Mean-field Ensemble Kalman Filter

Kody Law[1], Tembini[2], Raul Tempone[3]

[1] CEMSE Division, King Abdullah University of Science and Technology,
[2] CEMSE Division, King Abdullah University of Science and Technology,
[3] NYU Abu Dhabi.

kody.law@kaust.edu.sa, tembine@gmail.com,
raul.tempone@kaust.edu.sa

Abstract

A proof of convergence of the standard EnKF generalized to non-Gaussian state space models is provided. A density-based deterministic approximation of the mean-field limiting EnKF (MFEnKF) is proposed, consisting of a PDE solver and a quadrature rule. Given a certain minimal order of convergence κ between the two, this extends to the deterministic filter approximation, which is therefore asymptotically superior to standard EnKF for $d < 2\kappa$. The fidelity of approximation of the true distribution is also established using an extension of total variation metric to random measures. This is limited by a Gaussian bias term arising from non-linearity/non-Gaussianity of the model, which arises in both deterministic and standard EnKF. Numerical results support and extend the theory.

1. Setting

Let $\mathcal{K} : \mathbb{R}^d \to \mathcal{P}(\mathbb{R}^d)$. Consider the Markov chain

$$u_{j+1} \sim \mathcal{K}(u_j, \cdot),\ j \in \mathbb{N}, \quad u_0 \sim N(m_0, C_0). \tag{1}$$

Data $Y_k = \{y_j\}_{j=1}^k$ defined as

$$y_j = Hu_j + \eta_j,\ j \in \mathbb{Z}^+,$$

where $H : \mathbb{R}^d \to \mathbb{R}^m$ is linear and $\eta = \{\eta_j\}_{j\in\mathbb{Z}^+}$ is an i.i.d. sequence, independent of u_0 and the noise in $\mathcal{K}$, with $\eta_1 \sim N(0, \Gamma)$.

1.1 Filtering distribution

Distribution of $u_j|Y_j$ is *gold standard*. Likelihood function is

$$g(u, y) = \frac{\tilde g(u, y)}{\int \tilde g(u, y)du}, \qquad \tilde g(u, y) = e^{-\frac12|y-Hu|^2_\Gamma}, \tag{2}$$

with shorthand $g_j(u) = g(u, y_j)$. Define $\mathcal{C}_j$ as follows. For $u \sim \hat\rho$, $u|y_j \sim \rho$, where

$$\rho = \mathcal{C}_j\hat\rho := \frac{\hat\rho g_j}{\int \hat\rho g_j}. \tag{3}$$

Recursion for filtering density $\rho_j = \mathcal{C}_j\mathcal{K}\rho_{j-1}$

Prediction $\hat\rho_j = \mathcal{K}\rho_{j-1}$,

Update $\rho_j = \mathcal{C}_j\hat\rho_j$,

Optimal filter given by

$$\mathbb{E}(u_j|Y_j) = \operatorname{argmin}_{\hat u_j(Y_j)}\mathbb{E}|\hat u_j(Y_j) - u_j|^2.$$

1.2 Optimal Linear Filtering

$$m_j(y_j) = \operatorname{argmin}_{\{\hat u_j(y_j)=K_jy_j+b_j\}}\mathbb{E}|\hat u_j(y_j) - u_j|^2,$$

Optimizing with respect to K_j and b_j gives

$$m_j(y_j) = \mathbb{E}u_j + K_j(y_j - H\mathbb{E}u_j),$$
$$K_j = \mathbb{E}[(u_j - \mathbb{E}u_j)\otimes(y_j - \mathbb{E}y_j)]\mathbb{E}[(y_j - \mathbb{E}y_j)\otimes(y_j - \mathbb{E}y_j)]^{-1}.$$

2. ENKF

2.1 Mean-field EnKF

Prediction
$$\begin{cases} \widehat v_{j+1} &\sim \mathcal{K}(v_j,\cdot),\\ {}_{j+1} &= \mathbb{E}\widehat v_{j+1},\\ \widehat C_{j+1} &= \mathbb{E}(\widehat v_{j+1} - {}_{j+1})\otimes(\widehat v_{j+1} - {}_{j+1})\end{cases}$$

