SAND2017-10858C

Simulation-Based Optimal Bayesian Design for Sequential Experiments

(arXiv: 1604.08320)

Xun Huan (xhuan@sandia.gov) & Youssef M. Marzouk (ymarz@mit.edu)

Sandia National Laboratories Massachusetts Institute of Technology

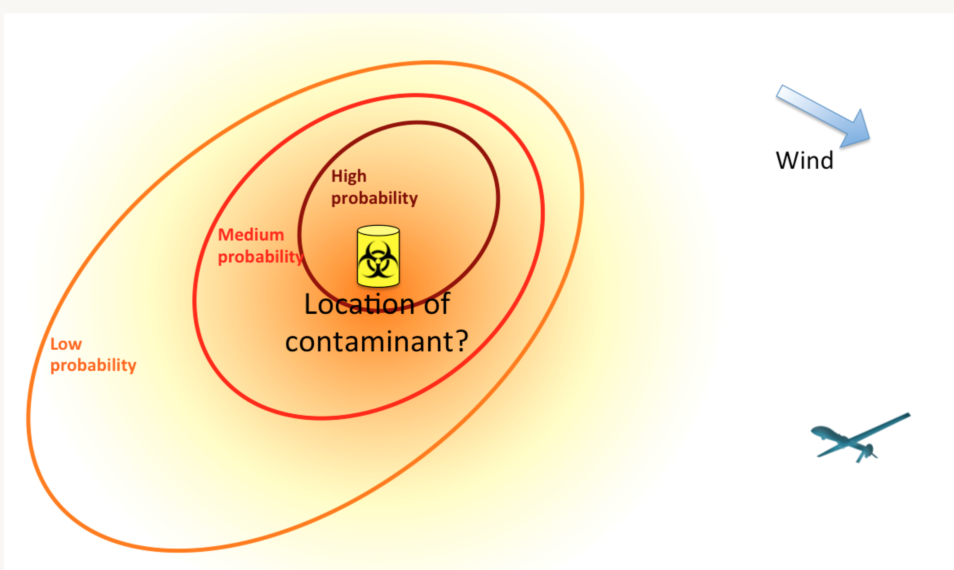
Summary

- *Optimal* design for sequential experiments requires **feedback** and **forward-looking**; batch (open-loop) and greedy (myopic) designs are provably suboptimal
- Optimality is formulated via **dynamic programming (DP)**
- Solving DP is expensive, requiring repeated Bayesian inferences
- Near-optimal policy is found numerically using backward induction coupled with adaptive regression (approximate dynamic programming)
- Fast approximate Bayesian inference is achieved using a transport map transforming the design-data-parameter joint distribution to a Gaussian

Goal: develop computational methods to find optimal policies for a sequence of experiments, accommodating:

- nonlinear models
- non-Gaussian distributions
- continuous variables
- information measure objectives

Motivation



Question: what sequence of measurement locations produces the best inference on the unknown source location?

