Gradient-based dimension reduction for solving Bayesian inverse problems

Ricardo Baptista[1]

Joint work with Youssef Marzouk[2] and Olivier Zahm[3]

[1]Computing+Mathematical Sciences
California Institute of Technology

[2]Center for Computational Science and Engineering
Massachusetts Institute of Technology

[3]INRIA and Université Grenoble Alpes

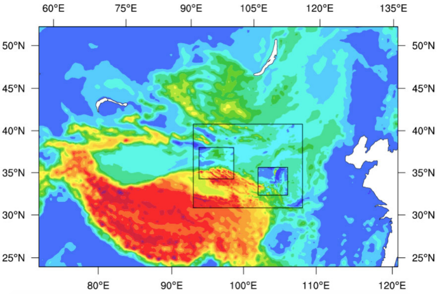
October 25, 2022

Goal: Solve Bayesian inference problems at scale

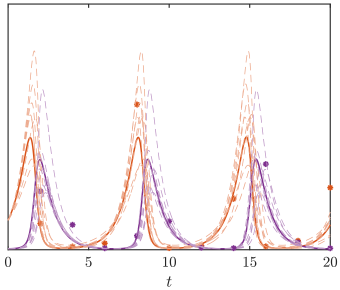
- Characterize posterior distribution of parameters $\mathbf{X}$ given data $\mathbf{Y}$

$$\pi_{\mathbf{X}|\mathbf{Y}} \propto \pi_{\mathbf{Y}|\mathbf{X}} \pi_{\mathbf{X}}$$

- **Applications**: inverse problems and data assimilation in geophysics, pharmacology, materials science, medical imaging, etc.



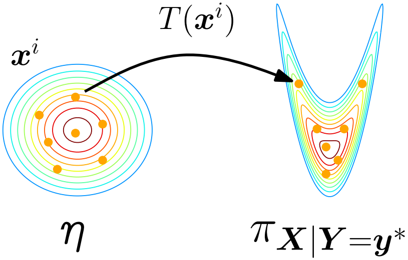
Wind forecasting [Source: NCAR]



Inference of population dynamics

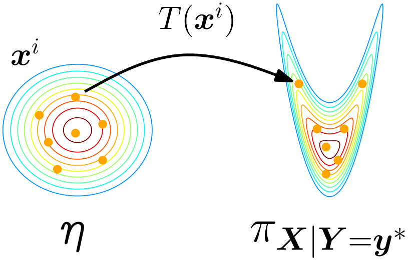
One approach: Characterize posterior using transport maps

Idea: Find map T that pushes forward reference distribution η (e.g., standard Normal) to posterior $\pi_{\mathbf{X}|\mathbf{Y}}$



One approach: Characterize posterior using transport maps

Idea: Find map T that pushes forward reference distribution η (e.g., standard Normal) to posterior $\pi_{\mathbf{X}|\mathbf{Y}}$



Advantages of invertible map:

1. Generate cheap and independent samples $\mathbf{x}^i \sim \eta \Leftrightarrow T_{\mathbf{y}^*}(\mathbf{x}^i) \sim \pi_{\mathbf{X}|\mathbf{y}^*}$
2. Evaluate the posterior density $\pi_{\mathbf{X}|\mathbf{y}^*}(\mathbf{x}) = \eta \circ T_{\mathbf{y}^*}^{-1}(\mathbf{x}) |\nabla T_{\mathbf{y}^*}^{-1}(\mathbf{x})|$

Block-triangular maps enable conditional sampling

Consider the map pushing forward $\eta_{\mathbf{Z}_1,\mathbf{Z}_2}$ to $\pi_{\mathbf{Y},\mathbf{X}} = \pi_{\mathbf{Y}}\pi_{\mathbf{X}|\mathbf{Y}}$:

$$T(\mathbf{y},\mathbf{x}) = \begin{bmatrix} T^{\mathcal{Y}}(\mathbf{y}) \\ T^{\mathcal{X}}(\mathbf{y},\mathbf{x}) \end{bmatrix}$$

- $T^{\mathcal{Y}}$ pushes forward $\eta_{\mathbf{Z}_1}$ to $\pi_{\mathbf{Y}}$
- $T^{\mathcal{X}}(\mathbf{y},\cdot)$ pushes forward $\eta_{\mathbf{Z}_2}$ to $\pi_{\mathbf{X}|\mathbf{y}}$ for any $\mathbf{y}$

Block-triangular maps enable conditional sampling

Consider the map pushing forward $\eta_{\mathbf{Z}_1,\mathbf{Z}_2}$ to $\pi_{\mathbf{Y},\mathbf{X}} = \pi_{\mathbf{Y}}\pi_{\mathbf{X}|\mathbf{Y}}$:

$$T(\mathbf{y},\mathbf{x}) = \begin{bmatrix} T^{\mathcal{Y}}(\mathbf{y}) \\ T^{\mathcal{X}}(\mathbf{y},\mathbf{x}) \end{bmatrix}$$

- $T^{\mathcal{Y}}$ pushes forward $\eta_{\mathbf{Z}_1}$ to $\pi_{\mathbf{Y}}$
- $T^{\mathcal{X}}(\mathbf{y},\cdot)$ pushes forward $\eta_{\mathbf{Z}_2}$ to $\pi_{\mathbf{X}|\mathbf{y}}$ for any $\mathbf{y}$

