm{T}} Q_t \bar{P}_t H_t^{\mathrm{T}} R_t^{-1} H_t \bar{P}_t Q_t^{\mathrm{T}} e_i &\leq \lambda_{\max} \left(H_t^{\mathrm{T}} R_t^{-1} H_t \right) \left(\lambda_t^i \right)^2 \\
e_i^{\mathrm{T}} Q_t \left(\tilde{C}_t R_t^{-1} H_t \bar{P}_t + \bar{P}_t H_t^{\mathrm{T}} R_t^{-1} \tilde{C}_t^{\mathrm{T}} \bar{P}_t^{+} \bar{P}_t \right) Q_t^{\mathrm{T}} e_i &\leq 2 \left| \tilde{C}_t R_t^{-1} H_t \right| \lambda_t^i \\
e_i^{\mathrm{T}} Q_t \left(\tilde{C}_t \tilde{C}_t^{\mathrm{T}} - \tilde{C}_t R_t^{-1} \tilde{C}_t^{\mathrm{T}} \bar{P}_t^{+} \bar{P}_t \right) Q_t^{\mathrm{T}} e_i &\geq \lambda_{\min} \left(\tilde{C}_t \left(I - R_t^{-1} \right) \tilde{C}_t^{\mathrm{T}} \right).
\end{aligned}
$$

Note that for the last inequality we also used that $\tilde{C}_t R_t^{-1} \tilde{C}_t^{\mathrm{T}}$ is positive semidefinite.

And thus we derive

$$\begin{aligned}\frac{\mathrm{d}\lambda_t^i}{\mathrm{d}t} &\geq -2\mathrm{Lip}(B)\ \sqrt{\mathrm{tr}\bar{P}_t}\sqrt{\lambda_t^i} \\ &\cdots + \inf_x \lambda_{\min}\left(C_t(x)C_t(x)^{\mathrm{T}}\right) + \lambda_{\min}\left(\tilde{C}_t\left(I - R_t^{-1}\right)\tilde{C}_t^{\mathrm{T}}\right) \\ &\cdots - \lambda_{\max}\left(H_t^{\mathrm{T}} R_t^{-1} H_t\right)\left(\lambda_t^i\right)^2 - 2\left|\tilde{C}_t R_t^{-1} H_t\right| \lambda_t^i.\end{aligned} \tag{23}$$

Let $\underline{\lambda}$ denote a solution of

$$\begin{aligned}\frac{\mathrm{d}\underline{\lambda}_t}{\mathrm{d}t} &\geq -2\mathrm{Lip}(B)\ \sqrt{\mathrm{tr}\bar{P}_t}\sqrt{\underline{\lambda}_t} \\ &\cdots + \frac{\inf_x \lambda_{\min}\left(C_t(x)C_t(x)^{\mathrm{T}}\right) + \lambda_{\min}\left(\tilde{C}_t\left(I - R_t^{-1}\right)\tilde{C}_t^{\mathrm{T}}\right)}{2} \\ &\cdots - \lambda_{\max}\left(H_t^{\mathrm{T}} R_t^{-1} H_t\right)\left(\underline{\lambda}_t\right)^2 - 2\left|\tilde{C}_t R_t^{-1} H_t\right| \underline{\lambda}_t.\end{aligned} \tag{24}$$

with initial condition $\underline{\lambda}_0 = \frac{\min_i \lambda_0^i}{2}$. The existence of solutions (which might not be unique) is guaranteed by the Peano theorem. Then by [Walter, II. Lemma, page 64], $\underline{\lambda}$ is a lower bound for all λ^i.
By assumption (20), the right-hand side of (24) is strictly bigger than zero, when $\underline{\lambda}_t = 0$. Thus for $t_0 > 0$ with $\underline{\lambda}_{t_0} = 0$, we have $\underline{\lambda}'(t_0) > 0$. According to [Walter]II. Lemma, page 64, this indeed implies $\underline{\lambda}_t > 0$ for all $t \geq 0$, if $\underline{\lambda}_0 > 0$, which is the case if $\bar{P}_0$ is regular. Since $\left|\bar{P}_t^{-1}\right|^2 = \sum_{i=1}^{d_x}\left(\lambda_t^i\right)^2$ this indeed gives us a bound $\underline{\Psi}(T) := \sqrt{d_x}\frac{1}{\inf_{t\leq T}\underline{\lambda}_t}$ for the inverse of the covariance.. □

Remark 12. *Note that if $\bar{P}_0$ is not regular, then one can still show that $\bar{P}_t^\epsilon$ is regular for all $t > 0$. To accomplish this one uses the proof above and argues by contradiction. As [Walter, II. Lemma, page 64] also implies that if there existed a $t_0 > 0$ with $\underline{\lambda}_{t_0} = 0$, then there would be a $\bar{t} > 0$ such that $\underline{\lambda}_t \leq 0$ for all $t \leq \bar{t}$. We know that λ^i is non-negative, thus it would hold that $\underline{\lambda}_t = 0$ for all $t \leq \bar{t}$. However by (24) and assumption (20) this would also mean that $\frac{\mathrm{d}\underline{\lambda}_t}{\mathrm{d}t} > 0$ on $(0, \bar{t})$, which is a contradiction to $\underline{\lambda}$ being constant on $(0, \bar{t})$. Therefore we indeed can condclude that $\underline{\lambda}_t > 0$ and thus $\lambda_t^i > 0$ for all $t > 0,\ i = 1, \cdots, d_x$.Therefore $\bar{P}_t^\epsilon$ is regular.*

For a regular matrix P we note that $\lim_{\epsilon\to 0} P^{+\epsilon,n} = P^{-1}$ for every $n \in \mathbb{N}$. Thus by theorem 11 and by formally taking the limit $\epsilon \to 0$ in (17) we obtain the equation

$$\begin{aligned}\mathrm{d}\bar{X}_t &= B_t\left(\bar{X}_t\right)\mathrm{d}t + C_t\left(\bar{X}_t\right)\mathrm{d}\bar{W}_t + \tilde{C}_t\left(\bar{X}_t\right)\mathrm{d}\bar{V}_t \\ &\cdots + \left(\bar{P}_t H_t^{\mathrm{T}} + \tilde{C}_t\left(\bar{X}_t\right)\right) R_t^{-1}\left(\mathrm{d}Y_t - \frac{H_t\left(\bar{X}_t + \bar{m}_t\right)}{2}\mathrm{d}t\right) \\ &\cdots - \left(\bar{P}_t H^{\mathrm{T}} + \tilde{C}_t\left(\bar{X}_t\right)\right) R_t^{-1}\tilde{C}_t^{\mathrm{T}}\left(\bar{X}_t\right)\bar{P}_t^{-1}\frac{\bar{X}_t - \bar{m}_t}{2}\mathrm{d}t.\end{aligned} \tag{25}$$

Just as for the regularized case $\epsilon > 0$ it is easy to show that the covariance matrix $\bar{P}$ of its solution $\bar{X}$ satisfies

$$\mathrm{tr}\bar{P}_t \leq \bar{\Psi}(T) \text{ and } \left|\bar{P}_t^{-1}\right| \leq \underline{\Psi}(T).$$

for all $t \in [0, T]$ We shall prove the well posedness of (25) in the next section. Section 5 is then devoted to the formalization of the limits $M \to \infty$ and $\epsilon \to 0$.

4 Well posedness of the mean-field Ensemble Kalman–Bucy filter

We now state the main theorem of this paper.

Theorem 13. *Assume that B is Lipschitz, C is a bounded Lipschitz function and that $\tilde{C}$ is constant. Furthermore we assume that $\bar{P}_0$ is regular. Then there exists a unique solution $\bar{X}$ of (25) on every fixed time intervall $[0, T]$.*

Since (25) is a McKean–Vlasov SDE, we can not simply use (19) to make a localization argument to derive the existence of solutions like we did for the approximating particle filter (15).

Instead we will make a fixed point argument just as we did in the linear case. However, since the equation of the covariance matrix (18) does not decouple form (15) for nonlinear B, we will not be able to simply guess the right fixed point. Instead we will use (19) to apply Banachs fixed point theorem.

The rest of this section is devoted to the proof of 13. Therefore we shall not mark the proof in a special way.

