An Optimal Transport Formulation of the Linear Feedback Particle Filter

American Control Conference (ACC),
Boston, MA, July 6-8, 2016

Amirhossein Taghvaei
Joint work with P. G. Mehta

Coordinated Science Laboratory
University of Illinois at Urbana-Champaign

July, 2016

ILLINOIS

Kalman Filter
Continuous time

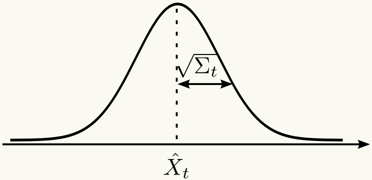
Model:

$$\mathrm{d}X_t = AX_t\,\mathrm{d}t + \mathrm{d}B_t$$
$$\mathrm{d}Z_t = CX_t\,\mathrm{d}t + \mathrm{d}W_t$$

Kalman Filter:

$\mathsf{P}(X_t|\mathcal{Z}_t)$ is Gaussian $N(\hat{X}_t, \Sigma_t)$,

$$\mathrm{d}\hat{X}_t = A\hat{X}_t\,\mathrm{d}t + \mathsf{K}_t(\,\mathrm{d}Z_t - C\hat{X}_t\,\mathrm{d}t)$$
$$\frac{\mathrm{d}\Sigma_t}{\mathrm{d}t} = \ldots \text{(Riccati equation)}$$



Feedback Particle Filter (FPF)

Continuous time

Model:

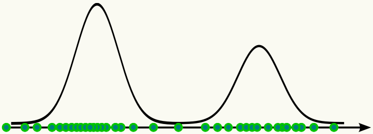
$$\mathrm{d}X_t = a(X_t)\,\mathrm{d}t + \mathrm{d}B_t$$
$$\mathrm{d}Z_t = h(X_t)\,\mathrm{d}t + \mathrm{d}W_t$$

Feedback Particle Filter:

$$\mathsf{P}(X_t|\mathcal{Z}_t) \approx \text{empirical dist. of } \{S^1, \ldots, S^N\},$$

$$\mathrm{d}S_t^i = a(S_t^i)\,\mathrm{d}t + \mathrm{d}B_t^i + \mathsf{K}_t(S_t^i) \circ (\,\mathrm{d}Z_t - \frac{h(S_t^i) + \hat{h}_t}{2}\,\mathrm{d}t)$$

$$S_t^i \sim \mathsf{P}(X_t|\mathcal{Z}_t) \text{ (Exactness)}$$



$\{S^1, \ldots, S^N\}$

Uniqueness issue: There are infinitely many ways to construct S_t^i s.t exactness is satisfied

T. Yang, R. S. Laugesen, P. G. Mehta, and S. P. Meyn. Multivariable feedback particle filter, *Automatica*, 2015

Feedback Particle Filter (FPF)

Continuous time

Model:

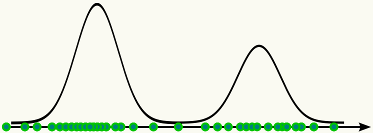
$$\mathrm{d}X_t = a(X_t)\,\mathrm{d}t + \mathrm{d}B_t$$
$$\mathrm{d}Z_t = h(X_t)\,\mathrm{d}t + \mathrm{d}W_t$$

Feedback Particle Filter:

$$\mathsf{P}(X_t|\mathcal{Z}_t) \approx \text{empirical dist. of } \{S^1,\ldots,S^N\},$$

$$\mathrm{d}S_t^i = a(S_t^i)\,\mathrm{d}t + \mathrm{d}B_t^i + \mathsf{K}_t(S_t^i)\circ(\mathrm{d}Z_t - \frac{h(S_t^i)+\hat{h}_t}{2}\,\mathrm{d}t)$$

$$S_t^i \sim \mathsf{P}(X_t|\mathcal{Z}_t) \text{ (Exactness)}$$



$\{S^1,\ldots,S^N\}$

Uniqueness issue: There are infinitely many ways to construct S_t^i s.t exactness is satisfied

T. Yang, R. S. Laugesen, P. G. Mehta, and S. P. Meyn. Multivariable feedback particle filter, *Automatica*, 2015

Feedback Particle Filter (FPF)

Continuous time

Model:

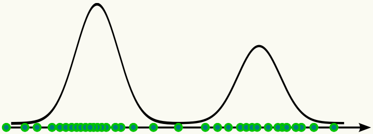
$$\mathrm{d}X_t = a(X_t)\,\mathrm{d}t + \mathrm{d}B_t$$
$$\mathrm{d}Z_t = h(X_t)\,\mathrm{d}t + \mathrm{d}W_t$$

