Uniform estimates for particle filters

P. Del Moral

Workshop, Singapore, Sept 7th 2018

Synthesis joint works with : A. Bishop, A. Guionnet, A. Jasra, K. Kamatani, A. Kurtzmann, L. Miclo, B. Rémillard, J. Tugaut.

Filtering and smoothing problems

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

Filtering and smoothing problems

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

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

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

with the historical processes

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

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

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

with the historical processes

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

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 ...

$$p(\boldsymbol{X_t} \mid \boldsymbol{Y_t}) \quad \propto \quad p(\boldsymbol{Y_t} \mid \boldsymbol{X_t})\, p(\boldsymbol{X_t})$$

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

In all cases . . .

$$p(\boldsymbol{X_t} | \boldsymbol{Y_t}) \quad \propto \quad p(\boldsymbol{Y_t} | \boldsymbol{X_t})\ p(\boldsymbol{X_t})$$

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

"$\propto$" $\Longrightarrow$ **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 . . .

$$p(\boldsymbol{X_t} \mid \boldsymbol{Y_t}) \quad \propto \quad p(\boldsymbol{Y_t} \mid \boldsymbol{X_t})\; p(\boldsymbol{X_t})$$

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

"$\propto$" $\Longrightarrow$ **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

↳ `Review on Kalman-Bucy (+ Bishop - SIAM-17)`
`Floquet representation (+ Bishop - Arxiv-18)`
`Stability random linear systems (+ Bishop - Arxiv-18)`

- Nonlinear filtering equation : $\rightsquigarrow$ Stable (forgets initial conditions)

When the signal is stable/mixing $\rightsquigarrow$ *Stabilize unstable signals/rates ?*

Important Observations & Questions

- Kalman/Riccati : Stable equations + Can stabilize unstable signals!

Under appropriate observability and controllability conditions

↳ `Review on Kalman-Bucy (+ Bishop - SIAM-17)`
`Floquet representation (+ Bishop - Arxiv-18)`
`Stability random linear systems (+ Bishop - Arxiv-18)`

- Nonlinear filtering equation : $\rightsquigarrow$ Stable (forgets initial conditions)

When the signal is stable/mixing $\rightsquigarrow$ *Stabilize unstable signals/rates ?*

↳ $\exists$ **Asymptotic results** ($\star$):
(discrete time [full obs./small noise + $p(x_0) \simeq q(x_0)$])
`Budhiraja-Ocone-99`, `Kleptsyna-Veretennikov-08`,
`Blackwell-Dubins-62` merging principle (Van Handel-08-09-10)

Mean field (approximate) particle samplers

- **(Nonlinear) Particle filters:**

Uniform estimates (w.r.t. time) when the signal is stable/mixing.

+ Guionnet [CRAS-99, IHP-01] ($\star\star$)
+ Miclo [Sem. Probab., Springer Lecture Notes -00]
Feynman-Kac formulae [Springer -04],
+ Hu-Wu, Fondations & trends in Machine learning -11
Mean field simulation for Monte Carlo integration [CRC -13]

. . . . . .

+ Jasra [SPA - 18]

Mean field (approximate) particle samplers

- **(Nonlinear) Particle filters:**

Uniform estimates (w.r.t. time) when the signal is stable/mixing.

+ Guionnet [CRAS-99, IHP-01] (⋆⋆)
+ Miclo [Sem. Probab., Springer Lecture Notes -00]
  Feynman-Kac formulae [Springer -04],
+ Hu-Wu, Fondations & trends in Machine learning -11
  Mean field simulation for Monte Carlo integration [CRC -13]

......

+ Jasra [SPA - 18]

Main tool from (⋆⋆):
Backward perturbation + propagations of local perturbations

Mean field (approximate) particle samplers

- **(Nonlinear) Particle filters:**

Uniform estimates (w.r.t. time) when the signal is stable/mixing.

+ Guionnet [CRAS-99, IHP-01] ($\star\star$)
+ Miclo [Sem. Probab., Springer Lecture Notes -00]
  Feynman-Kac formulae [Springer -04],
+ Hu-Wu, Fondations & trends in Machine learning -11
  Mean field simulation for Monte Carlo integration [CRC -13]

......

+ Jasra [SPA - 18]

Main tool from ($\star\star$)*:*
Backward perturbation + propagations of local perturbations

↳ **Unstable signal ?** $\rightsquigarrow$ no article based on the asympt. results ($\star$)**?**

Mean field (approximate) particle samplers

- **(Nonlinear) Particle filters:**

Uniform estimates (w.r.t. time) when the signal is stable/mixing.