Analysis
$$\begin{cases} S_{j+1} &= H\widehat C_{j+1}H^T + \Gamma\\ K_{j+1} &= \widehat C_{j+1}H^TS^{-1}_{j+1}\\ \tilde y_{j+1} &= y_{j+1} + \eta_{j+1}\\ v_{j+1} &= (I - K_{j+1}H)\widehat v_{j+1} + K_{j+1}\tilde y_{j+1}.\end{cases}$$

Here η_j are i.i.d. draws from $N(0, \Gamma)$

2.2 Finite ensemble EnKF

Prediction
$$\begin{cases} \widehat v^{(n)}_{j+1} &\sim \mathcal{K}(v^{(n)}_j,\cdot),\ n = 1, ..., N\\ \widehat m_{j+1} &= \frac1N\sum_{n=1}^N \widehat v^{(n)}_{j+1},\\ \widehat C_{j+1} &= \frac1N\sum_{n=1}^N(\widehat v^{(n)}_{j+1} - \widehat m_{j+1})\otimes(\widehat v^{(n)}_{j+1} - \widehat m_{j+1})\end{cases}$$

Analysis
$$\begin{cases} v^{(n)}_{j+1} &= (I - K_{j+1}H)\widehat v^{(n)}_{j+1} + K_{j+1}y^{(n)}_{j+1}\\ y^{(n)}_{j+1} &= y_{j+1} + \eta^{(n)}_{j+1}\end{cases}$$

3. Fokker-Planck Filters

$$du = F(u)dt + \sqrt{2b}dW, \quad u(0) = u_0,$$

$F : \mathbb{R}^d \to \mathbb{R}^d$ differentiable and Lipschitz, $b \in (0, \infty)$ constant.
Density governed by Fokker-Planck equation

$$\partial_t\rho = \mathcal{G}\rho, \quad \mathcal{G}\rho = \nabla\cdot(b\nabla\rho - F\rho), \quad \rho(u, 0) = \delta(u_0 - u),$$

3.1 FPF Algorithms

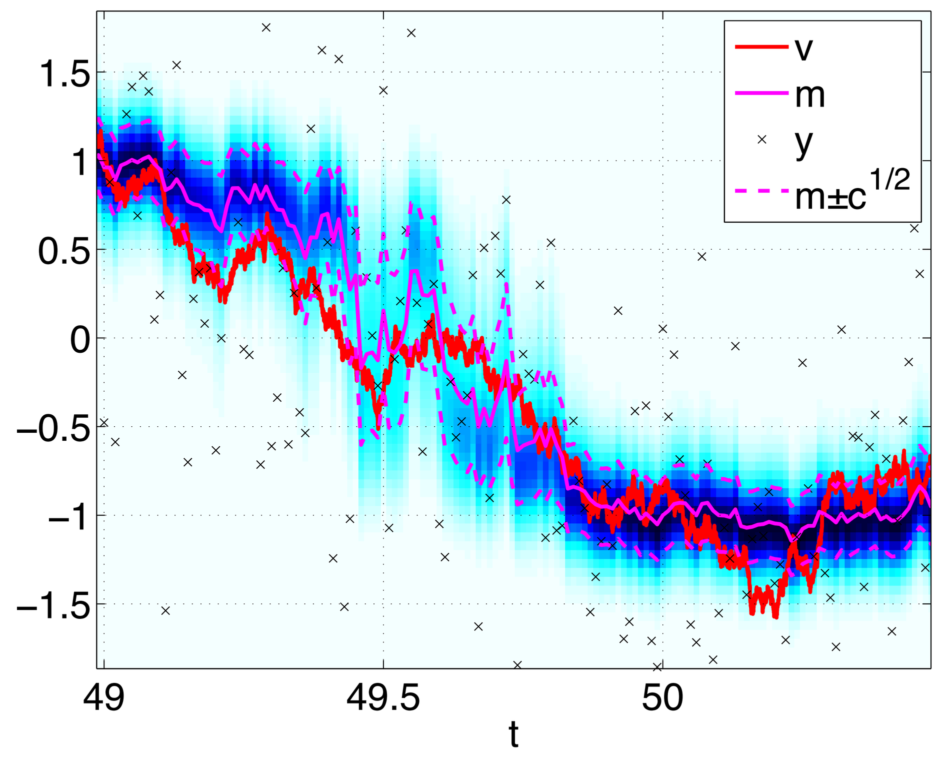

Figure 1: $F(u) = au(1-u^2)/(1+u^2)$, $m = \mathbb{E}(u|y)$, $c = \mathbb{E}[(u-m)^2|y]$.