Feedback Particle Filter:

$$\mathsf{P}(X_t|\mathcal{Z}_t) \approx \text{empirical dist. of } \{S^1,\ldots,S^N\},$$

$$\mathrm{d}S_t^i = a(S_t^i)\,\mathrm{d}t + \mathrm{d}B_t^i + \mathsf{K}_t(S_t^i)\circ(\,\mathrm{d}Z_t - \frac{h(S_t^i)+\hat{h}_t}{2}\,\mathrm{d}t)$$

$$S_t^i \sim \mathsf{P}(X_t|\mathcal{Z}_t) \text{ (Exactness)}$$



$\{S^1,\ldots,S^N\}$

Uniqueness issue: There are infinitely many ways to construct S_t^i s.t exactness is satisfied

T. Yang, R. S. Laugesen, P. G. Mehta, and S. P. Meyn. Multivariable feedback particle filter, *Automatica*, 2015

Model:

$$\mathrm{d}X_t = AX_t\,\mathrm{d}t + \mathrm{d}B_t, \quad X_0 \sim \mathcal{N}(\hat{X}_0, \Sigma_0),$$

$$\mathrm{d}Z_t = CX_t\,\mathrm{d}t + \mathrm{d}W_t,$$

Feedback Particle Filter:

$$\mathrm{d}S_t^i = AS_t^i\,\mathrm{d}t + \mathrm{d}B_t^i + \mathsf{K}_t(\,\mathrm{d}Z_t - \frac{CS_t^i + C\hat{S}_t}{2}\,\mathrm{d}t)$$

$\mathsf{K}_t := \Sigma_t C^T$ (Kalman Gain)

$\hat{S}_t$,Σ_t are mean and covariance of S_t^i

$S_t^i \sim N(\hat{X}_t, \Sigma_t)$ (Exactness)

- **Exactness:** The mean and covariance of S_t^i evolve according to Kalman filter equations
- **Non uniqueness:** For any skew symmetric matrix Ω_t, exactness is satisfied

Feedback Particle Filter: Linear Gaussian Case

Exactness and uniqueness

Model:

$$\mathrm{d}X_t = AX_t\,\mathrm{d}t + \mathrm{d}B_t, \quad X_0 \sim \mathcal{N}(\hat{X}_0, \Sigma_0),$$
$$\mathrm{d}Z_t = CX_t\,\mathrm{d}t + \mathrm{d}W_t,$$

Feedback Particle Filter:

$$\mathrm{d}S_t^i = AS_t^i\,\mathrm{d}t + \mathrm{d}B_t^i + \mathsf{K}_t(\,\mathrm{d}Z_t - \frac{CS_t^i + C\hat{S}_t}{2}\,\mathrm{d}t)$$

$\mathsf{K}_t := \Sigma_t C^T$ (Kalman Gain)

$\hat{S}_t$, Σ_t are mean and covariance of S_t^i

$$S_t^i \sim N(\hat{X}_t, \Sigma_t) \text{ (Exactness)}$$

- **Exactness:** The mean and covariance of S_t^i evolve according to Kalman filter equations
- **Non uniqueness:** For any skew symmetric matrix Ω_t, exactness is satisfied

Model:

$$dX_t = AX_t\,dt + dB_t, \quad X_0 \sim \mathcal{N}(\hat{X}_0, \Sigma_0),$$

$$dZ_t = CX_t\,dt + dW_t,$$

Feedback Particle Filter:

$$dS_t^i = AS_t^i\,dt + dB_t^i + \mathsf{K}_t(dZ_t - \frac{CS_t^i + C\hat{S}_t}{2}\,dt) + \Omega_t\Sigma_t^{-1}(S_t^i - \hat{S}_t)$$

$\mathsf{K}_t := \Sigma_t C^T$ (Kalman Gain)

$\hat{S}_t$,Σ_t are mean and covariance of S_t^i

$$S_t^i \sim N(\hat{X}_t, \Sigma_t) \text{ (Exactness)}$$

- **Exactness:** The mean and covariance of S_t^i evolve according to Kalman filter equations
- **Non uniqueness:** For any skew symmetric matrix Ω_t, exactness is satisfied

Objective of this talk

Model:

$$\mathrm{d}X_t = AX_t\,\mathrm{d}t + \mathrm{d}B_t, \quad X_0 \sim \mathcal{N}(\hat{X}_0, \Sigma_0),$$

$$\mathrm{d}Z_t = CX_t\,\mathrm{d}t + \mathrm{d}W_t,$$

Objective: Construct a unique process S_t^i s..t

$$S_t^i \sim N(\hat{X}_t, \Sigma_t)$$

where $\hat{X}_t$ and Σ_t are given by Kalman Filter.