+ Guionnet [CRAS-99, IHP-01] ($\star\star$)
+ Miclo [Sem. Probab., Springer Lecture Notes -00]
  Feynman-Kac formulae [Springer -04],
+ Hu-Wu, Fondations & trends in Machine learning -11
  Mean field simulation for Monte Carlo integration [CRC -13]

. . . . . .

+ Jasra [SPA - 18]

Main tool from ($\star\star$)*:*
Backward perturbation + propagations of local perturbations

↳ **Unstable signal ?** $\rightsquigarrow$ no article based on the asympt. results ($\star$)**?**

- **Today : Ensemble Kalman filters:**

Stable equations + Stabilize unstable signals + non asympt. rates !

Kalman-Bucy filter

$$\left\{ \begin{array}{lcl} dX_t & = & A\ X_t\ dt\ +\ R^{1/2}\ dW_t \\ dY_t & = & C\ X_t\ dt\ +\ \Sigma^{1/2}\ dV_t \end{array} \right. \Longrightarrow \eta_t = \mathcal{N}(\widehat{X}_t, P_t)$$

with the conditional mean/covariance matrix:

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

Kalman-Bucy filter

$$\left\{\begin{array}{lcl} dX_t &=& A\,X_t\,dt \;+\; R^{1/2}\,dW_t \\ dY_t &=& C\,X_t\,dt \;+\; \Sigma^{1/2}\,dV_t \end{array}\right. \Longrightarrow \eta_t = \mathcal{N}(\widehat{X}_t, P_t)$$

with the conditional mean/covariance matrix:

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

Evolution equations

$$\begin{array}{lcl} d\widehat{X}_t &=& A\,\widehat{X}_t\,dt + P_t\,C'\Sigma^{-1}\left(dY_t - C\widehat{X}_t\,dt\right) \\ \partial_t P_t &=& \mathrm{Ricc}(P_t) := AP_t + P_tA' - P_t\mathbf{S}P_t + R \quad \text{with} \quad \mathbf{S} := \mathbf{C}'\boldsymbol{\Sigma}\mathbf{C} \end{array}$$

Nonlinear Kalman-Bucy diffusion

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

$$\eta_t := \mathrm{Law}(\overline{X}_t \mid \boldsymbol{Y_t})$$

Nonlinear Kalman-Bucy diffusion

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

$$\eta_t := \mathrm{Law}(\overline{X}_t \mid \boldsymbol{Y_t})$$

Interaction with

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

Nonlinear Kalman-Bucy diffusion

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

$$\eta_t := \mathrm{Law}(\overline{X}_t \mid \boldsymbol{Y_t})$$

Interaction with

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

Consistency property

$$\eta_t := \mathcal{N}\left[\widehat{X}_t, P_t\right]$$

A couple of of McKean-Vlasov type diffusions

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

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

A couple of of McKean-Vlasov type diffusions

1) "Vanilla EnKF" ($\rightsquigarrow$ 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)

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

A couple of of McKean-Vlasov type diffusions

1) "Vanilla EnKF" ($\rightsquigarrow$ 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)

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

3) Pure transport equation (Reich-Cotter 13) [pb. "inversion" ?]

$$d\overline{X}_t = A\,\overline{X}_t dt$$

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

A couple of of McKean-Vlasov type diffusions

1) "Vanilla EnKF" ($\rightsquigarrow$ 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)

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

3) Pure transport equation (Reich-Cotter 13) [pb. "inversion" ?]

$$d\overline{X}_t = A\,\overline{X}_t dt$$

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

$\oplus\ \infty$ **many others, adding** $\mathcal{Q}_{\eta_t}\mathcal{P}_{\eta_t}^{-1}(\overline{X}_t - \eta_t(e))\ dt$ **for any** $\mathcal{Q}'_{\eta_t} = -\mathcal{Q}_{\eta_t}$**.**

The Ensemble Kalman-Bucy filter

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

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

with the rescaled particle covariance matrices

$$p_t := \left(1 + \frac{1}{N}\right)\ \mathcal{P}_{\eta_t^N} = \frac{1}{N} \sum_{1 \le i \le N+1} \left(\xi_t^i - m_t\right) \left(\xi_t^i - m_t\right)'$$

and the empirical measures

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

The Ensemble Kalman-Bucy filter

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

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

with the rescaled particle covariance matrices

$$p_t := \left(1 + \frac{1}{N}\right)\, \mathcal{P}_{\eta_t^N} = \frac{1}{N} \sum_{1 \le i \le N+1} \left(\xi_t^i - m_t\right) \left(\xi_t^i - m_t\right)'$$

and the empirical measures

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

Where are the Kalman & Riccati equations ?

