

Uniform estimates for particle filters

P. Del Moral

MCQMC, Rennes, July 4th 2018

Synthesis joint works with : A.N. Bishop, K. Kamatani, A. Kurtzmann, S. Pathiraja, B. Rémillard, J. Tugaut

Filtering and smoothing problems

$$\begin{cases} X_t & \text{Signal = Markov process on } \mathbb{R}^{r_1} \\ Y_t & \text{Observation (partial+noisy) on } \mathbb{R}^{r_2} \end{cases}$$

Filtering and smoothing problems

$$\begin{cases} X_t & \text{Signal} = \text{Markov process on } \mathbb{R}^{r_1} \\ Y_t & \text{Observation (partial+noisy) on } \mathbb{R}^{r_2} \end{cases}$$

Pb: Find/compute/sample/... sequentially

$$\eta_t = \text{Law}(X_t \mid Y_t) \quad \text{or the marginal} \quad \eta_t = \text{Law}(X_t \mid Y_t)$$

with the historical processes

$$X_t := (X_s)_{s \leq t} \quad \text{and} \quad Y_t := (Y_s)_{s \leq t}$$

Filtering and smoothing problems

$$\begin{cases} X_t & \text{Signal} = \text{Markov process on } \mathbb{R}^{r_1} \\ Y_t & \text{Observation (partial+noisy) on } \mathbb{R}^{r_2} \end{cases}$$

Pb: Find/compute/sample/... sequentially

$$\eta_t = \text{Law}(X_t \mid Y_t) \quad \text{or the marginal} \quad \eta_t = \text{Law}(X_t \mid Y_t)$$

with the historical processes

$$X_t := (X_s)_{s \leq t} \quad \text{and} \quad Y_t := (Y_s)_{s \leq t}$$

Some (consistent and sequential) solutions...

- ▶ Linear+Gauss model : Kalman filters \oplus Ensemble Kalman filters
- ▶ Nonlinear and/or non Gaussian models : Particle filters

In all cases ...

$$p(\mathbf{X}_t | \mathbf{Y}_t) \propto p(\mathbf{Y}_t | \mathbf{X}_t) p(\mathbf{X}_t)$$

⊂ *Bayes, Kallianpur-Striebel, Feynman-Kac, ...*

In all cases ...

$$p(\mathbf{X}_t | \mathbf{Y}_t) \propto p(\mathbf{Y}_t | \mathbf{X}_t) p(\mathbf{X}_t)$$

\subset Bayes, Kallianpur-Striebel, Feynman-Kac, ...

" \propto " \implies **Nonlinear evolution equations (in the space of probab.)**

- ▶ Kalman filters \supset Riccati equation
- ▶ In any case : Nonlinear filtering \rightsquigarrow Nonlinear Markov processes

In all cases ...

$$p(\mathbf{X}_t | \mathbf{Y}_t) \propto p(\mathbf{Y}_t | \mathbf{X}_t) p(\mathbf{X}_t)$$

\subset Bayes, Kallianpur-Striebel, Feynman-Kac, ...

" \propto " \implies **Nonlinear evolution equations (in the space of probab.)**

- ▶ Kalman filters \supset Riccati equation
- ▶ In any case : Nonlinear filtering \rightsquigarrow Nonlinear Markov processes

Nonlinear Markov process

Generator/transitions depending on the law of internal states

Mean field particle samplers

- ▶ Sample multiple (interacting) copies
- ▶ Use the empirical distribution when needed

Important Observations & Questions

- ▶ Kalman/Riccati : **Stable equations** + **Can stabilize unstable signals!**
Under appropriate observability and controllability conditions
- ▶ Nonlinear filtering equation : \rightsquigarrow **Stable (i.e. forget initial conditions)**
When the signal is stable/mixing \rightsquigarrow Stabilize unstable signals ?

Important Observations & Questions

- ▶ Kalman/Riccati : **Stable equations** + **Can stabilize unstable signals!**
Under appropriate observability and controllability conditions
- ▶ Nonlinear filtering equation : \rightsquigarrow **Stable (i.e. forget initial conditions)**
When the signal is stable/mixing \rightsquigarrow Stabilize unstable signals ?