Method: Optimal Transportation

Main idea: FPF is interpreted as transporting the prior to the posterior

Objective of this talk

Model:

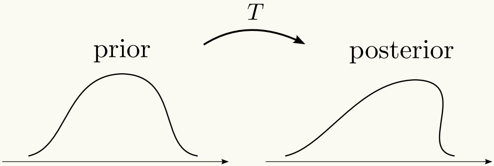
$$dX_t = AX_t\,dt + dB_t, \quad X_0 \sim \mathcal{N}(\hat{X}_0, \Sigma_0),$$
$$dZ_t = CX_t\,dt + dW_t,$$

Objective: Construct a unique process S_t^i s..t

$$S_t^i \sim N(\hat{X}_t, \Sigma_t)$$

where $\hat{X}_t$ and Σ_t are given by Kalman Filter.

Method: Optimal Transportation



Main idea: FPF is interpreted as transporting the prior to the posterior

Relevant Literature

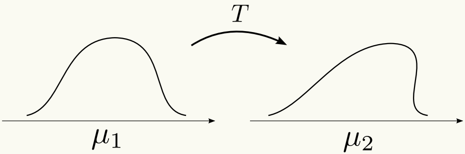
Control oriented particle filtering:

- S. K. Mitter, N. J. Newton, " *A variational approach to nonlinear estimation*", SIAM (2003).
- D. Crisan and J. Xiong. "*Approximate McKean-Vlasov representations for a class of SPDEs*", Stochastics (2009)
- F. Daum, J. Huang, "*Particle flow for nonlinear filters*", SPIE (2013).
- T. Yang, P. G. Mehta, S. P. Meyn. "*Feedback particle filter*", TAC (2013)
- K. Berntorp. "*Feedback particle filter: Application and evaluation*", FUSION (2015)

Optimal transportation for uncertainty propogation:

- Y. M. Marzouk, T. A. El Moselhi, "*Bayesian inference with optimal maps*", Journal of Computational Physics (2013)
- Y. Cheng and S. Reich. "*A mckean optimal transportation perspective on feynman-kac formulae with application to data assimilation*, (2013)
- F. Daum, J. Huang. "*Particle flow for nonlinear filters, Bayesian decisions and transport*". FUSION (2013)

Optimal Transportation
Quadratic cost



Problem: Given probability distributions μ_1 and μ_2:

$$\text{Minimize} \quad J(T) = \mathsf{E}\left[|T(X) - X|^2\right], \quad \text{over maps } T \text{ s.t}$$

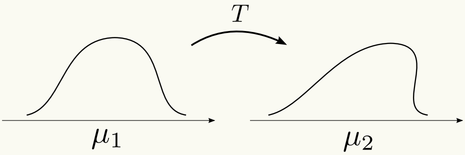
$$X \sim \mu_1, \quad T(X) \sim \mu_2$$

- If μ_1 is absoloutely continuous, there is a unique minimizer T^*
- The minimum value $J(T^*)$ is the Wasserstein distance between μ_1 and μ_2

C. Villani, Topics in optimal transportation. American Mathematical Soc., 2003

L. C. Evans, Partial differential equation and Monge-Kantorovich mass transfer, Current developments in mathematics, 1997

Optimal Transportation
Quadratic cost



Problem: Given probability distributions μ_1 and μ_2:

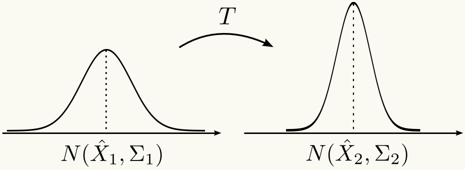
$$\text{Minimize} \quad J(T) = \mathsf{E}\left[|T(X) - X|^2\right], \quad \text{over maps } T \text{ s.t}$$

$$X \sim \mu_1, \quad T(X) \sim \mu_2$$

- If μ_1 is absoloutely continuous, there is a unique minimizer T^*
- The minimum value $J(T^*)$ is the Wasserstein distance between μ_1 and μ_2

C. Villani, Topics in optimal transportation. American Mathematical Soc., 2003
L. C. Evans, Partial differential equation and Monge-Kantorovich mass transfer, Current developments in mathematics, 1997

Optimal Transprtation: Gaussian case



Example: Optimal transport map between $N(\hat{X}_1, \Sigma_1)$ and $N(\hat{X}_2, \Sigma_2)$

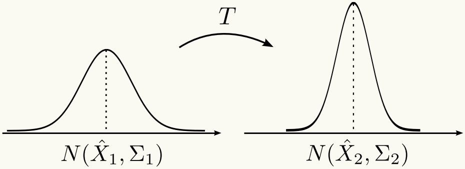
Scalar case : $$T^*(x) = \hat{X}_2 + \sqrt{\frac{\Sigma_2}{\Sigma_1}}(x - \hat{X}_1)$$