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

The EnKF (state estimate) equation

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

with an **r**-martingale $\overline{M}_t = \left(\overline{M}_t(k)\right)_{1\leq k\leq r}$ with angle brackets

$$\partial_t\langle\overline{M}\mid\otimes\mid\overline{M}\rangle_t = U + p_t V p_t.$$

With

1) $(U,V)=(R,S)$ and 2) $(U,V)=(R,0)$ and 3) $(U,V)=(0,0)$

Nb.:

$(U,V)=(0,0) \Longrightarrow$ direct consequence of contraction Kalman-Bucy filter

↳ Review on Kalman-Bucy (+ Bishop - SIAM-17)
Floquet representation (+ Bishop - Arxiv-18)

The particle/ensemble Riccati equation

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

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

The particle/ensemble Riccati equation

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

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

Angle brackets $=$ **Wick-type formula** (($\bullet \otimes\ \bullet)^\sharp$:= **entrywise)**

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

The particle/ensemble Riccati equation

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

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

Angle brackets $=$ **Wick-type formula** (($\cdot \otimes \cdot)^\sharp$:= **entrywise)**

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

Orthogonality property

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

Nb.:

$$(U,V) = (0,0) \Longrightarrow \text{direct consequence of contraction Riccati eq.}$$

↳ `Review on Kalman-Bucy (+ Bishop - SIAM-17)`
`Floquet representation (+ Bishop - Arxiv-18)`

In terms of random matrices with $\epsilon := \frac{2}{\sqrt{N}}$

$$\mathcal{W}_t = (\mathcal{W}_t(i,j))_{1\leq i,j\leq r} \quad \text{independent Brownian motions}$$

$$\downarrow$$

$$dp_t \overset{law}{=} [Ap_t + p_t A' + R - p_t S p_t]\ dt + \epsilon\ \left(p_t^{1/2} d\mathcal{W}_t\,(U + p_t V p_t)^{1/2}\right)_{sym}$$

- $S = 0 = V \rightsquigarrow$ Wishart process

In terms of random matrices with $\epsilon := \frac{2}{\sqrt{N}}$

$$\mathcal{W}_t = (\mathcal{W}_t(i,j))_{1 \le i,j \le r} \quad \text{independent Brownian motions}$$

$$\downarrow$$

$$dp_t \overset{law}{=} [Ap_t + p_t A' + R - p_t S p_t] \; dt + \epsilon \; \left(p_t^{1/2} d\mathcal{W}_t \, (U + p_t V p_t)^{1/2} \right)_{sym}$$

- $S = 0 = V \rightsquigarrow$ Wishart process
- $(A, R, S) = (\alpha \; I, \beta \; I, \gamma \; I)$ and $(U, V) = (I, 0)$

In terms of random matrices with $\epsilon := \frac{2}{\sqrt{N}}$

$$\mathcal{W}_t = (\mathcal{W}_t(i,j))_{1\leq i,j\leq r} \quad \text{independent Brownian motions}$$

$$\downarrow$$

$$dp_t \stackrel{law}{=} [Ap_t + p_t A' + R - p_t S p_t] \; dt + \epsilon \; \left(p_t^{1/2} d\mathcal{W}_t \, (U + p_t V p_t)^{1/2}\right)_{sym}$$

- $S = 0 = V \rightsquigarrow$ Wishart process
- $(A, R, S) = (\alpha \; I, \beta \; I, \gamma \; I)$ and $(U, V) = (I, 0)$

$$p_t \rightsquigarrow \textit{non colliding eigenvalues} \quad \lambda_r(t) < \ldots < \lambda_2(t) < \lambda_1(t)$$

satisfying the Dyson equation

$$d\lambda_i(t) =$$

$$\left[2\alpha\lambda_i(t) + \beta - \lambda_i(t)^2\gamma + \frac{\epsilon^2}{4} \sum_{j\neq i} \frac{\lambda_i(t) + \lambda_j(t)}{\lambda_i(t) - \lambda_j(t)}\right] dt + \epsilon \; \sqrt{\lambda_i(t)} \; dW_t^i$$