Remark 14. *The well posedness of (17) can be proven under exaclty the same assumtion in exactly the same way.*

Define for any differentiable spsd-valued function P with $P_0 = \bar{P}_0$ the stochastic process X^P to be the solution of

$$\mathrm{d}\bar{X}_t^P = B_t(\bar{X}_t^P)\mathrm{d}t + C_t(\bar{X}_t^P)\mathrm{d}W_t + \tilde{C}_t\mathrm{d}V_t + \left(P_t H_t^{\mathrm{T}} + \tilde{C}_t\right) R_t^{-1}\left(\mathrm{d}Y_t - \frac{H_t\left(\bar{X}_t^P + \bar{m}_t^P\right)}{2}\mathrm{d}t\right)$$

$$\cdots - \frac{P_t H_t^{\mathrm{T}} + \tilde{C}_t}{2} R_t^{-1}\tilde{C}_t^{\mathrm{T}} P_t^{-1}\left(\bar{X}_t^P - \bar{m}_t^P\right)\mathrm{d}t.$$

Hereby we denote $\bar{m}_t^P := \mathbb{E}_{(\bar{W},\bar{V})}\left[X_t^P\right]$. Since this is a McKean–Vlasov equation satisfying the usual Lipschitz conditions, there indeed exists a unique strong solution.

A solution of (25) is given by a fixed point

$$\mathbb{C}\mathrm{ov}_{(\bar W,\bar V)}\left[X_t^P\right] = P_t \text{ for all } t \in [0,T].$$

We will prove its existence and uniqueness on small time frames with the Banach fixed point theorem. To prove that the map

$$P \mapsto \mathbb{C}\mathrm{ov}_{(\bar W,\bar V)}\left[X^P\right]$$

is a strict contraction, a priori estimates on the covariance matrix $\mathbb{C}\mathrm{ov}_{(\bar W,\bar V)}\left[X^P\right]$, independent of P, will be key.

We note that the mean and covariance matrix of X^P satisfy the following two equations

$$\mathrm{d}m_t^P = \mathbb{E}_{(\bar W,\bar V)}\left[B_t\left(X_t^P\right)\right]\mathrm{d}t + P_t H_t^{\mathrm{T}} R_t^{-1}\left(\mathrm{d}Y_t - H_t m_t^P \mathrm{d}t\right), \tag{26}$$

and

$$\begin{aligned}\frac{\mathrm{d}\mathbb{C}\mathrm{ov}\left[\bar X_t^P\right]}{\mathrm{d}t} =& \ll B_t, \bar X_t^P \gg + \mathbb{E}_{(\bar W,\bar V)}\left[C_t\left(\bar X_t^P\right)C_t\left(\bar X_t^P\right)^{\mathrm{T}}\right] + \tilde C_t \tilde C_t^{\mathrm{T}} \\ &\cdots - \left(P_t H_t^{\mathrm{T}} + \tilde C_t\right) R_t^{-1}\frac{H_t \mathbb{C}\mathrm{ov}\left[\bar X_t^P\right]}{2} - \frac{\mathbb{C}\mathrm{ov}\left[\bar X_t^P\right] H_t^{\mathrm{T}}}{2} R_t^{-1}\left(H_t P_t + \tilde C_t^{\mathrm{T}}\right) \\ &\cdots - \frac{P_t H_t^{\mathrm{T}} + \tilde C_t}{2} R_t^{-1}\tilde C_t^{\mathrm{T}} P_t^{-1}\mathbb{C}\mathrm{ov}\left[\bar X_t^P\right] - \mathbb{C}\mathrm{ov}\left[\bar X_t^P\right] P_t^{-1}\tilde C_t R_t^{-1}\frac{H_t \bar P_t + \tilde C_t^{\mathrm{T}}}{2}.\end{aligned} \tag{27}$$

We note that if $H_t^{\mathrm{T}} R_t^{-1} H_t$ is not a scalar matrix, then we can not guarantee that $P_t H_t^{\mathrm{T}} R_t^{-1} H_t$ is positive definite, let alone symmetric. Thus we can not simply take the trace in (27), use the Gronwall lemma and expect to derive bounds independent of P. In other words, the covariance matrix is a Lyapunov function of (17), that, for general coefficients, is not robust to perturbations in the covariance matrix.

For the sake of simplicity, let us first concentrate on the case where (27) indeed delivers bounds independent of P, before we return to the general setting.

4.1 The case of uncorrelated, scalar observations

Let us assume in this subsection that $\tilde C = 0$ and that $H_t^{\mathrm{T}} R_t^{-1} H_t = \alpha_t I$ for some positive scalar function α.

In this case (27) reduces to

$$\begin{aligned}\frac{\mathrm{d}\mathbb{C}\mathrm{ov}\left[\bar X_t^P\right]}{\mathrm{d}t} =& \ll B_t, \bar X_t^P \gg + \mathbb{E}_{(\bar W,\bar V)}\left[C_t\left(\bar X_t^P\right)C_t\left(\bar X_t^P\right)^{\mathrm{T}}\right] \\ &- \frac{\alpha_t}{2}\left(P_t \mathbb{C}\mathrm{ov}_{(\bar W,\bar V)}\left[\bar X_t^P\right] + \mathbb{C}\mathrm{ov}_{(\bar W,\bar V)}\left[\bar X_t^P\right] P_t\right).\end{aligned} \tag{28}$$

We note that by the cyclical invariance of the trace we have

$$\begin{aligned}\operatorname{tr}\left[P_t \mathbb{Cov}_{(\bar{W},\bar{V})}\left[\bar{X}_t^P\right]\right] &= \mathbb{E}_{(\bar{W},\bar{V})}\left[\operatorname{tr}\left[P_t\left(X_t^P - m_t^P\right)\left(X_t^P - m_t^P\right)^{\mathrm{T}}\right]\right] \\ &= \mathbb{E}_{(\bar{W},\bar{V})}\left[\operatorname{tr}\left[\left(X_t^P - m_t^P\right)^{\mathrm{T}} P_t\left(X_t^P - m_t^P\right)\right]\right] \geq 0\end{aligned}$$

This gives us the differential inequality

$$\frac{\mathrm{d}\operatorname{tr}\mathbb{Cov}_{(\bar{W},\bar{V})}\left[\bar{X}_t^P\right]}{\mathrm{d}t} \leq 2\ \operatorname{Lip}(B_t)\ \mathbb{Cov}_{(\bar{W},\bar{V})}\left[\bar{X}_t^P\right] + \|C_t\|_\infty^2,$$

and thus we derive by the deterministic Gronwall inequality

$$\sup_{t\leq T}\ \operatorname{tr}\ \mathbb{Cov}_{(\bar{W},\bar{V})}\left[X_t^P\right] \leq \underbrace{\exp\left(2\int_0^T \operatorname{Lip}(B_t)\ \mathrm{d}t\right)\left(\operatorname{tr}\bar{P}_0 + \int_0^T \|C_t\|_\infty^2\ \mathrm{d}t\right)}_{=:\kappa_0(T)}.$$

This means that we can assume that $\sup_{t\leq T} P_t \leq \kappa_0(T)$. Using (28) we thus obtain

$$\sup_{t\leq T}\left|\frac{\mathrm{d}\operatorname{tr}\mathbb{Cov}\left[X_t^P\right]}{\mathrm{d}t}\right| \leq \sup_{t\leq T}\left(2\operatorname{Lip}(B_t) + \alpha_t\right)\kappa_0(T)^2 =: \kappa_1(T)$$

which lets us restrict the domain of potential fixed points to

$$\mathcal{X}_T := \left\{\ P \in C^1\left([0,T];\mathbb{R}^{d_x\times d_x}\right)\ :\ P_0 = \bar{P}_0,\ P_t \text{ is spsd},\ \|P\|_\infty \leq \kappa_0(T),\ \|\partial_t P\|_\infty \leq \kappa_1(T)\right\}. \tag{29}$$

As a closed subset of a Banach space, $\mathcal{X}$ is itself a complete metric space. Thus we can employ the Banach fixed point theorem and the following Lemma will be useful for deriving the needed contraction property.

Lemma 15. *For every fixed $T > 0$ there exists a constant $\kappa_Y(T)$, such that*

$$\sup_{t\leq T}\left|\int_0^t (P_s^1 - P_s^2) H_s^{\mathrm{T}} R_s^{-1}\mathrm{d}Y_s\right| \leq \kappa_Y(T)\left\|P^1 - P^2\right\|_{C^1([0,T];R^{d\times d})}$$

holds for all $P^1, P^2 \in C^1\left([0,T];\mathbb{R}^{d_x\times d_x}\right)$. Furthermore these constants satisfy $\kappa_Y(T) \xrightarrow{T\to 0} 0$.

Proof. We note that all $P^1, P^2 \in C^1\left([0,T];\mathbb{R}^{d_x\times d_x}\right)$ are processes of bounded variation and thus one gets by integration by parts

$$\int_0^t (P_s^1 - P_s^2) H_s^{\mathrm{T}} R_s^{-1}\mathrm{d}Y_s = \left(P_t^1 - P_t^1\right) H_t^{\mathrm{T}} R_t^{-1} Y_t - \int_0^t \frac{\mathrm{d}\ (P_s^1 - P_2^1) H_s^{\mathrm{T}} R_s^{-1}}{\mathrm{d}s} Y_s \mathrm{d}s,$$

where the boundary term at $t = 0$ disappears as $Y_0 = 0$.