Vector case : $$T^*(x) = \hat{X}_2 + F(x - \hat{X}_1)$$

$$F = \Sigma_2^{\frac{1}{2}}(\Sigma_2^{\frac{1}{2}}\Sigma_1\Sigma_2^{\frac{1}{2}})^{-1}\Sigma_2^{\frac{1}{2}}$$

C. Givens, et al., A class of Wasserstein metrics for probability distributions. Michigan Math. J, 1984.

Optimal Transprtation: Gaussian case



Example: Optimal transport map between $N(\hat{X}_1, \Sigma_1)$ and $N(\hat{X}_2, \Sigma_2)$

Scalar case : $T^*(x) = \hat{X}_2 + \sqrt{\frac{\Sigma_2}{\Sigma_1}}(x - \hat{X}_1)$

Vector case : $T^*(x) = \hat{X}_2 + F(x - \hat{X}_1)$

$$F = \Sigma_2^{\frac{1}{2}}(\Sigma_2^{\frac{1}{2}}\Sigma_1\Sigma_2^{\frac{1}{2}})^{-1}\Sigma_2^{\frac{1}{2}}$$

C. Givens, et al., A class of Wasserstein metrics for probability distributions. Michigan Math. J, 1984.

Optimal Transport formulation of the FPF
Time Stepping Procedure

Objective: Construct a unique process S_t that satisfy exactness,

$$S_t \sim \mathsf{P}(X_t|\mathcal{Z}_t).$$

Procedure:

1. Divide the interval $[0, t_f]$ into n time steps.
2. Construct a discrete time process $\{S_0, S_1, \dots, S_n\}$,

$$S_{k+1} = T_k^*(S_k), \quad S_0 \sim \mathsf{P}(X_0)$$

   where T_k^* is the optimal map between $\mathsf{P}(X_{t_k}|\mathcal{Z}_{t_k})$ and $\mathsf{P}(X_{t_{k+1}}|\mathcal{Z}_{t_{k+1}})$.
3. Take the continuous time limit

Objective: Construct a unique process S_t that satisfy exactness,

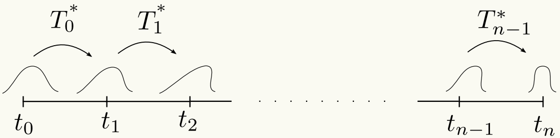
$$S_t \sim \mathsf{P}(X_t|\mathcal{Z}_t).$$

Procedure:

1. Divide the interval $[0, t_f]$ into n time steps.
2. Construct a discrete time process $\{S_0, S_1, \ldots, S_n\}$,

$$S_{k+1} = T_k^*(S_k), \quad S_0 \sim \mathsf{P}(X_0)$$

   where T_k^* is the optimal map between $\mathsf{P}(X_{t_k}|\mathcal{Z}_{t_k})$ and $\mathsf{P}(X_{t_{k+1}}|\mathcal{Z}_{t_{k+1}})$.
3. Take the continuous time limit



Objective: Construct a unique process S_t that satisfy exactness,

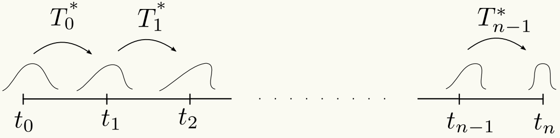
$$S_t \sim \mathsf{P}(X_t|\mathcal{Z}_t).$$

Procedure:

1. Divide the interval $[0, t_f]$ into n time steps.
2. Construct a discrete time process $\{S_0, S_1, \ldots, S_n\}$,

$$S_{k+1} = T_k^*(S_k), \quad S_0 \sim \mathsf{P}(X_0)$$

   where T_k^* is the optimal map between $\mathsf{P}(X_{t_k}|\mathcal{Z}_{t_k})$ and $\mathsf{P}(X_{t_{k+1}}|\mathcal{Z}_{t_{k+1}})$.
3. Take the continuous time limit



Objective: Construct a unique process S_t that satisfy exactness,

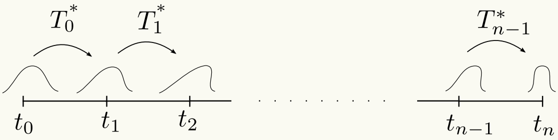
$$S_t \sim \mathsf{P}(X_t|\mathcal{Z}_t).$$

Procedure:

1. Divide the interval $[0, t_f]$ into n time steps.
2. Construct a discrete time process $\{S_0, S_1, \ldots, S_n\}$,

$$S_{k+1} = T_k^*(S_k), \quad S_0 \sim \mathsf{P}(X_0)$$

   where T_k^* is the optimal map between $\mathsf{P}(X_{t_k}|\mathcal{Z}_{t_k})$ and $\mathsf{P}(X_{t_{k+1}}|\mathcal{Z}_{t_{k+1}})$.
3. Take the continuous time limit



