Variational Optimal Transport Methods for Nonlinear Filtering

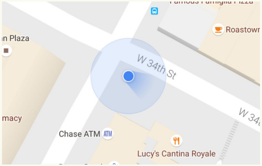
Presented at 7th Workshop on Cognition and Control
Univ. of Florida, Gainsville, Jan. 2024

Amirhossein Taghvaei

Department of Aeronautics & Astronautics
University of Washington, Seattle

Jan 26, 2024

Uncertainty is everywhere



Navigation

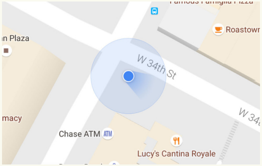
Weather forecast

COVID-19

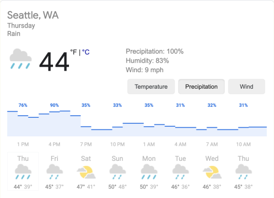
How to quantify uncertainty
and how to use data to reduce it

I. Stewart, Do Dice Play God? The Mathematics of Uncertainty. United States: Basic Books. 2019

Uncertainty is everywhere



Navigation



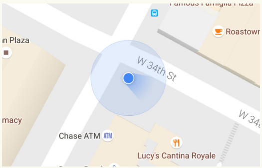
Weather forecast

COVID-19

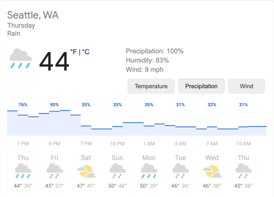
How to quantify uncertainty
and how to use data to reduce it

I. Stewart, Do Dice Play God? The Mathematics of Uncertainty. United States: Basic Books. 2019

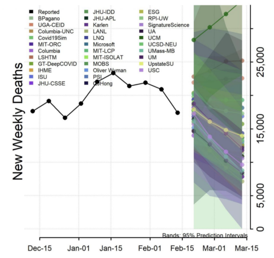
Uncertainty is everywhere



Navigation



Weather forecast

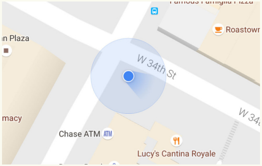


COVID-19

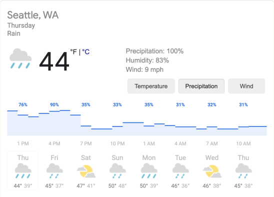
How to quantify uncertainty
and how to use data to reduce it

I. Stewart, Do Dice Play God? The Mathematics of Uncertainty. United States: Basic Books. 2019

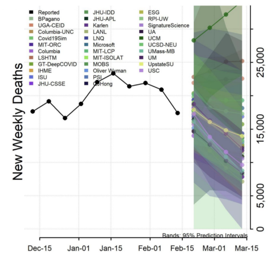
Uncertainty is everywhere



Navigation



Weather forecast



COVID-19

How to quantify uncertainty
and how to use data to reduce it

I. Stewart, Do Dice Play God? The Mathematics of Uncertainty. United States: Basic Books. 2019

Probability theory: (quantify uncertainty)

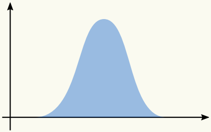
Optimal transport (OT) theory: (geometry for distributions)

This talk: application of OT to uncertainty quantification

Mathematics of uncertainty

Probability theory: (quantify uncertainty)

Probability distribution

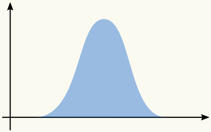


Optimal transport (OT) theory: (geometry for distributions)

This talk: application of OT to uncertainty quantification

Mathematics of uncertainty

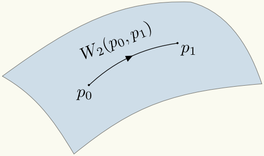
Probability theory: (quantify uncertainty)



Optimal transport (OT) theory: (geometry for distributions)

Nobel prize (1975)

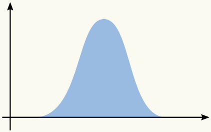
Fields medal (2010)



This talk: application of OT to uncertainty quantification

Mathematics of uncertainty

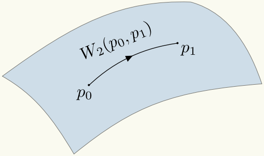
Probability theory: (quantify uncertainty)



Optimal transport (OT) theory: (geometry for distributions)

Nobel prize (1975)

Fields medal (2010)



This talk: application of OT to uncertainty quantification

Outline

- **Part I:** Bayes' law and importance sampling
- **Part II:** Conditioning with optimal transport maps
- **Part III:** Application to nonlinear filtering

Outline

- **Part I:** Bayes' law and importance sampling
- **Part II:** Conditioning with optimal transport maps
- **Part III:** Application to nonlinear filtering

Bayes' law

Problem:

- Hidden random variable X
- Observed random variable Y
- What is the conditional probability distribution of X given Y? (posterior)

Bayes' law: $P_{X|Y} = \frac{P_X P_{Y|X}}{P_Y}$

Simple to express, but difficult to implement, both intuitively and numerically

Bayes' law

Problem:

- Hidden random variable X
- Observed random variable Y
- What is the conditional probability distribution of X given Y? (posterior)

Bayes' law: $$P_{X|Y} = \frac{P_X P_{Y|X}}{P_Y}$$

Simple to express, but difficult to implement, both intuitively and numerically

Bayes' law

Problem:

- Hidden random variable X
- Observed random variable Y
- What is the conditional probability distribution of X given Y? (posterior)

Bayes' law: $$P_{X|Y} = \frac{P_X P_{Y|X}}{P_Y}$$

Simple to express, but difficult to implement, both intuitively and numerically

Two children puzzle:

- Smiths family has two children
- At least, one of them is a girl
- What is the probability that Smiths have two girls?
- What if you are told she is born on Tuesday?
- And her name is Florida.

I. Stewart, Do Dice Play God? The Mathematics of Uncertainty. United States: Basic Books. 2019

Two children puzzle:

- Smiths family has two children
- At least, one of them is a girl
- What is the probability that Smiths have two girls?
- What if you are told she is born on Tuesday?
- And her name is Florida.

I. Stewart, Do Dice Play God? The Mathematics of Uncertainty. United States: Basic Books. 2019

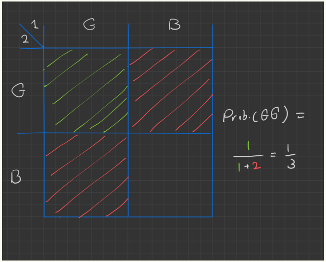
Two children puzzle:

- Smiths family has two children
- At least, one of them is a girl
- What is the probability that Smiths have two girls?
- What if you are told she is born on Tuesday?
- And her name is Florida.

I. Stewart, Do Dice Play God? The Mathematics of Uncertainty. United States: Basic Books. 2019

Two children puzzle:

- Smiths family has two children
- At least, one of them is a girl
- What is the probability that Smiths have two girls?
- What if you are told she is born on Tuesday?
- And her name is Florida.



I. Stewart, Do Dice Play God? The Mathematics of Uncertainty. United States: Basic Books. 2019

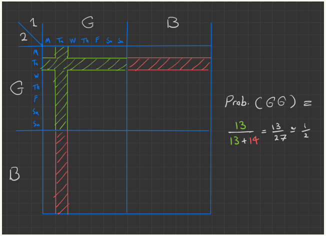
Two children puzzle:

- Smiths family has two children
- At least, one of them is a girl
- What is the probability that Smiths have two girls?
- What if you are told she is born on Tuesday?
- And her name is Florida.

I. Stewart, Do Dice Play God? The Mathematics of Uncertainty. United States: Basic Books. 2019

Two children puzzle:

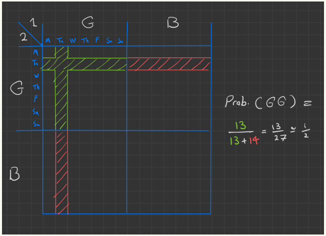
- Smiths family has two children
- At least, one of them is a girl
- What is the probability that Smiths have two girls?
- What if you are told she is born on Tuesday?
- And her name is Florida.



I. Stewart, Do Dice Play God? The Mathematics of Uncertainty. United States: Basic Books. 2019

Two children puzzle:

- Smiths family has two children
- At least, one of them is a girl
- What is the probability that Smiths have two girls?
- What if you are told she is born on Tuesday?
- And her name is Florida.

