Some theoretical aspects of Particle Filters and Ensemble Kalman Filters

P. Del Moral

Mathematical Foundations of Data Assimilation and Inverse Problems - FoCM Paris June 2023.

Brief review on filtering = X_t given observation $(Y_s)_{s\leqslant t}$

Brief review on filtering = X_t given observation $(Y_s)_{s\leqslant t}$

- **Linear+Gaussian world: Regression form., `Kalman filters`.**

Brief review on filtering = X_t given observation $(Y_s)_{s\leqslant t}$

- **Linear+Gaussian world: Regression form.,** `Kalman filters`**.**
- **Nonlinear/Non-Gaussian: Bayes rule, Nonlinear filtering eq.**

Brief review on filtering = X_t given observation $(Y_s)_{s \leqslant t}$

- **Linear+Gaussian world: Regression form., `Kalman filters`.**
- **Nonlinear/Non-Gaussian: Bayes rule, Nonlinear filtering eq.**
- **Application areas:**

  Data assimilation, forecasting, tracking, multiple objects tracking, machine learning...

Brief review on filtering = X_t given observation $(Y_s)_{s \leqslant t}$

- **Linear+Gaussian world: Regression form.,** `Kalman filters`**.**
- **Nonlinear/Non-Gaussian: Bayes rule, Nonlinear filtering eq.**
- **Application areas:**

  Data assimilation, forecasting, tracking, multiple objects tracking, machine learning...and much more $\rightsquigarrow$ Feynman-Kac sg, ground states Schrödinger sg, rare events, molecular chemistry, polymers,...
- **+ all celebrated application areas of Kalman filter:** localization/positioning/navigation/guidance systems (radar/sonar), regulation/control systems, dynamic Bayesian networks, hidden Markov chains, time series, health monitoring, nuclear medicine, brain computer interfaces,...

Brief review on filtering = X_t given observation $(Y_s)_{s \leqslant t}$

- **Linear+Gaussian world: Regression form.,** `Kalman filters`**.**
- **Nonlinear/Non-Gaussian: Bayes rule, Nonlinear filtering eq.**
- **Application areas:**

  Data assimilation, forecasting, tracking, multiple objects tracking, machine learning...and much more $\rightsquigarrow$ Feynman-Kac sg, ground states Schrödinger sg, rare events, molecular chemistry, polymers,...
- **+ all celebrated application areas of Kalman filter:** localization/positioning/navigation/guidance systems (radar/sonar), regulation/control systems, dynamic Bayesian networks, hidden Markov chains, time series, health monitoring, nuclear medicine, brain computer interfaces,...

Approximation/Computational methods

- Monte Carlo, IS, SIS, MCMC, **Particle filters, EnKF.**

Brief review on filtering = X_t given observation $(Y_s)_{s \leqslant t}$

- **Linear+Gaussian world: Regression form.,** `Kalman filters`**.**
- **Nonlinear/Non-Gaussian: Bayes rule, Nonlinear filtering eq.**
- **Application areas:**

  Data assimilation, forecasting, tracking, multiple objects tracking, machine learning...and much more $\rightsquigarrow$ Feynman-Kac sg, ground states Schrödinger sg, rare events, molecular chemistry, polymers,...
- **+ all celebrated application areas of Kalman filter:** localization/positioning/navigation/guidance systems (radar/sonar), regulation/control systems, dynamic Bayesian networks, hidden Markov chains, time series, health monitoring, nuclear medicine, brain computer interfaces,...

Approximation/Computational methods

- Monte Carlo, IS, SIS, MCMC, **Particle filters, EnKF.**
- Law of large numbers, Ergodic theo, **Stoch. Perturbation theory**.

Sample mean $\quad m_t := \frac{1}{N} \sum_{1 \leqslant i \leqslant N} X_t^i \quad$ with **iid** copies X_t^i of X_t

Sample mean $\quad m_t := \frac{1}{N} \sum_{1 \leqslant i \leqslant N} X_t^i \quad$ with **iid** copies X_t^i of X_t

$\rightsquigarrow$ **stochastic perturbation formulation**

$$\mathbf{m_t} := \mathbb{E}(\mathbf{X_t}) + \frac{1}{\sqrt{N}} \; V_t^N$$

$$\text{Sample mean} \quad m_t := \frac{1}{N} \sum_{1 \leqslant i \leqslant N} X_t^i \quad \text{with } \textbf{iid} \text{ copies } X_t^i \text{ of } X_t$$

⇝ **stochastic perturbation formulation**

$$\mathbf{m_t} := \mathbb{E}(\mathbf{X_t}) + \frac{1}{\sqrt{N}} \; V_t^N$$

with bias-variance perturbation control

$$\mathbb{E}(V_t^N) = 0 \quad \& \quad \mathbb{E}((V_t^N)^2) = \mathbb{E}((X_t - \mathbb{E}(X_t))^2)$$

$$\text{Sample mean} \quad m_t := \frac{1}{N}\sum_{1\leqslant i\leqslant N} X_t^i \quad \text{with } \textbf{iid} \text{ copies } X_t^i \text{ of } X_t$$

$\rightsquigarrow$ **stochastic perturbation formulation**

$$\mathbf{m_t} := \mathbb{E}(\mathbf{X_t}) + \frac{1}{\sqrt{N}}\ V_t^N$$

with bias-variance perturbation control

$$\mathbb{E}(V_t^N) = 0 \quad \& \quad \mathbb{E}((V_t^N)^2) = \mathbb{E}((X_t - \mathbb{E}(X_t))^2)$$

Key Observation:

$$X_t \text{ stable} \implies \sup_{t\geqslant 0} \mathbb{E}((m_t - \mathbb{E}(X_t))^2) \leqslant c/N$$

$$\text{Sample mean} \quad m_t := \frac{1}{N} \sum_{1 \leqslant i \leqslant N} X_t^i \quad \text{with } \textbf{iid} \text{ copies } X_t^i \text{ of } X_t$$

$\rightsquigarrow$ **stochastic perturbation formulation**

$$\mathbf{m_t} := \mathbb{E}(\mathbf{X_t}) + \frac{1}{\sqrt{N}} \; V_t^N$$

with bias-variance perturbation control

$$\mathbb{E}(V_t^N) = 0 \quad \& \quad \mathbb{E}((V_t^N)^2) = \mathbb{E}((X_t - \mathbb{E}(X_t))^2)$$

Key Observation:

$$X_t \text{ stable} \implies \sup_{t \geqslant 0} \mathbb{E}((m_t - \mathbb{E}(X_t))^2) \leqslant c/N$$

$$\lim_{t \to \infty} \mathbb{E}((X_t - \mathbb{E}(X_t))^2) = \infty \implies \sup_{t \geqslant 0} \mathbb{E}((m_t - \mathbb{E}(X_t))^2) = \infty$$

How to sample X_1 given $X_0 \sim \eta_0(dx_0)$?

$$X_1 = \int a(X_0, x_0)\ \eta_0(dx_0) + W_1$$

How to sample X_1 given $X_0 \sim \eta_0(dx_0)$?

$$X_1 \quad = \quad \int a(X_0, x_0)\ \eta_0(dx_0) + \ W_1 = \mathbb{E}(a(X_0, \overline{X}_0) \mid X_0) + W_1$$

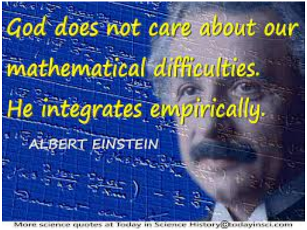
How to sample X_1 given $X_0 \sim \eta_0(dx_0)$?

$$\begin{aligned} X_1 &= \int a(X_0, x_0)\ \eta_0(dx_0) + \ W_1 = \mathbb{E}(a(X_0, \overline{X}_0) \mid X_0) + W_1 \\ &\overset{ex}{=} \int \sqrt{\log\left(1 + \|X_0 - x_0\|^2\right)}\ \eta_0(dx_0) + W_1 \end{aligned}$$

How to sample X_1 given $X_0 \sim \eta_0(dx_0)$?

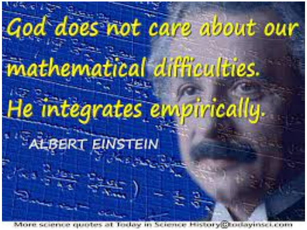
$$X_1 = \int a(X_0, x_0)\ \eta_0(dx_0) + W_1$$

Sol. based on (X_0^i, W^i) iid copies of (X_0, W_1)?



