An Ensemble Kalman-Bucy Filter for correlated observation noise

Sebastian Ertel, based on joint work with Wilhelm Stannat

arXiv: 2205.14253

June 1, 2022

Correlated noise framework

We consider the correlated noise framework of [BC09]

(S) $\mathrm{d}X_t = B(X_t)\mathrm{d}t + C\mathrm{d}W_t + \tilde{C}\mathrm{d}V_t$

(O) $\mathrm{d}Y_t = HX_t\mathrm{d}t + \mathrm{d}V_t,$

Our goal: Compute/approximate the posterior

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

Remark: Results can be generalized (colored observations noise, time inhomogeneity, non-constant diffusion, nonlinear observations etc.)

Ensemble Kalman-Bucy filter

For linear, Gaussian signals with uncorrelated observations ($\tilde{C} = 0$) the mean-field limit $\bar{X}$, adhering to

$$\mathrm{d}\bar{X}_t = B\left(\bar{X}_t\right)\mathrm{d}t + C\mathrm{d}\bar{W}_t + \bar{P}_t H^{\mathrm{T}}\left(\mathrm{d}Y_t - \frac{H\left(\bar{X}_t + \bar{m}_t\right)}{2}\mathrm{d}t\right)$$

$$\bar{m}_t := \mathbb{E}_{Y_t}\left[\bar{X}_t\right], \ \bar{P}_t := \mathbb{Cov}_{Y_t}\left[\bar{X}_t\right],$$

of the EnKBF

$$\mathrm{d}X^i_t = B(X^i_t)\mathrm{d}t + C\mathrm{d}W^i_t + P^M_t H^{\mathrm{T}}\left(\mathrm{d}Y_t - \frac{H\left(X^i_t + x^M_t\right)}{2}\mathrm{d}t\right)$$

$$x^M_t := \frac{1}{M}\sum_{j=1}^{M} X^j_t, \ P^M_t := \frac{1}{M}\sum_{j=1}^{M}\left(X^j_t - x^M_t\right)\left(X^j_t - x^M_t\right)^{\mathrm{T}}$$

achieves consistency $\mathrm{Law}(\bar{X}_t) := \bar{\eta}_t = \eta_t$.

Mean-field representations

Principle of EnKF: In the linear, Gaussian case the EnKF works as follows:

1. find a mean-field process $\left(\bar{X}_t\right)_{t\geq 0}$ such that for all $t \geq 0$:

$$\mathrm{Law}(\bar{X}_t) := \bar{\eta}_t = \eta_t \text{ (at least approximately)}$$

2. approximate $\bar{X}$ and its law/moments by an ensemble of (interacting) particles.

Problem: How do we find/choose such a $\bar{X}$?
$\implies$ use Kushner–Stratonovich equation (KSE).

We follow [PRS20], which unified e.g. the filters in [CX10],[YMM13]. Correlated observation noise also covered [NRR21].

The Kushner–Stratonovich equation

Notation: For all suitable functions f let

- $Lf := \sum_{i,j} \frac{(CC^{\mathrm{T}}+\tilde{C}\tilde{C}^{\mathrm{T}})_{ij}}{2} \partial_{x_i}\partial_{x_j} f - \sum_i B_i \partial_{x_i} f$ ...generator of X,
- $\eta_t(f) := \mathbb{E}[\, f(X_t) \mid Y_{0:t} \,]$.

Define the **innovation process** $(I_t)_{t\geq 0}$ by

$$\mathrm{d}I_t = \mathrm{d}Y_t - \eta_t(H)\mathrm{d}t. \tag{1}$$

The posterior η evolves according to the KSE

$$\mathrm{d}\eta_t = L^*\eta_t \mathrm{d}t + \left(\eta_t \left(Hx - \eta_t(H)\right)^{\mathrm{T}} - \nabla \cdot \left(\eta_t \tilde{C}\right)\right) \mathrm{d}I_t.$$

Why mean-field processes?