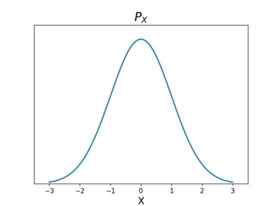


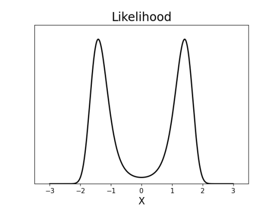
I. Stewart, Do Dice Play God? The Mathematics of Uncertainty. United States: Basic Books. 2019

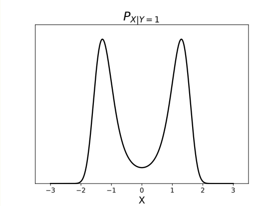
Example:

- $X \sim \mathcal{N}(0,1)$
- $Y = \frac{1}{2}X^2 + \epsilon W$
- $P_{X|Y=1} = ?$

Importance sampling (IS):







small noise regime: $\epsilon \to 0$

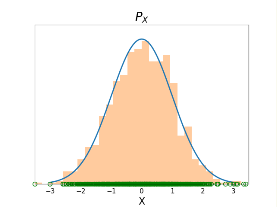
This is the main reason for the curse of dimensionality of IS-based particle filters

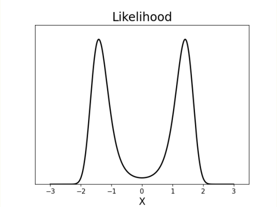
Example:

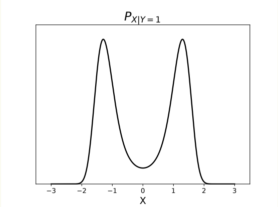
- $X \sim \mathcal{N}(0,1)$
- $Y = \frac{1}{2}X^2 + \epsilon W$
- $P_{X|Y=1} = ?$

Importance sampling (IS):

- $X^i \overset{\text{i.i.d}}{\sim} \mathcal{N}(0,1)$
- $w^i \propto P(Y=1|X^i)$
- $P_{X|Y=1} \approx \sum_{i=1}^{N} w^i \delta_{X^i}$







small noise regime: $\epsilon \to 0$

This is the main reason for the curse of dimensionality of IS-based particle filters

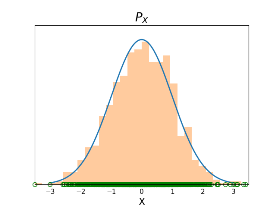
Bayes' law
When numerics fail

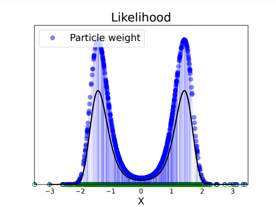
Example:

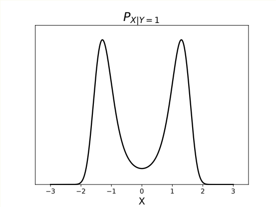
- $X \sim \mathcal{N}(0,1)$
- $Y = \frac{1}{2}X^2 + \epsilon W$
- $P_{X|Y=1} = ?$

Importance sampling (IS):

- $X^i \overset{\text{i.i.d}}{\sim} \mathcal{N}(0,1)$
- $w^i \propto P(Y=1|X^i)$
- $P_{X|Y=1} \approx \sum_{i=1}^{N} w^i \delta_{X^i}$







small noise regime: $\epsilon \to 0$

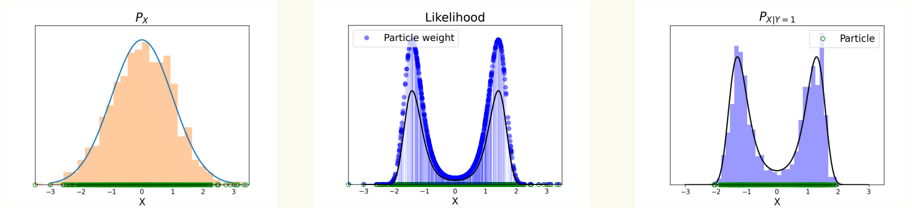
This is the main reason for the curse of dimensionality of IS-based particle filters

Example:

- $X \sim \mathcal{N}(0,1)$
- $Y = \frac{1}{2}X^2 + \epsilon W$
- $P_{X|Y=1} = ?$

Importance sampling (IS):

- $X^i \overset{\text{i.i.d}}{\sim} \mathcal{N}(0,1)$
- $w^i \propto P(Y=1|X^i)$
- $P_{X|Y=1} \approx \sum_{i=1}^{N} w^i \delta_{X^i}$



small noise regime: $\epsilon \to 0$

This is the main reason for the curse of dimensionality of IS-based particle filters

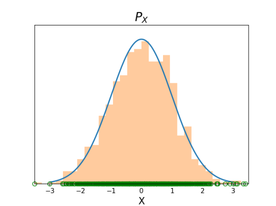
Bayes' law
When numerics fail

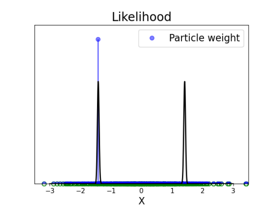
Example:

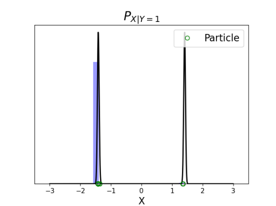
- $X \sim \mathcal{N}(0,1)$
- $Y = \frac{1}{2}X^2 + \epsilon W$
- $P_{X|Y=1} = ?$

Importance sampling (IS):

- $X^i \overset{\text{i.i.d}}{\sim} \mathcal{N}(0,1)$
- $w^i \propto P(Y=1|X^i)$
- $P_{X|Y=1} \approx \sum_{i=1}^{N} w^i \delta_{X^i}$







small noise regime: $\epsilon \to 0$

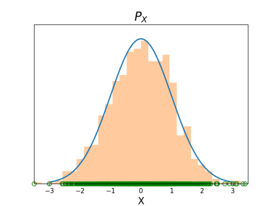
This is the main reason for the curse of dimensionality of IS-based particle filters

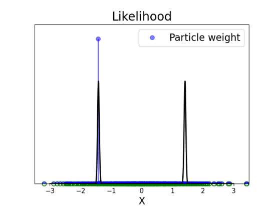
Example:

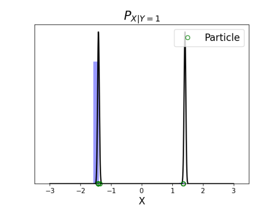
- $X \sim \mathcal{N}(0,1)$
- $Y = \frac{1}{2}X^2 + \epsilon W$
- $P_{X|Y=1} = ?$

Importance sampling (IS):

- $X^i \overset{\text{i.i.d}}{\sim} \mathcal{N}(0,1)$
- $w^i \propto P(Y=1|X^i)$
- $P_{X|Y=1} \approx \sum_{i=1}^{N} w^i \delta_{X^i}$







small noise regime: $\epsilon \to 0$

This is the main reason for the curse of dimensionality of IS-based particle filters

Curse of dimensionality in particle filters

- $X, Y \in \mathbb{R}^n$ with i.i.d. components.
- Exact posterior: π_{exact}
- IS approximation: $\pi_{\text{IS}}^{(N)}$
- Asymptotic limit as $N \to \infty$:

$$\lim_{N\to\infty} \sqrt{N} d(\pi_{\text{exact}}, \pi_{\text{IS}}^{(N)}) = C\gamma^n$$

  where $d(\cdot,\cdot)$ is the dual bounded metric.

- Good news: accurate as $N \to \infty$ (universal for any prior and likelihood)
- Bad news: error scales exponentially with the dimension n
- Remedy: exploit problem specific properties (e.g. spatial correlation decay in localization methods)
- Alternative method: replacing IS with control or coupling-based techniques

P. Del Moral, A.Guionnet. On the stability of interacting processes with applications to filtering and genetic algorithms. (2001)
P. Bickel, B. Li, and T. Bengtsson, Sharp failure rates for the bootstrap particle filter in high dimensions (2008).
P. Rebeschini and R. Van Handel, Can local particle filters beat the curse of dimensionality? The Annals of Applied Probability, (2015)

Curse of dimensionality in particle filters

- $X, Y \in \mathbb{R}^n$ with i.i.d. components.
- Exact posterior: π_{exact}
- IS approximation: $\pi_{\text{IS}}^{(N)}$
- Asymptotic limit as $N \to \infty$:

$$\lim_{N\to\infty} \sqrt{N} d(\pi_{\text{exact}}, \pi_{\text{IS}}^{(N)}) = C\gamma^n$$

where $d(\cdot,\cdot)$ is the dual bounded metric.

- Good news: accurate as $N \to \infty$ (universal for any prior and likelihood)
- Bad news: error scales exponentially with the dimension n
- Remedy: exploit problem specific properties (e.g. spatial correlation decay in localization methods)
- Alternative method: replacing IS with control or coupling-based techniques

P. Del Moral, A.Guionnet. On the stability of interacting processes with applications to filtering and genetic algorithms. (2001)
P. Bickel, B. Li, and T. Bengtsson, Sharp failure rates for the bootstrap particle filter in high dimensions (2008).
P. Rebeschini and R. Van Handel, Can local particle filters beat the curse of dimensionality? The Annals of Applied Probability, (2015)

Curse of dimensionality in particle filters

- $X, Y \in \mathbb{R}^n$ with i.i.d. components.
- Exact posterior: π_{exact}
- IS approximation: $\pi_{\text{IS}}^{(N)}$
- Asymptotic limit as $N \to \infty$:

$$\lim_{N\to\infty} \sqrt{N} d(\pi_{\text{exact}}, \pi_{\text{IS}}^{(N)}) = C\gamma^n$$

where $d(\cdot, \cdot)$ is the dual bounded metric.