Mean field (approximate) particle samplers

- ▶ (Nonlinear) Particle filters:
Uniform estimates (w.r.t. time) when the signal is stable/mixing
- ▶ Ensemble Kalman filters:
Stable equations + **Can stabilize unstable signals!**

Kalman-Bucy filter

$$\begin{cases} dX_t = A X_t dt + R^{1/2} dW_t \\ dY_t = C X_t dt + \Sigma^{1/2} dV_t \end{cases} \implies \eta_t = \mathcal{N}(\hat{X}_t, P_t)$$

with the conditional mean/covariance matrix:

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

Kalman-Bucy filter

$$\begin{cases} dX_t = A X_t dt + R^{1/2} dW_t \\ dY_t = C X_t dt + \Sigma^{1/2} dV_t \end{cases} \implies \eta_t = \mathcal{N}(\hat{X}_t, P_t)$$

with the conditional mean/covariance matrix:

$$\hat{X}_t := \mathbb{E}(X_t | \mathbf{Y}_t) \quad \text{and} \quad P_t := \mathbb{E} \left((X_t - \hat{X}_t) (X_t - \hat{X}_t)' \right)$$

Evolution equations

$$\begin{aligned} d\hat{X}_t &= A \hat{X}_t dt + P_t C' \Sigma^{-1} (dY_t - C \hat{X}_t dt) \\ \partial_t P_t &= \text{Ricc}(P_t) := AP_t + P_t A' - P_t \mathbf{S} P_t + R \quad \text{with} \quad \mathbf{S} := C' \Sigma C \end{aligned}$$

Nonlinear Kalman-Bucy diffusion

Nonlinear diffusions \bar{X}_t depending on the (conditional) distributions

$$\eta_t := \text{Law}(\bar{X}_t \mid \mathbf{Y}_t)$$

Nonlinear Kalman-Bucy diffusion

Nonlinear diffusions \bar{X}_t depending on the (conditional) distributions

$$\eta_t := \text{Law}(\bar{X}_t \mid \mathbf{Y}_t)$$

Interaction with

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

Nonlinear Kalman-Bucy diffusion

Nonlinear diffusions \bar{X}_t depending on the (conditional) distributions

$$\eta_t := \text{Law}(\bar{X}_t \mid \mathbf{Y}_t)$$

Interaction with

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

Consistency property

$$\eta_t := \mathcal{N}[\hat{X}_t, P_t]$$

A couple of McKean-Vlasov type diffusions

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

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

A couple of McKean-Vlasov type diffusions

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

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

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

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

The Ensemble Kalman-Bucy filter

(Case 1) Mean field interpretation $\rightsquigarrow N + 1$ interacting diffusions

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

with the rescaled particle covariance matrices

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

and the empirical measures

$$\eta_t^N := \frac{1}{N+1} \sum_{1 \leq i \leq N+1} \delta_{\xi_t^i} \quad \text{and the sample mean} \quad m_t := \frac{1}{N+1} \sum_{1 \leq i \leq N+1} \xi_t^i.$$

The Ensemble Kalman-Bucy filter

(Case 1) Mean field interpretation $\rightsquigarrow N + 1$ interacting diffusions

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

with the rescaled particle covariance matrices

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

and the empirical measures

$$\eta_t^N := \frac{1}{N+1} \sum_{1 \leq i \leq N+1} \delta_{\xi_t^i} \quad \text{and the sample mean} \quad m_t := \frac{1}{N+1} \sum_{1 \leq i \leq N+1} \xi_t^i.$$

Where are the Kalman & Riccati equations ?

Th: EnKF eq. [+ Tugaut (Arxiv-16/AAP-18)]

The EnKF (state estimate) equation

$$dm_t = A m_t dt + p_t C' \Sigma^{-1} (dY_t - C m_t dt) + \frac{1}{\sqrt{N+1}} d\bar{M}_t$$

with an \mathbf{r} -martingale $\bar{M}_t = (\bar{M}_t(k))_{1 \leq k \leq r}$ with angle brackets

$$\partial_t \langle \bar{M} | \otimes | \bar{M} \rangle_t = U + p_t V p_t.$$

With

$$1) (U, V) = (R, S) \quad \text{and} \quad 2) (U, V) = (R, 0)$$

The particle/ensemble Riccati equation

$$dp_t = \text{Ricc}(p_t) dt + \frac{1}{\sqrt{N}} dM_t$$

Symmetric matrix-valued martingale $M_t = (M_t(k, l))_{1 \leq k, l \leq r}$

The particle/ensemble Riccati equation

$$dp_t = \text{Ricc}(p_t) dt + \frac{1}{\sqrt{N}} dM_t$$

Symmetric matrix-valued martingale $M_t = (M_t(k, l))_{1 \leq k, l \leq r}$

Angle brackets = the Wick-type formula $((\cdot \otimes \cdot)^\sharp := \text{entrywise})$

$$\partial_t \langle M | \otimes | M \rangle_t^\sharp = p_t \otimes_{\text{sym}} (U + p_t V p_t)$$