Result: Optimal Transport FPF in Linear Gaussian case (scalar)

Model:

$$\mathrm{d}X_t = aX_t\,\mathrm{d}t + \mathrm{d}B_t, \quad X_0 \sim \mathcal{N}(\hat{X}_0, \Sigma_0),$$

$$\mathrm{d}Z_t = cX_t\,\mathrm{d}t + \mathrm{d}W_t,$$

Time stepping procedure:

FPF: $\mathrm{d}S_t = aS_t\,\mathrm{d}t + \mathrm{d}B_t + \mathsf{K}_t(\mathrm{d}Z_t - \frac{cS_t + c\hat{S}_t}{2}\,\mathrm{d}t)$

Opt. FPF: $\mathrm{d}S_t = aS_t\,\mathrm{d}t + \frac{1}{2\Sigma_t}(S_t - \hat{S}_t)\,\mathrm{d}t + \mathsf{K}_t(\mathrm{d}Z_t - \frac{cS_t + c\hat{S}_t}{2}\,\mathrm{d}t)$

- The difference is the replacement of the stochastic term $\mathrm{d}B_t$ with a deterministic term.

Result: Optimal Transport FPF in Linear Gaussian case (scalar)

Model:

$$\mathrm{d}X_t = aX_t\,\mathrm{d}t + \mathrm{d}B_t, \quad X_0 \sim \mathcal{N}(\hat{X}_0, \Sigma_0),$$
$$\mathrm{d}Z_t = cX_t\,\mathrm{d}t + \mathrm{d}W_t,$$

Time stepping procedure:

FPF: $$\mathrm{d}S_t = aS_t\,\mathrm{d}t + \mathrm{d}B_t + \mathsf{K}_t(\mathrm{d}Z_t - \frac{cS_t + c\hat{S}_t}{2}\,\mathrm{d}t)$$

Opt. FPF: $$\mathrm{d}S_t = aS_t\,\mathrm{d}t + \frac{1}{2\Sigma_t}(S_t - \hat{S}_t)\,\mathrm{d}t + \mathsf{K}_t(\mathrm{d}Z_t - \frac{cS_t + c\hat{S}_t}{2}\,\mathrm{d}t)$$

- The difference is the replacement of the stochastic term $\mathrm{d}B_t$ with a deterministic term.

Result: Optimal Transport FPF in Linear Gaussian case (scalar)

Model:

$$\mathrm{d}X_t = aX_t\,\mathrm{d}t + \mathrm{d}B_t, \quad X_0 \sim \mathcal{N}(\hat{X}_0, \Sigma_0),$$
$$\mathrm{d}Z_t = cX_t\,\mathrm{d}t + \mathrm{d}W_t,$$

Time stepping procedure:

FPF: $$\mathrm{d}S_t = aS_t\,\mathrm{d}t + \mathrm{d}B_t + \mathsf{K}_t(\mathrm{d}Z_t - \frac{cS_t + c\hat{S}_t}{2}\,\mathrm{d}t)$$

Opt. FPF: $$\mathrm{d}S_t = aS_t\,\mathrm{d}t + \frac{1}{2\Sigma_t}(S_t - \hat{S}_t)\,\mathrm{d}t + \mathsf{K}_t(\mathrm{d}Z_t - \frac{cS_t + c\hat{S}_t}{2}\,\mathrm{d}t)$$

- The difference is the replacement of the stochastic term $\mathrm{d}B_t$ with a deterministic term.

Model:

$$\mathrm{d}X_t = aX_t\,\mathrm{d}t + \mathrm{d}B_t, \quad X_0 \sim \mathcal{N}(\hat{X}_0, \Sigma_0),$$
$$\mathrm{d}Z_t = cX_t\,\mathrm{d}t + \mathrm{d}W_t,$$

Time stepping procedure:

FPF: $$\mathrm{d}S_t = aS_t\,\mathrm{d}t + \mathrm{d}B_t + \mathsf{K}_t(\mathrm{d}Z_t - \frac{cS_t + c\hat{S}_t}{2}\,\mathrm{d}t)$$

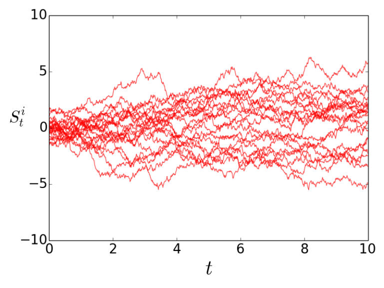
Opt. FPF: $$\mathrm{d}S_t = aS_t\,\mathrm{d}t + \frac{1}{2\Sigma_t}(S_t - \hat{S}_t)\,\mathrm{d}t + \mathsf{K}_t(\mathrm{d}Z_t - \frac{cS_t + c\hat{S}_t}{2}\,\mathrm{d}t)$$

- The difference is the replacement of the stochastic term $\mathrm{d}B_t$ with a deterministic term.