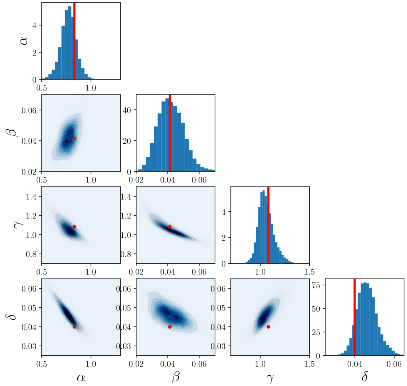
Recipe for amortized inference:
To characterize posterior $\pi_{\mathbf{X}|\mathbf{y}^*} \propto \pi_{\mathbf{y}^*|\mathbf{X}}\pi_{\mathbf{X}}$ given an observation $\mathbf{y}^*$:

- Simulate from the prior and likelihood model: $\mathbf{x}^i \sim \pi_{\mathbf{X}}$, $\mathbf{y}^i \sim \pi_{\mathbf{Y}|\mathbf{x}^i}$
- Estimate transport map $T^{\mathcal{X}}$ from joint samples $(\mathbf{x}^i,\mathbf{y}^i) \sim \pi_{\mathbf{X},\mathbf{Y}}$
- Simulate $\mathbf{x}^i = \widehat{T}^{\mathcal{X}}(\mathbf{y}^*,\mathbf{z}^i)$ for $\mathbf{z}^i \sim \eta_{\mathbf{Z}_2}$

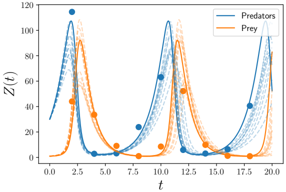
Related Work: Papamakarios & Murray, 2016; Lueckmann et al., 2017; Greenberg et al., 2019

Example: ODE parameter inference [Kovachki, B, et al., 2022]

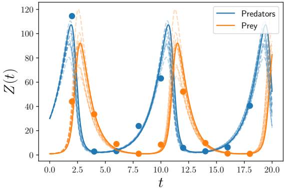
- Infer four parameters in Lotka–Volterra ODE with log-normal prior
- Observation: Noisy populations of two species at 9 times
- Inference is tractable without likelihood or prior evaluations



$\mathbf{x}|\mathbf{y}^*$ samples

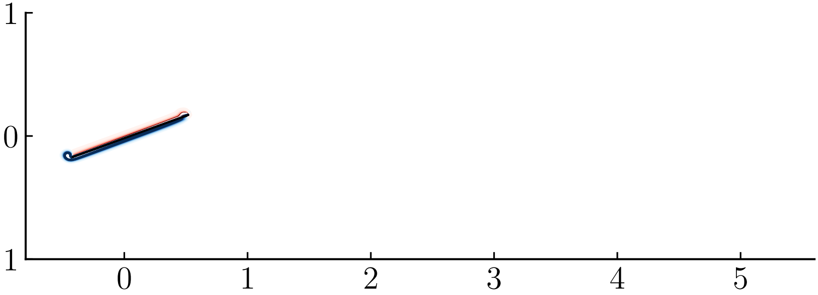


Transport predictive distribution



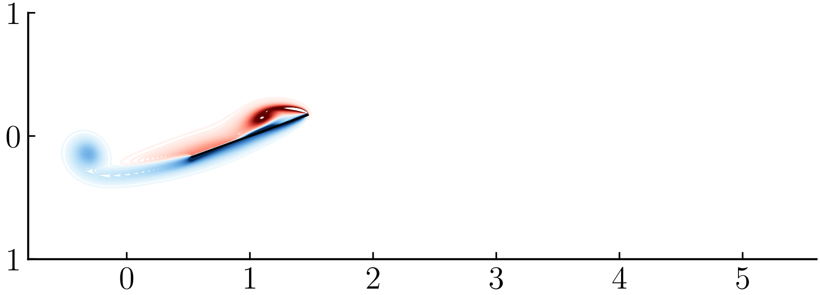
MCMC predictive distribution

Motivation: Estimating turbulent flow [Le Provost, B, et al., 2022]



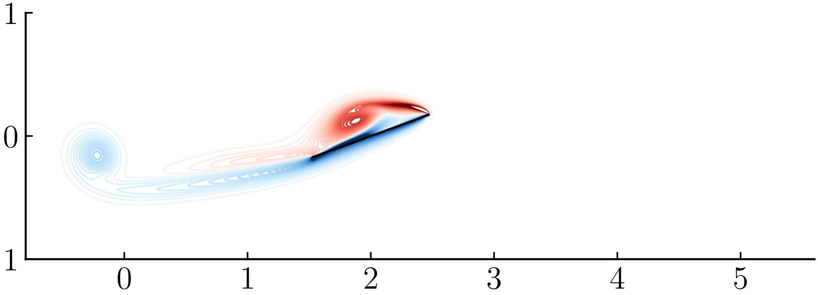
Vortex shedding around an aircraft wing

Motivation: Estimating turbulent flow [Le Provost, B, et al., 2022]