The 1d case $\rightsquigarrow$ Closed form Riccati semigroups

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

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

with

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

$$\Downarrow$$

$\forall t \geq \upsilon > 0$

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

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

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

$$\text{with} \quad \epsilon^2 U < 2R \quad \Longrightarrow \quad \text{origin repellent}$$

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

$$dp_t \overset{law}{=} \mathrm{Ricc}(p_t)dt + \epsilon \sqrt{p_t(U + p_t V p_t)}\; dW_t$$

$$\text{with} \quad \epsilon^2 U < 2R \quad \Longrightarrow \quad \text{origin repellent}$$

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

- $U \wedge V > 0 \rightsquigarrow$ Heavy tails

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

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

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

$$\text{with} \quad \epsilon^2 U < 2R \quad \Longrightarrow \quad \text{origin repellent}$$

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

- $U \wedge V > 0 \rightsquigarrow$ Heavy tails

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

- $U > V = 0 \rightsquigarrow$ 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 $\mathcal{P}_t^\epsilon$ (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 $0 \leq \epsilon \leq \epsilon_0$

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

Stability Markov transition semigroup $\mathcal{P}_t^{\epsilon}$ (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 $0 \leq \epsilon \leq \epsilon_0$

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

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

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

Stability Markov transition semigroup $\mathcal{P}_t^\epsilon$ (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 $0 \leq \epsilon \leq \epsilon_0$

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

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

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

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

$$d(\widehat{X}_t - X_t)$$

$$= (\mathbf{A} - \mathbf{P_t S})\,(\widehat{X}_t - X_t)dt - \boldsymbol{R}^{1/2}\,\boldsymbol{dW_t} + \boldsymbol{P_t C' \Sigma^{-1/2}\,dV_t}$$

Multivariate KB : Observability + Controllability

$$d(\widehat{X}_t - X_t)$$

$$= (\mathbf{A} - \mathbf{P_t S})\,(\widehat{X}_t - X_t)dt \,-\, \boldsymbol{R^{1/2}\, dW_t + P_t C' \Sigma^{-1/2}\, dV_t}$$

Steady state: $\exists! P_\infty > 0$ **s.t.** $\mathrm{Ricc}(P_\infty) = 0$ **and spectral abscissa**

$$\varsigma(\mathbf{A} - \mathbf{P_\infty S}) := \max\{\mathrm{Re}(\lambda) \;:\; \lambda \in \mathrm{Spec}(A - P_\infty S)\} < 0$$

Multivariate KB : Observability + Controllability

$$d(\widehat{X}_t - X_t)$$

$$= (\mathbf{A} - \mathbf{P_t S})\,(\widehat{X}_t - X_t)dt - \boldsymbol{R^{1/2}\, dW_t + P_t C' \Sigma^{-1/2}\, dV_t}$$

Steady state: $\exists! P_\infty > 0$ **s.t.** $\mathrm{Ricc}(P_\infty) = 0$ **and spectral abscissa**

$$\varsigma(\mathbf{A} - \mathbf{P_\infty S}) := \max\left\{\mathrm{Re}(\lambda)\ :\ \lambda \in \mathrm{Spec}(A - P_\infty S)\right\} < 0$$

$$\Downarrow$$

STABLE EVEN WHEN A is unstable.

$\rightsquigarrow$ `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*

Floquet representations

$$P_t = \phi_t(P_0) \quad \text{flow of the Riccati equation} \quad \partial_t P_t = \text{Ricc}(P_t)$$

$$\downarrow$$

Floquet representations

$$P_t = \phi_t(P_0) \quad \text{flow of the Riccati equation} \quad \partial_t P_t = \mathrm{Ricc}(P_t)$$

$$\downarrow$$

Exponential semigroups/Fundamental matrices:

$$E_{s,t}(P) = \exp \oint_s^t (A - \phi_u(P)S)\, du$$

Floquet representations

$$P_t = \phi_t(P_0) \quad \text{flow of the Riccati equation} \quad \partial_t P_t = \mathrm{Ricc}(P_t)$$

$$\downarrow$$

Exponential semigroups/Fundamental matrices:

$$E_{s,t}(P) = \exp \oint_s^t (A - \phi_u(P)S)\, du$$

in the sense that (with $E_{t,t}(P) = Id$)

$$\partial_t E_{s,t}(P) = (A - \phi_t(P)S) E_{s,t}(P) \quad \text{and} \quad \partial_s E_{s,t}(P) = -E_{s,t}(P)(A - \phi_s(P)S)$$