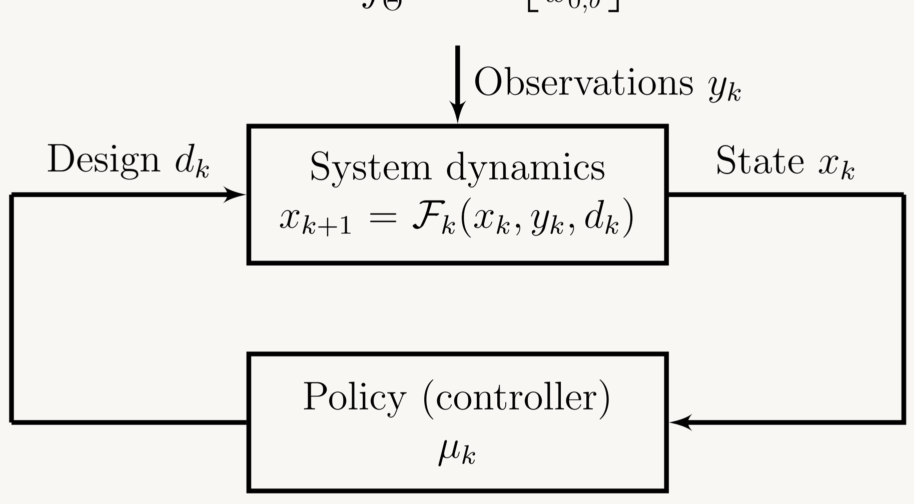
The Optimal Sequential Design Problem

- **Experiment:** $k = 0, \dots, N-1$, total N experiments, $N < \infty$
- **State:** $x_k = [x_{k,b}, x_{k,p}]$ contains all information needed for optimal future designs
  - *Belief state:* $x_{k,b}$ current state of uncertainty (e.g., posterior density $x_{k,b} = f(\theta|d_0, \dots, d_{k-1}, y_0, \dots, y_{k-1})$)
  - *Physical state:* $x_{k,p}$ deterministic variables (e.g., vehicle position)
- **Design:** $d_k = \mu_k(x_k)$; we seek good policy $\pi = \{\mu_0, \mu_1, \dots, \mu_{N-1}\}$
- **Observations:** y_k distributed according to likelihood $f(y_k|\theta, d_k)$ (e.g., $y_k = G(\theta, d_k) + \epsilon$, with ϵ Gaussian noise)
- **System dynamics:** $x_{k+1} = \mathcal{F}(x_k, d_k, y_k)$ state evolution
  - *Belief state:* Bayes' theorem

$$x_{k+1,b} = \frac{f(y_k|\theta, d_k) x_{k,b}}{f(y_k|d_0, \dots, d_k, y_0, \dots, y_{k-1})}$$

  - *Physical state:* physical process
- **Stage reward:** $g_k(x_k, y_k, d_k)$ experimental rewards/costs
- **Terminal reward:** $g_N(x_N)$ value of final state (e.g., information gain via Kullback-Leibler divergence)

$$g_N(x_N) = \int_\Theta x_{N,b} \ln \left[\frac{x_{N,b}}{x_{0,b}}\right] d\theta$$



Sequential optimal experimental design (sOED) problem:

Find π^* where

$$\pi^* = \underset{\pi=\{\mu_0,\dots,\mu_{N-1}\}}{\operatorname{argmax}} \; \mathbb{E}_{y_0,\dots,y_{N-1}|\pi}\left[\sum_{k=0}^{N-1} g_k(x_k, y_k, \mu_k(x_k)) + g_N(x_N)\right]$$

$$\text{s.t.} \quad x_{k+1} = \mathcal{F}_k(x_k, y_k, d_k), \forall k \qquad \mu_k(x_k) \in \mathcal{D}_k, \forall x_k, k$$

Approximate Dynamic Prgoramming

- Equivalent problem, different form— **dynamic program (Bellman equations):**

$$J_k(x_k) = \max_{d_k \in \mathcal{D}_k} \mathbb{E}_{y_k|x_k,d_k}\left[g_k(x_k, d_k, y_k) + J_{k+1}(\mathcal{F}_k(x_k, d_k, y_k))\right]$$
$$J_N(x_N) = g_N(x_N)$$

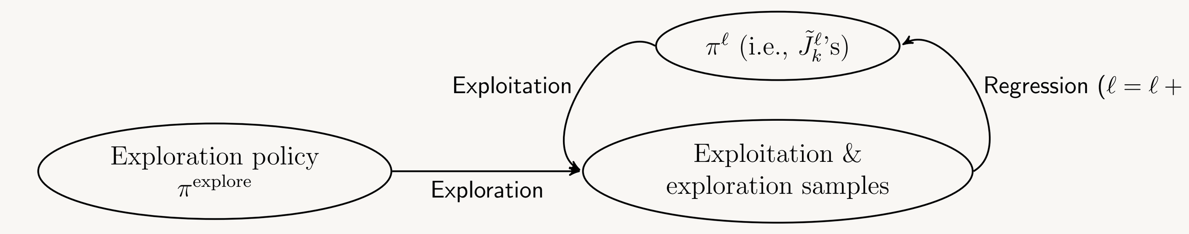
$k = 0, \dots, N-1$; $J_k(x_k)$ are called value functions
  - Optimal policy functions: $d_k^* = \mu_k^*(x_k)$
  - Rich existing work in *approximate dynamic programming*
- Represent policies: **one-step lookahead**

$$\mu_k(x_k) = \underset{d_k \in \mathcal{D}_k}{\operatorname{argmax}} \mathbb{E}_{y_k|x_k,d_k}\left[g_k(x_k, y_k, d_k) + \tilde{J}_{k+1}(\mathcal{F}_k(x_k, y_k, d_k))\right]$$