Vortex shedding around an aircraft wing

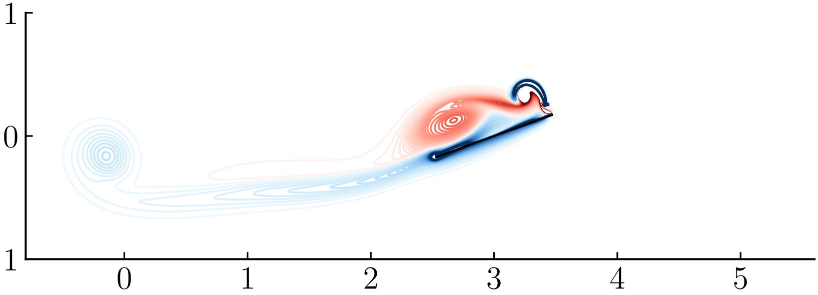
Motivation: Estimating turbulent flow [Le Provost, B, et al., 2022]



Vortex shedding around an aircraft wing

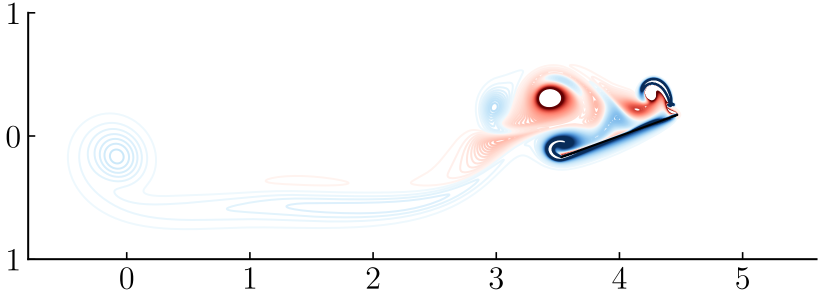
Tackling high-dimensional inverse problems

Motivation: Estimating turbulent flow [Le Provost, B, et al., 2022]



Vortex shedding around an aircraft wing

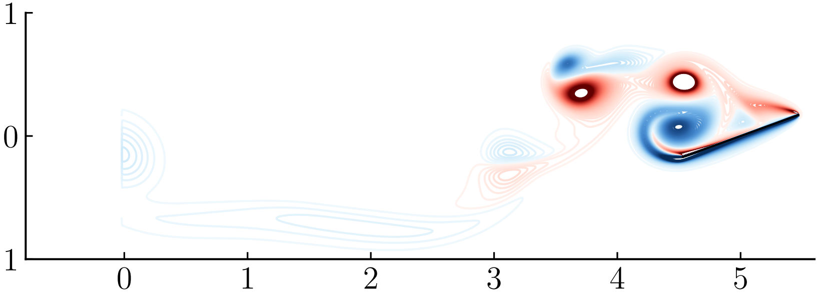
Motivation: Estimating turbulent flow [Le Provost, B, et al., 2022]



Vortex shedding around an aircraft wing

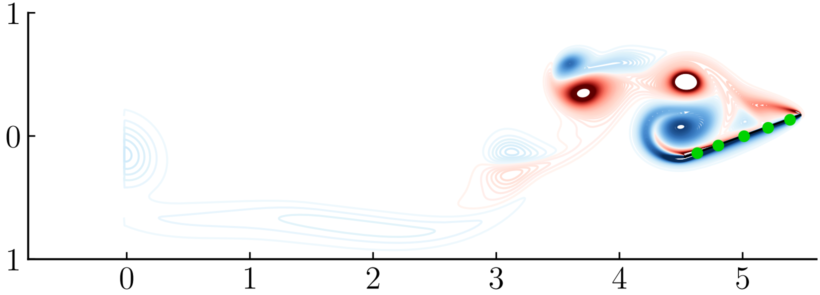
Tackling high-dimensional inverse problems

Motivation: Estimating turbulent flow [Le Provost, B, et al., 2022]



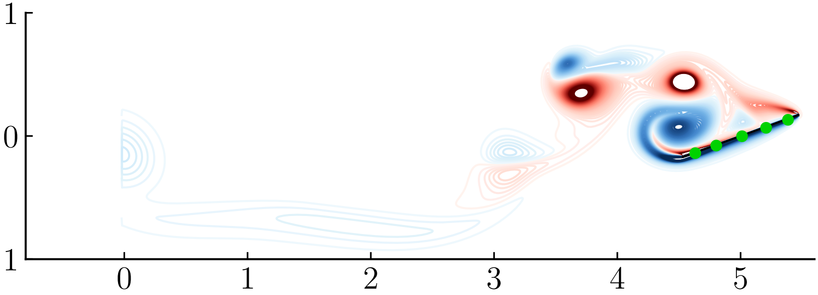
Vortex shedding around an aircraft wing

Motivation: Estimating turbulent flow [Le Provost, B, et al., 2022]



Vortex shedding around an aircraft wing

Motivation: Estimating turbulent flow [Le Provost, B, et al., 2022]



Vortex shedding around an aircraft wing

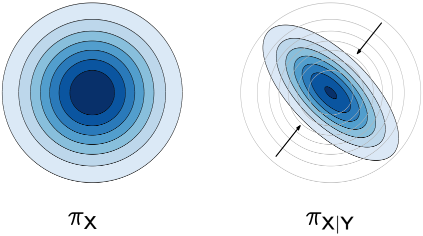
Challenge:

- High-dimensional states and observations $d = 180$ and $m = 50$
- States: Positions and strengths of point vortices $\mathbf{y}_t \in \mathbb{R}^d$
- Observation: Pressure along airfoil $\mathbf{y}_t \in \mathbb{R}^m$

Jointly reducing parameters and observation dimensions

Main ideas

- Only part of the parameters is informed by observations
- Only part of the observations is relevant to the parameters



Related work: State-space projections [Cui et al., 2014, Zahm et al., 2018], Observation-space projections [Giraldi et al., 2018]

Decomposition of parameters and observations

- Decompose $\mathbf{X} \in \mathbb{R}^d, \mathbf{Y} \in \mathbb{R}^m$ using orthogonal subspaces

$$\mathbf{X} = U_r^T \mathbf{X}_r + U_\perp^T \mathbf{X}_\perp, \qquad \mathbf{X}_r \in \mathbb{R}^r \text{ is informed by } \mathbf{Y}$$

$$\mathbf{Y} = V_s^T \mathbf{Y}_s + V_\perp^T \mathbf{Y}_\perp, \qquad \mathbf{Y}_s \in \mathbb{R}^s \text{ is informative of } \mathbf{X}$$

Decomposition of parameters and observations

- Decompose $\mathbf{X} \in \mathbb{R}^d, \mathbf{Y} \in \mathbb{R}^m$ using orthogonal subspaces

$$\mathbf{X} = U_r^T \mathbf{X}_r + U_\perp^T \mathbf{X}_\perp, \qquad \mathbf{X}_r \in \mathbb{R}^r \text{ is informed by } \mathbf{Y}$$
$$\mathbf{Y} = V_s^T \mathbf{Y}_s + V_\perp^T \mathbf{Y}_\perp, \qquad \mathbf{Y}_s \in \mathbb{R}^s \text{ is informative of } \mathbf{X}$$

- Consider the class of posterior density approximations

$$\widehat{\pi}_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y}) = \widehat{\pi}_{\mathbf{X}_r|\mathbf{Y}_s}(\mathbf{x}_r|\mathbf{y}_s)\pi_{\mathbf{X}_\perp|\mathbf{X}_r}(\mathbf{x}_\perp|\mathbf{x}_r) \propto \widehat{\pi}_{\mathbf{Y}_s|\mathbf{X}_r}(\mathbf{y}_s|\mathbf{x}_r)\pi_{\mathbf{X}}(\mathbf{x})$$

Decomposition of parameters and observations

- Decompose $\mathbf{X} \in \mathbb{R}^d, \mathbf{Y} \in \mathbb{R}^m$ using orthogonal subspaces

$$\mathbf{X} = U_r^T \mathbf{X}_r + U_\perp^T \mathbf{X}_\perp, \qquad \mathbf{X}_r \in \mathbb{R}^r \text{ is informed by } \mathbf{Y}$$
$$\mathbf{Y} = V_s^T \mathbf{Y}_s + V_\perp^T \mathbf{Y}_\perp, \qquad \mathbf{Y}_s \in \mathbb{R}^s \text{ is informative of } \mathbf{X}$$

- Consider the class of posterior density approximations

$$\widehat{\pi}_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y}) = \widehat{\pi}_{\mathbf{X}_r|\mathbf{Y}_s}(\mathbf{x}_r|\mathbf{y}_s)\pi_{\mathbf{X}_\perp|\mathbf{X}_r}(\mathbf{x}_\perp|\mathbf{x}_r) \propto \widehat{\pi}_{\mathbf{Y}_s|\mathbf{X}_r}(\mathbf{y}_s|\mathbf{x}_r)\pi_{\mathbf{X}}(\mathbf{x})$$

- **Goal**: Find U_r, V_s with $r(\epsilon) \ll d$ and $s(\epsilon) \ll m$ such that

$$\mathbb{E}_{\mathbf{Y}}[\mathrm{D}_{\mathrm{KL}}(\pi_{\mathbf{X}|\mathbf{Y}}||\widehat{\pi}_{\mathbf{X}|\mathbf{Y}})] \leq \epsilon$$

Decomposition of parameters and observations

- Decompose $\mathbf{X} \in \mathbb{R}^d, \mathbf{Y} \in \mathbb{R}^m$ using orthogonal subspaces

$$\mathbf{X} = U_r^T \mathbf{X}_r + U_\perp^T \mathbf{X}_\perp, \qquad \mathbf{X}_r \in \mathbb{R}^r \text{ is informed by } \mathbf{Y}$$
$$\mathbf{Y} = V_s^T \mathbf{Y}_s + V_\perp^T \mathbf{Y}_\perp, \qquad \mathbf{Y}_s \in \mathbb{R}^s \text{ is informative of } \mathbf{X}$$

- Consider the class of posterior density approximations

$$\widehat{\pi}_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y}) = \widehat{\pi}_{\mathbf{X}_r|\mathbf{Y}_s}(\mathbf{x}_r|\mathbf{y}_s)\pi_{\mathbf{X}_\perp|\mathbf{X}_r}(\mathbf{x}_\perp|\mathbf{x}_r) \propto \widehat{\pi}_{\mathbf{Y}_s|\mathbf{X}_r}(\mathbf{y}_s|\mathbf{x}_r)\pi_{\mathbf{X}}(\mathbf{x})$$