The particle/ensemble Riccati equation

$$dp_t = \text{Ricc}(p_t) dt + \frac{1}{\sqrt{N}} dM_t$$

Symmetric matrix-valued martingale $M_t = (M_t(k, l))_{1 \leq k, l \leq r}$

Angle brackets = the Wick-type formula $((\cdot \otimes \cdot)^\sharp := \text{entrywise})$

$$\partial_t \langle M | \otimes | M \rangle_t^\sharp = p_t \otimes_{\text{sym}} (U + p_t V p_t)$$

Orthogonality property

$$\forall 1 \leq k, l, l' \leq r \quad \langle M(k, l), \overline{M}(l') \rangle_t = 0.$$

The 1d case \rightsquigarrow Closed form Riccati semigroups

Deterministic Riccati P_t on \mathbb{R}_+ : $\text{Ricc}(\varpi_{\pm}) = 0$ for

$$S\varpi_- := A - \lambda/2 < 0 < S\varpi_+ := A + \lambda/2$$

with

$$\lambda = 2\sqrt{A^2 + RS}$$

\Downarrow

$\forall t \geq v > 0$

$$\left[|P_t - \varpi_+| \vee \exp\left(2 \int_0^t [A - P_s S] ds\right) \right] \leq c_v \exp(-\lambda t)$$

Stochastic Riccati flow $p_t \in \mathbb{R}_+$ with $\epsilon = 2/\sqrt{N}$:

$$dp_t \stackrel{\text{law}}{=} \text{Ricc}(p_t)dt + \epsilon \sqrt{p_t(U + p_t V p_t)} dW_t$$

with $\epsilon^2 U < 2R \implies$ origin repellent

Stochastic Riccati flow $p_t \in \mathbb{R}_+$ with $\epsilon = 2/\sqrt{N}$:

$$dp_t \stackrel{\text{law}}{=} \text{Ricc}(p_t)dt + \epsilon \sqrt{p_t(U + p_t V p_t)} dW_t$$

with $\epsilon^2 U < 2R \implies$ origin repellent

Reversible measures $\pi_\epsilon(dx)$ on \mathbb{R}_+ :

► $U \wedge V > 0 \rightsquigarrow$ **Heavy tails**

$$\propto \frac{x^{\frac{2}{\epsilon^2} \frac{R}{U} - 1}}{[U + Vx^2]^{1 + \frac{1}{\epsilon^2} (\frac{R}{U} + \frac{S}{V})}} \exp \left[\frac{4}{\epsilon^2} \frac{A}{\sqrt{UV}} \tan^{-1} \left(x \sqrt{\frac{V}{U}} \right) \right] dx.$$

Stochastic Riccati flow $p_t \in \mathbb{R}_+$ with $\epsilon = 2/\sqrt{N}$:

$$dp_t \stackrel{\text{law}}{=} \text{Ricc}(p_t)dt + \epsilon \sqrt{p_t(U + p_t V p_t)} dW_t$$

with $\epsilon^2 U < 2R \implies$ origin repellent

Reversible measures $\pi_\epsilon(dx)$ on \mathbb{R}_+ :

▶ $U \wedge V > 0 \rightsquigarrow$ **Heavy tails**

$$\propto \frac{x^{\frac{2}{\epsilon^2} \frac{R}{U} - 1}}{[U + Vx^2]^{1 + \frac{1}{\epsilon^2} (\frac{R}{U} + \frac{S}{V})}} \exp \left[\frac{4}{\epsilon^2} \frac{A}{\sqrt{UV}} \tan^{-1} \left(x \sqrt{\frac{V}{U}} \right) \right] dx.$$

▶ $U > V = 0 \rightsquigarrow$ **Gaussian-type tails**

$$\propto x^{\frac{2}{\epsilon^2} \frac{R}{U} - 1} \exp \left[-\frac{S}{U\epsilon^2} \left(x - 2 \frac{A}{S} \right)^2 \right] dx.$$

Stability Markov transition semigroup \mathcal{P}_t^ϵ (of p_t)