  - policy π implicitly parameterized by $\tilde{J}_1(x_1), \dots, \tilde{J}_N(x_N)$
  - if $\tilde{J}_{k+1}(x_{k+1}) = J_{k+1}(x_{k+1})$, then $\mu_k(x_k) = \mu_k^*(x_k)$ is optimal: therefore want $\tilde{J}_{k+1}(x_{k+1})$ close to $J_{k+1}(x_{k+1})$
- Represent $\tilde{J}_k$'s: **linear architecture** $\tilde{J}_k(x_k) = r_k^\top \phi_k(x_k)$; ϕ_k features (selected from heuristics: posterior moments), r_k weights
- Construct $\tilde{J}_k$'s using **approximate value iteration (backward induction with regression):**

$$\tilde{J}_k(x_k) = \mathcal{P}\left\{\max_{d_k \in \mathcal{D}_k} \mathbb{E}_{y_k|x_k,d_k}\left[g_k(x_k, d_k, y_k) + \tilde{J}_{k+1}(\mathcal{F}_k(x_k, d_k, y_k))\right]\right\}$$

Start with $\tilde{J}_N(x_N) \equiv g_N(x_N)$, and proceed backwards $k = N-1, \dots, 1$
$\mathcal{P}$ is regression operator
- Efficient regression point selection
  Exploration: random designs, "try something new", regularization
  Exploitation: applies current policy approximation, leverages current understanding of good policy, places more points at frequently-visited states
- Procedure to iteratively adapt regression as $\tilde{J}_k$'s improved

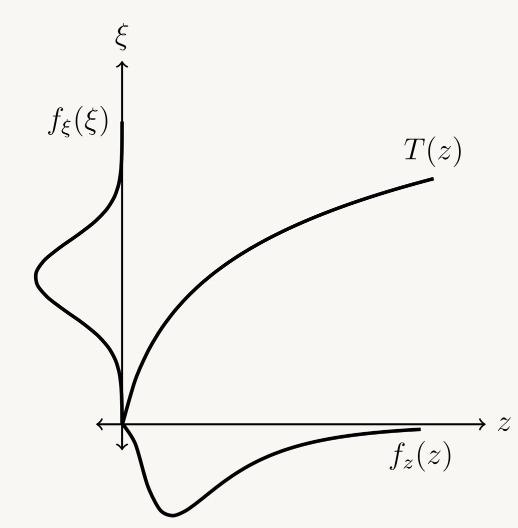


Fast Approximate Bayesian Inference: Transport Maps

- **How to numerically represent belief state (posterior)...**
  - for general non-Gaussian continuous random variables
  - in a finite-dimensional manner
  - to easily perform Bayesian inference repeatedly in evaluating

$$\tilde{J}_k(x_k) = \mathcal{P}\left\{\max_{d_k \in \mathcal{D}_k} \mathbb{E}_{y_k|x_k,d_k}\left[g_k(x_k, d_k, y_k) + \tilde{J}_{k+1}(\mathcal{F}_k(x_k, d_k, y_k))\right]\right\}$$

  $\Rightarrow$ potentially MILLIONS of repeated Bayesian inferences
  - Need approach that can quickly perform **many** Bayesian inferences
- **Transport maps:**



  - $\xi \sim$ reference distribution, $z \sim$ target distribution
  - Equivalence in distribution $\xi \overset{i.d.}{=} T(z)$
  - Knothe-Rosenblatt (KR) map: defined by conditional distributions, triangular and monotone, exists and is unique
  - Easy to construct: a benign convex optimization problem
- **Fast approximate inference:** posterior is conditioning the joint distribution

$$f(\theta|y) = \frac{f(y|\theta)f(\theta)}{f(y)} = \frac{f(y,\theta)}{f(y)}$$

For one experiment: KR map from (d, y, θ) to $\xi \sim \mathcal{N}(0, I)$ is:

$$\begin{bmatrix}\xi_1\\ \xi_2\\ \xi_3\end{bmatrix} = \begin{bmatrix}T_d(d)\\ T_{y|d}(d,y)\\ T_{\theta|y,d}(d,y,\theta)\end{bmatrix} = \begin{bmatrix}\Phi^{-1}(F(d))\\ \Phi^{-1}(F(y|d))\\ \Phi^{-1}(F(\theta|y,d))\end{bmatrix}$$

KR map of posterior given realizations d^* and y^* is:

$$T_{\theta|y^*,d^*}(\theta) = \Phi^{-1}(F(\theta|y^*,d^*))$$

This is precisely $T_{\theta|y,d}$ substituted with d^* and y^*: $T_{\theta|y,d}(d^*, y^*, \theta)$!