- **Goal**: Find U_r, V_s with $r(\epsilon) \ll d$ and $s(\epsilon) \ll m$ such that

$$\mathbb{E}_{\mathbf{Y}}[\mathrm{D}_{\mathrm{KL}}(\pi_{\mathbf{X}|\mathbf{Y}}||\widehat{\pi}_{\mathbf{X}|\mathbf{Y}})] \leq \epsilon$$

- **Result**: Sample posterior by building lower dimensional maps:
  1. Construct map $T^{\mathcal{X}}(\mathbf{y}_s, \mathbf{x}_r)$ to sample $\mathbf{X}_r^i \sim \pi_{\mathbf{X}_r|\mathbf{Y}_s}$
  2. Join with conditional prior samples $\mathbf{X}_\perp^i \sim \pi_{\mathbf{X}_\perp|\mathbf{x}_r^i}$

Decomposition of parameters and observations

Approach: Minimize error of closest approximation $\pi^*_{\mathbf{Y}|\mathbf{X}} := \pi_{\mathbf{Y}_s|\mathbf{X}_r}\pi_{\mathbf{X}}$

$$\mathbb{E}_{\mathbf{Y}}[D_{KL}(\pi_{\mathbf{X}|\mathbf{Y}}||\pi^*_{\mathbf{X}|\mathbf{Y}})] = I(\mathbf{X}_\perp, \mathbf{Y}|\mathbf{X}_r) + I(\mathbf{Y}_\perp, \mathbf{X}|\mathbf{Y}_s) - I(\mathbf{Y}_\perp, \mathbf{Y}_\perp|\mathbf{Y}_s, \mathbf{X}_r)$$
$$\leq \underbrace{I(\mathbf{X}_\perp, \mathbf{Y}|\mathbf{X}_r)}_{\text{function}(U_\perp)} + \underbrace{I(\mathbf{Y}_\perp, \mathbf{X}|\mathbf{Y}_s)}_{\text{function}(V_\perp)}$$

Recall: Conditional mutual information (CMI) $I(\mathbf{A}, \mathbf{B}|\mathbf{C}) = 0$ if $\mathbf{A} \perp\!\!\!\perp \mathbf{B}|\mathbf{C}$

Decomposition of parameters and observations

Approach: Minimize error of closest approximation $\pi^*_{\mathbf{Y}|\mathbf{X}} := \pi_{\mathbf{Y}_s|\mathbf{X}_r}\pi_{\mathbf{X}}$

$$\mathbb{E}_{\mathbf{Y}}[\mathrm{D}_{\mathrm{KL}}(\pi_{\mathbf{X}|\mathbf{Y}}||\pi^*_{\mathbf{X}|\mathbf{Y}})] = I(\mathbf{X}_\perp, \mathbf{Y}|\mathbf{X}_r) + I(\mathbf{Y}_\perp, \mathbf{X}|\mathbf{Y}_s) - I(\mathbf{Y}_\perp, \mathbf{Y}_\perp|\mathbf{Y}_s, \mathbf{X}_r)$$
$$\leq \underbrace{I(\mathbf{X}_\perp, \mathbf{Y}|\mathbf{X}_r)}_{\text{function}(U_\perp)} + \underbrace{I(\mathbf{Y}_\perp, \mathbf{X}|\mathbf{Y}_s)}_{\text{function}(V_\perp)}$$

Recall: Conditional mutual information (CMI) $I(\mathbf{A}, \mathbf{B}|\mathbf{C}) = 0$ if $\mathbf{A} \perp\!\!\!\perp \mathbf{B}|\mathbf{C}$

Idea: For non-Gaussian π, minimize tractable upper bounds for CMI

Theorem [B, Marzouk, et al., 2021]

If $\pi_{\mathbf{X},\mathbf{Y}}$ satisfies a conditional log-Sobolev inequality with constant C_π,

$$I(\mathbf{X}_\perp, \mathbf{Y}|\mathbf{X}_r) \leq C_\pi^2 \mathbb{E}_\pi \|\nabla_{\mathbf{y},\mathbf{x}} \log \pi_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) U_\perp\|_F^2$$
$$I(\mathbf{Y}_\perp, \mathbf{X}|\mathbf{Y}_s) \leq C_\pi^2 \mathbb{E}_\pi \|V_\perp^T \nabla_{\mathbf{y},\mathbf{x}} \log \pi_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})\|_F^2$$

Example: subspaces for Gaussian likelihood models

Let $\mathbf{Y} = G(\mathbf{X}) + \boldsymbol{\epsilon}$ where $\text{Cov}(\mathbf{X}) = \text{I}_d$ and $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \text{I}_m)$.

Example: subspaces for Gaussian likelihood models

Let $\mathbf{Y} = G(\mathbf{X}) + \boldsymbol{\epsilon}$ where $\mathrm{Cov}(\mathbf{X}) = \mathrm{I}_d$ and $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \mathrm{I}_m)$.