Floquet representations

$$P_t = \phi_t(P_0) \quad \text{flow of the Riccati equation} \quad \partial_t P_t = \mathrm{Ricc}(P_t)$$

$$\downarrow$$

Exponential semigroups/Fundamental matrices:

$$E_{s,t}(P) = \exp \oint_s^t (A - \phi_u(P)S)\, du$$

in the sense that (with $E_{t,t}(P) = Id$)

$$\partial_t E_{s,t}(P) = (A - \phi_t(P)S)E_{s,t}(P) \quad \text{and} \quad \partial_s E_{s,t}(P) = -E_{s,t}(P)(A - \phi_s(P)S)$$

Nb.:

$$P = P_\infty \Longrightarrow E_{s,t}(P_\infty) = e^{(t-s)(A - P_\infty S)} \quad \text{with} \quad A - P_\infty S \quad \text{stable}$$

Floquet representations 2/2

Theo.: (+ Bishop - Arxiv-18)

$$E_t(P) := \exp \oint_0^t (\mathbf{A} - \phi_{\mathbf{s}}(\mathbf{P})\mathbf{S})\, ds = e^{t(A - P_\infty S)}\ \mathbb{C}_t(P)^{-1}$$

with

$$\sup_{t \geq 0} \|\mathbb{C}_t(P)^{-1}\| \leq c\ (1 + \|Q\|)$$

Cor.:

$$\|\phi_t(P_1) - \phi_t(P_2)\| \leq \|e^{t(A - P_\infty S)}\|\ (1 + \|P_1\|^2 + \|P_2\|^2)\ \|P_1 - P_2\|$$

$\oplus$ **same type of estimates for the time varying linear process**

$$d(\widehat{X}_t - X_t)$$

$$= (\mathbf{A} - \mathbf{P_t S})\ (\widehat{X}_t - X_t)\ dt - \boldsymbol{R^{1/2}\ dW_t + P_t C' \Sigma^{-1/2}\ dV_t}$$

Multivariate : EnKF

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

$$\Downarrow$$

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

Multivariate : EnKF

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

$$\Downarrow$$

$$d(m_t - X_t) = (A - \mathbf{p_t} S)\,(m_t - X_t)\,dt - \boldsymbol{R}^{1/2}\boldsymbol{dW_t} + \boldsymbol{p_t}\;\boldsymbol{C'\Sigma}^{-1/2}\;\boldsymbol{dV_t} + \frac{\mathbf{d\overline{M}_t}}{\sqrt{\mathbf{N}}}$$

Observations:

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

  DRIVEN BY A STOCH. MATRIX-RICCATI DIFFUSION $\mathbf{p_t}$

Multivariate : EnKF

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

$$\Downarrow$$

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

Observations:

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

  DRIVEN BY A STOCH. MATRIX-RICCATI DIFFUSION $\mathbf{p_t}$

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

  $\exists p \ :\ \lambda_{max}((A - pS)_{sym}) > 0 \quad$ even if A stable in dimension ≥ 2

Multivariate : EnKF

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

$$\Downarrow$$

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

Observations:

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

  DRIVEN BY A STOCH. MATRIX-RICCATI DIFFUSION $\mathbf{p_t}$

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

  $\exists p \ : \ \lambda_{max}((A - pS)_{sym}) > 0$ even if A stable in dimension ≥ 2

- Always under-biased

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

When A stable and $S > 0$

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

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

for any matrix norm (ex.: the spectral of the Frobenius norm).

Under only : Observability + Controllability

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),...

$\rightsquigarrow$ **more details in the talk by A.N. Bishop this conference $\oplus$ workshop**

Some refs:

- `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
- `Arxiv-18 (+ Bishop):`
  Stability of Matrix Riccati equations.

Nonlinear models

Extended Kalman-Bucy-filters

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

with the "stochastic" Riccati equation:

$$\partial_t P_t = \partial A(\widehat{X}_t) P_t + P_t\ \partial A(\widehat{X}_t)' + R - P_t S P_t$$

Nonlinear models

Extended Kalman-Bucy-filters

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

with the "stochastic" Riccati equation:

$$\partial_t P_t = \partial A(\widehat{X}_t) P_t + P_t\ \partial A(\widehat{X}_t)' + R - P_t S P_t$$