KSE is a nonlinear, nonlocal (S)PDE $\Longrightarrow$ represent the solution via McKean–Vlasov SDE

This motivates the choice

$$\mathrm{d}\bar{X}_t = B(\bar{X}_t)\mathrm{d}t + C\mathrm{d}\bar{W}_t + \tilde{C}\mathrm{d}\bar{V}_t + a(\bar{X}_t, \bar{\eta}_t)\mathrm{d}t + K(\bar{X}_t, \bar{\eta}_t)\mathrm{d}Y_t,$$

with

- i.i.d. copies $\bar{W}$ and $\bar{V}$ of W and V
- $\mathrm{Law}\left(\bar{X}_t \mid Y_{0:t}\right) = \bar{\eta}_t$

Different interpretation of $\bar{\eta}$

Another way to define $\bar{\eta}$ suitably is, that from now on, all integrals (expectations, covariances, etc.) shall be computed from the joint law of $\bar{W}$ and $\bar{V}$. Thus for any (suitable) function f we have

$$\bar{\eta}_t(f) := \int f(\bar{X}_t)\, \mathbb{P}^{\bar{W}}(\mathrm{d}\bar{w})\, \mathbb{P}^{\bar{V}}(\mathrm{d}\bar{v}),$$

and we are looking for $\bar{X}$ such that

$$\bar{\eta}_t(f) = \eta_t(f).$$

Note that $\bar{W}$, $\bar{V}$ and Y are independent.

Representation via McKean–Vlasov SDEs

Goal: Find a and K such that $\bar{\eta}_t = \eta_t \implies \bar{\eta}$ satisfies the KSE.

Compare the KSE

$$\mathrm{d}\eta_t = L^*\eta_t \mathrm{d}t + \left(\eta_t \left(Hx - \eta_t(H)\right)^{\mathrm{T}} - \nabla \cdot \left(\eta_t \tilde{C}\right)\right)(dY_t - \bar{\eta}_t(H)\mathrm{d}t)$$

and the Fokker-Planck equation of $\bar{X}$

$$\begin{aligned}\mathrm{d}\bar{\eta}_t = L^*\bar{\eta}_t \mathrm{d}t - \nabla \cdot (\bar{\eta}_t K(\cdot, \bar{\eta}_t))\,\mathrm{d}Y_t - \nabla \cdot (\bar{\eta}_t a(\cdot, \bar{\eta}_t))\,\mathrm{d}t \\ \cdots + \frac{1}{2}\sum_{i,j} \partial_{x_i}\partial_{x_j}\left(\bar{\eta}_t K(\cdot, \bar{\eta}_t) K(\cdot, \bar{\eta}_t)^{\mathrm{T}}\right)\mathrm{d}t.\end{aligned}$$

Consistency conditions for K

Comparing the $\mathrm{d}Y_t$ terms in both equations, we see that

$$K = K^0 + \tilde{C} \tag{2}$$

with

$$-\operatorname{div}(\bar{\eta}_t K^0(\cdot, \bar{\eta}_t)) = (H - \bar{\eta}_t(H))^{\mathrm{T}}\, \bar{\eta}_t. \tag{3}$$

Thus K is unique modulo $\ker\left[\operatorname{div}(\bar{\eta}_t \cdot)\right]$.

Interpretation of the gain term

Writing (3) in flux form

$$\int_{\partial D} \bar{\eta}_t(-\nu_D)^{\mathrm{T}} K^0(\cdot, \bar{\eta}_t)\mathrm{d}s = \int_D H^{\mathrm{T}} x \bar{\eta}_t(x)\mathrm{d}x - \bar{\eta}_t(H^{\mathrm{T}}),$$

for arbitrary domain D, we see that

K is a velocity

such that

flux $K\eta =$ the difference to expected observation.