- Good news: accurate as $N \to \infty$ (universal for any prior and likelihood)
- Bad news: error scales exponentially with the dimension n
- Remedy: exploit problem specific properties (e.g. spatial correlation decay in localization methods)
- Alternative method: replacing IS with control or coupling-based techniques

P. Del Moral, A.Guionnet. On the stability of interacting processes with applications to filtering and genetic algorithms. (2001)
P. Bickel, B. Li, and T. Bengtsson, Sharp failure rates for the bootstrap particle filter in high dimensions (2008).
P. Rebeschini and R. Van Handel, Can local particle filters beat the curse of dimensionality? The Annals of Applied Probability, (2015)

Curse of dimensionality in particle filters

- $X, Y \in \mathbb{R}^n$ with i.i.d. components.
- Exact posterior: π_{exact}
- IS approximation: $\pi_{\text{IS}}^{(N)}$
- Asymptotic limit as $N \to \infty$:

$$\lim_{N\to\infty} \sqrt{N} d(\pi_{\text{exact}}, \pi_{\text{IS}}^{(N)}) = C\gamma^n$$

  where $d(\cdot,\cdot)$ is the dual bounded metric.
- Good news: accurate as $N \to \infty$ (universal for any prior and likelihood)
- Bad news: error scales exponentially with the dimension n
- Remedy: exploit problem specific properties (e.g. spatial correlation decay in localization methods)
- Alternative method: replacing IS with control or coupling-based techniques

P. Del Moral, A.Guionnet. On the stability of interacting processes with applications to filtering and genetic algorithms. (2001)
P. Bickel, B. Li, and T. Bengtsson, Sharp failure rates for the bootstrap particle filter in high dimensions (2008).
P. Rebeschini and R. Van Handel, Can local particle filters beat the curse of dimensionality? The Annals of Applied Probability, (2015)

Curse of dimensionality in particle filters

- $X, Y \in \mathbb{R}^n$ with i.i.d. components.
- Exact posterior: π_{exact}
- IS approximation: $\pi_{\text{IS}}^{(N)}$
- Asymptotic limit as $N \to \infty$:

$$\lim_{N\to\infty} \sqrt{N} d(\pi_{\text{exact}}, \pi_{\text{IS}}^{(N)}) = C\gamma^n$$

  where $d(\cdot,\cdot)$ is the dual bounded metric.
- Good news: accurate as $N \to \infty$ (universal for any prior and likelihood)
- Bad news: error scales exponentially with the dimension n
- Remedy: exploit problem specific properties (e.g. spatial correlation decay in localization methods)
- Alternative method: replacing IS with control or coupling-based techniques

P. Del Moral, A.Guionnet. On the stability of interacting processes with applications to filtering and genetic algorithms. (2001)
P. Bickel, B. Li, and T. Bengtsson, Sharp failure rates for the bootstrap particle filter in high dimensions (2008).
P. Rebeschini and R. Van Handel, Can local particle filters beat the curse of dimensionality? The Annals of Applied Probability, (2015)

Curse of dimensionality in particle filters

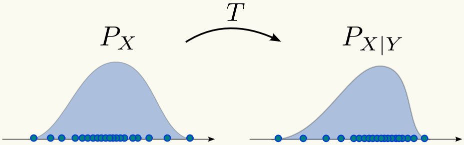
- $X, Y \in \mathbb{R}^n$ with i.i.d. components.
- Exact posterior: π_{exact}
- IS approximation: $\pi_{\text{IS}}^{(N)}$
- Asymptotic limit as $N \to \infty$:

$$\lim_{N\to\infty} \sqrt{N} d(\pi_{\text{exact}}, \pi_{\text{IS}}^{(N)}) = C\gamma^n$$

  where $d(\cdot,\cdot)$ is the dual bounded metric.
- Good news: accurate as $N \to \infty$ (universal for any prior and likelihood)
- Bad news: error scales exponentially with the dimension n
- Remedy: exploit problem specific properties (e.g. spatial correlation decay in localization methods)
- Alternative method: replacing IS with control or coupling-based techniques

P. Del Moral, A.Guionnet. On the stability of interacting processes with applications to filtering and genetic algorithms. (2001)
P. Bickel, B. Li, and T. Bengtsson, Sharp failure rates for the bootstrap particle filter in high dimensions (2008).
P. Rebeschini and R. Van Handel, Can local particle filters beat the curse of dimensionality? The Annals of Applied Probability, (2015)

Coupling/Control viewpoint

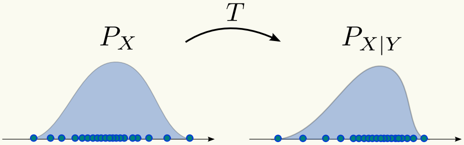


$$X^i \sim P_X \;\longrightarrow\; T(X^i, y) \sim P_{X|Y=y}$$

Example:

How to numerically find the map T in a general setting?

Coupling/Control viewpoint



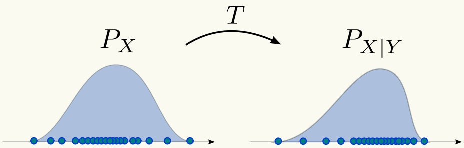
$$X^i \sim P_X \;\longrightarrow\; T(X^i, y) \sim P_{X|Y=y}$$

Example:

- Consider $Y = X$. Then, $P_{X|Y=y} = \delta_y$ is represented by the map $T(x, y) = y$
- Consider jointly Gaussian (X, Y). Then $P_{X|Y=y}$ is represented by the (stochastic) map $X \mapsto X + K(y - Y)$

How to numerically find the map T in a general setting?

Coupling/Control viewpoint



$$X^i \sim P_X \;\longrightarrow\; T(X^i, y) \sim P_{X|Y=y}$$

Example:

- Consider $Y = X$. Then, $P_{X|Y=y} = \delta_y$ is represented by the map $T(x, y) = y$
- Consider jointly Gaussian (X, Y). Then $P_{X|Y=y}$ is represented by the (stochastic) map $X \mapsto X + K(y - Y)$

How to numerically find the map T in a general setting?

- Particle flow filters [Daum et. al. 2010]
- A dynamical systems framework for data assimilation [Reich. 2011]
- Mean-field control approach [Yang, Mehta, Meyn, 2011]
  $\rightarrow$ Feedback Particle Filter (FPF)
- Posterior Matching via optimal transportation [Ma & Coleman, 2011]
- Bayesian inference with optimal maps [El Moselhy & Marzouk, 2012]

Recent surveys:

- Spantini et. al. (2022). Coupling techniques for nonlinear ensemble filtering. SIAM Review
- A. Taghvaei, and P. G. Mehta, (2023). A survey of feedback particle filter and related controlled interacting particle systems (CIPS). Annual Reviews in Control

This talk: Optimal Transport (OT) Method

- A. Taghvaei, and B. Hosseini, (2022). An optimal transport formulation of Bayes' law for nonlinear filtering algorithms. IEEE Conference on Decision and Control (CDC)
- M. Al-Jarrah, B. Hosseini, and A. Taghvaei (2023). Optimal transport particle filters. IEEE Conference on Decision and Control (CDC)

Literature survey

Control and coupling techniques for filtering and Bayesian inference

- Particle flow filters [Daum et. al. 2010]
- A dynamical systems framework for data assimilation [Reich. 2011]
- Mean-field control approach [Yang, Mehta, Meyn, 2011]
  $\rightarrow$ Feedback Particle Filter (FPF)
- Posterior Matching via optimal transportation [Ma & Coleman, 2011]
- Bayesian inference with optimal maps [El Moselhy & Marzouk, 2012]

Recent surveys:

- Spantini et. al. (2022). Coupling techniques for nonlinear ensemble filtering. SIAM Review
- A. Taghvaei, and P. G. Mehta, (2023). A survey of feedback particle filter and related controlled interacting particle systems (CIPS). Annual Reviews in Control

This talk: Optimal Transport (OT) Method

- A. Taghvaei, and B. Hosseini, (2022). An optimal transport formulation of Bayes' law for nonlinear filtering algorithms. IEEE Conference on Decision and Control (CDC)
- M. Al-Jarrah, B. Hosseini, and A. Taghvaei (2023). Optimal transport particle filters. IEEE Conference on Decision and Control (CDC)