Full FPF
- (1) Discretize the density at time j over space.
- (2) Evolve with accurate time-stepper to time $j + 1$.
- (3) Update using integration rule, return to step (2).

MFEnKF
- (1), (2) Same as above.
- (3) Obtain mean and covariance of the *prediction*.
- (4) Change variables $u \to (I - K_jH)u$ using interpolation.
- (5) Convolve with the density of $K_jN(y_j, \Gamma)$ on range of K_j.

MFEnKF-G1
- (1),(2),(3) Same as in Algorithm Full FPF.
- (4) Approximate the *updated* distribution by the Gaussian.

MFEnKF-G2
- (1),(2) Same as above.
- (3) Approximate the *predicting* density by a Gaussian.
- (4) Use Kalman update formula for the mean and covariance.

Approximated in d−dimensions with N d.o.f. to $\mathcal{O}(N^{-\kappa/d})$: $\kappa = \min\{\kappa_1, \kappa_2\}$, PDE method order κ_1, quadrature order κ_2.

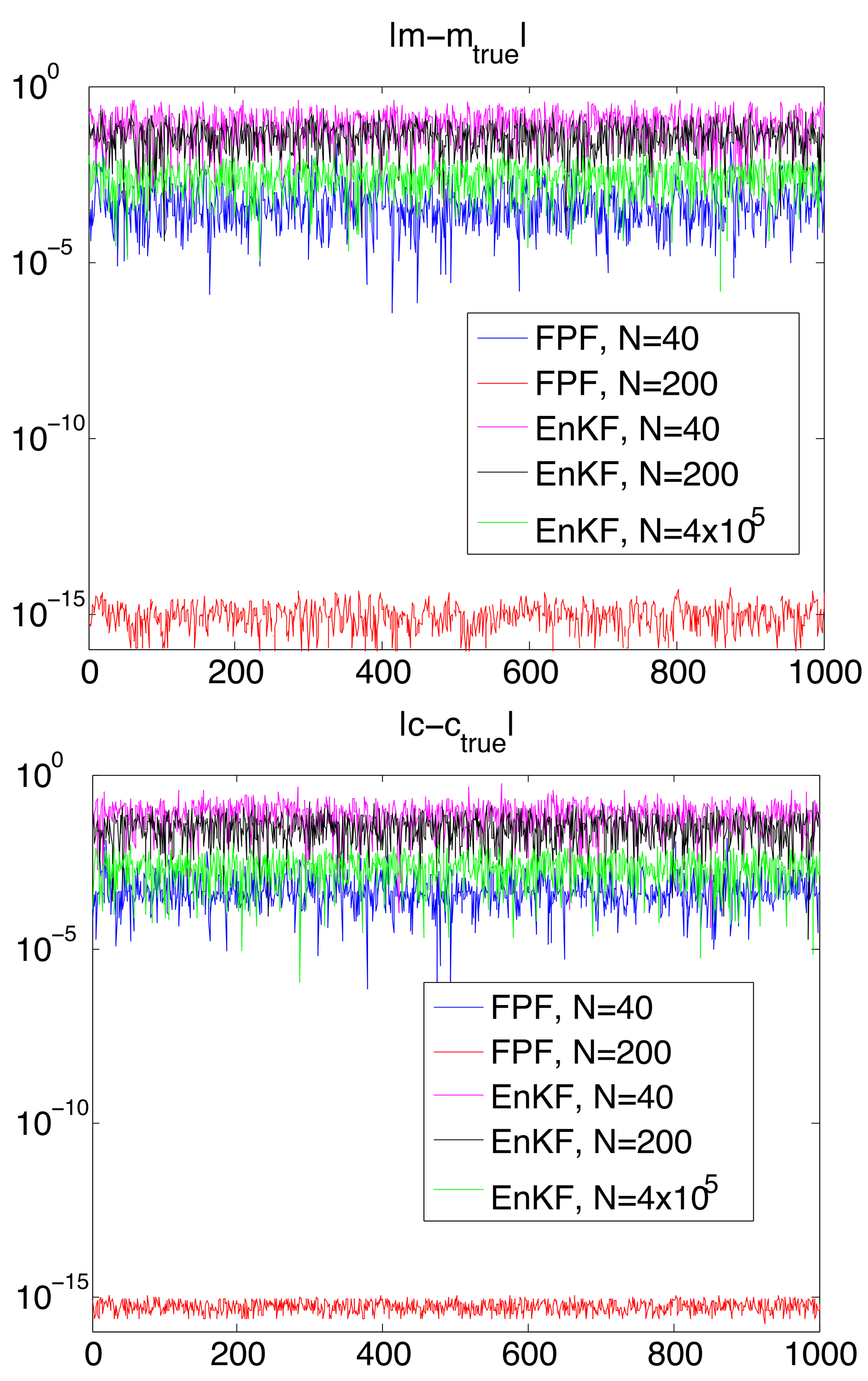

Figure 2: $F(u) = -au$, *Error of FPF MFEnKF-G2 vs. EnKF.*

3.2 ENKF Converges to MFENKF

Let $\pi, f : \mathbb{R}^d \to \mathbb{R}$ and $|f|_L = |f|_\infty + \sup_{x\neq x'}\frac{|f(x)-f(x')|}{|x-x'|}$. Define

$$\|\pi\| = \sup_{|f|_L\le1}\sqrt{\mathbb{E}\left|\int f\pi\right|^2},$$

Now define

$$d(\pi, p) = \|\pi - p\|. \tag{4}$$

- ρ_j – filtering density.
- π_j – the mean-field EnKF density.