Th [+ Bishop, Kamatani, Rémillard Arxiv-17] $\forall A, R \wedge S > 0$

- $\exists \zeta, \epsilon_0 > 0$ and some Wasserstein distance \mathbb{D} s.t. for any $0 \leq \epsilon \leq \epsilon_0$

$$\mathbb{D}(\mu_1 \mathcal{P}_t^\epsilon, \mu_2 \mathcal{P}_t^\epsilon) \leq \exp(-\lambda (1 - \epsilon^2 \zeta) t) \mathbb{D}(\mu_1, \mu_2)$$

Stability Markov transition semigroup \mathcal{P}_t^ϵ (of p_t)

Th [+ Bishop, Kamatani, Rémillard Arxiv-17] $\forall A, R \wedge S > 0$

- $\exists \zeta, \epsilon_0 > 0$ and some Wasserstein distance \mathbb{D} s.t. for any $0 \leq \epsilon \leq \epsilon_0$

$$\mathbb{D}(\mu_1 \mathcal{P}_t^\epsilon, \mu_2 \mathcal{P}_t^\epsilon) \leq \exp(-\lambda (1 - \epsilon^2 \zeta) t) \mathbb{D}(\mu_1, \mu_2)$$

- $\forall n \geq 1 \exists \zeta_n, \epsilon_n > 0$ for any $0 \leq \epsilon \leq \epsilon_n$

$$\mathbb{E} \left[\exp \left[n \int_0^t (A - p_s S) ds \right] \right]^{1/n} \leq c_Q \exp(-\lambda (1 - \epsilon^2 \zeta_n) t)$$

Stability Markov transition semigroup \mathcal{P}_t^ϵ (of p_t)

Th [+ Bishop, Kamatani, Rémillard Arxiv-17] $\forall A, R \wedge S > 0$

- $\exists \zeta, \epsilon_0 > 0$ and some Wasserstein distance \mathbb{D} s.t. for any $0 \leq \epsilon \leq \epsilon_0$

$$\mathbb{D}(\mu_1 \mathcal{P}_t^\epsilon, \mu_2 \mathcal{P}_t^\epsilon) \leq \exp(-\lambda (1 - \epsilon^2 \zeta) t) \mathbb{D}(\mu_1, \mu_2)$$

- $\forall n \geq 1 \exists \zeta_n, \epsilon_n > 0$ for any $0 \leq \epsilon \leq \epsilon_n$

$$\mathbb{E} \left[\exp \left[n \int_0^t (A - p_s S) ds \right] \right]^{1/n} \leq c_Q \exp(-\lambda (1 - \epsilon^2 \zeta_n) t)$$

Some extensions

Case 2: Poincaré inequalities (and $\mathbb{L}_2(\pi_\epsilon)$ -contractions), ...

Consequences

Uniform estimates for state estimates + particle Riccati diffusions, ...

Multivariate KB : Observability + Controllability

$$\begin{aligned} & d(\bar{X}_t - X_t) \\ &= (\mathbf{A} - \mathbf{P}_t \mathbf{S}) (\bar{X}_t - X_t) dt + R^{1/2} d(\bar{W}_t - W_t) + P_t C' \Sigma^{-1/2} d[V_t - \bar{V}_t] \end{aligned}$$

Multivariate KB : Observability + Controllability

$$\begin{aligned} & d(\bar{X}_t - X_t) \\ &= (\mathbf{A} - \mathbf{P}_t \mathbf{S}) (\bar{X}_t - X_t) dt + R^{1/2} d(\bar{W}_t - W_t) + P_t C' \Sigma^{-1/2} d[V_t - \bar{V}_t] \end{aligned}$$

Steady state: $\exists P > 0$ such that $\text{Ricc}(P) = 0$ and spectral abscissa

$$\zeta(\mathbf{A} - \mathbf{P}\mathbf{S}) := \max \{ \text{Re}(\lambda) : \lambda \in \text{Spec}(\mathbf{A} - \mathbf{P}\mathbf{S}) \} < 0$$

Multivariate KB : Observability + Controllability

$$\begin{aligned} & d(\bar{X}_t - X_t) \\ &= (\mathbf{A} - \mathbf{P}_t \mathbf{S}) (\bar{X}_t - X_t) dt + R^{1/2} d(\bar{W}_t - W_t) + P_t C' \Sigma^{-1/2} d[V_t - \bar{V}_t] \end{aligned}$$

Steady state: $\exists! P > 0$ such that $\text{Ric}(P) = 0$ and spectral abscissa

$$\varsigma(\mathbf{A} - \mathbf{P}\mathbf{S}) := \max \{ \text{Re}(\lambda) : \lambda \in \text{Spec}(\mathbf{A} - \mathbf{P}\mathbf{S}) \} < 0$$



STABLE EVEN WHEN A is unstable.