Since both $H_s^{\mathrm{T}} R_s^{-1}$ is a $C^1\left([0,T];\mathbb{R}^{d_x\times d_y}\right)$ function we immediatley derive the desired inequality with

$$\kappa_Y(T) := \sup_{t\le T}\left|H_t^{\mathrm{T}} R_t^{-1} Y_t\right| + \int_0^T \left|H_t^{\mathrm{T}} R_t^{-1} Y_t\right| + \left|\frac{\mathrm{d}H_t^{\mathrm{T}} R_t^{-1}}{\mathrm{d}t} Y_t\right| \,\mathrm{d}t.$$

By the continuity of Y we also immediatley see that $\kappa_Y(T) \xrightarrow{T\to 0} 0$. □

As a corollary one also immediatley derives for any $P \in \mathcal{X}_T$

$$\sup_{t\le T}\left|\int_0^t P_s H_s^{\mathrm{T}} R_s^{-1} \mathrm{d}Y_s\right| \le \kappa_Y(T)\sqrt{\kappa_0(T)^2 + \kappa_1(T)^2}. \tag{30}$$

Using the Lipschitz continuity of B we see that

$$\begin{aligned}\left|m_t^P\right| \le \left|m_0^P\right| + \int_0^t \left|B_s(0)\right| \mathrm{d}s + \int_0^t \mathrm{Lip}(B_s)\mathbb{E}_Y\left[\left|X_s^P - m_s^P\right|\right] \mathrm{d}s \cdots +\\ \cdots + \left|\int_0^t P_s H_s^{\mathrm{T}} R_s^{-1} \mathrm{d}Y_s\right| + \int_0^t \big(\underbrace{|P_s|}_{\le \kappa_0(T)} \left|H_s^{\mathrm{T}} R_s^{-1} H_s\right| + \mathrm{Lip}(B_s)\big)\left|m_s^P\right| \mathrm{d}s.\end{aligned}$$

Applying the deterministic Gronwall inequality together with

$$\mathbb{E}_{(\bar W,\bar V)}\left[\left|X_s^P - m_s^P\right|\right] \le \sqrt{\mathbb{E}_{(\bar W,\bar V)}\left[\left|X_s^P - m_s^P\right|^2\right]} = \sqrt{\mathrm{tr}\ \mathbb{Cov}_{(\bar W,\bar V)}\left[X_s^P\right]} \le \sqrt{\kappa_0(T)}$$

and (30), we obtain

$$\begin{aligned}\sup_{t\le T}\left|m_t^P\right| \le \kappa_m(T) := \exp\left(\int_0^t \left(\kappa_0(T)\left|H_s^{\mathrm{T}} R_s^{-1} H_s\right| + \mathrm{Lip}(B_s)\right)\mathrm{d}s\right)\cdots\\ \cdots\left(\left|m_0^P\right| + \int_0^t \left|B_s(0)\right| + \mathrm{Lip}(B_s)\sqrt{\kappa_0(T)}\mathrm{d}s + \kappa_Y(T)\sqrt{\kappa_0(T)^2 + \kappa_1(T)^2}\right).\end{aligned}$$

Now we are able to prove the desired contraction property. To this end let $P^1, P^2 \in \mathcal{X}$ be given and denote by X^1, X^2 the corresponding solutions and by m^1, m^2 their means. Then it is clear that

$$\begin{aligned}X_t^1 - X_t^2 = \int_0^t \left(B_s(X_s^1) - B_s(X_s^2)\right)\mathrm{d}s + \int_0^t \left(P_s^1 - P_s^2\right) H_s^{\mathrm{T}} R_s^{-1} \mathrm{d}Y_s\\ \cdots - \int_0^t \left(P_s^1 - P_s^2\right) H_s^{\mathrm{T}} R_s^{-1} H_s \frac{X_s^1 + m_s^1}{2}\mathrm{d}s\\ \cdots - \int_0^t P_s^2 H_s^{\mathrm{T}} R_s^{-1} H_s \frac{X_s^1 - X_s^2 + m_s^1 - m_s^2}{2}\mathrm{d}s + \int_0^t \left(C_s(X_s^1) - C_s(X_s^2)\right)\mathrm{d}W_s\end{aligned}$$

Using the estimates $\left|\sum_{i=1}^k a_i\right|^2 \leq k \sum_{i=1}^k |a_i|^2$ and $\left|\int_0^t f(t)\mathrm{d}t\right|^2 \leq t\int_0^t |f(t)|^2 \,\mathrm{d}t$, we get

$$\begin{aligned}\left|X_t^1 - X_t^2\right|^2 &\leq 5t\int_0^t \left|B_s(X_s^1) - B_s(X_s^2)\right|^2 \mathrm{d}s + 5\left|\int_0^t \left(P_s^1 - P_s^2\right) H_s^{\mathrm{T}} R_s^{-1} \mathrm{d}Y_s\right|^2 \\ &\cdots + 5t\int_0^t \left|P_s^1 - P_s^2\right|^2 \left|H_s^{\mathrm{T}} R_s^{-1} H_s\right|^2 \frac{|X_s^1|^2 + |m_s^1|^2}{2}\mathrm{d}s \\ &\cdots + 5t\int_0^t \left|P_s^2\right|^2 \left|H_s^{\mathrm{T}} R_s^{-1} H_s\right|^2 \frac{\left|X_s^1 - X_s^2\right|^2 + \left|m_s^1 - m_s^2\right|^2}{2}\mathrm{d}s \\ &\cdots + 5\left|\int_0^t \left(C_s(X_s^1) - C_s(X_s^2)\right)\mathrm{d}W_s\right|^2\end{aligned}$$

Note that

$\left|m_s^1 - m_s^2\right|^2 \leq \mathbb{E}_{(\bar{W},\bar{V})}\left[\left|X_s^1 - X_s^2\right|^2\right]$ and $\mathbb{E}_{(\bar{W},\bar{V})}\left[|X_s^1|^2\right] + |m_s^1|^2 \leq \kappa_0(T) + 2\kappa_m(T)^2$.

Using the Lipschitz continuity of B and C, Itô isometry and we thus derive

$$\begin{aligned}&\mathbb{E}_{(\bar{W},\bar{V})}\left[\left|X_t^1 - X_t^2\right|^2\right] \\ &\leq 5\int_0^t \left(T\mathrm{Lip}(B_s)^2 + T\left|P_s^2\right|^2 \left|H_s^{\mathrm{T}} R_s^{-1} H_s\right|^2 + \mathrm{Lip}(C_s)^2\right) \mathbb{E}_{(\bar{W},\bar{V})}\left[\left|X_s^1 - X_s^2\right|^2\right] \mathrm{d}s \\ &\cdots + 5\left(\kappa_Y(T)^2 + 5T\frac{\kappa_0(T) + 2\kappa_m(T)^2}{2}\int_0^t \left|H_s^{\mathrm{T}} R_s^{-1} H_s\right|^2 \mathrm{d}s\right) \left\|P^1 - P^2\right\|_{C^1([0,T])}^2 .\end{aligned}$$

Thus we derive from the Gronwall lemma

$$\begin{aligned}\sup_{t\leq T}\mathbb{E}_{(\bar{W},\bar{V})}\left[\left|X_t^1 - X_t^2\right|^2\right] &\leq \exp\left(5\int_0^T T\mathrm{Lip}(B_s)^2 + T\left|P_s^2\right|^2 \left|H_s^{\mathrm{T}} R_s^{-1} H_s\right|^2 + \mathrm{Lip}(C_s)^2 \mathrm{d}s\right) \\ &\cdots 5\left(\kappa_Y(T)^2 + 5T\frac{\kappa_0(T) + 2\kappa_m(T)^2}{2}\int_0^t \left|H_s^{\mathrm{T}} R_s^{-1} H_s\right|^2 \mathrm{d}s\right) \left\|P^1 - P^2\right\|_{C^1([0,T])}^2 \\ &:= \kappa_{\mathrm{contr}}(T) \left\|P^1 - P^2\right\|_{C^1([0,T])}^2 .\end{aligned}$$