- Particle flow filters [Daum et. al. 2010]
- A dynamical systems framework for data assimilation [Reich. 2011]
- Mean-field control approach [Yang, Mehta, Meyn, 2011]
  $\rightarrow$ Feedback Particle Filter (FPF)
- Posterior Matching via optimal transportation [Ma & Coleman, 2011]
- Bayesian inference with optimal maps [El Moselhy & Marzouk, 2012]

Recent surveys:

- Spantini et. al. (2022). Coupling techniques for nonlinear ensemble filtering. SIAM Review
- A. Taghvaei, and P. G. Mehta, (2023). A survey of feedback particle filter and related controlled interacting particle systems (CIPS). Annual Reviews in Control

This talk: Optimal Transport (OT) Method

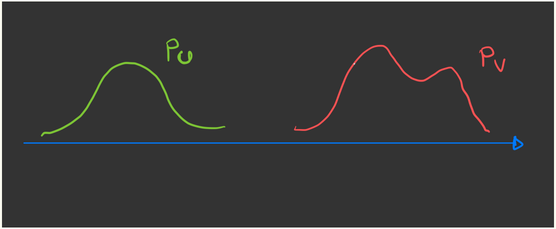
- A. Taghvaei, and B. Hosseini, (2022). An optimal transport formulation of Bayes' law for nonlinear filtering algorithms. IEEE Conference on Decision and Control (CDC)
- M. Al-Jarrah, B. Hosseini, and A. Taghvaei (2023). Optimal transport particle filters. IEEE Conference on Decision and Control (CDC)

Outline

- **Part I:** Bayes' law and importance sampling
- **Part II:** Conditioning with optimal transport maps
- **Part III:** Application to nonlinear filtering

Outline

- **Part I:** Bayes' law and importance sampling
- **Part II:** Conditioning with optimal transport maps
- **Part III:** Application to nonlinear filtering

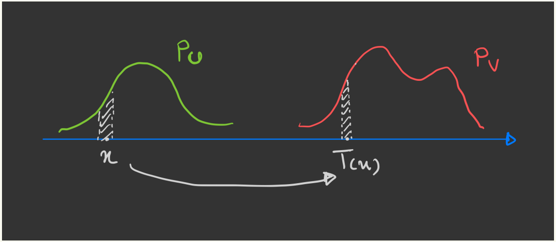


- Given two random variables $U \sim P_U$ and $V \sim P_V$
- find a map $x \mapsto T(x)$ that transports P_U to P_V, i.e. $T_{\#}P_U = P_V$ or $T(U) \stackrel{d}{=} V$
- with minimal transportation cost $\|T(x) - x\|^2$

Brenier's result

If P_U admits (Lebesgue) density, the optimal map $\overline{T} = \nabla \overline{f}$ where $\overline{f}$ minimizes

$$\min_{f \text{ is convex}} \mathbb{E}[f(U) + f^*(V)]$$

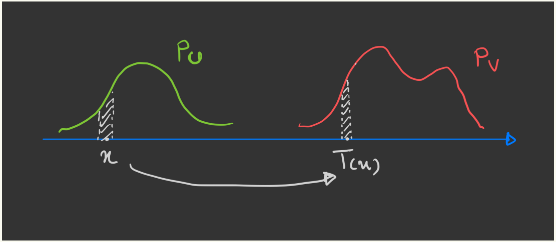


- Given two random variables $U \sim P_U$ and $V \sim P_V$
- find a map $x \mapsto T(x)$ that transports P_U to P_V, i.e. $T_{\#}P_U = P_V$ or $T(U) \stackrel{d}{=} V$
- with minimal transportation cost $\|T(x) - x\|^2$

Brenier's result

If P_U admits (Lebesgue) density, the optimal map $\overline{T} = \nabla \overline{f}$ where $\overline{f}$ minimizes

$$\min_{f \text{ is convex}} \mathbb{E}[f(U) + f^*(V)]$$

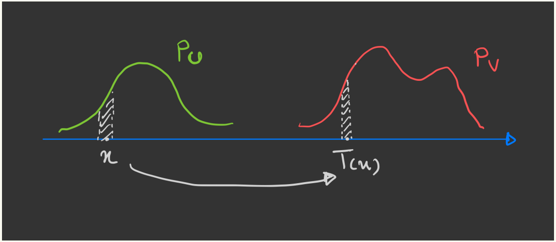


- Given two random variables $U \sim P_U$ and $V \sim P_V$
- find a map $x \mapsto T(x)$ that transports P_U to P_V, i.e. $T_{\#}P_U = P_V$ or $T(U) \overset{d}{=} V$
- with minimal transportation cost $\|T(x) - x\|^2$

Brenier's result

If P_U admits (Lebesgue) density, the optimal map $\overline{T} = \nabla \overline{f}$ where $\overline{f}$ minimizes

$$\min_{f \text{ is convex}} \mathbb{E}[f(U) + f^*(V)]$$



- Given two random variables $U \sim P_U$ and $V \sim P_V$
- find a map $x \mapsto T(x)$ that transports P_U to P_V, i.e. $T_\# P_U = P_V$ or $T(U) \overset{d}{=} V$
- with minimal transportation cost $\|T(x) - x\|^2$

Brenier's result

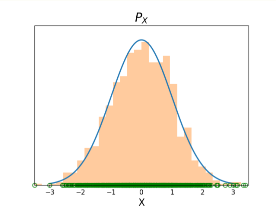
If P_U admits (Lebesgue) density, the optimal map $\overline{T} = \nabla \overline{f}$ where $\overline{f}$ minimizes