How to sample X_1 given $X_0 \sim \eta_0(dx_0)$?

$$X_1 = \int a(X_0, x_0)\ \eta_0(dx_0) + W_1$$

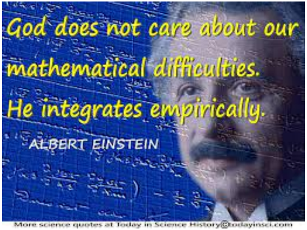


Sol. based on (X_0^i, W^i) iid copies of (X_0, W_1)?

$$X_1^i = \int a(X_0^i, x_0)\ \eta_0^{\mathbf{N}}(dx_0) + W_1^i \quad \text{with} \quad \eta_0^{\mathbf{N}} := \frac{1}{N}\sum_{1\leqslant j\leqslant N} \delta_{X_0^j}$$

How to sample X_1 given $X_0 \sim \eta_0(dx_0)$?

$$X_1 = \int a(X_0, x_0)\ \eta_0(dx_0) + W_1$$



Sol. based on (X_0^i, W^i) iid copies of (X_0, W_1)?

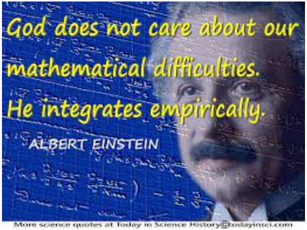
$$X_1^i = \int a(X_0^i, x_0)\ \eta_0^{\mathbf{N}}(dx_0) + W_1^i \quad \text{with} \quad \eta_0^{\mathbf{N}} := \frac{1}{N} \sum_{1 \leqslant j \leqslant N} \delta_{X_0^j}$$

$$\Updownarrow$$

$$X_1^i = \frac{1}{N} \sum_{1 \leqslant j \leqslant N} a(X_0^i, \mathbf{X_0^j}) + W_1^i$$

[$\subset$ Mean field particle sampler]

How to sample X_1 given $X_0 \sim \eta_0(dx_0)$?



$$X_1 = \int a(X_0, x_0)\ \eta_0(dx_0) + \ W_1$$

Sol. based on (X_0^i, W^i) iid copies of (X_0, W_1)?

$$X_1^i = \int a(X_0^i, x_0)\ \eta_0^{\mathbf{N}}(dx_0) + \ W_1^i \quad \text{with} \quad \eta_0^{\mathbf{N}} := \frac{1}{N} \sum_{1 \leqslant j \leqslant N} \delta_{X_0^j}$$

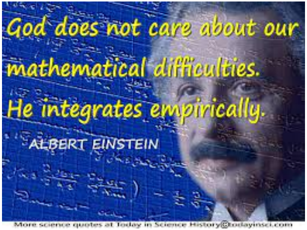
$$\Updownarrow$$

$$X_1^i = \frac{1}{N} \sum_{1 \leqslant j \leqslant N} a(X_0^i, \mathbf{X_0^j}) + \ W_1^i$$

[$\subset$ Mean field particle sampler]

Note: Running cost N^2 and X_1^i NOT iid

How to sample X_1 given $X_0 \sim \eta_0(dx_0)$?



$$X_1 = \int a(X_0, x_0)\ \eta_0(dx_0) + \ W_1$$

Sol. based on (X_0^i, W^i) iid copies of (X_0, W_1)?

$$X_1^i = \int a(X_0^i, x_0)\ \eta_0^{\mathbf{N}}(dx_0) + \ W_1^i \quad \text{with} \quad \eta_0^{\mathbf{N}} := \frac{1}{N}\sum_{1\leqslant j\leqslant N} \delta_{X_0^j}$$

$$\Updownarrow$$

$$X_1^i = \frac{1}{N}\sum_{1\leqslant j\leqslant N} a(X_0^i, \mathbf{X_0^j}) + \ W_1^i$$

[⊂ Mean field particle sampler]

Note: Running cost N^2 and X_1^i NOT iid
but "almost iid" (propagation (initial) chaos)...

A brief review on sampling $\overline{X}_{t+dt}$ given $\overline{X}_t \sim \eta_t(dx_t)$

Continuous time version = McKean-Vlasov/Interacting diffusions

$$d\overline{X}_t = \overline{X}_{t+dt} - \overline{X}_t = \left(\int b(\overline{X}_t, x_t)\ \eta_t(dx_t) \right)\ dt + \underbrace{\sqrt{dt}\ N(0,1)}_{W_{t+dt} - W_t = dW_t}$$

A brief review on sampling $\overline{X}_{t+dt}$ given $\overline{X}_t \sim \eta_t(dx_t)$

Continuous time version = McKean-Vlasov/Interacting diffusions

$$d\overline{X}_t = \overline{X}_{t+dt} - \overline{X}_t = \left(\int b(\overline{X}_t, x_t)\ \eta_t(dx_t) \right)\ dt + \underbrace{\sqrt{dt}\ N(0,1)}_{W_{t+dt} - W_t = dW_t}$$

Continuous/Discrete time versions : Nonlinear/Interacting diffusions/jumps/accept-reject nonlinear Markov chains,. . .

A brief review on sampling $\overline{X}_{t+dt}$ given $\overline{X}_t \sim \eta_t(dx_t)$

Continuous time version = McKean-Vlasov/Interacting diffusions

$$d\overline{X}_t = \overline{X}_{t+dt} - \overline{X}_t = \left(\int b(\overline{X}_t, x_t)\ \eta_t(dx_t) \right)\ dt + \underbrace{\sqrt{dt}\ N(0,1)}_{W_{t+dt}-W_t=dW_t}$$

Continuous/Discrete time versions : Nonlinear/Interacting diffusions/jumps/accept-reject nonlinear Markov chains,...

$\rightsquigarrow$ **Particle filters, GA, SMC, DMC,..., EnKF,...**

1. **Find some possibly nonlinear process**

$$\overline{X}_t \sim \mathrm{Law}(X_t \mid Y_0, \ldots, Y_t)$$

2. **Then apply mean field particle methodology**

Say N = precision (nb of samples/particles, time steps $\Delta t = 1/N, \ldots$)

Maths literature abounds with fancy bounds/theo of the type:

"Theorem": Mean error/bias/variance/estimate at time $t \leqslant e^{7t}/N$

Say N = precision (nb of samples/particles, time steps $\Delta t = 1/N, \ldots$)

Maths literature abounds with fancy bounds/theo of the type:

"Theorem": Mean error/bias/variance/estimate at time $t \leqslant e^{7t}/N$

or of type $e^{c_1 t}/N$, $c_2(t)$ or sometimes $\forall t\ \exists c$ s.t ....

Say N = precision (nb of samples/particles, time steps $\Delta t = 1/N, \dots$)

Maths literature abounds with fancy bounds/theo of the type:

"Theorem": Mean error/bias/variance/estimate at time $t \leqslant e^{7t}/N$

or of type $e^{c_1 t}/N$, $c_2(t)$ **or sometimes** $\forall t \; \exists c$ **s.t ....**

BUT $t = 30 \Longrightarrow e^{7t} > 2600 \times$ **Nb particles** $\in$ **visible universe**

Say N = precision (nb of samples/particles, time steps $\Delta t = 1/N, \ldots$)

Maths literature abounds with fancy bounds/theo of the type:

"Theorem": Mean error/bias/variance/estimate at time $t \leqslant e^{7t}/N$

or of type $e^{c_1 t}/N$, $c_2(t)$ or sometimes $\forall t \; \exists c$ s.t ....

BUT $t = 30 \Longrightarrow e^{7t} > 2600 \times$ Nb particles $\in$ visible universe

Impossible to run such particle algorithm

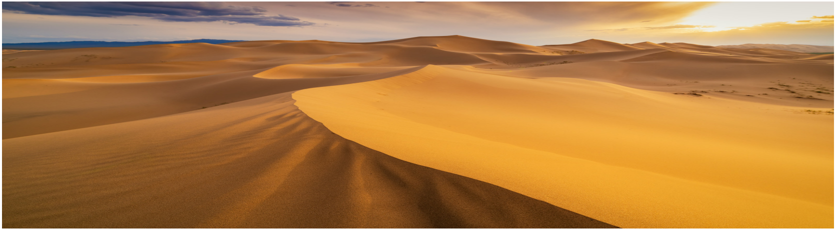
Say N = precision (nb of samples/particles, time steps $\Delta t = 1/N$,...)

Maths literature abounds with fancy bounds/theo of the type:

"Theorem": Mean error/bias/variance/estimate at time $t \leqslant e^{7t}/N$

or $\Rightarrow e^{c_1 t}/N$, $c_2(t)$ **with** $0 < c_i, c_j(t) < \infty$.