Numerical Example

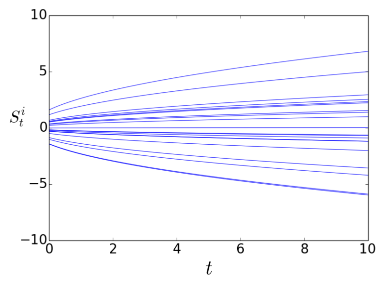
Monte Carlo

$$dS_t^i = dB_t^i,$$
$$S_0^i \sim N(0,1)$$



Optimal transport

$$dS_t^i = \frac{1}{2\Sigma_t}(S_t^i - \hat{S}_t)\,dt,$$
$$S_0^i \sim N(0,1)$$



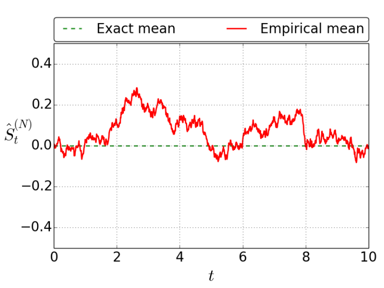
Particles trajectory in one simulation

$$S_t^i \sim N(0, 1+t)$$

Numerical Example

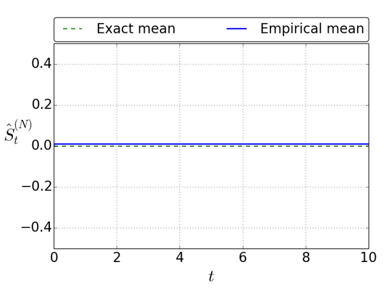
Monte Carlo

$$dS_t^i = dB_t^i,$$
$$S_0^i \sim N(0,1)$$



Optimal transport

$$dS_t^i = \frac{1}{2\Sigma_t}(S_t^i - \hat{S}_t)\,dt,$$
$$S_0^i \sim N(0,1)$$



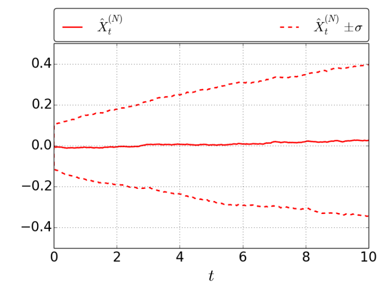
Empirical mean of particles in one simulation

$$\hat{S}_t^{(N)} := \frac{1}{N}\sum_{i=1}^{N} S_t^i$$

Numerical Example

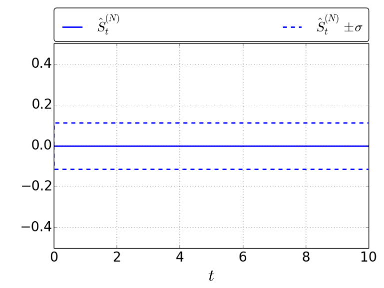
Monte Carlo

$$\mathrm{d}S_t^i = \mathrm{d}B_t^i,$$
$$S_0^i \sim N(0,1)$$



Optimal transport

$$\mathrm{d}S_t^i = \frac{1}{2\Sigma_t}(S_t^i - \hat{S}_t)\,\mathrm{d}t,$$
$$S_0^i \sim N(0,1)$$



Simulation variance as the number of particles vary

$$\mathrm{Var}(\hat{S}_t^{(N)}) = \frac{c}{N}$$

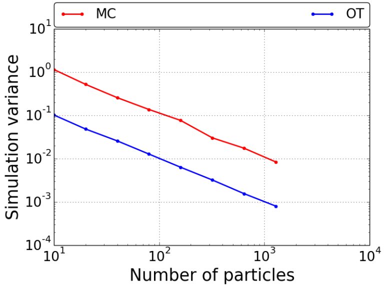
Numerical Example

Monte Carlo

$$\mathrm{d}S_t^i = \mathrm{d}B_t^i,$$
$$S_0^i \sim N(0,1)$$

Optimal transport

$$\mathrm{d}S_t^i = \frac{1}{2\Sigma_t}(S_t^i - \hat{S}_t)\,\mathrm{d}t,$$
$$S_0^i \sim N(0,1)$$