$$\min_{f \text{ is convex}} \mathbb{E}[f(U) + f^*(V)]$$

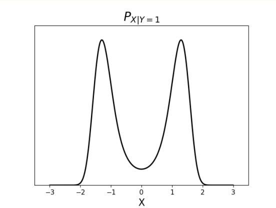
Conditioning with optimal transport map

Illustrative example

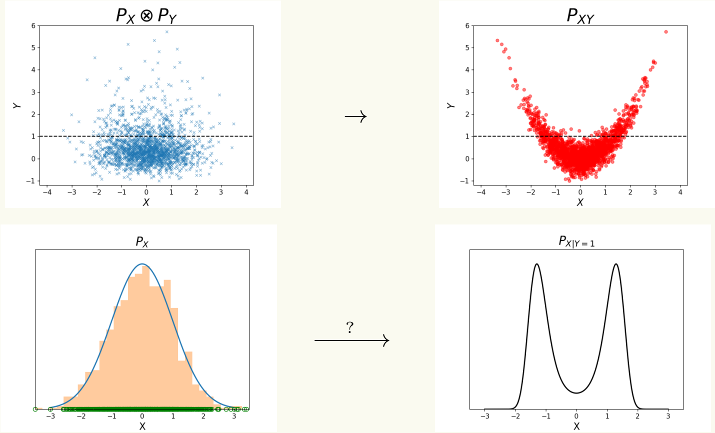
$\rightarrow$



$\xrightarrow{?}$

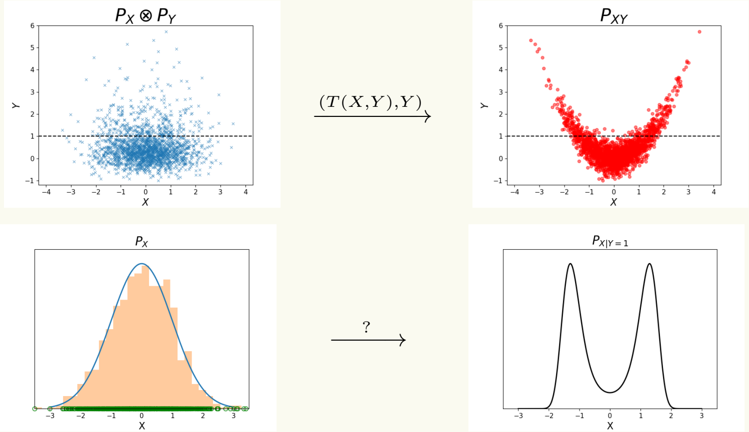


Conditioning with optimal transport map
Illustrative example



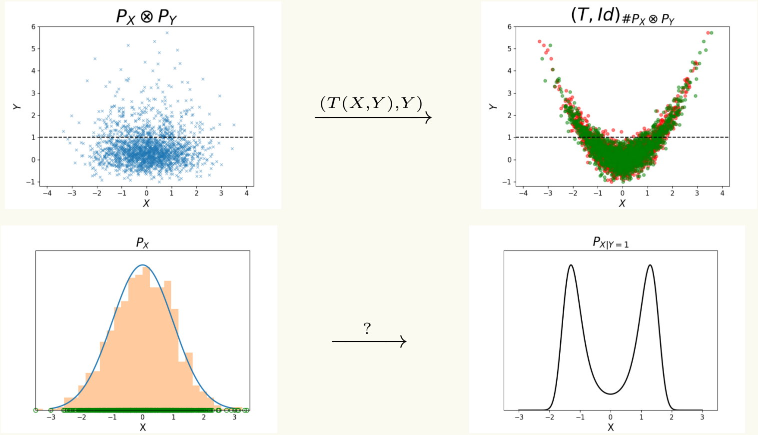
Conditioning with optimal transport map

Illustrative example

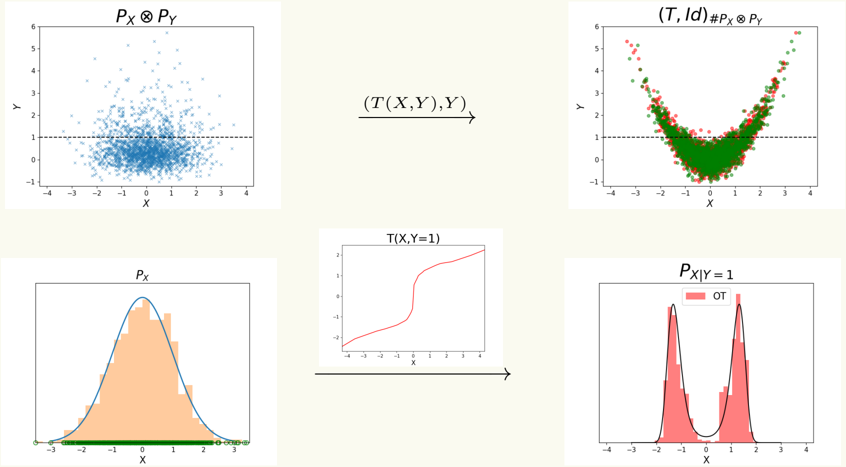


Conditioning with optimal transport map

Illustrative example

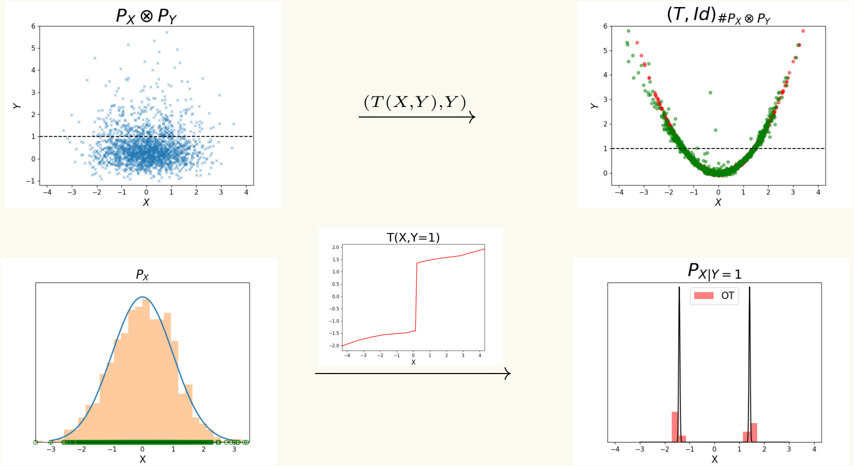


Conditioning with optimal transport map
Illustrative example



Conditioning with optimal transport map

Illustrative example



small noise limit

Bayes law: $$P_{X|Y} = \frac{P_X P_{Y|X}}{P_Y}$$

$$= \nabla_x \bar{f}(\cdot;Y) \# P_X$$

where $$\bar{f} = \arg\min_{f \in L^1(\mathcal{X}\times\mathcal{Y})} \mathbb{E}_{(X,Y)\sim P_X \otimes P_Y}[f(X;Y)] + \mathbb{E}_{(X,Y)\sim P_{XY}}[f^\star(X;Y)]$$

Computational properties:

- Only requires samples $(X_i, Y_i) \sim P_{XY}$ (data-driven/simulation based)
- Enables construction of “approximate” posterior distributions
- Allows application of ML tools (stochastic optimization and neural nets)

Conditioning with optimal transport map
Variational formulation of the Bayes' law

Bayes law:
$$P_{X|Y} = \frac{P_X P_{Y|X}}{P_Y}$$
$$= \nabla_x \bar{f}(\cdot; Y)\# P_X$$

where $\bar{f} = \underset{f \in L^1(\mathcal{X}\times\mathcal{Y})}{\arg\min} \; \mathbb{E}_{(X,Y)\sim P_X \otimes P_Y}[f(X;Y)] + \mathbb{E}_{(X,Y)\sim P_{XY}}[f^\star(X;Y)]$

Computational properties:

- Only requires samples $(X_i, Y_i) \sim P_{XY}$ (data-driven/simulation based)
- Enables construction of "approximate" posterior distributions
- Allows application of ML tools (stochastic optimization and neural nets)

Bayes law: $$P_{X|Y} = \frac{P_X P_{Y|X}}{P_Y}$$
$$= \nabla_x \bar{f}(\cdot;Y)\#P_X$$

where $$\bar{f} = \underset{f \in L^1(\mathcal{X}\times\mathcal{Y})}{\arg\min}\ \mathbb{E}_{(X,Y)\sim P_X \otimes P_Y}[f(X;Y)] + \mathbb{E}_{(X,Y)\sim P_{XY}}[f^\star(X;Y)]$$

Computational properties:

- Only requires samples $(X_i, Y_i) \sim P_{XY}$ (data-driven/simulation based)
- Enables construction of "approximate" posterior distributions
- Allows application of ML tools (stochastic optimization and neural nets)

Bayes law: $$P_{X|Y} = \frac{P_X P_{Y|X}}{P_Y} = \nabla_x \bar{f}(\cdot;Y)\#P_X$$

where $$\bar{f} = \underset{f \in L^1(\mathcal{X}\times\mathcal{Y})}{\arg\min}\ \mathbb{E}_{(X,Y)\sim P_X\otimes P_Y}[f(X;Y)] + \mathbb{E}_{(X,Y)\sim P_{XY}}[f^\star(X;Y)]$$

Computational properties:

- Only requires samples $(X_i, Y_i) \sim P_{XY}$ (data-driven/simulation based)
- Enables construction of "approximate" posterior distributions
- Allows application of ML tools (stochastic optimization and neural nets)

- $(X,Y) \sim P_{X,Y}$ and $(\bar{X},Y) \sim P_X \otimes P_Y$
- Variational problem: $\min_f J(f, P_{X,Y}) := \mathbb{E}[f(\bar{X},Y) + f^*(X,Y)]$

(Conditional) Brenier's theorem

- (Well-posedness) If P_X admits (Lebesgue) density, then, there exists a unique function $\overline{f}$ that solves the variational problem and

$$\nabla \overline{f}(\cdot, y) \# P_X = P_{X|Y=y}, \quad \text{a.e} \quad y$$

- (Sensitivity) Let f be a possibly non-optimal function. Assume $x \mapsto f(x,y)$ is convex and β-smooth for all y. Then,

$$d(\nabla f(\cdot, Y) \# P_X, P_{X|Y}) \leq \sqrt{2\beta \left(J(f) - J(\overline{f})\right)}.$$

Carlier, G., Chernozhukov, V., and Galichon, A. (2016). Vector quantile regression: an optimal trans- port approach. The Annals of Statistics,

- $(X, Y) \sim P_{X,Y}$ and $(\bar{X}, Y) \sim P_X \otimes P_Y$
- Variational problem: $\min_f J(f, P_{X,Y}) := \mathbb{E}[f(\bar{X}, Y) + f^*(X, Y)]$

(Conditional) Brenier's theorem

- (Well-posedness) If P_X admits (Lebesgue) density, then, there exists a unique function $\overline{f}$ that solves the variational problem and

$$\nabla \overline{f}(\cdot, y)\#P_X = P_{X|Y=y}, \quad \text{a.e} \quad y$$

- (Sensitivity) Let f be a possibly non-optimal function. Assume $x \mapsto f(x, y)$ is convex and β-smooth for all y. Then,

$$d(\nabla f(\cdot, Y)\#P_X, P_{X|Y}) \leq \sqrt{2\beta \left(J(f) - J(\overline{f})\right)}.$$

Carlier, G., Chernozhukov, V., and Galichon, A. (2016). Vector quantile regression: an optimal trans- port approach. The Annals of Statistics,

- $(X,Y) \sim P_{X,Y}$ and $(\bar{X},Y) \sim P_X \otimes P_Y$
- Variational problem: $\min_f J(f, P_{X,Y}) := \mathbb{E}[f(\bar{X},Y) + f^*(X,Y)]$

(Conditional) Brenier's theorem

- (Well-posedness) If P_X admits (Lebesgue) density, then, there exists a unique function $\overline{f}$ that solves the variational problem and

$$\nabla \overline{f}(\cdot, y)\#P_X = P_{X|Y=y}, \quad \text{a.e} \quad y$$

- (Sensitivity) Let f be a possibly non-optimal function. Assume $x \mapsto f(x,y)$ is convex and β-smooth for all y. Then,

$$d(\nabla f(\cdot, Y)\#P_X, P_{X|Y}) \leq \sqrt{2\beta\left(J(f) - J(\overline{f})\right)}.$$

Carlier, G., Chernozhukov, V., and Galichon, A. (2016). Vector quantile regression: an optimal trans- port approach. The Annals of Statistics,

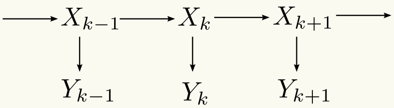
Outline

- **Part I:** Bayes' law and importance sampling
- **Part II:** Conditioning with optimal transport maps
- **Part III:** Application to nonlinear filtering

Outline

- **Part I:** Bayes' law and importance sampling
- **Part II:** Conditioning with optimal transport maps
- **Part III:** Application to nonlinear filtering

Nonlinear filtering problem



- X_t is the state (unknown)
- Y_t is the observation

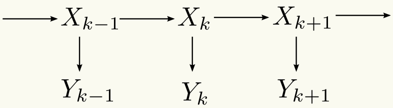
Questions: Given history of observation $Y_{1:t} := \{Y_1, \ldots, Y_t\}$,

- What is the most likely value of X_t?
- What is the probability of $X_t \in A$?
- What is the best m.s.e estimate for X_t?
- ...