Where clearly $\kappa_{\mathrm{contr}}(T) \xrightarrow{T\to 0} 0$. By employing

$$\left|\mathbb{Cov}_{(\bar{W},\bar{V})}\left[X_t^1\right] - \mathbb{Cov}_{(\bar{W},\bar{V})}\left[X_t^2\right]\right| \leq 2\sqrt{\mathrm{tr}\mathbb{Cov}_{(\bar{W},\bar{V})}\left[X_t^1\right] + \mathrm{tr}\mathbb{Cov}_{(\bar{W},\bar{V})}\left[X_t^2\right]}\sqrt{\mathbb{E}_{(\bar{W},\bar{V})}\left[|X_t^1 - X_t^2|^2\right]}$$

we therefore obtain

$$\sup_{t\leq T}\left|\mathbb{Cov}_{(\bar{W},\bar{V})}\left[X_t^1\right] - \mathbb{Cov}_{(\bar{W},\bar{V})}\left[X_t^2\right]\right| \leq 2\sqrt{2\kappa_0(T)\kappa_{\mathrm{contr}}(T)}\left\|P^1 - P^2\right\|_{C^1([0,T])},$$

Using the differential equation for the evolution of the covariance matrix (28) and the previously derived bounds, one also derives the existence of $q(T) > 0$

with $q(T) \xrightarrow{T\to 0} 0$ such that

$$\sup_{t\le T}\left|\frac{\mathrm{d}\mathbb{C}\mathrm{ov}_{(\bar W,\bar V)}\left[X_t^1\right]}{\mathrm{d}t}-\frac{\mathrm{d}\mathbb{C}\mathrm{ov}_{(\bar W,\bar V)}\left[X_t^2\right]}{\mathrm{d}t}\right|\le q(T)\left\|P^1-P^2\right\|_{C^1([0,T])}.$$

Thus we have proven the contraction property for sufficiently small $T>0$ and therefore the existence and uniqueness of solutions to (17) for small time domains. To prove this fact for arbitrary time domains one now simply uses a standard glueing argument.

4.2 The general case

As we have already pointed out, for general H and R, we can not expect the trace of the covariance matrix tr $\mathbb{C}\mathrm{ov}_{(\bar W,\bar V)}\left[X^P\right]$ to be bounded independent of P.

To make up for this, we partially localize the dynamics in a determinstic way. To this end define for arbitrary $k\in\mathbb{N}$

$$\rho_k:\mathbb{R}\to[0,1]:x\mapsto\left(\mathbb{1}_{[-1/2,k-1/2]]}*\rho\right)(x), \tag{31}$$

where ρ is a standard mollifier.

With this we define the process $\left(X_t^k\right)_{t\in[0,T]}$ to be the solution of

$$\begin{aligned}\mathrm{d}\bar X_t^k &= B_t(\bar X_t^k)\mathrm{d}t + C_t(\bar X_t^k)\mathrm{d}W_t+\tilde C_t\mathrm{d}V_t\\ &\cdots+\rho_k(|\bar P_t^k|^2)\left(\bar P_t^kH_t^{\mathrm{T}}+\tilde C_t\right)R_t^{-1}\left(\mathrm{d}Y_t-\frac{H_t\left(\bar X_t^k+\bar m_t^k\right)}{2}\mathrm{d}t\right)\\ &\cdots-\rho_k(|\bar P_t^k|^2)\frac{\bar P_t^kH_t^{\mathrm{T}}+\tilde C_t}{2}R_t^{-1}\tilde C_t^{\mathrm{T}}\rho_k\left(\left|\left(\bar P_t^k\right)^{-1}\right|^2\right)\left(\bar P_t^k\right)^{-1}\left(\bar X_t^k-\bar m_t^k\right)\mathrm{d}t,\end{aligned} \tag{32}$$

where $\bar m_t^k:=\mathbb{E}_{(\bar W,\bar V)}\left[\bar X_t^k\right]$ and $\bar P_t^k:=\mathbb{C}\mathrm{ov}_{(\bar W,\bar V)}\left[\bar X_t^k\right]$.

Note that (32) still falls outside of the standard framework for the analysis of McKean–Vlasov equation, as the product of a bounded and an unbounded Lipschitz function may still not be Lipschitz, and thus the coefficients in (32) therefore are only locally Lipschitz.

Howeber the existence and uniqueness of such a solution X^k for every $k\in\mathbb{N}$ can be proven just as in subsection 4.1 by making a fixed point argument with respect to the covariance matrix $\bar P^k$ as the involvement of ρ_k bounds the covariance matrix independent of the argument in the fixed point map. Let $X^{P,k}$ be the solution of the linear equation

$$\begin{aligned} \mathrm{d}X_t^{P,k} &= B_t\left(X_t^{P,k}\right)\mathrm{d}t + C_t\left(X_t^{P,k}\right)\mathrm{d}W_t + \tilde{C}_t\mathrm{d}V_t \\ &\cdots + \rho_k(|P_t|^2)\left(P_tH_t^{\mathrm{T}} + \tilde{C}_t\right)R_t^{-1}\left(\mathrm{d}Y_t - \frac{H_t\left(X^{P,k} + \bar{m}_t^{P,k}\right)}{2}\mathrm{d}t\right) \\ &\cdots - \rho_k(|P_t|^2)\frac{P_tH_t^{\mathrm{T}} + \tilde{C}_t}{2}R_t^{-1}\tilde{C}_t^{\mathrm{T}}\rho_k\left(\left|(P_t)^{-1}\right|^2\right)P_t^{-1}\left(X^{P,k} - \bar{m}_t^{P,k}\right)\mathrm{d}t, \end{aligned}$$

for any given spsd-matrix valued function P.

In this case we have

$$\begin{aligned} \frac{\mathrm{d}\mathbb{Cov}\left[X_t^{P,k}\right]}{\mathrm{d}t} =& \ll B_t, X_t^{P,k} \gg -\rho_k(|P_t|^2)\left(P_tH_t^{\mathrm{T}} + \tilde{C}_t\right)R_t^{-1}\frac{H_t\mathbb{Cov}\left[X_t^{P,k}\right]}{2} \\ &\cdots - \frac{\mathbb{Cov}\left[X_t^{P,k}\right]H_t^{\mathrm{T}}}{2}R_t^{-1}\left(H_tP_t + \tilde{C}_t^{\mathrm{T}}\right)\rho_k(|P_t|^2) \\ &\cdots - \rho_k(|P_t|^2)\frac{P_tH_t^{\mathrm{T}} + \tilde{C}_t}{2}R_t^{-1}\tilde{C}_t^{\mathrm{T}}\rho_k(|P_t^{-1}|^2)P_t^{-1}\mathbb{Cov}\left[X_t^{P,k}\right] \\ &\cdots - \mathbb{Cov}\left[X_t^{P,k}\right]P_t^{-1}\rho_k(|P_t^{-1}|^2)\tilde{C}_tR_t^{-1}\frac{H_t\bar{P}_t + \tilde{C}_t^{\mathrm{T}}}{2}\rho_k(|P_t|^2) \\ &\cdots + \mathbb{E}_{(\bar{W},\bar{V})}\left[C\left(X_t^{P,k}\right)C\left(X_t^{P,k}\right)^{\mathrm{T}}\right] + \tilde{C}\tilde{C}^{\mathrm{T}} \end{aligned} \tag{33}$$

Since $\mathbb{1}_{[0,k-1]]} \leq \rho_k \leq \mathbb{1}_{[-1,k]]}$ it holds that

$$\begin{aligned} \left|\rho_k(|P_t|^2)\left(P_tH_t^{\mathrm{T}} + \tilde{C}_t\right).\right| &\leq \sqrt{k}|\left|H_t^{\mathrm{T}}\right| + \left|\tilde{C}_t\right| \\ \left|\rho_k(|P_t^{-1}|^2)P_t^{-1}\right| &\leq \sqrt{k} \end{aligned}$$

and one can derive for every fixed k the boundedness of tr $\mathbb{Cov}_{(\bar{W},\bar{V})}\left[X_t^{P,k}\right]$ independent of P by using the Gronwall inequality. As both

$$P \mapsto \rho_k(|P|^2)\left(PH^{\mathrm{T}} + \tilde{C}\right) \text{ and } P \mapsto \rho_k(|P|^2)\frac{PH^{\mathrm{T}} + \tilde{C}}{2}\tilde{C}^{\mathrm{T}}\rho_k\left(\left|(P)^{-1}\right|^2\right)P^{-1}$$

are smooth functions with compact support (and therefore Lipschitz) one can now derive the existence and uniqueness of X^k for every fixed k just as in subsection 4.1.