Decrease in simulation variance

Model:

$$\mathrm{d}X_t = AX_t\,\mathrm{d}t + \mathrm{d}B_t, \quad X_0 \sim \mathcal{N}(\hat{X}_0, \Sigma_0),$$
$$\mathrm{d}Z_t = CX_t\,\mathrm{d}t + \mathrm{d}W_t,$$

Time stepping procedure:

FPF: $\mathrm{d}S_t = AS_t\,\mathrm{d}t + \mathrm{d}\tilde{B}_t + \mathsf{K}_t(\mathrm{d}Z_t - \frac{CS_t + C\hat{S}_t}{2}\,\mathrm{d}t),$

Opt. FPF: $\mathrm{d}S_t = AS_t\,\mathrm{d}t + \frac{\Sigma_t^{-1}}{2}(S_t - \hat{S}_t)\,\mathrm{d}t + \mathsf{K}_t(\mathrm{d}Z_t - \frac{CS_t + C\hat{S}_t}{2}\,\mathrm{d}t) + \Omega_t\Sigma_t^{-1}(S_t - \hat{S}_t)\,\mathrm{d}t,$

- Ω_t is the (skew symmetric) solution to the matrix equation:

$$\Omega_t\Sigma_t^{-1} + \Sigma_t^{-1}\Omega_t = A^T - A + \frac{1}{2}(\mathsf{K}_t C - C^T\mathsf{K}_t^T)$$

Model:

$$\mathrm{d}X_t = AX_t\,\mathrm{d}t + \mathrm{d}B_t, \quad X_0 \sim \mathcal{N}(\hat{X}_0, \Sigma_0),$$
$$\mathrm{d}Z_t = CX_t\,\mathrm{d}t + \mathrm{d}W_t,$$

Time stepping procedure:

FPF: $\mathrm{d}S_t = AS_t\,\mathrm{d}t + \mathrm{d}\tilde{B}_t + \mathsf{K}_t(\mathrm{d}Z_t - \frac{CS_t + C\hat{S}_t}{2}\,\mathrm{d}t),$

Opt. FPF: $$\mathrm{d}S_t = AS_t\,\mathrm{d}t + \frac{\Sigma_t^{-1}}{2}(S_t - \hat{S}_t)\,\mathrm{d}t + \mathsf{K}_t(\mathrm{d}Z_t - \frac{CS_t + C\hat{S}_t}{2}\,\mathrm{d}t) + \Omega_t\Sigma_t^{-1}(S_t - \hat{S}_t)\,\mathrm{d}t,$$

- Ω_t is the (skew symmetric) solution to the matrix equation:

$$\Omega_t\Sigma_t^{-1} + \Sigma_t^{-1}\Omega_t = A^T - A + \frac{1}{2}(\mathsf{K}_tC - C^T\mathsf{K}_t^T)$$

Model:

$$\mathrm{d}X_t = AX_t\,\mathrm{d}t + \mathrm{d}B_t, \quad X_0 \sim \mathcal{N}(\hat{X}_0, \Sigma_0),$$
$$\mathrm{d}Z_t = CX_t\,\mathrm{d}t + \mathrm{d}W_t,$$

Time stepping procedure:

FPF: $$\mathrm{d}S_t = AS_t\,\mathrm{d}t + \mathrm{d}\tilde{B}_t + \mathsf{K}_t(\mathrm{d}Z_t - \frac{CS_t + C\hat{S}_t}{2}\,\mathrm{d}t),$$

Opt. FPF: $$\mathrm{d}S_t = AS_t\,\mathrm{d}t + \frac{\Sigma_t^{-1}}{2}(S_t - \hat{S}_t)\,\mathrm{d}t + \mathsf{K}_t(\mathrm{d}Z_t - \frac{CS_t + C\hat{S}_t}{2}\,\mathrm{d}t) + \Omega_t\Sigma_t^{-1}(S_t - \hat{S}_t)\,\mathrm{d}t,$$

- Ω_t is the (skew symmetric) solution to the matrix equation:

$$\Omega_t\Sigma_t^{-1} + \Sigma_t^{-1}\Omega_t = A^T - A + \frac{1}{2}(\mathsf{K}_tC - C^T\mathsf{K}_t^T)$$

Model:

$$\mathrm{d}X_t = AX_t\,\mathrm{d}t + \mathrm{d}B_t, \quad X_0 \sim \mathcal{N}(\hat{X}_0, \Sigma_0),$$
$$\mathrm{d}Z_t = CX_t\,\mathrm{d}t + \mathrm{d}W_t,$$

Time stepping procedure:

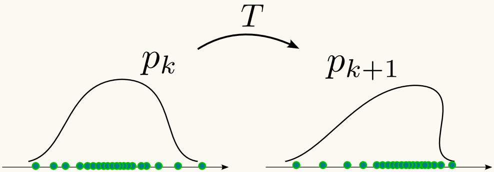
FPF: $$\mathrm{d}S_t = AS_t\,\mathrm{d}t + \mathrm{d}\tilde{B}_t + \mathsf{K}_t(\mathrm{d}Z_t - \frac{CS_t + C\hat{S}_t}{2}\,\mathrm{d}t),$$

Opt. FPF: $$\mathrm{d}S_t = AS_t\,\mathrm{d}t + \frac{\Sigma_t^{-1}}{2}(S_t - \hat{S}_t)\,\mathrm{d}t + \mathsf{K}_t(\mathrm{d}Z_t - \frac{CS_t + C\hat{S}_t}{2}\,\mathrm{d}t) + \Omega_t\Sigma_t^{-1}(S_t - \hat{S}_t)\,\mathrm{d}t,$$

- Ω_t is the (skew symmetric) solution to the matrix equation:

$$\Omega_t\Sigma_t^{-1} + \Sigma_t^{-1}\Omega_t = A^T - A + \frac{1}{2}(\mathsf{K}_tC - C^T\mathsf{K}_t^T)$$

Final Slide
Summary



Optimal transport formulation of FPF (Linear Gaussian setting):

- Connecting optimal transportation and continuous time filtering
- Resolve the uniqueness issue
- Reduce the simulation variance

Ongoing work:

- Extend the formulation to nonlinear setting
- Optimal transport for continuous-discrete time filtering

Exactness and uniqueness (back up)

Particles update: $\mathrm{d}S_t^i = A\hat{S}_t\,\mathrm{d}t + \mathsf{K}_t(\,\mathrm{d}Z_t - C\hat{S}_t\,\mathrm{d}t) \quad \rightarrow \quad$ mean ✓

$$+\,(A - \frac{1}{2}\mathsf{K}_t C + \Omega_t\Sigma_t^{-1})(S_t^i - \hat{S}_t)\,\mathrm{d}t + \,\mathrm{d}B_t \quad \rightarrow \quad \text{covariance ✓}$$

Lyapunov equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\Sigma_t = (A - \frac{1}{2}\mathsf{K}_t C + \Omega_t\Sigma_t^{-1})\Sigma_t + \Sigma_t(A^T - \frac{1}{2}C^T\mathsf{K}_t^T + \Sigma_t^{-1}\Omega_t^T) + I$$

$$= A\Sigma_t + \Sigma_t A^T - \frac{1}{2}\mathsf{K}_t C\Sigma_t - \frac{1}{2}\Sigma_t C^T\mathsf{K}_t^T + \Omega_t + \Omega_t^T + I$$

$$= A\Sigma_t + \Sigma_t A^T - \Sigma_t C^T C\Sigma_t + I$$

Finite N approximation

$$\hat{S}_t \approx \frac{1}{N}\sum_{i=1}^{N} S_t^i,$$

$$\Sigma_t \approx \frac{1}{N}\sum_{i=1}^{N}(S_t^i - \hat{S}_t)^2$$

Example 1:

$$\mathrm{d}S_t = ? \quad \text{s.t} \quad S_t \sim N(0, 1+t), \text{ with } X_0 \sim N(0,1)$$

Solution:

$$\mathrm{d}S_t = \mathrm{d}W_t$$

$\{W_t\}$ is standard Wiener process.

$$\mathrm{d}S_t = \frac{1}{2\Sigma_t} S_t \, \mathrm{d}t$$

Σ_t is variance of S_t.

Example 2:

$$\mathrm{d}S_t = ? \quad \text{s.t} \quad S_t \sim N(0, \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}), \text{ with } X_0 \sim N(0, \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix})$$

Solution:

$$\mathrm{d}S_t = \begin{bmatrix} 0 & 2 \\ -1 & 0 \end{bmatrix} S_t \, \mathrm{d}t$$

$$\mathrm{d}S_t = 0$$

Example 1:

$$\mathrm{d}S_t = ? \quad \text{s.t} \quad S_t \sim N(0, 1+t), \text{ with } X_0 \sim N(0,1)$$

Solution:

$$\mathrm{d}S_t = \mathrm{d}W_t$$

$\{W_t\}$ is standard Wiener process.

$$\mathrm{d}S_t = \frac{1}{2\Sigma_t} S_t \,\mathrm{d}t$$

Σ_t is variance of S_t.

Example 2:

$$\mathrm{d}S_t = ? \quad \text{s.t} \quad S_t \sim N(0, \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}), \text{ with } X_0 \sim N(0, \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix})$$

Solution:

$$\mathrm{d}S_t = \begin{bmatrix} 0 & 2 \\ -1 & 0 \end{bmatrix} S_t \,\mathrm{d}t$$

$$\mathrm{d}S_t = 0$$