- $\pi^N_j \approx \pi_j$ – the deterministic approximation (MFEnKF).
- $\hat\pi^N_j \approx \pi_j$ – the standard Monte-Carlo EnKF.

THEOREM I: It is true that $d(\hat\pi^N_j, \pi_j) = \mathcal{O}(N^{-1/2})$.
Proof: See [1], Sec. 4.3.1, for extension of [2].

4. MFENKF is *not* the Posterior

$$d\left(\hat\pi^N_j, \rho_j\right) \le \underbrace{d\left(\hat\pi^N_j, \pi_j\right)}_{\text{ensemble error}} + \underbrace{d\left(\pi_j, \rho_j\right)}_{\text{Gaussian error}},$$

$$d\left(\pi^N_j, \rho_j\right) \le \underbrace{d\left(\pi^N_j, \pi_j\right)}_{\text{discretization error}} + \underbrace{d\left(\pi_j, \rho_j\right)}_{\text{Gaussian error}}.$$

Assume (i) $\alpha \le g(u, y) \le \alpha^{-1}\ \ \forall\, u \in \mathbb{R}^d,\ y \in \mathbb{R}^m,\ j \in \mathbb{Z}$. And (ii) for time h, $\mathcal{K}_h\pi_j = \pi_j + \mathcal{O}(h)$ and $\mathcal{K}_h\rho_j = \rho_j + \mathcal{O}(h)$. Then
THEOREM II: Given observation increment h, and under assumptions above, as $h \to 0$ and $N \to \infty$

$$d\left(\pi^N_j, \rho_j\right) = \mathcal{O}(N^{-\kappa/d} + \lambda h) \quad \text{and} \quad d\left(\hat\pi^N_j, \rho_j\right) = \mathcal{O}(N^{-1/2} + \lambda h)$$

where $\lambda = 0$ if F is linear and $\lambda = 1$ if f is nonlinear.