BUT $t \leqslant 6 \Longrightarrow$ **Eventually use all sand grains on earth**



Particle Filters discrete time models

Personal "crude" $c_1 \ e^{c_2 t}/\sqrt{N}$ mean error style estimates

Particle Filters discrete time models

Personal "crude" $c_1\ e^{c_2 t}/\sqrt{N}$ mean error style estimates

- Particle filters = Genetic Algo = Diffusion Monte Carlo =...
  (unbiasedness properties + first rigorous, MPRF 96)
- General particle methodology, AAP 98, LPD+CLT+...SPA 98,...

Particle Filters discrete time models

Personal "crude" $c_1\ e^{c_2 t}/\sqrt{N}$ mean error style estimates

- Particle filters = Genetic Algo = Diffusion Monte Carlo =...
  (unbiasedness properties + first rigorous, MPRF 96)
- General particle methodology, AAP 98, LPD+CLT+...SPA 98,...

Time uniform estimates-stable signals (first time unif. mean field particle)

- Stab Particle filters/GA (+ Guionnet CRAS 99, IHP 98/01)
- Feynman-Kac/particle filters (+ Miclo, Sem Proba 00)+......
- $\supset$ **New approach: stochastic perturbation $\sim$ stability limiting process**

Particle Filters discrete time models

Personal "crude" $c_1 \, e^{c_2 t}/\sqrt{N}$ mean error style estimates

- Particle filters = Genetic Algo = Diffusion Monte Carlo =...
  (unbiasedness properties + first rigorous, MPRF 96)
- General particle methodology, AAP 98, LPD+CLT+...SPA 98,...

Time uniform estimates-stable signals
(first time-uniform estimates for mean field particle systems)

- Stab Particle filters/GA (+ Guionnet CRAS 99, IHP 98/01)
- Feynman-Kac/particle filters (+ Miclo, Sem Proba 00)+......
- $\supset$ **New approach: stochastic perturbation $\sim$ stability limiting process**

$\rightsquigarrow$ **Stability Markov & Positive/Feynman-Kac semigroups**

- Lyapunov sg. & illustrations + Arnaudon, Ouhabaz SAA 23
- Stability positive sg. + Horton AAP23

Stochastic perturbation $\sim$ stability limiting process

Applications to Measure-valued processes & diffusion flows

- Interacting jumps + Arnaudon EJP-20.
- Interacting diffusions + Arnaudon AAP-20.
- Interpolation diffusion flows + Singh SPA-22 (CRAS-20).

Stochastic perturbation $\sim$ stability limiting process

Applications to Measure-valued processes & diffusion flows

- Interacting jumps + Arnaudon EJP-20.
- Interacting diffusions + Arnaudon AAP-20.
- Interpolation diffusion flows + Singh SPA-22 (CRAS-20).

$\rightsquigarrow$ **Linear-Gaussian world - possibly unstable/transient "signals"**

- Stability Kalman-Bucy filters + Bishop SIAM-17.
- Inflation/localisation + Bishop, S. Pathiraja SPA-17.
- Harmonic Oscillator ($Y = 0$)+ Horton CIMP-23/Arxiv21.

KEY OBS:
Innovation/Weights/Penalties/likelihoods... stabilizing effects !!

Uniform estimates for EnKF continuous time models

For stable signals

- AAP-18/Arxiv-16 (Unif. EnKBF)+ Tugaut.
- SIAM-17/Arxiv-16 (Unif. Extended EnKBF)+ Kurtzmann, Tugaut.
- EJP-18/Arxiv-16 (Stability Extended KBF)+ Kurtzmann, Tugaut.

Uniform estimates for EnKF continuous time models

For stable signals

- AAP-18/Arxiv-16 (Unif. EnKBF)+ Tugaut.
- SIAM-17/Arxiv-16 (Unif. Extended EnKBF)+ Kurtzmann, Tugaut.
- EJP-18/Arxiv-16 (Stability Extended KBF)+ Kurtzmann, Tugaut.

Ensemble KB filters for possibly unstable/transient signals

- AAP-19 (1d-case)+ Bishop, Kamatani, Rémillard.
- IHP-19 (Perturbation Stoch. Riccati)+ Bishop, A. Niclas.
- EJP-19 (Stability Stoch. Riccati)+ Bishop.
- $\rightsquigarrow$ MCSS-23 Review article (+ Bishop, Arxiv 20)

Continuous time **Linear+Gaussian filtering problem**

$$\left\{\begin{array}{rcl} dX_t &=& A\,X_t\,dt + R^{1/2}\,dW_t \in \mathbb{R}^r \\ dY_t &=& C\,X_t\,dt + \Sigma^{1/2}\,dV_t \end{array}\right. \quad \rightsquigarrow \mathcal{Y}_t := \sigma(Y_s,\ s \leqslant t).$$

Continuous time **Linear+Gaussian filtering problem**

$$\left\{\begin{array}{rcl} dX_t &=& A\ X_t\ dt\ +\ R^{1/2}\ dW_t\ \in \mathbb{R}^r \\ dY_t &=& C\ X_t\ dt\ +\ \Sigma^{1/2}\ dV_t \qquad \rightsquigarrow \mathcal{Y}_t := \sigma(Y_s,\ s \leqslant t). \end{array}\right.$$

Optimal $\mathbb{L}_2$-estimate = Kalman-Bucy filter

$$\widehat{X}_t := \mathbb{E}(X_t \mid \mathcal{Y}_t) \quad \text{and} \quad P_t := \mathbb{E}\left(\left(X_t - \widehat{X}_t\right)\left(X_t - \widehat{X}_t\right)'\right)$$

Continuous time **Linear+Gaussian filtering problem**

$$\begin{cases} dX_t &= A\ X_t\ dt\ +\ R^{1/2}\ dW_t\ \in \mathbb{R}^r \\ dY_t &= C\ X_t\ dt\ +\ \Sigma^{1/2}\ dV_t \end{cases} \rightsquigarrow \mathcal{Y}_t := \sigma(Y_s,\ s \leqslant t).$$

Optimal $\mathbb{L}_2$-estimate = Kalman-Bucy filter

$$\widehat{X}_t := \mathbb{E}(X_t \mid \mathcal{Y}_t) \quad \text{and} \quad P_t := \mathbb{E}\left(\left(X_t - \widehat{X}_t\right)\left(X_t - \widehat{X}_t\right)'\right)$$

State estimate

$$d\widehat{X}_t = A\ \widehat{X}_t\ dt + P_t\ C'\Sigma^{-1}\ \left(dY_t - C\widehat{X}_t\ dt\right)$$

Continuous time **Linear+Gaussian filtering problem**

$$\begin{cases} dX_t &= A\ X_t\ dt\ +\ R^{1/2}\ dW_t\ \in \mathbb{R}^r \\ dY_t &= C\ X_t\ dt\ +\ \Sigma^{1/2}\ dV_t \qquad \rightsquigarrow \mathcal{Y}_t := \sigma(Y_s,\ s \leqslant t). \end{cases}$$

Optimal $\mathbb{L}_2$-estimate = Kalman-Bucy filter

$$\widehat{X}_t := \mathbb{E}(X_t \mid \mathcal{Y}_t) \quad \text{and} \quad P_t := \mathbb{E}\left(\left(X_t - \widehat{X}_t\right)\left(X_t - \widehat{X}_t\right)'\right)$$

State estimate

$$d\widehat{X}_t = A\ \widehat{X}_t\ dt + P_t\ C'\Sigma^{-1}\ \left(dY_t - C\widehat{X}_t\ dt\right)$$

with the gain given by the matrix Riccati equation

$$\partial_t P_t = \mathrm{Ricc}(P_t) := AP_t + P_tA' - P_t\mathbf{S}P_t + R \quad \text{with} \quad \mathbf{S} := \mathbf{C}'\mathbf{\Sigma}^{-1}\mathbf{C}$$

Reformulation $\rightsquigarrow$ Nonlinear Kalman-Bucy diffusion

$\Longleftrightarrow$ **McKean-Vlasov type diffusions $\overline{X}_t$ such that (given $\mathcal{Y}_t$)**

$$\eta_t := \mathrm{Law}(\overline{X}_t) \ = \ \mathcal{N}\left[\widehat{X}_t, P_t\right]$$

$\rightsquigarrow$ **Interacting with their conditional mean and covariance matrices**

$$\mathcal{P}_{\eta_t} = \eta_t\left[(e-\eta_t(e))(e-\eta_t(e))'\right] \quad \text{with} \quad e(x) := x.$$