↪ SIAM Control & Opt.-17 \oplus Arxiv-18 (+ Bishop)
Review on the stability of Kalman-Bucy filters and Riccati matrix semigroups \oplus Floquet representation of exponential semigroups

Multivariate : EnKF

$(m_t, X_t, p_t) = (\text{sample mean, true signal, sample covariance})$

↓

$$d(m_t - X_t) = (A - p_t S) (m_t - X_t) dt + p_t C' \Sigma^{-1/2} dV_t + \frac{d\bar{M}_t}{\sqrt{N}}$$

Multivariate : EnKF

$(m_t, X_t, p_t) = (\text{sample mean, true signal, sample covariance})$

↓

$$d(m_t - X_t) = (A - p_t S) (m_t - X_t) dt + p_t C' \Sigma^{-1/2} dV_t + \frac{d\bar{M}_t}{\sqrt{N}}$$

Observations:

- ▶ **Time varying** \oplus **stochastic type** Ornstein-Uhlenbeck diffusion

DRIVEN BY A STOCH. MATRIX-RICCATI DIFFUSION p_t

Multivariate : EnKF

$(m_t, X_t, p_t) = (\text{sample mean, true signal, sample covariance})$

↓

$$d(m_t - X_t) = (A - p_t S) (m_t - X_t) dt + p_t C' \Sigma^{-1/2} dV_t + \frac{d\bar{M}_t}{\sqrt{N}}$$

Observations:

- ▶ **Time varying** \oplus **stochastic type** Ornstein-Uhlenbeck diffusion

DRIVEN BY A STOCH. MATRIX-RICCATI DIFFUSION p_t

- ▶ The matrix $(A - pS)$ may be ill-conditioned in the sense that

$$\exists p : \lambda_{\max}((A - pS)_{sym}) > 0 > \lambda_{\min}((A - pS)_{sym})$$

Multivariate : EnKF

$(m_t, X_t, p_t) = (\text{sample mean, true signal, sample covariance})$

↓

$$d(m_t - X_t) = (A - p_t S) (m_t - X_t) dt + p_t C' \Sigma^{-1/2} dV_t + \frac{dM_t}{\sqrt{N}}$$

Observations:

- ▶ **Time varying** \oplus **stochastic type** Ornstein-Uhlenbeck diffusion

DRIVEN BY A STOCH. MATRIX-RICCATI DIFFUSION p_t

- ▶ The matrix $(A - pS)$ may be ill-conditioned in the sense that

$$\exists p : \lambda_{\max}((A - pS)_{sym}) > 0 > \lambda_{\min}((A - pS)_{sym})$$

- ▶ Always under-biased

$$\forall t > 0 \quad 0 < p_t \quad \text{but} \quad 0 < \mathbb{E}(p_t) < P_t$$

When A stable and $S > 0$

Theo [+ Tugaut (Arxiv-16/AAP-18)] $\forall n \geq 1 \exists N_n \geq 1 : \forall N \geq N_n$

$$\sup_{t \geq 0} \left[\mathbb{E}(\|p_t - P_t\|^n)^{1/n} \vee \mathbb{E}(\|m_t - \hat{X}_t\|^n)^{1/n} \right] < c_n / \sqrt{N}$$

Under : Observability + Controllability

Some additional and more refined theorems (N large enough)

Uniform moments + Uniform Riccati estimates + stability and invariant measures Riccati diffusions + non asymptotic CLT rates + Bias-Taylor type expansions + Robustness and Perturbations analysis (inflation, masking, shrinkage, projections),...

- ▶ SPA-17 (+ Bishop, Pathiraja):
Uniform robustness properties : inflation and localisation techniques.
- ▶ Arxiv-17 (+ Bishop, Niclas):
Exact propagation of chaos expansions matrix Riccati diffusions, non asymptotic bias + CLT.
- ▶ Arxiv-18 (+ Bishop):
Stability of stochastic Ornstein-Uhlenbeck diffusions
- ▶ in a few days/weeks (+ Bishop):
Stability of Matrix Riccati equations.