Informed state space [Cui et al., 2020]

- $U_r = [u_1, \ldots, u_r]$ where $(\lambda_{\mathbf{X},i}, u_i)$ are leading eigen-pairs of

$$H_{\mathbf{X}} = \int \nabla G(\mathbf{x})^T \nabla G(\mathbf{x}) d\pi_{\mathbf{X}}(\mathbf{x})$$

Informative observations space

- $V_s = [v_1, \ldots, v_s]$ where $(\lambda_{\mathbf{Y},j}, v_j)$ are leading eigen-pairs of

$$H_{\mathbf{Y}} = \int \nabla G(\mathbf{x}) \nabla G(\mathbf{x})^T d\pi_{\mathbf{X}}(\mathbf{x})$$

Example: subspaces for Gaussian likelihood models

Let $\mathbf{Y} = G(\mathbf{X}) + \boldsymbol{\epsilon}$ where $\mathrm{Cov}(\mathbf{X}) = \mathrm{I}_d$ and $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \mathrm{I}_m)$.

Informed state space [Cui et al., 2020]

- $U_r = [u_1, \ldots, u_r]$ where $(\lambda_{\mathbf{X},i}, u_i)$ are leading eigen-pairs of

$$H_{\mathbf{X}} = \int \nabla G(\mathbf{x})^T \nabla G(\mathbf{x}) \mathrm{d}\pi_{\mathbf{X}}(\mathbf{x})$$

Informative observations space

- $V_s = [v_1, \ldots, v_s]$ where $(\lambda_{\mathbf{Y},j}, v_j)$ are leading eigen-pairs of

$$H_{\mathbf{Y}} = \int \nabla G(\mathbf{x}) \nabla G(\mathbf{x})^T \mathrm{d}\pi_{\mathbf{X}}(\mathbf{x})$$

Corollary: Error bound for posterior approximation

$$\mathbb{E}_{\mathbf{Y}}[\mathrm{D}_{\mathrm{KL}}(\pi_{\mathbf{X}|\mathbf{Y}} || \pi^*_{\mathbf{X}|\mathbf{Y}})] \leq C_\pi^2 \Big(\sum_{i>r} \lambda_{\mathbf{X},i} + \sum_{j>s} \lambda_{\mathbf{Y},j}\Big)$$

Generalization of linear dimension reduction

Let $\mathbf{Y} = \mathbf{G}\mathbf{X} + \boldsymbol{\epsilon}$ where $\text{Cov}(\mathbf{X}) = \mathsf{I}_d$ and $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \mathsf{I}_m)$.

Diagnostic matrices:

$$H_{\mathbf{X}} = \mathbf{G}^T\mathbf{G}, \qquad H_{\mathbf{Y}} = \mathbf{G}\mathbf{G}^T$$

Proposition

After a rotation, eigenvectors of $H_{\mathbf{X}}$ and $H_{\mathbf{Y}}$ reduce to solution of canonical correlation analysis (CCA)

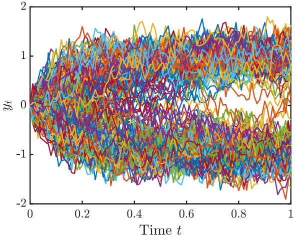
$$\text{Cov}(\mathbf{X},\mathbf{Y})\text{Cov}(\mathbf{Y})^{-1}\text{Cov}(\mathbf{X},\mathbf{Y})^T u_i = \lambda_{\mathbf{X},i}/(1+\lambda_{\mathbf{X},i})u_i$$

$$\text{Cov}(\mathbf{Y},\mathbf{X})\text{Cov}(\mathbf{X})^{-1}\text{Cov}(\mathbf{Y},\mathbf{X})^T v_j = \lambda_{\mathbf{Y},j}/(1+\lambda_{\mathbf{Y},i})v_j$$

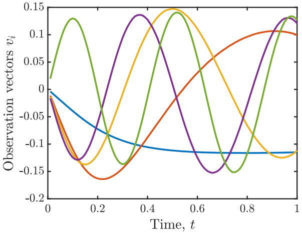
Takeaway: Gradient-based diagnostic matrices generalize CCA for nonlinear forward models

Conditioned diffusion problem

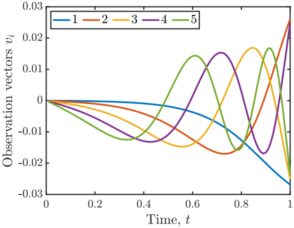
- Particle follows SDE: $du_t = f(u_t)dt + dX_t$ with drift $f(u) = \beta u(1 - u^2)/(1 + u^2)$ and Brownian motion X
- Infer driving force x given noisy state observations $y_{t_i} = u_{t_i} + \epsilon_i$
- Discretized parameters $\mathbf{X}$ and observations $\mathbf{Y}$ have dimension 100