Consistency conditions for a

Using (2) to simplify

$$\sum_{i,j} \partial_{x_i} \partial_{x_j} \left(\bar{\eta}_t K(\cdot, \bar{\eta}_t) K(\cdot, \bar{\eta}_t)^{\mathrm{T}} \right),$$

one derives that

$$a(\cdot, \bar{\eta}_t) = -\frac{K(\cdot, \bar{\eta}_t)\,(H + \bar{\eta}_t(H))}{2} + \frac{\left(\left(K(\cdot, \bar{\eta}_t) \cdot \nabla \right) K^{\mathrm{T}}(\cdot, \bar{\eta}_t) \right)^{\mathrm{T}}}{2}$$

$$\cdots + \frac{K(\cdot, \bar{\eta}_t)\,\mathrm{div}\left(\bar{\eta}_t \tilde{C} \right)^{\mathrm{T}}}{2\ \bar{\eta}_t} + \Omega_t^0$$

for some $\Omega_t^0 \in \ker\left[\mathrm{div}(\bar{\eta}_t \cdot)\right]$.

Representation via McKean–Vlasov SDEs II

Since $\tilde{C}$ is constant, we note that

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

Thus $\bar{X}$ satisfies the McKean–Vlasov SDE

$$\begin{aligned} \mathrm{d}\bar{X}_t &= B(\bar{X}_t)\mathrm{d}t + C\mathrm{d}\bar{W}_t + \tilde{C}\mathrm{d}\bar{V}_t \\ &\cdots + K(\bar{X}_t, \bar{\eta}_t)\left(\mathrm{d}Y_t - \frac{H\bar{X}_t + \bar{\eta}_t(H)}{2}\right) \\ &\cdots + \frac{\left(\left(K\left(\cdot, \bar{\eta}_t\right) \cdot \nabla\right) K^{\mathrm{T}}\left(\cdot, \bar{\eta}_t\right)\right)^{\mathrm{T}}}{2}\mathrm{d}t \\ &\cdots + \frac{K(\bar{X}_t, \bar{\eta}_t)\tilde{C}\nabla \log \bar{\eta}_t}{2}\mathrm{d}t + \Omega_t^0 \mathrm{d}t. \end{aligned} \tag{4}$$

Consistent mean-field representation in the Gaussian case

For $\bar{\eta}_t = \mathcal{N}\left(\bar{m}_t, \bar{P}_t\right)$ it is easy to show that one can choose

$$K^0\left(x, \bar{\eta}\right) = \bar{P}_t H^{\mathrm{T}}$$

Thus $\bar{X}$ is given by equation

$$\begin{aligned}
\mathrm{d}\bar{X}_t &= B\left(\bar{X}_t\right)\mathrm{d}t + C\mathrm{d}\bar{W}_t + \tilde{C}\mathrm{d}\bar{V}_t \\
&\cdots + \left(\bar{P}_t H^{\mathrm{T}} + \tilde{C}\right)\left(\mathrm{d}Y_t - \frac{H\left(\bar{X}_t + \bar{m}_t\right)}{2}\mathrm{d}t\right) \\
&\cdots - \left(\bar{P}_t H^{\mathrm{T}} + \tilde{C}\right)\tilde{C}^{\mathrm{T}}\bar{P}_t^{-1}\frac{\bar{X}_t - \bar{m}_t}{2}\mathrm{d}t.
\end{aligned} \tag{5}$$

Justification in the non Gaussian case

Integration by parts shows

$$\mathbb{E}_Y\left[K^0\left(\bar{X}_t, \bar{\eta}_t\right)\right] = \mathbb{C}\mathrm{ov}_Y\left[\bar{X}_t\right] H^{\mathrm{T}}. \tag{6}$$

Thus the EnKBF is a universal 0-order approximation of consistent mean-field filters, for example w.r.t.

- Karhunen-Loeve expansion
- polynomial projections.

Well-posedness of the mean-field EnKBF - part I

The EnKBF (5) is a McKean–Vlasov equation with locally Lipschitz coefficients.