2 classes of McKean-Vlasov type diffusions

1) "Vanilla EnKF" ($\rightsquigarrow$ (corrected) discrete time - Evensen 94)

$$d\overline{X}_t = A\,\overline{X}_t dt + R^{1/2}\,d\overline{W}_t + \mathcal{P}_{\eta_t} C'\Sigma^{-1}\left[dY_t - \left(C\,\overline{X}_t dt + \Sigma^{1/2}\,d\overline{V}_t\right)\right]$$

2 classes of McKean-Vlasov type diffusions

1) "Vanilla EnKF" ($\rightsquigarrow$ (corrected) discrete time - Evensen 94)

$$d\overline{X}_t = A\,\overline{X}_t dt + R^{1/2}\,d\overline{W}_t + \mathcal{P}_{\eta_t} C'\Sigma^{-1}\left[dY_t - \left(C\,\overline{X}_t dt + \Sigma^{1/2}\,d\overline{V}_t\right)\right]$$

2) "Deterministic EnKF" ($\rightsquigarrow$ discrete time - Sakov-Oke 08)

$$d\overline{X}_t = A\,\overline{X}_t dt + R^{1/2}\,d\overline{W}_t + \mathcal{P}_{\eta_t} C'\Sigma^{-1}\left[dY_t - C\left(\frac{\overline{X}_t + \eta_t(e)}{2}\right)dt\right]$$

2 classes of McKean-Vlasov type diffusions

1) "Vanilla EnKF" ($\rightsquigarrow$ (corrected) discrete time - Evensen 94)

$$d\overline{X}_t = A\,\overline{X}_t dt \;+\; R^{1/2}\, d\overline{W}_t + \mathcal{P}_{\eta_t} C'\Sigma^{-1}\,\left[dY_t - \left(C\,\overline{X}_t dt + \Sigma^{1/2}\, d\overline{V}_t\right)\right]$$

2) "Deterministic EnKF" ($\rightsquigarrow$ discrete time - Sakov-Oke 08)

$$d\overline{X}_t = A\,\overline{X}_t dt \;+\; R^{1/2}\, d\overline{W}_t + \mathcal{P}_{\eta_t} C'\Sigma^{-1}\,\left[dY_t - C\,\left(\frac{\overline{X}_t + \eta_t(e)}{2}\right)\, dt\right]$$

Nonlinear models:

Tempting to replace "A x" and "C x" by $\mathbf{A}(\mathbf{x}), \mathbf{C}(\mathbf{x})$ (often done)

2 classes of McKean-Vlasov type diffusions

1) "Vanilla EnKF" ($\rightsquigarrow$ (corrected) discrete time - Evensen 94)

$$d\overline{X}_t = A\,\overline{X}_t dt \;+\; R^{1/2}\,d\overline{W}_t + \mathcal{P}_{\eta_t} C'\Sigma^{-1}\left[dY_t - \left(C\,\overline{X}_t dt + \Sigma^{1/2}\,d\overline{V}_t\right)\right]$$

2) "Deterministic EnKF" ($\rightsquigarrow$ discrete time - Sakov-Oke 08)

$$d\overline{X}_t = A\,\overline{X}_t dt \;+\; R^{1/2}\,d\overline{W}_t + \mathcal{P}_{\eta_t} C'\Sigma^{-1}\left[dY_t - C\,\left(\frac{\overline{X}_t + \eta_t(e)}{2}\right)\,dt\right]$$

Nonlinear models:

Tempting to replace "A x" and "C x" by $\mathbf{A}(\mathbf{x}), \mathbf{C}(\mathbf{x})$ (often done)

BUT <u>NOT</u> CONSISTENT WITH THE OPTIMAL FILTER

Example: Extended Kalman-Bucy Ensemble filter ($A(x)$)

$$\left\{\begin{array}{rcl} d\widehat{X}_t & = & A(\widehat{X}_t)\ dt + P_t C' \Sigma^{-1}\ \left(dY_t - C\widehat{X}_t\ dt\right) \\ \\ \partial_t P_t & = & \partial A(\widehat{X}_t)\ P_t + P_t\ \partial A(\widehat{X}_t)' - P_t S P_t + R \end{array}\right.$$

Example: Extended Kalman-Bucy Ensemble filter ($A(x)$)

$$\left\{ \begin{array}{rcl} d\widehat{X}_t & = & A(\widehat{X}_t)\ dt + P_t C'\Sigma^{-1}\ \left(dY_t - C\widehat{X}_t\ dt\right) \\ \\ \partial_t P_t & = & \partial A(\widehat{X}_t)\ P_t + P_t\ \partial A(\widehat{X}_t)' - P_t S P_t + R \end{array} \right.$$

Nonlinear/McKean-Vlasov-type diffusion

$$\begin{array}{rcl} d\overline{X}_t & = & \mathcal{A}\left(\overline{X}_t, \eta_t(e)\right)\ dt + R^{1/2}\ d\overline{W}_t \\ \\ & & +\mathcal{P}_{\eta_t} C'\Sigma^{-1}\ \left(dY_t - \left(C\overline{X}_t\ dt + \Sigma^{1/2}\ d\overline{V}_t\right)\right) \end{array}$$

$$\mathcal{A}(x, m) \quad := \quad A(m) + \partial A(m)\ (x - m)$$

Example: Extended Kalman-Bucy Ensemble filter ($A(x)$)

$$\begin{cases} d\widehat{X}_t &= A(\widehat{X}_t)\ dt + P_t C' \Sigma^{-1}\ \left(dY_t - C\widehat{X}_t\ dt\right) \\ \partial_t P_t &= \partial A(\widehat{X}_t)\ P_t + P_t\ \partial A(\widehat{X}_t)' - P_t S P_t + R \end{cases}$$

Nonlinear/McKean-Vlasov-type diffusion

$$\begin{aligned} d\overline{X}_t &= \mathcal{A}\left(\overline{X}_t, \eta_t(e)\right)\ dt + R^{1/2}\ d\overline{W}_t \\ &\quad + \mathcal{P}_{\eta_t} C' \Sigma^{-1}\ \left(dY_t - \left(C\overline{X}_t\ dt + \Sigma^{1/2}\ d\overline{V}_t\right)\right) \end{aligned}$$

$$\mathcal{A}(x, m) := A(m) + \partial A(m)\ (x - m)$$

Refs:

- SIAM-17/Arxiv-16(Extended EnKBF)+ Kurtzmann, Tugaut.
- EJP-18 (Stability Extended KBF)+ Kurtzmann, Tugaut.

The Ensemble Kalman-Bucy filter

(Ex. Case 1) Mean field sampler $\rightsquigarrow$ $N+1$ interacting diffusions

$$d\xi_t^i = A\,\xi_t^i dt + R^{1/2} d\overline{W}_t^i + p_t C' \Sigma^{-1} \left[dY_t - \left(C\xi_t^i dt + \Sigma^{1/2}\, d\overline{V}_t^i \right) \right]$$

with the rescaled particle covariance matrices

$$p_t := \frac{1}{N} \sum_{1 \leqslant i \leqslant N+1} \left(\xi_t^i - m_t \right) \left(\xi_t^i - m_t \right)'$$

and the sample mean

$$m_t := \frac{1}{N+1} \sum_{1 \leqslant i \leqslant N+1} \xi_t^i.$$

The Ensemble Kalman-Bucy filter

(Ex. Case 1) Mean field sampler $\rightsquigarrow$ $N+1$ interacting diffusions

$$d\xi_t^i = A\,\xi_t^i dt + R^{1/2} d\overline{W}_t^i + p_t C' \Sigma^{-1} \left[dY_t - \left(C\xi_t^i dt + \Sigma^{1/2}\, d\overline{V}_t^i \right) \right]$$

with the rescaled particle covariance matrices

$$p_t := \frac{1}{N} \sum_{1 \leqslant i \leqslant N+1} \left(\xi_t^i - m_t \right) \left(\xi_t^i - m_t \right)'$$

and the sample mean

$$m_t := \frac{1}{N+1} \sum_{1 \leqslant i \leqslant N+1} \xi_t^i.$$

where are the Kalman-Bucy filter and the Riccati equations ?