Sample realizations of y_t



$U_{1:5}$ from PCA



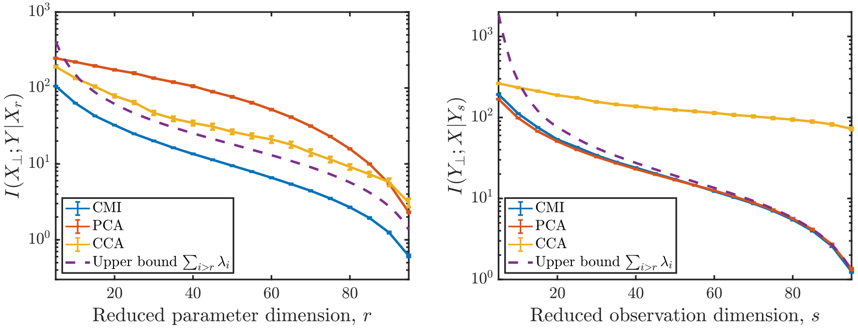
$U_{1:5}$ from CMI

Takeaway: CMI-based eigenvectors are more relevant for inference

CMI-based subspaces are more relevant for inference

Conditioned diffusion problem

- Particle follows SDE: $du_t = f(u_t)dt + dX_t$ with drift $f(u) = \beta u(1-u^2)/(1+u^2)$ and Brownian motion X
- Infer driving force x given noisy state observations $y_{t_i} = u_{t_i} + \epsilon_i$
- Discretized parameters $\mathbf{X}$ and observations $\mathbf{Y}$ have dimension 100

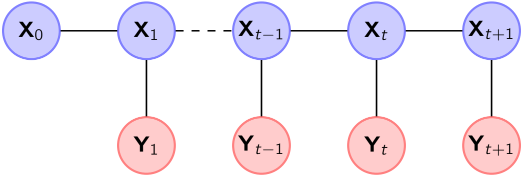


Takeaway: CMI-based subspaces minimize posterior approximation error

Back to turbulent flows

Sequential Bayesian inference:

- States: Biot-Savart dynamics $\pi_{\mathbf{X}_t|\mathbf{X}_{t-1}}$
- Observations: Poisson equation with additive noise $\pi_{\mathbf{Y}_t|\mathbf{X}_t}$

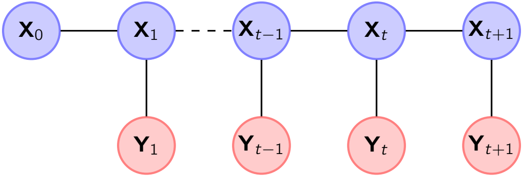


Goal: Recursively characterize filtering distributions $\pi_{\mathbf{X}_t|\mathbf{y}_1^*,\ldots,\mathbf{y}_t^*}$

Back to turbulent flows

Sequential Bayesian inference:

- States: Biot-Savart dynamics $\pi_{\mathbf{X}_t|\mathbf{X}_{t-1}}$
- Observations: Poisson equation with additive noise $\pi_{\mathbf{Y}_t|\mathbf{X}_t}$



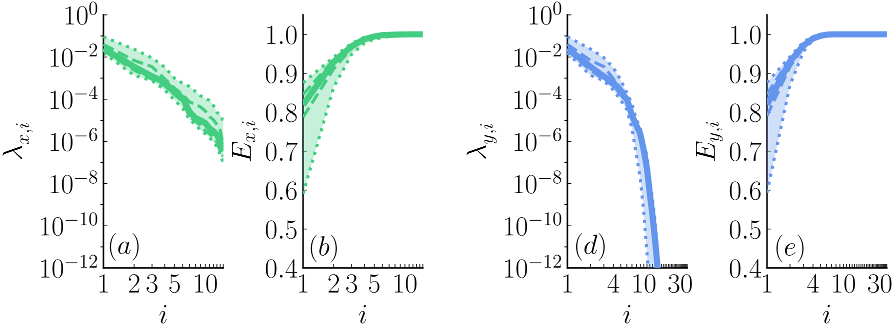
Goal: Recursively characterize filtering distributions $\pi_{\mathbf{X}_t|\mathbf{y}_1^*,\ldots,\mathbf{y}_t^*}$

Recursive approach: At each time t

- Use model dynamics to predict state from $\pi_{\mathbf{X}_t|\mathbf{y}_1^*,\ldots,\mathbf{y}_{t-1}^*}$ (i.e., prior)
- Solve inverse problem for $\pi_{\mathbf{X}_t|\mathbf{y}_1^*,\ldots,\mathbf{y}_t^*}$ given observation $\mathbf{y}_t^*$

State and observation diagnostic matrices are often low-rank

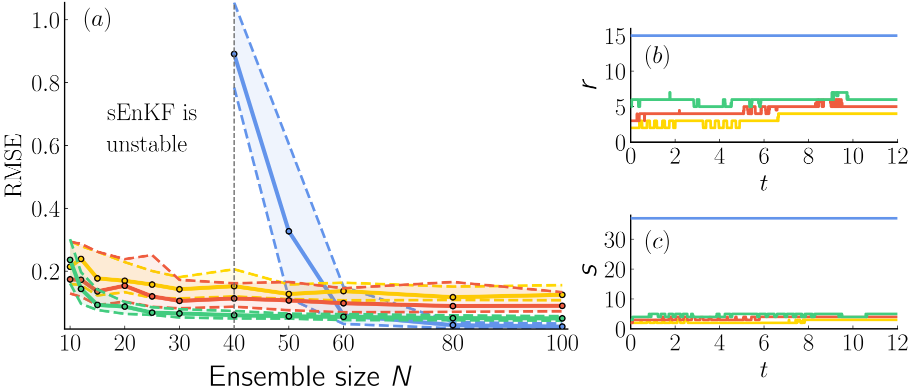
Spectra and energy of $H_{\mathbf{X}}$, $H_{\mathbf{Y}}$