Proving well-posedness for:

- **SDEs:**

$$\text{locally Lipschitz} \xrightarrow{\text{stopping time}} \text{global Lipschitz}$$

- **McKean–Vlasov:**

$$\text{locally Lipschitz} \quad \text{~~stopping time~~} \quad \text{changed dynamics}$$

Counter examples showing non uniqueness of locally Lipschitz McKean–Vlasov equations exist [S87].

Well-posedness of the mean-field EnKBF - part II

Basic idea: fixed point argument w.r.t. the covariance $\bar{P}$.

For linear signals $\bar{P}$ decouples from (5) via Kalman–Bucy equations $\implies$ use solution as the argument in fixed point equation [CDMJR21].

Not possible for nonlinear signals (no decoupled characterization of the fixed point).

[CNNR21] proved well posedness for a different version of the EnKBF without the inverse and under the assumption that H is bounded.

Well-posedness of the mean-field EnKBF - part III

Theorem

Assume that $\bar{P}_0$ is regular and that

$$\lambda_{\min}\left(CC^{\mathrm{T}}\right) > 0.$$

Then there exists a unique solution $\bar{X}$ of (5).

Main tool: Spectral bounds for $\bar{P}$

$$\begin{aligned}
\frac{\mathrm{d}\lambda_t^i}{\mathrm{d}t} &\geq -2\mathrm{Lip}(B)\,\sqrt{\mathrm{tr}\bar{P}_t}\sqrt{\lambda_t^i} + \lambda_{\min}\left(CC^{\mathrm{T}}\right) \\
&\quad \cdots - \lambda_{\max}\left(H^{\mathrm{T}}H\right)\left(\lambda_t^i\right)^2 - 2\left|\tilde{C}_t R_t^{-1} H_t\right|\lambda_t^i \\
\frac{\mathrm{d}\lambda_t^i}{\mathrm{d}t} &\leq 2\mathrm{Lip}(B)\,\sqrt{\mathrm{tr}\bar{P}_t}\sqrt{\lambda_t^i} + \lambda_{\max}\left(CC^{\mathrm{T}}\right) \\
&\quad \cdots - \lambda_{\min}\left(H^{\mathrm{T}}H\right)\left(\lambda_t^i\right)^2 + 2\left|\tilde{C}_t R_t^{-1} H_t\right|\lambda_t^i.
\end{aligned}$$

Well-posedness of the mean-field EnKBF - part IV

For $H = I$ and $\tilde{C} = 0$ upper bounds are robust w.r.t. perturbations of $\bar{P}$ in the dynamics.

Our proof relies on the linearity of H.

nonlinear, Lipschitz continuous H + nonlinear signal dynamics not covered by existing literature.

EnKBF for correlated observation noise

A canonical way to approximate (5) uses the interacting particle system $X^i,\ i = 1, \cdots, M$ determined by

$$\begin{aligned} \mathrm{d}X_t^i &= B(X_t^i)\mathrm{d}t + C\mathrm{d}W_t^i + \tilde{C}\mathrm{d}V_t^i \\ &\cdots + \left(P_t^M H^{\mathrm{T}} + \tilde{C}\right)\left(\mathrm{d}Y_t - \frac{H\left(X_t^i + x_t^M\right)}{2}\mathrm{d}t\right) \\ &\cdots - \left(P_t^M H^{\mathrm{T}} + \tilde{C}\right)\tilde{C}^{\mathrm{T}}\left(P_t^M\right)^{+}\frac{X_t^i - x_t^m}{2}\mathrm{d}t \end{aligned} \tag{7}$$

Problem: $\left(P_t^M\right)^{+}\left(X_t^i - x_t^M\right)$ may develop singularities.