Perturbation theorem(s):

$$
\begin{aligned}
dm_t &= A\ m_t\ dt + p_t\ C'\Sigma^{-1}\ (dY_t - Cm_t\ dt) + \frac{1}{\sqrt{N+1}}\ dM_t \\
dp_t &= \mathrm{Ricc}(p_t)\ dt + \frac{1}{\sqrt{N}}\ d\overline{M}_t
\end{aligned}
$$

Perturbation theorem(s):

$$
\begin{aligned}
dm_t &= A\, m_t\, dt + p_t\, C'\Sigma^{-1}\, (dY_t - Cm_t\, dt) + \frac{1}{\sqrt{N+1}}\, dM_t \\
dp_t &= \mathrm{Ricc}(p_t)\, dt + \frac{1}{\sqrt{N}}\, d\overline{M}_t
\end{aligned}
$$

- OK for Extended Kalman $A(m_t)$ with $\mathrm{Ricc}_{m_t}(p_t)$ (cf. SIAM-17)

Perturbation theorem(s):

$$
\begin{aligned}
dm_t &= A\ m_t\ dt + p_t\ C'\Sigma^{-1}\ (dY_t - Cm_t\ dt) + \frac{1}{\sqrt{N+1}}\ dM_t \\
dp_t &= \mathrm{Ricc}(p_t)\ dt + \frac{1}{\sqrt{N}}\ d\overline{M}_t
\end{aligned}
$$

- OK for Extended Kalman $A(m_t)$ with $\mathrm{Ricc}_{m_t}(p_t)$ (cf. SIAM-17)
- Orthogonal martingales $(M_t, \overline{M}_t)$ and $m_0 \perp p_0$ (well known iid).

Perturbation theorem(s):

$$dm_t = A\, m_t\, dt + p_t\, C'\Sigma^{-1}\, (dY_t - Cm_t\, dt) + \frac{1}{\sqrt{N+1}}\, dM_t$$

$$dp_t = \mathrm{Ricc}(p_t)\, dt + \frac{1}{\sqrt{N}}\, d\overline{M}_t$$

- OK for Extended Kalman $A(m_t)$ with $\mathrm{Ricc}_{m_t}(p_t)$ (cf. SIAM-17)
- Orthogonal martingales $(M_t, \overline{M}_t)$ and $m_0 \perp p_0$ (well known iid).
- $\mathbf{S} = \mathbf{0} \Longrightarrow \mathrm{Ricc}(p) := Ap + pA' + R$ **Wishart process.**

Perturbation theorem(s):

$$dm_t = A\ m_t\ dt + p_t\ C'\Sigma^{-1}\ (dY_t - Cm_t\ dt) + \frac{1}{\sqrt{N+1}}\ dM_t$$

$$dp_t = \mathrm{Ricc}(p_t)\ dt + \frac{1}{\sqrt{N}}\ d\overline{M}_t$$

- OK for Extended Kalman $A(m_t)$ with $\mathrm{Ricc}_{m_t}(p_t)$ (cf. SIAM-17)
- Orthogonal martingales $(M_t, \overline{M}_t)$ and $m_0 \perp p_0$ (well known iid).
- $\mathbf{S} = \mathbf{0} \Longrightarrow \mathrm{Ricc}(p) := Ap + pA' + R$ **Wishart process.**
- *In one dimension* [AAP19 + Bishop-Kamatani-Rémillard]:

Perturbation theorem(s):

$$dm_t = A\, m_t\, dt + p_t\, C'\Sigma^{-1}\, (dY_t - Cm_t\, dt) + \frac{1}{\sqrt{N+1}}\, dM_t$$

$$dp_t = \mathrm{Ricc}(p_t)\, dt + \frac{1}{\sqrt{N}}\, d\overline{M}_t$$

- OK for Extended Kalman $A(m_t)$ with $\mathrm{Ricc}_{m_t}(p_t)$ (cf. SIAM-17)
- Orthogonal martingales $(M_t, \overline{M}_t)$ and $m_0 \perp p_0$ (well known iid).
- $\mathbf{S} = \mathbf{0} \Longrightarrow \mathrm{Ricc}(p) := Ap + pA' + R$ **Wishart process.**
- *In one dimension* [AAP19 + Bishop-Kamatani-Rémillard]:
  Vanilla EnKF with $S > 0$ p_t **heavy tailed invariant measure**.

Perturbation theorem(s):

$$dm_t = A\ m_t\ dt + p_t\ C'\Sigma^{-1}\ (dY_t - Cm_t\ dt) + \frac{1}{\sqrt{N+1}}\ dM_t$$

$$dp_t = \mathrm{Ricc}(p_t)\ dt + \frac{1}{\sqrt{N}}\ d\overline{M}_t$$

- OK for Extended Kalman $A(m_t)$ with $\mathrm{Ricc}_{m_t}(p_t)$ (cf. SIAM-17)
- Orthogonal martingales $(M_t, \overline{M}_t)$ and $m_0 \perp p_0$ (well known iid).
- $\mathbf{S} = \mathbf{0} \Longrightarrow \mathrm{Ricc}(p) := Ap + pA' + R$ **Wishart process.**
- *In one dimension* [AAP19 + Bishop-Kamatani-Rémillard]:
  Vanilla EnKF with $S > 0$ p_t **heavy tailed invariant measure**.
  Deterministic EnKF p_t **Gaussian tailed invariant measure**.
- **Stoch. perturbation** $\Rightarrow$ **time-uniform estimates** $\forall(A, R, S)$**.**

Perturbation theorem(s):

$$dm_t = A\ m_t\ dt + p_t\ C'\Sigma^{-1}\ (dY_t - Cm_t\ dt) + \frac{1}{\sqrt{N+1}}\ dM_t$$

$$dp_t = \mathrm{Ricc}(p_t)\ dt + \frac{1}{\sqrt{N}}\ d\overline{M}_t$$

- OK for Extended Kalman $A(m_t)$ with $\mathrm{Ricc}_{m_t}(p_t)$ (cf. SIAM-17)
- Orthogonal martingales $(M_t, \overline{M}_t)$ and $m_0 \perp p_0$ (well known iid).
- $\mathbf{S} = \mathbf{0} \Longrightarrow \mathrm{Ricc}(p) := Ap + pA' + R$ **Wishart process.**
- *In one dimension* [AAP19 + Bishop-Kamatani-Rémillard]:
  Vanilla EnKF with $S > 0$ p_t **heavy tailed invariant measure**.
  Deterministic EnKF p_t **Gaussian tailed invariant measure**.
- **Stoch. perturbation** $\Rightarrow$ **time-uniform estimates** $\forall (A, R, S)$**.**

$$d(m_t - X_t) = (A - p_t S)\ (m_t - X_t)dt$$

$$+p_t\ C'\Sigma^{-1/2}dV_t - R^{1/2}dW_t + \frac{1}{\sqrt{N+1}}\ dM_t$$

HAPPY-ENDING STORY?

HAPPY-ENDING STORY? In practice discrete time (harder!):

HAPPY-ENDING STORY? In practice discrete time (harder!):

• **Discrete time EnKF**
(Uniform estimates?

HAPPY-ENDING STORY? In practice discrete time (harder!):

- **Discrete time EnKF**
(Uniform estimates? Effective dimensions/Stab. theory?)

HAPPY-ENDING STORY? **In practice discrete time (harder!):**

- **Discrete time EnKF**
(Uniform estimates? Effective dimensions/Stab. theory?)

- **or Particle filters (a.k.a. SMC/DMC/Genetic/. . .)**

HAPPY-ENDING STORY? **In practice <u>discrete time</u> (harder!):**

- **Discrete time EnKF**
(Uniform estimates? Effective dimensions/Stab. theory?)

- **or Particle filters (a.k.a. SMC/DMC/Genetic/...)**
Since 30 years only for stable/ergodic signals?

HAPPY-ENDING STORY? In practice discrete time (harder!):

- Discrete time EnKF

(Uniform estimates? Effective dimensions/Stab. theory?)

- or Particle filters (a.k.a. SMC/DMC/Genetic/...)
  Since 30 years only for stable/ergodic signals?

HAPPY-ENDING STORY? **In practice discrete time (harder!):**

- **Discrete time EnKF**

(Uniform estimates? Effective dimensions/Stab. theory?)

- **or Particle filters (a.k.a. SMC/DMC/Genetic/...)**
  Since 30 years only for stable/ergodic signals?