Adaptive rank algorithm

- Use energy $E_i = \sum_{j=1}^{i} \lambda_j / \sum_j \lambda_j$ to select reduced dimensions
- For example, choose r such that $E_r > 0.99$

Low-rank filter is stable for small ensemble sizes

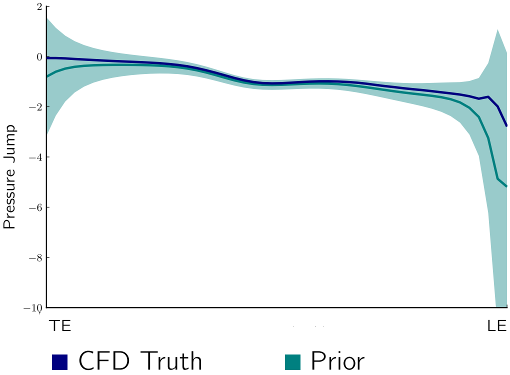


Observations:

- RMSE is stable for small N for different energy ratios
- Reduced dimensions r, s do not increase over time

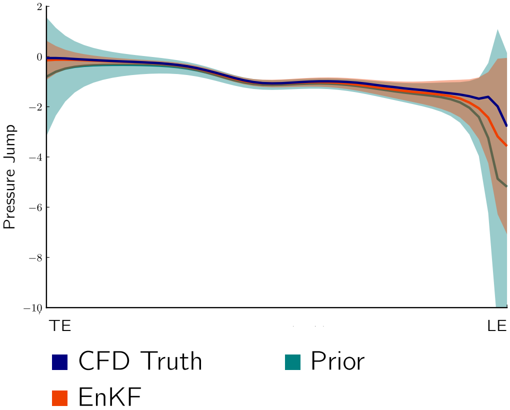
Low-rank filter improves pressure estimation

- Estimate flow around the airfoil at 20° angle of attack and $Re = 500$ subject to force actuation mimicking gusts



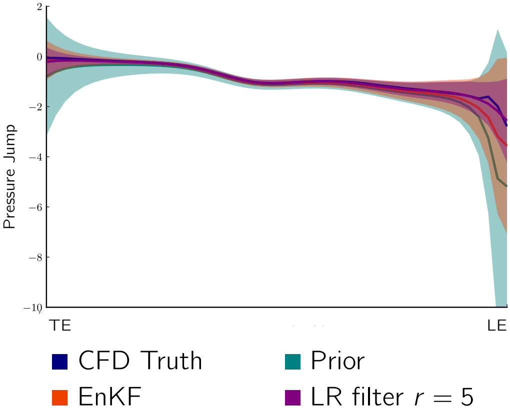
Low-rank filter improves pressure estimation

- Estimate flow around the airfoil at 20° angle of attack and $Re = 500$ subject to force actuation mimicking gusts



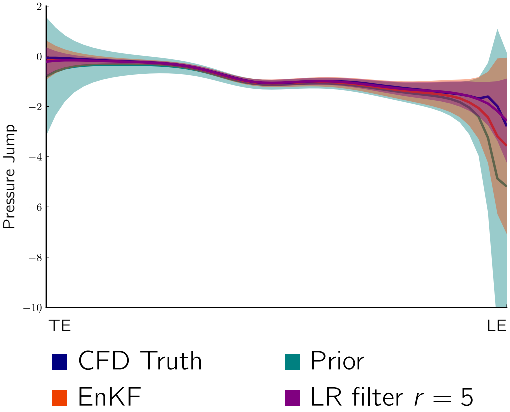
Low-rank filter improves pressure estimation

- Estimate flow around the airfoil at 20° angle of attack and $Re = 500$ subject to force actuation mimicking gusts



Low-rank filter improves pressure estimation

- Estimate flow around the airfoil at 20° angle of attack and $Re = 500$ subject to force actuation mimicking gusts



- Posterior predictive distribution has lower bias and spread at the leading edge

Conclusion and outlook

Conclusion and outlook

Main idea: Dimension reduction of parameters and observations

- Detect subspaces using gradients of the observation model
- Provide error guarantees on posterior approximation
- Stable tracking of turbulent flows with small ensemble sizes

Conclusion and outlook

Main idea: Dimension reduction of parameters and observations

- Detect subspaces using gradients of the observation model
- Provide error guarantees on posterior approximation
- Stable tracking of turbulent flows with small ensemble sizes

Future work

- Gradient-free identification of low-dimensional structure (e.g., using score estimation methods [Song et al., 2019])
- Other sources of structure, e.g., conditional independence

Conclusion and outlook

Main idea: Dimension reduction of parameters and observations

- Detect subspaces using gradients of the observation model
- Provide error guarantees on posterior approximation
- Stable tracking of turbulent flows with small ensemble sizes

Future work

- Gradient-free identification of low-dimensional structure (e.g., using score estimation methods [Song et al., 2019])
- Other sources of structure, e.g., conditional independence

References: arXiv:2203.05120, arXiv:2207.08670

Thank You

Supported by the U.S. Department of Energy