Well posedness of the EnKBF

Theorem

We assume that P_0^M is regular and that for all $t > 0$

$$\lambda_{\min}\left(CC^{\mathrm{T}}\right) - \frac{2}{M-1}\left(1 + \sqrt{\dim(X)}\right)\left(|C|^2 + \left|\tilde{C}\right|^2\right) > 0. \quad (8)$$

Then there exists a unique strong solution to (7).

The proof uses bounds for both P^M and $\left(P^M\right)^+$ similar to the ones derived in the mean-field system.

Dynamics of the spectral decomposition cant be used due to missing differentiablity.

Note that the regularity of P_0^M implies $M > \dim(X)$.

Computing the Pseudoinverse

The inflation term can be computed in linear complexity using the recursion found in [Kov79]

$$\left(P^M\right)^+ = \left(\frac{(M-2)P^{M-1} + \hat{X}\hat{X}^\mathrm{T}}{M-1}\right)^+$$

$$= (M-1)\frac{\left(P^{M-1}\right)^+}{M-2} + (M-1)\frac{\frac{\left(P^{M-1}\right)^+}{M-2}\hat{X}\hat{X}^\mathrm{T}\frac{\left(P^{M-1}\right)^+}{M-2}}{1+\hat{X}^\mathrm{T}\frac{\left(P^{M-1}\right)^+}{M-2}\hat{X}}$$

$$\cdots + (M-1)\frac{\hat{X}_\perp\hat{X}_\perp^\mathrm{T}}{\left|\hat{X}_\perp\right|^4}$$

with

$$\hat{X} := X^M - x^M$$

$$\hat{X}_\perp := \hat{X}_\perp - P^{M-1}\left(P^{M-1}\right)^+\hat{X}_\perp$$

Propagation of chaos

Theorem

Let $\bar{X}^i,\ i = 1, \cdots, M$ be i.i.d. copies of $\bar{X}$ and define the error term $r^i_t := X^i_t - \bar{X}^i_t$, then

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

in probability (and almost surely w.r.t. Y).

We can derive implicit rates

$$\sup_{M \in \mathbb{N}} \sqrt{M} \sqrt{\mathbb{E}\left[\sup_{t \le T \wedge \zeta_\kappa} \frac{1}{M} \sum_{i=1}^{M} \left|r^i_t\right|^2\right]} \le C(\kappa, T) < +\infty,$$

where ξ_κ is a hitting time of level κ for both P^M and its inverse.

A. Bain, D. Crisan
Fundamentals of Stochastic Filtering
Book, 2009, Volume 60, ISBN : 978-0-387-76895-3

M. Coghi, T. Nilssen, N. Nüsken, S. Reich
Rough McKean–Vlasov dynamics for robust ensemble Kalman filtering,
arXiv preprint arXiv:2107.06621, 2021

D. Crisan, P. Del Moral, A. Jasra, H. Ruzayqat
Log-Normalization Constant Estimation using the Ensemble Kalman-Bucy Filter with Application to High-Dimensional Models
arXiv:2101.11460, 2021.

D. Crisan, J. Xiong *Approximate McKean–Vlasov representations for a class of SPDEs*
Stochastics, Vol. 82, 2010.

P. Kovanic *On the Pseudoinverse of a Sum of Symmetric Matrices with Applications to Estimation*, Kybernetika, Vol. 15, No. 5, 1979.

Nikolas Nüsken, Sebastian Reich, Paul J. Rozdeba *State and Parameter Estimation from Observed Signal Increments*
Entropy, Vol. 21, no. 5, 505, 2019.

S. Pathiraja, S. Reich, W. Stannat
McKean–Vlasov SDEs in nonlinear filtering
arXiv:2007:12658

M. Scheutzow
Uniqueness and non-uniqueness of solutions of Vlasov-McKean equations
j. Austral. Math. Soc. (Series A) 43 (1987),246-256

T. Yang, P. Mehta, S. Meyn *Feedback particle filter*
IEEE Transactions on Automatic Control, Vol. 58, 2465–2480, 2013.