This talk = answer these 2 questions for 1D linear/Gaussian

- `1d-discrete time/EnKF (+ Horton, Arxiv 21, AAP 23)`
- $\rightsquigarrow$ `EnKF Review article (+ Bishop, Arxiv 20, MCSS 23)`

Linear+Gaussian+discrete 1*d*-filtering problem

$$\begin{cases} X_{n+1} = A\,X_n + B\,W_{n+1} & X_0 \sim \mathcal{N}(\widehat{X}_0^-, P_0) \\ Y_n = C\,X_n + D\,V_n & n \in \mathbb{N} := \{0, 1, 2 \ldots\} \end{cases}$$

Linear+Gaussian+discrete 1*d*-filtering problem

$$\begin{cases} X_{n+1} = A\,X_n + B\,W_{n+1} & X_0 \sim \mathcal{N}(\widehat{X}_0^-, P_0) \\ Y_n = C\,X_n + D\,V_n & n \in \mathbb{N} := \{0,1,2\ldots\} \end{cases}$$

$$\Downarrow \quad \mathcal{Y}_n := (Y_0, \ldots, Y_n)$$

One-step predictor & Optimal filter = Gaussian

$$\mathrm{Law}(X_n \mid \mathcal{Y}_{n-1}) = \mathcal{N}(\widehat{X}_n^-, P_n) \quad \& \quad \mathrm{Law}(X_n \mid \mathcal{Y}_n) = \mathcal{N}(\widehat{X}_n, \widehat{P}_n)$$

Linear+Gaussian+discrete 1d-filtering problem

$$\begin{cases} X_{n+1} = A\,X_n + B\,W_{n+1} & X_0 \sim \mathcal{N}(\widehat{X}_0^-, P_0) \\ Y_n = C\,X_n + D\,V_n & n \in \mathbb{N} := \{0,1,2\ldots\} \end{cases}$$

$$\Downarrow \quad \mathcal{Y}_n := (Y_0, \ldots, Y_n)$$

One-step predictor & Optimal filter = Gaussian

$$\mathrm{Law}(X_n \mid \mathcal{Y}_{n-1}) = \mathcal{N}(\widehat{X}_n^-, P_n) \quad \& \quad \mathrm{Law}(X_n \mid \mathcal{Y}_n) = \mathcal{N}(\widehat{X}_n, \widehat{P}_n)$$

$\rightsquigarrow$ **Kalman filter (1960s') = Gauss-Legendre regression (1800s')**

$$(\widehat{X}_n^-, P_n) \xrightarrow{\text{updating}} (\widehat{X}_n, \widehat{P}_n) \xrightarrow{\text{prediction}} (\widehat{X}_{n+1}^-, P_{n+1})$$

Particle filters = GA = SMC = DMC = ...

$$\left(\xi_n^{i-}\right)_{1\leqslant i\leqslant N} \in \mathbb{R}^N \xrightarrow{\text{Selection}} \left(\xi_n^{j}\right)_{1\leqslant j\leqslant N} \xrightarrow{\text{Mutation}} \left(\xi_{n+1}^{i-}\right)_{1\leqslant i\leqslant N}$$

Particle filters = GA = SMC = DMC = ...

$$\left(\xi_n^{i-}\right)_{1\leqslant i\leqslant N} \in \mathbb{R}^N \xrightarrow{\text{Selection}} \left(\xi_n^{j}\right)_{1\leqslant j\leqslant N} \xrightarrow{\text{Mutation}} \left(\xi_{n+1}^{i-}\right)_{1\leqslant i\leqslant N}$$

Selection/ **(Vanilla)** **Mutation**:

$$\xi_n^{j} \sim \sum_{1\leqslant i\leqslant N} \frac{e^{-(Y_n - C\xi_n^{i-})^2/(2D^2)}}{\sum_{1\leqslant j\leqslant N} e^{-(Y_n - C\xi_n^{j-})^2/(2D^2)}}\, \delta_{\xi_n^{i-}} \quad \text{and set} \quad \xi_{n+1}^{j-} := A\,\xi_n^{j} + B\,W_{n+1}^{j}$$

Particle filters = GA = SMC = DMC = ...

$$\left(\xi_n^{i-}\right)_{1\leqslant i\leqslant N} \in \mathbb{R}^N \xrightarrow{\text{Selection}} \left(\xi_n^j\right)_{1\leqslant j\leqslant N} \xrightarrow{\text{Mutation}} \left(\xi_{n+1}^{i-}\right)_{1\leqslant i\leqslant N}$$

Selection/ **(Vanilla)** **Mutation**:

$$\xi_n^j \sim \sum_{1\leqslant i\leqslant N} \frac{e^{-(Y_n - C\xi_n^{i-})^2/(2D^2)}}{\sum_{1\leqslant j\leqslant N} e^{-(Y_n - C\xi_n^{j-})^2/(2D^2)}} \, \delta_{\xi_n^{i-}} \quad \text{and set} \quad \xi_{n+1}^{j-} := A\,\xi_n^j + B\,W_{n+1}^j$$

Sample means $\simeq$ **Conditional expectations:**

$$\forall n \in \mathbb{N} \qquad \widehat{X}_n^{\mathrm{PF}} := \frac{1}{N} \sum_{1\leqslant i\leqslant N} \xi_n^i \simeq_{N\to\infty} \widehat{X}_n$$

Particle filters = GA = SMC = DMC = ...

$$\left(\xi_n^{i-}\right)_{1\leqslant i\leqslant N} \in \mathbb{R}^N \xrightarrow{\text{Selection}} \left(\xi_n^j\right)_{1\leqslant j\leqslant N} \xrightarrow{\text{Mutation}} \left(\xi_{n+1}^{i-}\right)_{1\leqslant i\leqslant N}$$

Selection/ **(Vanilla)** **Mutation**:

$$\xi_n^j \sim \sum_{1\leqslant i\leqslant N} \frac{e^{-(Y_n - C\xi_n^{i-})^2/(2D^2)}}{\sum_{1\leqslant j\leqslant N} e^{-(Y_n - C\xi_n^{j-})^2/(2D^2)}} \, \delta_{\xi_n^{i-}} \quad \text{and set} \quad \xi_{n+1}^{j-} := A\,\xi_n^j + B\, W_{n+1}^j$$

Sample means $\simeq$ Conditional expectations:

$$\forall n \in \mathbb{N} \qquad \widehat{X}_n^{\mathrm{PF}} := \frac{1}{N} \sum_{1\leqslant i\leqslant N} \xi_n^i \simeq_{N\to\infty} \widehat{X}_n$$

BUT for any $A > 1$

$$\xi_0^{i-} > \frac{|B|}{A-1} \sqrt{2\log N} \Longrightarrow \lim_{n\to\infty} \mathbb{E}\left[\left|\widehat{X}_n^{\mathrm{PF}} - \widehat{X}_n\right|\right] = +\infty$$

Particle filters = GA = SMC = DMC = ...

$$\left(\xi_n^{i-}\right)_{1\leqslant i\leqslant N} \in \mathbb{R}^N \xrightarrow{\text{Selection}} \left(\xi_n^j\right)_{1\leqslant j\leqslant N} \xrightarrow{\text{Mutation}} \left(\xi_{n+1}^{i-}\right)_{1\leqslant i\leqslant N}$$

Selection/ **(Vanilla)** **Mutation**:

$$\xi_n^j \sim \sum_{1\leqslant i\leqslant N} \frac{e^{-(Y_n-C\xi_n^{i-})^2/(2D^2)}}{\sum_{1\leqslant j\leqslant N} e^{-(Y_n-C\xi_n^{j-})^2/(2D^2)}}\, \delta_{\xi_n^{i-}} \quad \text{and set} \quad \xi_{n+1}^{j-} := A\,\xi_n^j + B\,W_{n+1}^j$$

Sample means $\simeq$ Conditional expectations:

$$\forall n\in\mathbb{N} \qquad \widehat{X}_n^{\mathrm{PF}} := \frac{1}{N}\sum_{1\leqslant i\leqslant N} \xi_n^i \simeq_{N\to\infty} \widehat{X}_n$$

BUT for any $A>1$

$$\xi_0^{i-} > \frac{|B|}{A-1}\sqrt{2\log N} \Longrightarrow \lim_{n\to\infty} \mathbb{E}\left[\left|\widehat{X}_n^{\mathrm{PF}} - \widehat{X}_n\right|\right] = +\infty$$

and for $|A|<1$?