McKean-Vlasov interpretation

$$\begin{aligned} d\overline{X}_t &= \mathcal{A}\left(\overline{X}_t, \eta_t(e)\right)\ dt + R^{1/2}\ d\overline{W}_t \\ &\quad + \mathcal{P}_{\eta_t} C' R_2^{-1}\ \left[dY_t - \left(C\overline{X}_t\ dt + \Sigma^{1/2}\ d\overline{V}_t\right)\right] \end{aligned}$$

with the drift function

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

Extended Ensemble Kalman-Bucy-filters

En-EKF = Mean field particle model

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

with the sample means m_t *and covariance matrices* p_t and the drift

$$\mathcal{A}(\xi_t^i, m_t) := A[m_t] + \underbrace{\partial A[m_t]\ (\xi_t^i - m_t)}_{\text{Repulsion/Attraction w.r.t. } m_t}$$

Some illustrations

Langevin type signal processes

$$R = \sigma^2 \; Id \quad \text{and} \quad (A, \partial A) = (-\partial \mathcal{V}, -\partial^2 \mathcal{V})$$

Non quadratic potential $(q \in \mathbb{R}^r, \mathcal{Q}_1, \mathcal{Q}_2 \geq 0)$

$$\mathcal{V}(x) = \frac{1}{2} \; \langle \mathcal{Q}_1 x, x \rangle + \langle q, x \rangle + \frac{1}{3} \; \langle \mathcal{Q}_2 x, x \rangle^{3/2}$$

Interacting diffusion gradient flows

$$\mathcal{V}(x) = \sum_{1 \leq i \leq r} \mathcal{U}_1(x_i) + \sum_{1 \leq i \neq j \leq r} \mathcal{U}_2(x_i, x_j)$$

for some convex confining potential $\mathcal{U}_i : \mathbb{R}^i \mapsto [0, \infty[$

Regularity conditions

Full observation $S = s\ Id$ and

$$\begin{aligned} -\lambda_{\partial A} &:= \sup_{x\in\mathbb{R}^r} \lambda_{max}(\partial A(x) + \partial A(x)') < 0 \\ \|\partial A(x) - \partial A(y)\| &\leq \kappa_{\partial A}\ \|x-y\| \end{aligned}$$

Examples: Langevin signal-diffusion

$$(\lambda_{\partial A}, \kappa_{\partial A}) = \beta\ \left(2^{-1}\lambda_{min}(\mathcal{Q}_1),\ 2\lambda_{max}^{3/2}(\mathcal{Q}_2)\right).$$

more generally $\partial^2\mathcal{V} \geq v\ Id \oplus$ Lipschitz condition

Stability theorem

$$(\overline{X}_t, \overline{Z}_t) := \text{McKean-Vlasov starting at} \quad (\overline{X}_0, \overline{Z}_0)$$

Stability theorem

$$(\overline{X}_t, \overline{Z}_t) := \text{McKean-Vlasov starting at} \quad (\overline{X}_0, \overline{Z}_0)$$

$$\Downarrow$$

Theo [+Kurtzmann-Tugaut]

When $\lambda_{\partial A}$ is sufficiently large we have

$$\mathbb{W}_2(\mathrm{Law}(\overline{X}_t), \mathrm{Law}(\overline{Z}_t)) \leq c \ \exp[-t\ \lambda] \quad \text{for some} \quad \lambda > 0.$$

$\exists$ *more explicit description in terms of* $(R, S, \kappa_{\partial A})$.

Propagation of chaos

$$\mathbb{P}_t^N := \mathrm{Law}(m_t, p_t) \qquad \mathbb{P}_t := \mathrm{Law}(\widehat{X}_t, P_t)$$

and

$$\mathbb{Q}_t^N := \mathrm{Law}(\xi_t^1) \qquad \mathbb{Q}_t := \mathrm{Law}(\overline{X}_t)$$

Propagation of chaos

$$\mathbb{P}_t^N := \mathrm{Law}(m_t, p_t) \qquad \mathbb{P}_t := \mathrm{Law}(\widehat{X}_t, P_t)$$

and

$$\mathbb{Q}_t^N := \mathrm{Law}(\xi_t^1) \qquad \mathbb{Q}_t := \mathrm{Law}(\overline{X}_t)$$

$$\Downarrow$$

Theo [+Kurtzmann-Tugaut]

When $\lambda_{\partial A}$ is sufficiently large, $\exists \beta \in]0, 1/2]$ s.t.

$$\sup_{t \geq 0} \mathbb{W}_2\left(\mathbb{P}_t^N, \mathbb{P}_t\right) \vee \sup_{t \geq 0} \mathbb{W}_2\left(\mathbb{Q}_t^N, \mathbb{Q}_t\right) \leq c\ N^{-\beta}$$

as soon as $\mathrm{tr}(P_0)$ is not too large and N large enough...