For the sake of brevity let us define for every k

$$\bar{\rho}_k := \rho_k\left(|\bar{P}_t^k|^2\right) \text{ and } \bar{\rho}_{-k} := \rho_k\left(\left|\left(\bar{P}_t^k\right)^{-1}\right|^2\right)$$

equation (33) gives us

$$\frac{\mathrm{d}\mathrm{tr}\bar{P}_t^k}{\mathrm{d}t} = \ll B_t, \bar{X}_t^k \gg -\bar{\rho}_k \bar{P}_t^k H_t^{\mathrm{T}} R_t^{-} H \bar{P}_t^k - \bar{\rho}_k \left(1+\bar{\rho}_{-k}\right) \frac{\tilde{C}_t R_t^{-1} H_t \bar{P}_t^k + \bar{P}_t^k H^{\mathrm{T}} R_t^{-1} \tilde{C}_t}{2}$$
$$\cdots + \mathbb{E}_{(\bar{W},\bar{V})}\left[C\left(X_t^{P,k}\right) C\left(X_t^{P,k}\right)^{\mathrm{T}}\right] + \tilde{C}_t \left(1 - \bar{\rho}_k \bar{\rho}_{-k} R_t^{-1}\right) \tilde{C}_t^{\mathrm{T}} \tag{34}$$

and thus one derives just as in section 3

$$\frac{\mathrm{d}\mathrm{tr}\bar{P}_t^k}{\mathrm{d}t} \leq \left(2\mathrm{Lip}\left(B_t\right) + d_x \bar{\rho}_k \left(1+\bar{\rho}_{-k}\right) \left|\tilde{C}_t R_t^{-1} H_t\right|\right) \ \mathrm{tr}\bar{P}_t^k$$
$$\cdots + \|C_t\|_\infty^2 + \left|\tilde{C}_t\right|^2 + d_x \bar{\rho}_k \bar{\rho}_{-k} \left|\tilde{C}_t R_t^{-1} \tilde{C}_t^{\mathrm{T}}\right|$$
$$\leq 2\left(\mathrm{Lip}\left(B_t\right) + d_x \left|\tilde{C}_t R_t^{-1} H_t\right|\right) \ \mathrm{tr}\bar{P}_t^k + \|C_t\|_\infty^2 + \left|\tilde{C}_t\right|^2 + d_x \left|\tilde{C}_t R_t^{-1} \tilde{C}_t^{\mathrm{T}}\right|,$$

which is the same differential inequality that also holds for $\bar{P}$. Therefore we derive by the Gronwall lemma that $\sup_{t\leq T} \mathrm{tr}\bar{P}_t^k \leq \bar{\Psi}(T)$ (the constant is defined in (19)) and we see that if

$$k \geq \bar{\Psi}(T)^2 + 1,$$

then $\bar{\rho}_k = \rho_k(|\bar{P}_t^k|^2) = 1$ for all $t \in [0,T]$.

Similarly we can bound $\lambda_{\min}\left(\bar{P}_t^k\right)$ from below. Using (34), we see by employing the same inequalities as in lemma 11, that for the i-th eigenvalue $\lambda^{k,i}$ of $\bar{P}^k$ the differential inequality

$$\frac{\mathrm{d}\lambda_t^{k,i}}{\mathrm{d}t} \geq -2\mathrm{Lip}(B) \sqrt{\bar{\Psi}(T)}\sqrt{\lambda_t^{k,i}} - \bar{\rho}_k \lambda_{\max}\left(H_t^{\mathrm{T}} R_t^{-1} H_t\right) \left(\lambda_t^i\right)^2$$
$$\cdots - 2\bar{\rho}_k \left(1+\bar{\rho}_{-k}\right) \lambda_{\max}\left(\mathrm{Sym}\left(\tilde{C}_t R_t^{-1} H_t\right)\right) \lambda_t^i$$
$$\cdots + \inf_x \lambda_{\min}\left(C_t(x) C_t(x)^{\mathrm{T}}\right) + \lambda_{\min}\left(\tilde{C}_t \left(I - \bar{\rho}_k \bar{\rho}_{-k} R_t^{-1}\right) \tilde{C}_t^{\mathrm{T}}\right)$$

holds. Since $0 \leq \bar{\rho}_k, \bar{\rho}_{-k} \leq 1$, we have due to the variational characterization of the smallest eigenvalue

$$\lambda_{\min}\left(\tilde{C}_t \left(I - \bar{\rho}_k \bar{\rho}_{-k} R_t^{-1}\right) \tilde{C}_t^{\mathrm{T}}\right) = \min_{|v|=1} v^{\mathrm{T}} \tilde{C}_t \tilde{C}_t^{\mathrm{T}} v - \bar{\rho}_k \bar{\rho}_{-k} v^{\mathrm{T}} \tilde{C}_t R_t^{-1} \tilde{C}_t^{\mathrm{T}} v$$
$$\geq \min_{|v|=1} v^{\mathrm{T}} \tilde{C}_t \tilde{C}_t^{\mathrm{T}} v - v^{\mathrm{T}} \tilde{C}_t R_t^{-1} \tilde{C}_t^{\mathrm{T}} v \geq \lambda_{\min}\left(\tilde{C}_t \left(I - R_t^{-1}\right) \tilde{C}_t^{\mathrm{T}}\right).$$

This lets us derive the same inequality as in the proof of Lemma 11

$$\frac{\mathrm{d}\lambda_t^{k,i}}{\mathrm{d}t} \geq -2\mathrm{Lip}(B) \sqrt{\bar{\Psi}(T)}\sqrt{\lambda_t^{k,i}} - \lambda_{\max}\left(H_t^{\mathrm{T}} R_t^{-1} H_t\right) \left(\lambda_t^i\right)^2 - 4\lambda_{\max}\left(\mathrm{Sym}\left(\tilde{C}_t R_t^{-1} H_t\right)\right) \lambda_t^i$$
$$\cdots + \inf_x \lambda_{\min}\left(C_t(x) C_t(x)^{\mathrm{T}}\right) + \lambda_{\min}\left(\tilde{C}_t \left(I - R_t^{-1}\right) \tilde{C}_t^{\mathrm{T}}\right),$$

This inequality allows us to bound $\lambda_{\min}\left(\bar{P}_t^k\right)$ from below, independenty of k. To see this, let $\underline{\lambda}^i$ be a solution of the initial value problem

$$\begin{cases} \frac{\mathrm{d}\underline{\lambda}_t^i}{\mathrm{d}t} & = -2\mathrm{Lip}(B)\ \sqrt{\bar{\Psi}(T)}\sqrt{\underline{\lambda}_t^i} - \lambda_{\max}\left(H_t^{\mathrm{T}} R_t^{-1} H_t\right)\left(\underline{\lambda}_t^i\right)^2 - 4\lambda_{\max}\left(\mathrm{Sym}\left(\tilde{C}_t R_t^{-1} H_t\right)\right)\underline{\lambda}_t^i \\ & \quad \cdots + \frac{\inf_x \lambda_{\min}\left(C_t(x)C_t(x)^{\mathrm{T}}\right) + \lambda_{\min}\left(\tilde{C}_t\left(I - R_t^{-1}\right)\tilde{C}_t^{\mathrm{T}}\right)}{2} \\ \underline{\lambda}_0^i & = \frac{\bar{\lambda}_0^i}{2}, \end{cases}$$

where $\bar{\lambda}_0^i$ denotes the i-th eigenvector of $\bar{P}_0$. Note that the Peano theorem guarentees the existence of such a solution, while its uniqueness is not guaranteed. Thus we have $\left(\lambda^{k,i}\right)'(t_0) > \left(\underline{\lambda}^i\right)'(t_0)$, if $\lambda_{t_0}^{k,i} = \bar{\lambda}_{t_0}^i$. Now [Walter, II. Lemma, page 64] guarantees that $\lambda_t^{k,i} \geq \bar{\lambda}_t^i$ and furthermore it also guarantees that $\bar{\lambda}_t^i > 0$ for all times t. Therefore we have found a uniform (both in k and in i) lower bound.

This lets us conclude that for $k \in \mathbb{N}$ large enough such that $k \geq \frac{1}{d_x\left(\min_{t\leq T}\bar{\lambda}_t^i\right)^2}$, we have $\rho_k\left(\left|\left(\bar{P}_t^k\right)^{-1}\right|^2\right) = 1$ and thus $\bar{X}^k$ is the unique solution of (25) on $[0, T]$.

5 Propagation of Chaos

In this section we aim to prove that the particles defined by the regularized EnKBF (15) indeed converge to the solution of (17) for $M \to \infty$ and $\epsilon \to 0$. A proof for the uncorrelated case with constant signal volatility $C_t(x) = C$, $\tilde{C}_t = 0$ can be found in [LaSt].