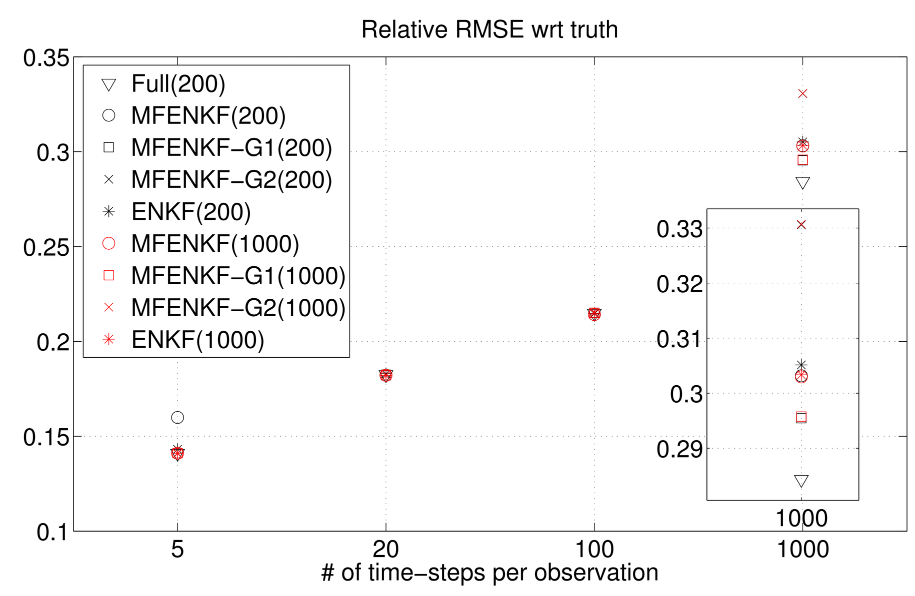

Figure 3: *Nonlinear case, RMSE of mean with respect to the true (unconditioned) signal.*

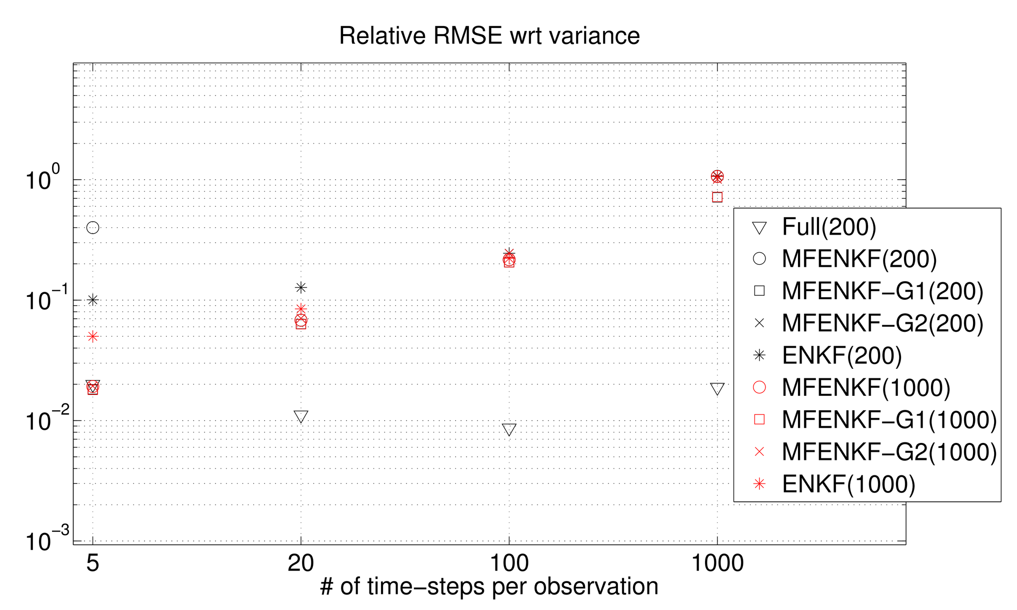

Figure 4: *Nonlinear case, RMSE of covariance with respect to the true posterior covariance.*

Acknowledgements

Research reported in this publication was supported by the King Abdullah University of Science and Technology (KAUST). K. J.H. Law, H. Tembine and R. Tempone are members of the KAUST SRI Center for Uncertainty Quantification.

References

[1] K. J.H. Law, H. Tembine, R. Tempone. "DETERMINISTIC METHODS FOR FILTERING, PART I: MEAN-FIELD ENSEMBLE KALMAN FILTERING." http://arxiv.org/abs/1409.0628.

[2] F. LeGland, V. Monbet, V.D Tran. "Large sample asymptotics for the ensemble kalman filter." The Oxford Handbook of Nonlinear Filtering, 598-631, 2011.