Answer: given by the conditional distribution $\pi_t = P_{X_t|Y_{1:t}}$ (posterior, belief)

Nonlinear filtering: numerical approximation of the posterior π_t for all t.

Nonlinear filtering problem



- X_t is the state (unknown)
- Y_t is the observation

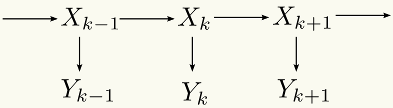
Questions: Given history of observation $Y_{1:t} := \{Y_1, \ldots, Y_t\}$,

- What is the most likely value of X_t?
- What is the probability of $X_t \in A$?
- What is the best m.s.e estimate for X_t?
- ...

Answer: given by the conditional distribution $\pi_t = P_{X_t|Y_{1:t}}$ (posterior, belief)

Nonlinear filtering: numerical approximation of the posterior π_t for all t.

Nonlinear filtering problem



- X_t is the state (unknown)
- Y_t is the observation

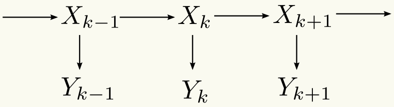
Questions: Given history of observation $Y_{1:t} := \{Y_1, \ldots, Y_t\}$,

- What is the most likely value of X_t?
- What is the probability of $X_t \in A$?
- What is the best m.s.e estimate for X_t?
- ...

Answer: given by the conditional distribution $\pi_t = P_{X_t|Y_{1:t}}$ (posterior, belief)

Nonlinear filtering: numerical approximation of the posterior π_t for all t.

Nonlinear filtering problem



- X_t is the state (unknown)
- Y_t is the observation

Questions: Given history of observation $Y_{1:t} := \{Y_1, \ldots, Y_t\}$,

- What is the most likely value of X_t?
- What is the probability of $X_t \in A$?
- What is the best m.s.e estimate for X_t?
- ...

Answer: given by the conditional distribution $\pi_t = P_{X_t|Y_{1:t}}$ (posterior, belief)

Nonlinear filtering: numerical approximation of the posterior π_t for all t.

Filtering equations

- $\pi_t := \mathsf{P}(X_t | Y_{1:t})$
- Two important operations:

$$\text{Propagation:} \quad \pi \xrightarrow{\text{dynamics}} \mathcal{A}\pi$$

$$\text{Conditioning:} \quad \pi \xrightarrow{\text{Bayes law}} B_y(\pi)$$

- Recursive update law for the posterior

$$\pi_{t-1} \xrightarrow{\text{dynamics}} \mathcal{A}\pi_{t-1} \xrightarrow{\text{Bayes law}} B_{Y_t}(\mathcal{A}\pi_{t-1}) =: \mathcal{T}_{t,t-1}(\pi_{t-1})$$

- (Exponential) filter stability : $\exists \lambda \in (0,1)$ s.t.

$$d(\mathcal{T}_{t,0}(\pi_0), \mathcal{T}_{t,0}(\tilde{\pi}_0)) \leq C\lambda^k d(\pi_0, \tilde{\pi}_0), \quad \forall \pi_0, \tilde{\pi}_0.$$

Filtering equations

- $\pi_t := \mathsf{P}(X_t | Y_{1:t})$
- Two important operations:

$$\text{Propagation:} \quad \pi \xrightarrow{\text{dynamics}} \mathcal{A}\pi$$

$$\text{Conditioning:} \quad \pi \xrightarrow{\text{Bayes law}} B_y(\pi)$$

- Recursive update law for the posterior

$$\pi_{t-1} \xrightarrow{\text{dynamics}} \mathcal{A}\pi_{t-1} \xrightarrow{\text{Bayes law}} B_{Y_t}(\mathcal{A}\pi_{t-1}) =: \mathcal{T}_{t,t-1}(\pi_{t-1})$$

- (Exponential) filter stability : $\exists \lambda \in (0,1)$ s.t.

$$d(\mathcal{T}_{t,0}(\pi_0), \mathcal{T}_{t,0}(\tilde{\pi}_0)) \leq C\lambda^k d(\pi_0, \tilde{\pi}_0), \quad \forall \pi_0, \tilde{\pi}_0.$$

Filtering equations

- $\pi_t := \mathsf{P}(X_t | Y_{1:t})$
- Two important operations:

$$\text{Propagation:} \quad \pi \xrightarrow{\text{dynamics}} \mathcal{A}\pi$$

$$\text{Conditioning:} \quad \pi \xrightarrow{\text{Bayes law}} B_y(\pi)$$

- Recursive update law for the posterior

$$\pi_{t-1} \xrightarrow{\text{dynamics}} \mathcal{A}\pi_{t-1} \xrightarrow{\text{Bayes law}} B_{Y_t}(\mathcal{A}\pi_{t-1}) =: \mathcal{T}_{t,t-1}(\pi_{t-1})$$

- (Exponential) filter stability : $\exists \lambda \in (0,1)$ s.t.

$$d(\mathcal{T}_{t,0}(\pi_0), \mathcal{T}_{t,0}(\tilde{\pi}_0)) \leq C\lambda^k d(\pi_0, \tilde{\pi}_0), \quad \forall \pi_0, \tilde{\pi}_0.$$

Filtering equations

- $\pi_t := \mathsf{P}(X_t | Y_{1:t})$
- Two important operations:

$$\text{Propagation:} \quad \pi \xrightarrow{\text{dynamics}} \mathcal{A}\pi$$

$$\text{Conditioning:} \quad \pi \xrightarrow{\text{Bayes law}} B_y(\pi)$$

- Recursive update law for the posterior

$$\pi_{t-1} \xrightarrow{\text{dynamics}} \mathcal{A}\pi_{t-1} \xrightarrow{\text{Bayes law}} B_{Y_t}(\mathcal{A}\pi_{t-1}) =: \mathcal{T}_{t,t-1}(\pi_{t-1})$$

- (Exponential) filter stability : $\exists \lambda \in (0,1)$ s.t.

$$d(\mathcal{T}_{t,0}(\pi_0), \mathcal{T}_{t,0}(\tilde{\pi}_0)) \leq C\lambda^k d(\pi_0, \tilde{\pi}_0), \quad \forall \pi_0, \tilde{\pi}_0.$$

Optimal Transport Filter
No dynamics setting (for simplicity)

Filter design steps:

exact posterior: $\pi_t = \mathcal{B}_{Y_t}(\pi_{t-1})$

mean-field process: $\bar{X}_t = \nabla \overline{f}_t(\bar{X}_{t-1}, Y_t)$

particle system: $X_t^i = \nabla \hat{f}_t(X_{t-1}^i, Y_t)$

Variational problem:

$$= \arg\min_{f} J(f; P_{X_t, Y_t})$$

$$\hat{f}_t = \arg\min_{f \in \mathcal{F}} J(f; \frac{1}{N}\sum_{i=1}^{N} \delta_{(X_t^i, Y_t^i)})$$

Posterior approximation:

$$\pi_t \approx \hat{\pi}_t^{(N)} = \frac{1}{N}\sum_{i=1}^{N} \delta_{X_t^i}$$

Optimal Transport Filter
No dynamics setting (for simplicity)

Filter design steps:

exact posterior: $\pi_t = \nabla \overline{f}_t(\cdot, Y_t)_{\#}\pi_{t-1}$

mean-field process: $\bar{X}_t = \nabla \overline{f}_t(\bar{X}_{t-1}, Y_t)$

particle system: $X_t^i = \nabla \hat{f}_t(X_{t-1}^i, Y_t)$

Variational problem:

$$\overline{f}_t = \arg\min_f \; J(f, P_{X_t, Y_t})$$

$$\hat{f}_t = \arg\min_{f \in \mathcal{F}} \; J(f, \frac{1}{N}\sum_{i=1}^N \delta_{(X_t^i, Y_t^i)})$$