5.1 Propagation of chaos for a regularized EnKBF

We consider M independent copies $\tilde{X}^i$, $i = 1, \cdots, M$ of the solution $\tilde{X}$ to (17) satisfying

$$\begin{aligned} \mathrm{d}\tilde{X}_t^i &= B_t(\tilde{X}_t^i)\mathrm{d}t + C_t(\tilde{X}_t^i)\mathrm{d}W_t^i + \tilde{C}_t\mathrm{d}V_t^i \\ &\quad \cdots + \left(\tilde{P}_t H_t^{\mathrm{T}} + \tilde{C}_t\right) R_t^{-1}\left(\mathrm{d}Y_t - \frac{H_t\left(\tilde{X}_t^i + \tilde{m}_t\right)}{2}\mathrm{d}t\right) \\ &\quad \cdots - \frac{\tilde{P}_t H_t^{\mathrm{T}} + \tilde{C}_t}{2} R_t^{-1}\tilde{C}_t^{\mathrm{T}}\tilde{P}_t^{+\epsilon,\nu}\left(\tilde{X}_t^i - \tilde{m}_t\right)\mathrm{d}t. \end{aligned}$$

As noted in remark 14, this equation is well posed.
.jk

Theorem 16. *Define the error term $r_t^i := X_t^i - \tilde{X}_t^i$, then the following result holds*

$$\sup_{t\le T} \frac{1}{M}\sum_{i=1}^{M} |r_t^i|^2 \xrightarrow{M\to\infty} 0$$

in probability (and almost surely with respect to Y).

Proof. For the sake of brevity in formulas let us define $\psi_t(P) := \frac{PH_t^{\mathrm{T}}+\tilde{C}_t}{2}R_t^{-1}\tilde{C}_t^{\mathrm{T}}P^{+\epsilon,\nu}$. We note that since both X^i and $\tilde{X}^i$ share the same initial condition, we have

$$\begin{aligned}
r_t^i &= \int_0^t \left(B_s(X_s^i) - B_s(\tilde{X}_s^i)\right)\mathrm{d}s + \int_0^t \left(C_s(X_s^i) - C_s(\tilde{X}_s^i)\right)\mathrm{d}W_s^i\\
&\cdots + \int_0^t \left(P_s^M - \tilde{P}_s\right) H_s^{\mathrm{T}} R_s^{-1}\left(\mathrm{d}Y_s - \frac{H_s(\tilde{X}_s^i + \tilde{m}_s)}{2}\mathrm{d}s\right)\\
&\cdots - \frac{1}{2}\int_0^t \left(\tilde{P}_s H_s^{\mathrm{T}} + \tilde{C}_s\right) R_s^{-1} H_s \left(\left(X_t^i - \tilde{X}_t^i\right) + \left(x_s^M - \tilde{m}_s\right)\right)\mathrm{d}s\\
&\cdots + \int_0^t \psi_s(\tilde{P}_s)\left(\left(X_t^i - \tilde{X}_t^i\right) - \left(x_s^M - \tilde{m}_s\right)\right)\mathrm{d}s\\
&\cdots - \int_0^t \left(\psi_s(P_s^M) - \psi_s(\tilde{P}_s)\right)\left(X_t^i - x_s^M\right)\mathrm{d}s.
\end{aligned}$$

Using $\mathrm{d}Y_s = H\left(X_s^{\mathrm{ref}}\right)\mathrm{d}s + \Gamma dV_s$ this can be rewritten as

$$\begin{aligned}
r_t^i &= \int_0^t \mathrm{d}\left(X_s^i - \tilde{X}_s^i\right)\\
&= \int_0^t \left(B(X_s^i) - B(\tilde{X}_s^i) - \frac{1}{2}\tilde{P}_s H^{\mathrm{T}} H\left(X_t^i - \tilde{X}_t^i\right)\right)\mathrm{d}s + \cdots\\
&\cdots + \int_0^t \left(C(X_s^i) - C(\tilde{X}_s^i)\right)\mathrm{d}W_s^i + \int_0^t \left(P_s^M - \tilde{P}_s\right) H^{\mathrm{T}}\Gamma\mathrm{d}V_s + \cdots\\
&\cdots + \int_0^t \left(P_s^M - \tilde{P}_s\right) H^{\mathrm{T}} H\left(X_s^{\mathrm{ref}} - \frac{\tilde{X}_s^i + \tilde{m}_s}{2}\mathrm{d}s\right)\mathrm{d}s - \cdots\\
&\cdots - \frac{1}{2}\int_0^t \left(\tilde{P}_s H^{\mathrm{T}} + \tilde{C}\right) H\left(x_s^M - \tilde{m}_s\right)\mathrm{d}s + \cdots\\
&\cdots + \int_0^t \psi(\tilde{P}_s)\left(\left(X_t^i - \tilde{X}_t^i\right) - \left(x_s^M - \tilde{m}_s\right)\right)\mathrm{d}s - \cdots\\
&\cdots - \int_0^t \left(\psi(P_s^M) - \psi(\tilde{P}_s)\right)\left(X_t^i - x_s^M\right)\mathrm{d}s
\end{aligned}$$

Using Itô's rule, the Lipschitz properties of the coefficients, as well as the fact that $2a \cdot b \leq |a|^2 + |b|^2$, we get

$$\frac{1}{M}\sum_{i=1}^{M}\left|r_t^i\right|^2 \leq \int_0^t \left(2\ \mathrm{Lip}(B_s) + \mathrm{Lip}(C_s)^2 + \left|\tilde{P}_s H_s^{\mathrm{T}} R_s^{-1} H_s\right| + 2\left|\psi_s\left(\tilde{P}_s\right)\right| + 2\right)\frac{1}{M}\sum_{i=1}^{M}\left|r_t^i\right|^2 \mathrm{d}s$$

$$\cdots + \frac{2}{M}\sum_{i=1}^{M}\int_0^t r_s^i \cdot \left(C_s(X_s^i) - C_s(\tilde{X}_s^i)\right)\mathrm{d}W_s^i + \frac{2}{M}\sum_{i=1}^{M}\int_0^t r_s^i \cdot \left(P_s^M - \tilde{P}_s\right) H_s^{\mathrm{T}} \Gamma_s \mathrm{d}V_s$$

$$\cdots + \int_0^t \left|P_s^M - \tilde{P}_s\right|^2 \left|H_s^{\mathrm{T}} R_s^{-1} H_s\right|^2 \left(\left|X_s^{\mathrm{ref}}\right|^2 + \frac{\frac{1}{M}\sum_{i=1}^{M}\left|\tilde{X}_s^i\right|^2 + |\tilde{m}_s|^2}{2}\right)\mathrm{d}s$$

$$\cdots - \int_0^t \underbrace{\left(\frac{1}{M}\sum_{i=1}^{M} r_s^i\right)}_{=x_s^M - \tilde{x}_s^M} \cdot \left(\tilde{P}_s H_s^{\mathrm{T}} R_s^{-1} H_s + \tilde{C}_s R_s^{-1} H_s + 2\psi_s\left(\tilde{P}_s\right)\right)\left(x_s^M - \tilde{m}_s\right)\mathrm{d}s$$

$$\cdots + \int_0^t \left|\psi_s(P_s^M) - \psi_s(\tilde{P}_s)\right|^2 \underbrace{\left(\frac{1}{M}\sum_{i=1}^{M}\left|X_s^i - x_s^M\right|^2\right)}_{=\mathrm{tr} P_s^M}\mathrm{d}s$$

$$\cdots + \int_0^t \left|H_s^{\mathrm{T}} R_s^{-1} H_s\right| \left|P_s^M - \tilde{P}_s\right|^2 \mathrm{d}s$$

First we note that

$$-\left(x_s^M - \tilde{x}_s^M\right) \cdot \left(\tilde{P}_s H_s^{\mathrm{T}} R_s^{-1} H_s + \tilde{C}_s R_s^{-1} H_s + 2\psi_s\left(\tilde{P}_s\right)\right)\left(x_s^M - \tilde{m}_s\right)$$

$$\leq \frac{1}{M}\sum_{i=1}^{M}\left|r_s^i\right|^2 + \left|\tilde{P}_s H_s^{\mathrm{T}} R_s^{-1} H_s + \tilde{C}_s R_s^{-1} H_s + 2\psi_s\left(\tilde{P}_s\right)\right|^2 \left|\tilde{x}_s^M - \tilde{m}_s\right|^2$$

and therefore if we denote

$$\mathfrak{lm}_t^M := \frac{2}{M}\sum_{i=1}^{M}\int_0^t r_s^i \cdot \left(C_s(X_s^i) - C_s(\tilde{X}_s^i)\right)\mathrm{d}W_s^i + \frac{2}{M}\sum_{i=1}^{M}\int_0^t r_s^i \cdot \left(P_s^M - \tilde{P}_s\right) H_s^{\mathrm{T}} R_s^{-1} \Gamma_s \mathrm{d}V_s,$$