Particle filters = GA = SMC = DMC = ...

$$\left(\xi_n^{i-}\right)_{1\leqslant i\leqslant N} \in \mathbb{R}^N \xrightarrow{\text{Selection}} \left(\xi_n^j\right)_{1\leqslant j\leqslant N} \xrightarrow{\text{Mutation}} \left(\xi_{n+1}^{i-}\right)_{1\leqslant i\leqslant N}$$

Selection/ **(Vanilla)** **Mutation**:

$$\xi_n^j \sim \sum_{1\leqslant i\leqslant N} \frac{e^{-(Y_n - C\xi_n^{i-})^2/(2D^2)}}{\sum_{1\leqslant j\leqslant N} e^{-(Y_n - C\xi_n^{j-})^2/(2D^2)}} \,\delta_{\xi_n^{i-}} \quad \text{and set} \quad \xi_{n+1}^{j-} := A\,\xi_n^j + B\,W_{n+1}^j$$

Sample means $\simeq$ Conditional expectations:

$$\forall n \in \mathbb{N} \qquad \widehat{X}_n^{\mathrm{PF}} := \frac{1}{N}\sum_{1\leqslant i\leqslant N} \xi_n^i \simeq_{N\to\infty} \widehat{X}_n$$

BUT for any $A > 1$

$$\xi_0^{i-} > \frac{|B|}{A-1}\sqrt{2\log N} \Longrightarrow \lim_{n\to\infty} \mathbb{E}\left[\left|\widehat{X}_n^{\mathrm{PF}} - \widehat{X}_n\right|\right] = +\infty$$

and for $|A| < 1$**?**$\rightsquigarrow$ **Time-uniform estimates (new coming article)**

Particle filters = GA = SMC = DMC = ...

$$\left(\xi_n^{i-}\right)_{1\leqslant i\leqslant N} \in \mathbb{R}^N \xrightarrow{\text{Selection}} \left(\xi_n^{j}\right)_{1\leqslant j\leqslant N} \xrightarrow{\text{Mutation}} \left(\xi_{n+1}^{i-}\right)_{1\leqslant i\leqslant N}$$

Selection/ **(Vanilla)** **Mutation**:

$$\xi_n^{j} \sim \sum_{1\leqslant i\leqslant N} \frac{e^{-(Y_n - C\xi_n^{i-})^2/(2D^2)}}{\sum_{1\leqslant j\leqslant N} e^{-(Y_n - C\xi_n^{j-})^2/(2D^2)}}\, \delta_{\xi_n^{i-}} \quad \text{and set} \quad \xi_{n+1}^{j-} := A\,\xi_n^{j} + B\, W_{n+1}^{j}$$

Sample means $\simeq$ Conditional expectations:

$$\forall n \in \mathbb{N} \qquad \widehat{X}_n^{\mathrm{PF}} := \frac{1}{N} \sum_{1\leqslant i\leqslant N} \xi_n^{i} \simeq_{N\to\infty} \widehat{X}_n$$

BUT for any $A > 1$

$$\xi_0^{i-} > \frac{|B|}{A-1}\sqrt{2\log N} \Longrightarrow \lim_{n\to\infty} \mathbb{E}\left[\left|\widehat{X}_n^{\mathrm{PF}} - \widehat{X}_n\right|\right] = +\infty$$

and for $|A| < 1$? $\rightsquigarrow$ **Time-uniform estimates (new coming article)**
$\rightsquigarrow$ **find stable twisted guiding mutations (up to change of weights/probab)**

(Conditional) Mean equations

$$\left\{\begin{array}{rcl} \widehat{X}_n & = & \widehat{X}_n^- + Gain_n \left(Y_n - C\widehat{X}_n^-\right) \quad \text{with} \quad Gain_n := CP_n/(C^2P_n + D^2) \\ \widehat{X}_{n+1}^- & = & A\,\widehat{X}_n \end{array}\right.$$

Offline Riccati equations

$$\left\{\begin{array}{rcl} \widehat{P}_n & = & (1 - G_nC)P_n = P_n/(1 + SP_n) \quad \text{with} \quad S := (C/D)^2 \\ & & \\ P_{n+1} & = & A^2\widehat{P}_n + R \quad \text{with} \quad R = B^2 \end{array}\right.$$

(Conditional) Mean equations

$$\left\{\begin{array}{rcl}\widehat{X}_n & = & \widehat{X}_n^- + Gain_n\left(Y_n - C\widehat{X}_n^-\right) \quad \text{with} \quad Gain_n := CP_n/(C^2P_n + D^2)\\ \widehat{X}_{n+1}^- & = & A\,\widehat{X}_n\end{array}\right.$$

Offline Riccati equations

$$\left\{\begin{array}{rcl}\widehat{P}_n & = & (1 - G_nC)P_n = P_n/(1 + SP_n) \quad \text{with} \quad S := (C/D)^2\\ P_{n+1} & = & A^2\widehat{P}_n + R \quad \text{with} \quad R = B^2\end{array}\right.$$

$$\rightsquigarrow P_{n+1} = \phi(P_n) := \frac{aP_n + b}{cP_n + d} \quad \text{with} \quad (a, b, c, d) := \left(A^2 + RS, R, S, 1\right)$$

Reminder:

P_n and $(\widehat{X}_n - X_n)$ are stable for any A (Kalman/Bucy-Stab Theory) !

Conditional-Nonlinear Markov chain ($\mathfrak{Perfect\ Sampler}$)

$$\left\{\begin{array}{rcl} \widehat{\mathfrak{X}}_n & = & \mathfrak{X}_n + \mathfrak{gain}_n \left(Y_n - (C\,\mathfrak{X}_n + D\,\mathfrak{V}_n)\right) \quad \text{with} \quad \mathfrak{gain}_n := C\mathfrak{P}_n/(C^2\mathfrak{P}_n + D^2) \\ \mathfrak{X}_{n+1} & = & A\,\widehat{\mathfrak{X}}_n + B\,\mathfrak{W}_{n+1}. \end{array}\right.$$

$(\mathfrak{V}_n, \mathfrak{W}_n)$ copies of (V_n, W_n) and $\mathfrak{P}_n$ variance of the state $\mathfrak{X}_n$.

Conditional-Nonlinear Markov chain (Perfect Sampler)

$$\begin{cases} \widehat{\mathfrak{X}}_n = \mathfrak{X}_n + \mathfrak{gain}_n \left(Y_n - (C\,\mathfrak{X}_n + D\,\mathfrak{V}_n) \right) & \text{with} \quad \mathfrak{gain}_n := C\mathfrak{P}_n/(C^2\mathfrak{P}_n + D^2) \\ \mathfrak{X}_{n+1} = A\,\widehat{\mathfrak{X}}_n + B\,\mathfrak{W}_{n+1}. \end{cases}$$

$(\mathfrak{V}_n, \mathfrak{W}_n)$ copies of (V_n, W_n) and $\mathfrak{P}_n$ variance of the state $\mathfrak{X}_n$.

$$\Downarrow$$

Consistency property:

$$\mathfrak{X}_n \sim \mathcal{N}\left(\widehat{X}_n^-, P_n = \mathfrak{P}_n\right) \quad \text{and} \quad \widehat{\mathfrak{X}}_n \sim \mathcal{N}\left(\widehat{X}_n, \widehat{P}_n\right).$$

Conditional-Nonlinear Markov chain ($\mathfrak{Perfect\ Sampler}$)

$$\begin{cases} \widehat{\mathfrak{X}}_n &= \mathfrak{X}_n + \mathfrak{gain}_n \left(Y_n - (C\,\mathfrak{X}_n + D\,\mathfrak{V}_n)\right) \quad \text{with} \quad \mathfrak{gain}_n := C\mathfrak{P}_n/(C^2\mathfrak{P}_n + D^2) \\ \mathfrak{X}_{n+1} &= A\,\widehat{\mathfrak{X}}_n + B\,\mathfrak{W}_{n+1}. \end{cases}$$

$(\mathfrak{V}_n, \mathfrak{W}_n)$ copies of (V_n, W_n) and $\mathfrak{P}_n$ variance of the state $\mathfrak{X}_n$.

$$\Downarrow$$

Consistency property:

$$\mathfrak{X}_n \sim \mathcal{N}\left(\widehat{X}_n^-, P_n = \mathfrak{P}_n\right) \quad \text{and} \quad \widehat{\mathfrak{X}}_n \sim \mathcal{N}\left(\widehat{X}_n, \widehat{P}_n\right).$$

Nb: *The Sakov & Oke (a.k.a. deterministic EnKF) discrete time versions are* **NOT consistent**.