Posterior approximation:

$$\pi_t \approx \hat{\pi}_t^{(N)} = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}$$

Optimal Transport Filter
No dynamics setting (for simplicity)

Filter design steps:

$$\text{exact posterior:} \quad \pi_t = \nabla \overline{f}_t(\cdot, Y_t)_{\#}\pi_{t-1}$$

$$\text{mean-field process:} \quad \bar{X}_t = \nabla \overline{f}_t(\bar{X}_{t-1}, Y_t)$$

$$\text{particle system:} \quad X_t^i = \nabla \hat{f}_t(X_{t-1}^i, Y_t)$$

Variational problem:

$$\overline{f}_t = \arg\min_{f} J(f, P_{X_t,Y_t})$$

$$\hat{f}_t = \arg\min_{f\in\mathcal{F}} J(f, \frac{1}{N}\sum_{i=1}^N \delta_{(X_t^i, Y_t^i)})$$

Posterior approximation:

$$\pi_t \approx \hat{\pi}_t^{(N)} = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}$$

Optimal Transport Filter

No dynamics setting (for simplicity)

Filter design steps:

$$\text{exact posterior:} \quad \pi_t = \nabla \overline{f}_t(\cdot, Y_t)_{\#}\pi_{t-1}$$

$$\text{mean-field process:} \quad \bar{X}_t = \nabla \overline{f}_t(\bar{X}_{t-1}, Y_t)$$

$$\text{particle system:} \quad X_t^i = \nabla \hat{f}_t(X_{t-1}^i, Y_t)$$

Variational problem:

$$\overline{f}_t = \arg\min_f \; J(f, P_{X_t, Y_t})$$

$$\hat{f}_t = \arg\min_{f \in \mathcal{F}} \; J(f, \frac{1}{N}\sum_{i=1}^N \delta_{(X_t^i, Y_t^i)})$$

Posterior approximation:

$$\pi_t \approx \hat{\pi}_t^{(N)} = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}$$

Theorem

Assume

1. The exact filter is exponentially stable
2. Uniform bound $\epsilon_{\mathcal{F},N}$ on the optimality gap $J(\hat{f}_t) - J(\overline{f}_t)$
3. The function $\hat{f}_t(\cdot, y)$ is convex and β-smooth for all t and y.
4. Particles are resampled at each step

Then,

$$d(\frac{1}{N}\sum_{i=1}^{N}\delta_{X_t^i}, \pi_t) \leq C\left(\sqrt{2\beta\epsilon_{\mathcal{F},N}} + \frac{1}{\sqrt{N}}\right), \quad \forall t.$$

- Optimality gap $\epsilon_{\mathcal{F},N}$ has bias-variance decomposition

$$\epsilon_{\mathcal{F},N} \leq \underbrace{\epsilon_{\mathcal{F}}}_{\text{approx. theory}} + \underbrace{\frac{C_{\mathcal{F}}}{\sqrt{N}}}_{\text{statistical generalization}}$$

Theorem

Assume

1. The exact filter is exponentially stable
2. Uniform bound $\epsilon_{\mathcal{F},N}$ on the optimality gap $J(\hat{f}_t) - J(\overline{f}_t)$
3. The function $\hat{f}_t(\cdot, y)$ is convex and β-smooth for all t and y.
4. Particles are resampled at each step

Then,

$$d(\frac{1}{N}\sum_{i=1}^{N}\delta_{X_t^i}, \pi_t) \leq C\left(\sqrt{2\beta\epsilon_{\mathcal{F},N}} + \frac{1}{\sqrt{N}}\right), \quad \forall t.$$

- Optimality gap $\epsilon_{\mathcal{F},N}$ has bias-variance decomposition

$$\epsilon_{\mathcal{F},N} \leq \underbrace{\epsilon_{\mathcal{F}}}_{\text{approx. theory}} + \underbrace{\frac{C_{\mathcal{F}}}{\sqrt{N}}}_{\text{statistical generalization}}$$

Theorem

Assume

1. The exact filter is exponentially stable
2. Uniform bound $\epsilon_{\mathcal{F},N}$ on the optimality gap $J(\hat{f}_t) - J(\overline{f}_t)$
3. The function $\hat{f}_t(\cdot, y)$ is convex and β-smooth for all t and y.
4. Particles are resampled at each step

Then,

$$d(\frac{1}{N}\sum_{i=1}^{N}\delta_{X_t^i}, \pi_t) \leq C\left(\sqrt{2\beta\epsilon_{\mathcal{F},N}} + \frac{1}{\sqrt{N}}\right), \quad \forall t.$$

- Optimality gap $\epsilon_{\mathcal{F},N}$ has bias-variance decomposition

$$\epsilon_{\mathcal{F},N} \leq \underbrace{\epsilon_{\mathcal{F}}}_{\text{approx. theory}} + \underbrace{\frac{C_{\mathcal{F}}}{\sqrt{N}}}_{\text{statistical generalization}}$$

Theorem

Assume

1. The exact filter is exponentially stable
2. Uniform bound $\epsilon_{\mathcal{F},N}$ on the optimality gap $J(\hat{f}_t) - J(\overline{f}_t)$
3. The function $\hat{f}_t(\cdot, y)$ is convex and β-smooth for all t and y.
4. Particles are resampled at each step

Then,

$$d(\frac{1}{N}\sum_{i=1}^{N}\delta_{X_t^i}, \pi_t) \leq C\left(\sqrt{2\beta\epsilon_{\mathcal{F},N}} + \frac{1}{\sqrt{N}}\right), \quad \forall t.$$

- Optimality gap $\epsilon_{\mathcal{F},N}$ has bias-variance decomposition

$$\epsilon_{\mathcal{F},N} \leq \underbrace{\epsilon_{\mathcal{F}}}_{\text{approx. theory}} + \underbrace{\frac{C_{\mathcal{F}}}{\sqrt{N}}}_{\text{statistical generalization}}$$

Theorem

Assume

1. The exact filter is exponentially stable
2. Uniform bound $\epsilon_{\mathcal{F},N}$ on the optimality gap $J(\hat{f}_t) - J(\overline{f}_t)$
3. The function $\hat{f}_t(\cdot, y)$ is convex and β-smooth for all t and y.
4. Particles are resampled at each step

Then,

$$d(\frac{1}{N}\sum_{i=1}^{N}\delta_{X_t^i}, \pi_t) \leq C\left(\sqrt{2\beta\epsilon_{\mathcal{F},N}} + \frac{1}{\sqrt{N}}\right), \quad \forall t.$$

- Optimality gap $\epsilon_{\mathcal{F},N}$ has bias-variance decomposition

$$\epsilon_{\mathcal{F},N} \leq \underbrace{\epsilon_{\mathcal{F}}}_{\text{approx. theory}} + \underbrace{\frac{C_{\mathcal{F}}}{\sqrt{N}}}_{\text{statistical generalization}}$$

Theorem

Assume

1. The exact filter is exponentially stable
2. Uniform bound $\epsilon_{\mathcal{F},N}$ on the optimality gap $J(\hat{f}_t) - J(\overline{f}_t)$
3. The function $\hat{f}_t(\cdot, y)$ is convex and β-smooth for all t and y.
4. Particles are resampled at each step

Then,

$$d(\frac{1}{N}\sum_{i=1}^{N}\delta_{X_t^i}, \pi_t) \leq C\left(\sqrt{2\beta\epsilon_{\mathcal{F},N}} + \frac{1}{\sqrt{N}}\right), \quad \forall t.$$

- Optimality gap $\epsilon_{\mathcal{F},N}$ has bias-variance decomposition

$$\epsilon_{\mathcal{F},N} \leq \underbrace{\epsilon_{\mathcal{F}}}_{\text{approx. theory}} + \underbrace{\frac{C_{\mathcal{F}}}{\sqrt{N}}}_{\text{statistical generalization}}$$

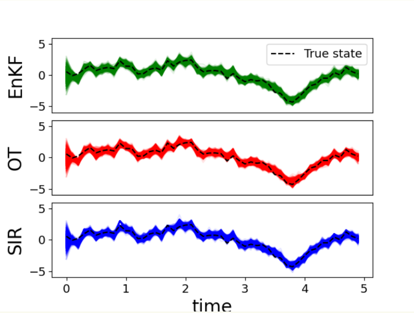
Optimal Transport Filter
Numerical example

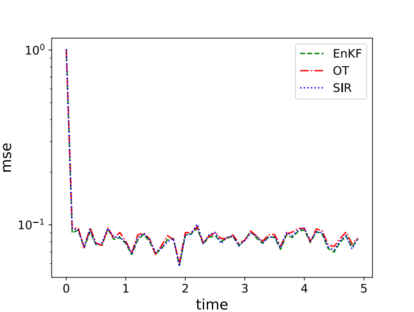
$$X_t = (1-\alpha)X_{t-1} + \sigma_V V_t, \quad X_0 \sim \mathcal{N}(0, I_n),$$
$$Y_t = X_t + \sigma_W W_t,$$