we derive

$$\frac{1}{M}\sum_{i=1}^{M}\left|r_t^i\right|^2 \leq \int_0^t \left(2\ \mathrm{Lip}(B_s) + \mathrm{Lip}(C_s)^2 + \left|\tilde{P}_s H_s^{\mathrm{T}} R_s^{-1} H_s\right| + 2\left|\psi_s\left(\tilde{P}_s\right)\right| + 3\right)\frac{1}{M}\sum_{i=1}^{M}\left|r_t^i\right|^2 \mathrm{d}s$$

$$\cdots + \int_0^t \left|P_s^M - \tilde{P}_s\right|^2 \left|H_s^{\mathrm{T}} R_s^{-1} H_s\right|^2 \left(\left|X_s^{\mathrm{ref}}\right|^2 + \frac{\frac{1}{M}\sum_{i=1}^{M}\left|\tilde{X}_s^i\right|^2 + |\tilde{m}_s|^2}{2} + 1\right)\mathrm{d}s$$

$$\cdots + \int_0^t \left|\tilde{P}_s H_s^{\mathrm{T}} R_s^{-1} H_s + \tilde{C}_s R_s^{-1} H_s + \psi_s\left(\tilde{P}_s\right)\right|^2 \left|\tilde{x}_s^M - \tilde{m}_s\right|^2 \mathrm{d}s$$

$$\cdots + \int_0^t \left|\psi_s(P_s^M) - \psi_s(\tilde{P}_s)\right|^2\ \mathrm{tr} P_s^M\ \mathrm{d}s + \mathfrak{lm}_t^M.$$

Clearly ψ_s is locally Lipschitz if both arguments are regular. Let us denote its Lipschitz constant on $\left\{ A \in \mathbb{R}^{d_x \times d_y} \mid |A| \leq \kappa \right\}$ by $\mathrm{Lip}_{\mathrm{loc}}(\psi, \kappa)$. This gives us the bound

$$\left|\psi(P_s^M) - \psi(\tilde{P}_s)\right| \leq \mathrm{Lip}_{\mathrm{loc}}\left(\psi_s, \max\left\{\left|P_s^M\right|, \left|\tilde{P}_s\right|\right\}\right)\left|P_t^M - \tilde{P}_t\right|.$$

Now we note that

$$\frac{1}{M}\sum_{i=1}^{M}\left|\tilde{X}_s^i\right|^2 \leq \frac{2}{M}\sum_{i=1}^{M}\left|\tilde{X}_s^i - \tilde{x}_s^m\right|^2 + 2\left|\tilde{x}_s^m\right|^2 \leq \frac{2(M-1)}{M}\left|\tilde{P}_s^M\right|^2 + 2\left|\tilde{x}_s^m\right|^2,$$

and that

$$\left|P_t^M - \tilde{P}_t\right| \leq \left(\sqrt{\frac{M}{M-1}} + \frac{M}{M-1}\right)\sqrt{\mathrm{tr}P_s^M + \mathrm{tr}\tilde{P}_s^M}\left(\frac{1}{M}\sum_{i=1}^{M}\left|r_s^i\right|^2\right)^{1/2} + \left|\tilde{P}_t^M - \tilde{P}_t\right|.$$

Therefore we derive

$$\frac{1}{M}\sum_{i=1}^{M}\left|r_t^i\right|^2 \leq \int_0^t \mathcal{L}_s^1 \frac{1}{M}\sum_{i=1}^{M}\left|r_t^i\right|^2 \mathrm{d}s + \int_0^t \mathcal{L}_s^2\left|\tilde{P}_s^M - \tilde{P}_s\right|^2 \mathrm{d}s + \int_0^t \mathcal{L}_s^3\left|\tilde{x}_s^M - \tilde{m}_s\right|^2 \mathrm{d}s + \mathfrak{m}_t^M,$$

where

$$\mathcal{L}_s^1 := 2\,\mathrm{Lip}(B_s) + \mathrm{Lip}(C_s)^2 + \left|\tilde{P}_s H_s^{\mathrm{T}} R_s^{-1} H_s\right| + 2\left|\psi_s\left(\tilde{P}_s\right)\right| + 3 + \cdots$$

$$\cdots + \left(1 + \mathrm{Lip}_{\mathrm{loc}}\left(\psi_s, \max\left\{\left|P_s^M\right|, \left|\tilde{P}_s\right|\right\}\right)\right)^2\left(\sqrt{\frac{M}{M-1}} + \frac{M}{M-1}\right)^2\left(\mathrm{tr}P_s^M + \mathrm{tr}\tilde{P}_s^M\right)\mathcal{L}_s^2$$

$$\mathcal{L}_s^2 := 2\left(\left|H_s^{\mathrm{T}} R_s^{-1} H_s\right|^2 + 1\right)\left(\left|X_s^{\mathrm{ref}}\right|^2 + \frac{(M-1)}{M}\left|\tilde{P}_s^M\right|^2 + \left|\tilde{x}_s^m\right|^2 + \frac{\left|\tilde{m}_s\right|^2}{2} + 1\right)$$

$$\mathcal{L}_s^3 := \left|\tilde{P}_s\right|^2\left|H_s^{\mathrm{T}} R_s^{-1} H_s\right|^2.$$

From this point on, the proof of convergence coincides with [LaSt].

□

5.2 Uniform approximation property of the regularized equation

It is easy to see that $\tilde{P}$, the covariance matrix of $\tilde{X}$ the solution of (??) satisfies the same upper and lower bounds a $\bar{P}$, i.e.

$$\mathrm{tr}\tilde{P}_t \leq \bar{\Psi}(T) \text{ and } \lambda_{\min}\left(\tilde{P}_t\right) \geq \underline{\Psi}(T) \text{ for all } t \in [0, T].$$

The difference $\bar{X} - \tilde{X}$ satisfies the following integral inequality

$$\begin{aligned}
\left|\bar{X}_t - \tilde{X}_t\right|^2 \leq 8t \int_0^t \mathrm{Lip}(B)^2 \left|\bar{X}_s - \tilde{X}_s\right|^2 \mathrm{d}s + \left|\int_0^t \mathrm{Lip}(B)^2 \left(C(\bar{X}_s) - C(\tilde{X}_s)\right) \mathrm{d}W_s\right|^2 \\
\cdots + \left|\int_0^t \mathrm{Lip}(B)^2 \left(\bar{P}_s - \tilde{P}_s\right) H_s^{\mathrm{T}} R_s^{-1} \mathrm{d}Y_s\right|^2 \\
\cdots + 8t \int_0^t \left|\bar{P}_s - \tilde{P}_s\right|^2 \left|H_s^{\mathrm{T}} R_s^{-1}\right|^2 \left(\left(\left|H_s\right|^2 + \left|\tilde{C}_s\right|^2 \left|\bar{P}_s^{-1}\right|^2\right) \left|\bar{X}_s - \bar{m}_s\right|^2 + 2\left|H_s\right|^2 \left|\bar{m}_s\right|^2\right) \mathrm{d}s \\
\cdots + 8t \int_0^t \left(\left|\tilde{P}_s H_s^{\mathrm{T}} R^{-1} H_s\right|^2 + \left|\left(\tilde{P}_s H_s^{\mathrm{T}} + \tilde{C}_s\right) R^{-1} \tilde{C}_s^{\mathrm{T}} \tilde{P}_s^{+\epsilon,\nu}\right|^2\right) \left(\left|\bar{X}_s - \tilde{X}_s\right|^2 + \left|\bar{m}_s - \tilde{m}_s\right|^2\right) \mathrm{d}s \\
\cdots + 8t \int_0^t \left|\tilde{P}_s H_s^{\mathrm{T}} + \tilde{C}_s\right|^2 \left|R_s^{-1}\right|^2 \left|\tilde{C}_s\right|^2 \frac{\left|\bar{X}_s - \bar{m}_s\right|^2}{2} \left|\bar{P}_s^{-1} - \bar{P}_s^{+\epsilon,\nu}\right|^2 \mathrm{d}s \\
\cdots + 8t \int_0^t \left|\tilde{P}_s H_s^{\mathrm{T}} + \tilde{C}_s\right|^2 \left|R_s^{-1}\right|^2 \left|\tilde{C}_s\right|^2 \frac{\left|\bar{X}_s - \bar{m}_s\right|^2}{2} \left|\bar{P}_s^{+\epsilon,\nu} - \tilde{P}_s^{+\epsilon,\nu}\right| \mathrm{d}s.
\end{aligned}$$

Note that since the spectrum of $\bar{P}$ and $\tilde{P}$ can be uniformely bounded from above and since $\left|\bar{P}_s^{-1} - \bar{P}_s^{+\epsilon,\nu}\right| \lesssim \epsilon$, one easily derives by the Gronwall lemma that there exists a constant $C(T) > 0$ only depending on the timeframe T, such that

$$\mathbb{E}\left[\sup_{t \leq T} \left|\bar{X}_t - \tilde{X}_t\right|^2\right] \leq C(T)\epsilon$$