EnKF = Mean field interacting particle sampler

$$\left\{\begin{array}{rcl}\widehat{\xi}^i_n & = & \xi^i_n + g_n\left(Y_n - (C\xi^i_n + D\mathcal{V}^i_n)\right) \quad \text{with} \quad g_n := Cp_n/(C^2p_n + D^2)\\ \\ \xi^i_{n+1} & = & A\,\widehat{\xi}^i_n + B\,\mathcal{W}^i_{n+1} \qquad i \in \{1,\ldots,N+1\}\end{array}\right.$$

EnKF = Mean field interacting particle sampler

$$\begin{cases} \widehat{\xi}_n^i &= \xi_n^i + g_n\left(Y_n - (C\xi_n^i + D\mathcal{V}_n^i)\right) \quad \text{with} \quad g_n := C p_n/(C^2 p_n + D^2) \\ \xi_{n+1}^i &= A\,\widehat{\xi}_n^i + B\,\mathcal{W}_{n+1}^i \qquad i \in \{1,\ldots,N+1\} \end{cases}$$

$(\mathcal{V}_n^i, \mathcal{W}_n^i)$ *copies of* (V_n, W_n) *and re-scaled sample variance*

$$p_n := \frac{1}{N} \sum_{1 \leqslant i \leqslant N+1} (\xi_n^i - m_n)^2$$

with the sample mean

$$m_n := \frac{1}{N+1} \sum_{1 \leqslant i \leqslant N+1} \xi_n^i$$

Perturbation theo. *(Up to a change of probability space)*

$$\left\{\begin{array}{rcl} \widehat{m}_n & = & m_n + g_n\,(Y_n - Cm_n) + \dfrac{1}{\sqrt{N+1}}\,\widehat{\upsilon}_n \\[2ex] \widehat{p}_n & = & (1 - g_n C)\;p_n + \dfrac{1}{\sqrt{N}}\,\widehat{\nu}_n \end{array}\right. \qquad \left\{\begin{array}{rcl} m_{n+1} & = & A\;\widehat{m}_n + \dfrac{1}{\sqrt{N+1}}\,\upsilon_{n+1} \\[2ex] p_{n+1} & = & A^2\;\widehat{p}_n + R + \dfrac{1}{\sqrt{N}}\,\nu_{n+1}. \end{array}\right.$$

Perturbation theo. *(Up to a change of probability space)*

$$\begin{cases} \widehat{m}_n = m_n + g_n\,(Y_n - Cm_n) + \frac{1}{\sqrt{N+1}}\,\widehat{\upsilon}_n \\ \widehat{p}_n = (1 - g_n C)\;p_n + \frac{1}{\sqrt{N}}\,\widehat{\nu}_n \end{cases} \qquad \begin{cases} m_{n+1} = A\;\widehat{m}_n + \frac{1}{\sqrt{N+1}}\,\upsilon_{n+1} \\ p_{n+1} = A^2\;\widehat{p}_n + R + \frac{1}{\sqrt{N}}\,\nu_{n+1}. \end{cases}$$

local perturbations υ_n, ν_n and $\widehat{\nu}_n, \widehat{\upsilon}_n$ in terms of χ^2

Perturbation theo. *(Up to a change of probability space)*

$$\begin{cases} \widehat{m}_n &= m_n + g_n\,(Y_n - C m_n) + \frac{1}{\sqrt{N+1}}\,\widehat{\upsilon}_n \\ \widehat{p}_n &= (1 - g_n C)\,p_n + \frac{1}{\sqrt{N}}\,\widehat{\nu}_n \end{cases} \quad \begin{cases} m_{n+1} &= A\,\widehat{m}_n + \frac{1}{\sqrt{N+1}}\,\upsilon_{n+1} \\ p_{n+1} &= A^2\,\widehat{p}_n + R + \frac{1}{\sqrt{N}}\,\nu_{n+1}. \end{cases}$$

local perturbations υ_n, ν_n and $\widehat{\nu}_n, \widehat{\upsilon}_n$ in terms of χ^2

Corollary: p_n Markov chain $\mathcal{K}(p, dq)$ = Stochastic Riccati eq.

$$p_{n+1} = \phi(p_n) + \frac{1}{\sqrt{N}}\,\delta_{n+1} \quad \text{with} \quad \delta_{n+1} := A^2\,\widehat{\nu}_n + \nu_{n+1}.$$

Stab. theo. (cf. 10 pages section 2 in (+Arnaudon, Ouhabaz SAA 23))

Theo 1: $\exists$ Lyapunov fct. $\mathcal{U}(p) = 1 + u(p + 1/p)$ s.t. $\beta_{\mathcal{U}}(\mathcal{K}) < 1$

Stab. theo. (cf. 10 pages section 2 in (+Arnaudon, Ouhabaz SAA 23))

Theo 1: $\exists$ Lyapunov fct. $\mathcal{U}(p) = 1 + u(p + 1/p)$ s.t. $\beta_{\mathcal{U}}(\mathcal{K}) < 1$

$\Longleftrightarrow$

$$\|\mu_1\mathcal{K}^n - \mu_2\mathcal{K}^n\|_{\mathcal{U}} \leqslant \beta_{\mathcal{U}}(\mathcal{K})^n \|\mu_1 - \mu_2\|_{\mathcal{U}}.$$

Stab. theo. (cf. 10 pages section 2 in (+Arnaudon, Ouhabaz SAA 23))

Theo 1: $\exists$ Lyapunov fct. $\mathcal{U}(p) = 1 + u(p + 1/p)$ s.t. $\beta_{\mathcal{U}}(\mathcal{K}) < 1$

$\Longleftrightarrow$

$$\|\mu_1 \mathcal{K}^n - \mu_2 \mathcal{K}^n\|_{\mathcal{U}} \leqslant \beta_{\mathcal{U}}(\mathcal{K})^n \|\mu_1 - \mu_2\|_{\mathcal{U}}.$$

Theo 2: Same expo. stab. **for the pair** (p_n, M_n) with

$$M_n := m_n - X_n = \frac{A}{1 + Sp_{n-1}} M_{n-1} + \text{perturbations}$$

Stab. theo. (cf. 10 pages section 2 in (+Arnaudon, Ouhabaz SAA 23))

Theo 1: $\exists$ Lyapunov fct. $\mathcal{U}(p) = 1 + u(p + 1/p)$ s.t. $\beta_{\mathcal{U}}(\mathcal{K}) < 1$

$\Longleftrightarrow$

$$\|\mu_1 \mathcal{K}^n - \mu_2 \mathcal{K}^n\|_{\mathcal{U}} \leqslant \beta_{\mathcal{U}}(\mathcal{K})^n \|\mu_1 - \mu_2\|_{\mathcal{U}}.$$

Theo 2: Same expo. stab. **for the pair** (p_n, M_n) with

$$M_n := m_n - X_n = \frac{A}{1 + Sp_{n-1}} M_{n-1} + \text{perturbations}$$

Key/difficulty: **Estimates/Expo. decays $\underline{\forall A}$ of random products**

$$\mathcal{E}_{l,n} := \prod_{l \leqslant k \leqslant n} \frac{A}{1 + Sp_k}$$

OK for $|A| < 1$ but also for $\underline{|A| \geqslant 1}$ unstable/effective dimension/...

Time uniform estimates for any A

Theo 1 [(Under) Bias]: $\forall k \geqslant 1 \ \exists \iota_k < \infty$ s.t. $\forall N \geqslant 1 \ \forall n \geqslant 0$

$$0 \leqslant P_n - \mathbb{E}(p_n) \leqslant \iota_1/N$$

Time uniform estimates for any A

Theo 1 [(Under) Bias]: $\forall k \geqslant 1\ \exists \iota_k < \infty$ s.t. $\forall N \geqslant 1\ \forall n \geqslant 0$

$$0 \leqslant P_n - \mathbb{E}(p_n) \leqslant \iota_1 / N$$

& Uniform control of the bias $(N \geqslant N_k)$

$$\mathbb{E}\left(\left| \mathbb{E}(\widehat{m}_n \mid \mathcal{Y}_n) - \widehat{X}_n \right|^k \right)^{1/k} \leqslant \iota_k / N$$

Time uniform estimates for any A

Theo 2 [$\mathbb{L}_k$-mean errors]: $\forall k \geqslant 1\ \exists \iota_k < \infty$ s.t. $\forall N \geqslant 1\ \forall n \geqslant 0$

$$\mathbb{E}\left(|\widehat{m}_n - \widehat{X}_n|^k\right)^{1/k} \vee \mathbb{E}\left(|p_n - P_n|^k\right)^{1/k} \leqslant \iota_k/\sqrt{N}.$$

(state estimates for $N \geqslant N_k$)

Time uniform estimates for any A

Theo 2 [$\mathbb{L}_k$-mean errors]: $\forall k \geqslant 1\ \exists \iota_k < \infty$ s.t. $\forall N \geqslant 1\ \forall n \geqslant 0$

$$\mathbb{E}\left(|\widehat{m}_n - \widehat{X}_n|^k\right)^{1/k} \vee \mathbb{E}\left(|p_n - P_n|^k\right)^{1/k} \leqslant \iota_k/\sqrt{N}.$$

(state estimates for $N \geqslant N_k$)

Many other results/fairly complete analysis: multivariate central limit theorems, exponential decays random products,...