- Ensemble Kalman filter (EnKF)
- sequential importance re-sampling (SIR)
- Optimal Transport (OT)

Optimal Transport Filter
Numerical example

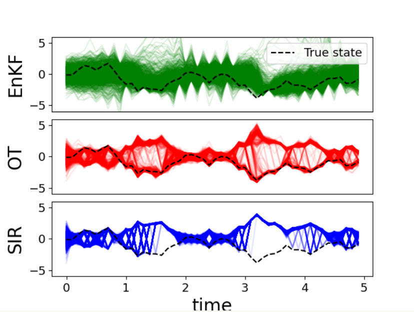
$$X_t = (1-\alpha)X_{t-1} + \sigma_V V_t, \quad X_0 \sim \mathcal{N}(0, I_n),$$
$$Y_t = X_t + \sigma_W W_t,$$

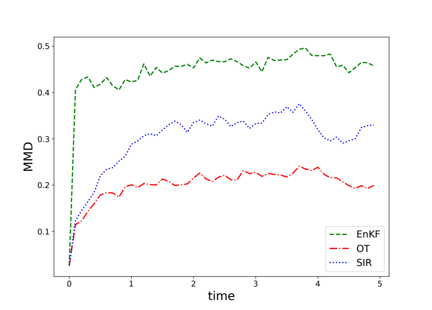




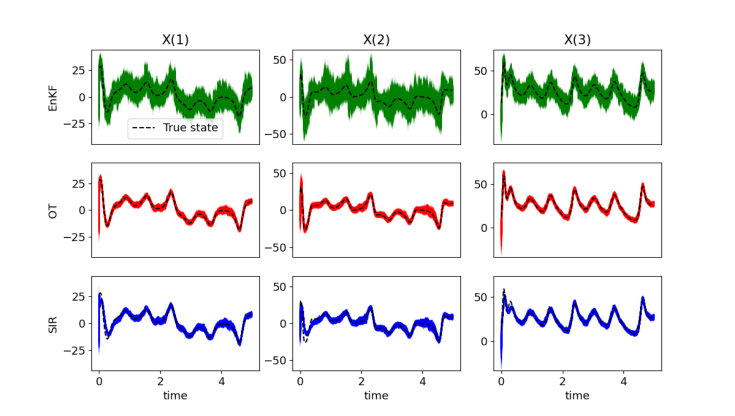
- Ensemble Kalman filter (EnKF)
- sequential importance re-sampling (SIR)
- Optimal Transport (OT)

$$X_t = (1-\alpha)X_{t-1} + \sigma_V V_t, \quad X_0 \sim \mathcal{N}(0, I_n),$$
$$Y_t = X_t^2 + \sigma_W W_t,$$



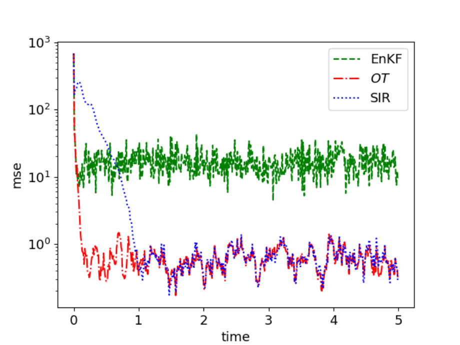


- Ensemble Kalman filter (EnKF)
- sequential importance re-sampling (SIR)
- Optimal Transport (OT)



- Trajectory of the particles
- mean-squared error (mse) in estimating the state

Optimal Transport Filter
Numerical example: Lorenz 63



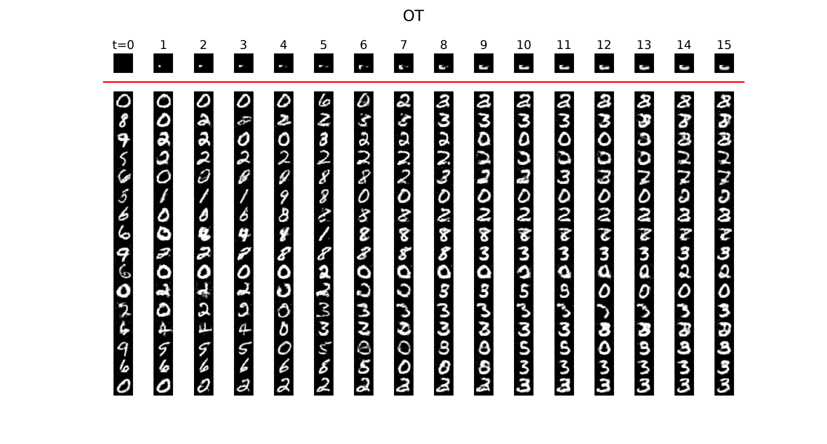
- Trajectory of the particles
- mean-squared error (mse) in estimating the state

Numerical example: Image in-painting

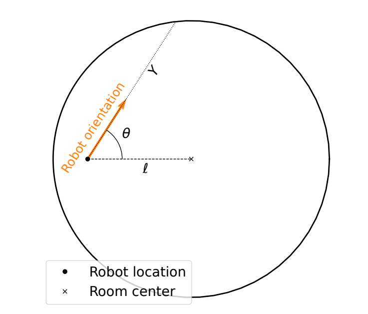
$$X \sim N(0, I_{100}),$$
$$Y_t = h(G(X), c_t) + W_t,$$
$$G : \mathbb{R}^{100} \to \mathbb{R}^{28\times 28} \text{(pre-trained generator)}$$

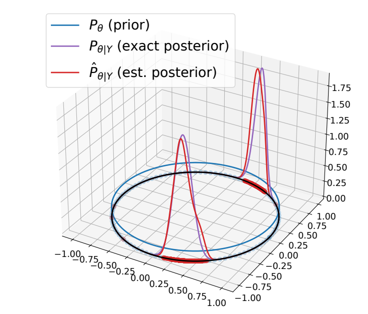
Numerical example: Image in-painting

$$X \sim N(0, I_{100}),$$
$$Y_t = h(G(X), c_t) + W_t,$$
$$G : \mathbb{R}^{100} \to \mathbb{R}^{28\times 28} \text{(pre-trained generator)}$$



Numerical example: Attitude estimation

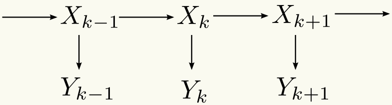




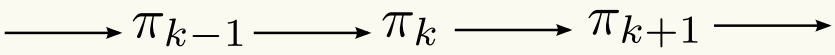
D. Grange, M. Al-Jarrah, R. Baptista, A. Taghvaei, T. Georgiou, S. Phillips, A. Tannenbaum, Computational optimal transport and filtering on Riemannian manifolds, IEEE Control Systems Letters, 2023

Summary

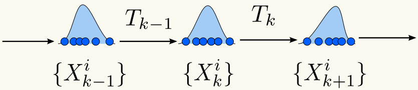
- **Mathematical model:**



- **Nonlinear filtering:** compute the posterior $\pi_k = \mathsf{P}(X_k | Y_{1:k})$



- **OT approach:**



- **Variational problem:**

$$T_k = \nabla_x \bar{f}_k, \quad \text{where} \quad \bar{f}_k = \arg\min_{f \in \mathcal{F}} J^{(N)}(f; \{(X^i_k, Y^i_k)\})$$

References

- A. Taghvaei, B. Hosseini, An optimal transport formulation of Bayes' law for nonlinear filtering algorithms, IEEE Conference on Decision and Control (CDC), 2022
- M. Al-Jarrah, B. Hosseini, A. Taghvaei, Optimal Transport Particle Filters, IEEE Conference on Decision and Control (CDC), 2023
- M. Al-Jarrah, N. Jin, B. Hosseini, A. Taghvaei, Optimal Transport-based Nonlinear Filtering in High-dimensional Settings, arXiv:2310.13886
- B. Hosseini, AW. Hsu, A. Taghvaei, Conditional Optimal Transport on Function Spaces, arXiv:2311.05672
- D. Grange, M. Al-Jarrah, R. Baptista, A. Taghvaei, T. Georgiou, S. Phillips, A. Tannenbaum, Computational optimal transport and filtering on Riemannian manifolds, IEEE Control Systems Letters, 2023
- D. Grange, R. Baptista, A. Taghvaei, A. Tannenbaum, S. Phillips, Distributed Nonlinear Filtering using Triangular Transport Maps, IEEE American Control Conference (ACC), 2024

Acknowledgments

Mohammad Al-Jarrah

Jenny Jin

Michele Martino

Bamdad Hosseini

Allen Tannenbaum

NSF