5.3 Associated stochastic partial differential equation

Let ϕ be an arbitrary smooth test function. Using Itô's rule we derive for the law $\bar{\eta}_t$ of $\bar{X}_t$ defined by equation (17)

$$\begin{aligned}
\mathrm{d}\bar{\eta}_t(\phi) = \bar{\eta}_t(L_t\phi)\mathrm{d}t + \bar{\eta}_t\left(\nabla\phi \cdot \left(\bar{P}_t H_t^{\mathrm{T}} + \tilde{C}_t\right)\right) R_t^{-1} \mathrm{d}Y_t \\
\cdots - \bar{\eta}_t\left(\nabla\phi \cdot \left(\bar{P}_t H_t^{\mathrm{T}} + \tilde{C}_t\right) R_t^{-1} \frac{H(\cdot + \bar{m}_t)}{2}\right) \mathrm{d}t \\
\cdots - \bar{\eta}_t\left(\nabla\phi \cdot \left(\bar{P}_t H_t^{\mathrm{T}} + \tilde{C}_t\right) R_t^{-1} \tilde{C}_t^{\mathrm{T}} \bar{P}_t^{-1} \frac{\cdot - \bar{m}_t}{2}\right) \mathrm{d}t \\
\cdots + \frac{1}{2}\bar{\eta}_t\left(\mathrm{tr}\left[\phi'' \left(\bar{P}_t H_t^{\mathrm{T}} + \tilde{C}_t\right) R_t^{-1} \left(H_t \bar{P}_t + \tilde{C}_t^{\mathrm{T}}\right)\right]\right) \mathrm{d}t,
\end{aligned}$$

which may not coincide with the KSE. In the Gaussian case however we already know that this indeed coincides with the KSE. This can be shown by using integration by parts.

6 Generalization to other EnKBFs

Note that for the uncorrelated case $\tilde{C} = 0$ the solution of (15) is not the only particle system that is known as a EnKBF. Indeed in the literature it is sometimes referred to as the deterministic EnKBF [CrDMJaRu], while the classical EnKBF is given by

$$\mathrm{d}\bar{X}_t^i = B(\bar{X}_t^i)\mathrm{d}t + C(\bar{X}_t^i)\mathrm{d}W_t^i + \bar{P}_t^M H^{\mathrm{T}} R^{-1}\left(\mathrm{d}Y_t - H\bar{X}_t^i\mathrm{d}t - \Gamma\mathrm{d}V_t^i\right),$$

for $i = 1, \cdots, M$. Its mean-field limit is then formally given by $\bar{X}$, the solution of the McKean–Vlasov SDE

$$\mathrm{d}\bar{X}_t = B(\bar{X}_t)\mathrm{d}t + C(\bar{X}_t)\mathrm{d}W_t + \bar{P}_t H^{\mathrm{T}} R^{-1}\left(\mathrm{d}Y_t - H\bar{X}_t\mathrm{d}t - \Gamma\mathrm{d}V_t\right). \tag{35}$$

The for first two moments of $\bar{X}$ suffice to the equations

$$\mathrm{d}\bar{m}_t = \mathbb{E}_Y\left[B\left(\bar{X}_t\right)\right]\mathrm{d}t + \bar{P}_t\left(\mathrm{d}Y_t - H\bar{m}_t\mathrm{d}t\right) \tag{36}$$

and

$$\begin{aligned}\frac{\mathrm{d}\bar{P}_t}{\mathrm{d}t} &= \ll B, \bar{X}_t \gg + \mathbb{E}_Y\left[C(\bar{X}_t)C(\bar{X}_t)^{\mathrm{T}}\right] - 2\bar{P}_t H^{\mathrm{T}} R^{-1} H\bar{P}_t\bar{P}_t H^{\mathrm{T}} R^{-1}\underbrace{\Gamma\Gamma^{\mathrm{T}}}_{=R} H\bar{P}_t \\ &= \ll B, \bar{X}_t \gg + \mathbb{E}_Y\left[C(\bar{X}_t)C(\bar{X}_t)^{\mathrm{T}}\right] - \bar{P}_t H^{\mathrm{T}} R^{-1} H\bar{P}_t.\end{aligned}$$

Thus it is easy to see that the well posedness of (35) can be proven just as it was done for (25).

Another version of the EnKBF that is often considered in the literature for the uncorrelated noise case $\tilde{C} = 0$, is given by

$$\mathrm{d}\bar{X}_t^i = B(\bar{X}_t^i)\mathrm{d}t + CC^{\mathrm{T}}\left(\bar{P}_t^M\right)^+\left(\bar{X}_t^i - \bar{x}_t^M\right)\mathrm{d}t + \bar{P}_t^M H^{\mathrm{T}} R^{-1}\left(\mathrm{d}Y_t - \frac{H\left(\bar{X}_t^i + \bar{x}_t^M\right)}{2}\mathrm{d}t\right),$$

for $i = 1, \cdots, M$. Formally the mean-field limit $\bar{X}$ is then given by the equation

$$\mathrm{d}\bar{X}_t = B(\bar{X}_t)\mathrm{d}t + CC^{\mathrm{T}}\bar{P}_t^{-1}\left(\bar{X}_t - \bar{m}_t\right)\mathrm{d}t + \bar{P}_t H^{\mathrm{T}} R^{-1}\left(\mathrm{d}Y_t - \frac{H\left(\bar{X}_t + \bar{m}_t\right)}{2}\mathrm{d}t\right). \tag{37}$$

The mean $\bar{m}$ of $\bar{X}$ also satisfies (36), while the covariance satisfies

$$\begin{aligned}\frac{\mathrm{d}\bar{P}_t}{\mathrm{d}t} &= \ll B, \bar{X}_t \gg + CC^{\mathrm{T}}\bar{P}_t^{-1}\bar{P}_t + \bar{P}_t\bar{P}_t^{-1}CC^{\mathrm{T}} - \bar{P}_t H^{\mathrm{T}} R^{-1} H\bar{P}_t \\ &= \ll B, \bar{X}_t \gg + 2CC^{\mathrm{T}} - \bar{P}_t H^{\mathrm{T}} R^{-1} H\bar{P}_t.\end{aligned}$$

Showing existence and uniqueness of solutions to (37) can thus be done just as in the proof of theorem 13.

References

[BaCr] ALAN BAIN, DAN CRISAN
Fundamentals of stochastic filtering,
Stochastic Modelling and Applied Probability, 60. Springer, New York, 2009. xiv+390 pp. ISBN: 978-0-387-76895-3

[CrDMJaRu] DAN CRISAN, PIERRE DEL MORAL, AJAY JASRA, HAMZA RUZAYQAT
Log-Normalization Constant Estimation using the Ensemble Kalman-Bucy Filter with Application to High-Dimensional Models
,
arXiv preprint arXiv:2101.11460, 2021

[CoNiNuRe] MICHELE COGHI, TORSTEIN NILSSEN, NIKOLAS NÜSKEN, SEBASTIAN REICH
Rough McKean–Vlasov dynamics for robust ensemble Kalman filtering,
arXiv preprint arXiv:2107.06621, 2021

[DiEi] LUCA DIECI, TIMO EIROLA
On smooth decompositions of matrices,
SIAM J. Matrix Anal. Appl. 20 (1999), no. 3, 800–819

[Jaz] ANDREW H. JAZWINSKI
Stochastic Processes and Filtering Theory,
New York: Academic, 1st Edition, 1970, 376 pp.

[NueReiRoz] NIKOLAS NÜSKEN, SEBASTIAN REICH, PAUL J. ROZDEBA
State and Parameter Estimation from Observed Signal Increments,
Entropy, 21(5):505, 2019. doi: 10.3390/ e21050505

[LaSt] THERESA LANGE, WILHELM STANNAT
Mean field limit of Ensemble Square Root filters - discrete and continuous time,
Foundations of Data Science, Vol. 3, no. 3, 563-588, 2021

[PaStRe] SAHANI PATHIRAJA, WILHELM STANNAT, SEBASTIAN REICH
McKean-Vlasov SDEs in nonlinear filtering,
SIAM J. Control Optim. 59 (2021), no. 6, 4188–4215. 978-0-387-76895-3

[Scheu] MICHAEL SCHEUTZOW
A stochastic Gronwall lemma,
Infin. Dimens. Anal. Quantum Probab. Relat. Top. 16 (2013), no. 2, 1350019, 4 pp.

[Walter] WOLFGANG WALTER
Differential and integral inequalities.,

Translated from the German by Lisa Rosenblatt and Lawrence Shampine Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 55 Springer-Verlag, New York-Berlin 1970 x+352 pp. 34.90 (35.00)