- Inference for multiple experiments by building a joint distribution map:

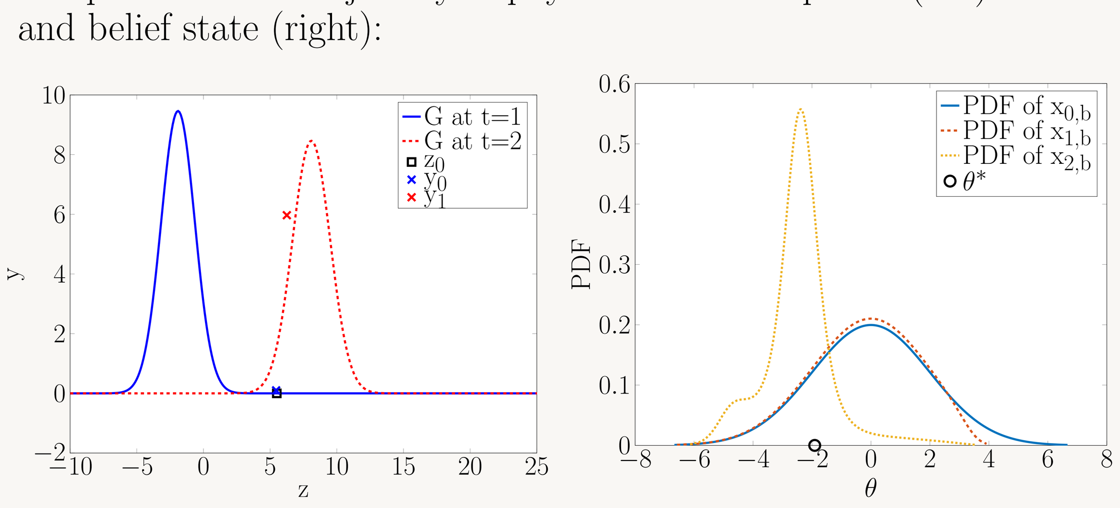
$k=0$	1	2	...
$\xi_{d_0} = T_{d_0}(d_0)$	$\xi_{d_0} = T_{d_0}(d_0)$	$\xi_{d_0} = T_{d_0}(d_0)$	
$\xi_{y_0} = T_{y_0}(d_0, y_0)$	$\xi_{y_0} = T_{y_0}(d_0, y_0)$	$\xi_{y_0} = T_{y_0}(d_0, y_0)$	
$\xi_{\theta_0} = T_{\theta_0}(d_0, y_0, \theta)$	$\xi_{d_1} = T_{d_1}(d_0, y_0, d_1)$	$\xi_{d_1} = T_{d_1}(d_0, y_0, d_1)$	
	$\xi_{y_1} = T_{y_1}(d_0, y_0, d_1, y_1)$	$\xi_{y_1} = T_{y_1}(d_0, y_0, d_1, y_1)$	
	$\xi_{\theta_1} = T_{\theta_1}(d_0, y_0, d_1, y_1, \theta)$	$\xi_{d_2} = T_{d_2}(d_0, y_0, d_1, y_1, d_2)$	
		$\xi_{y_2} = T_{y_2}(d_0, y_0, d_1, y_1, d_2, y_2)$	
		$\xi_{\theta_2} = T_{\theta_2}(d_0, y_0, d_1, y_1, d_2, y_2, \theta)$	

  - Posterior map after Bayesian inference on $k+1$ experiments is the n_θ-dimensional $T_{\theta_k|d_0^*,y_0^*,\dots,d_k^*,y_k^*}(\theta)$
  - Components grouped by the red rectangular boxes are identical; concatenate unique parts and construct overall map in one shot
  - Map constructed using exploration and exploitation samples

Results: Source Inversion in 1D Space

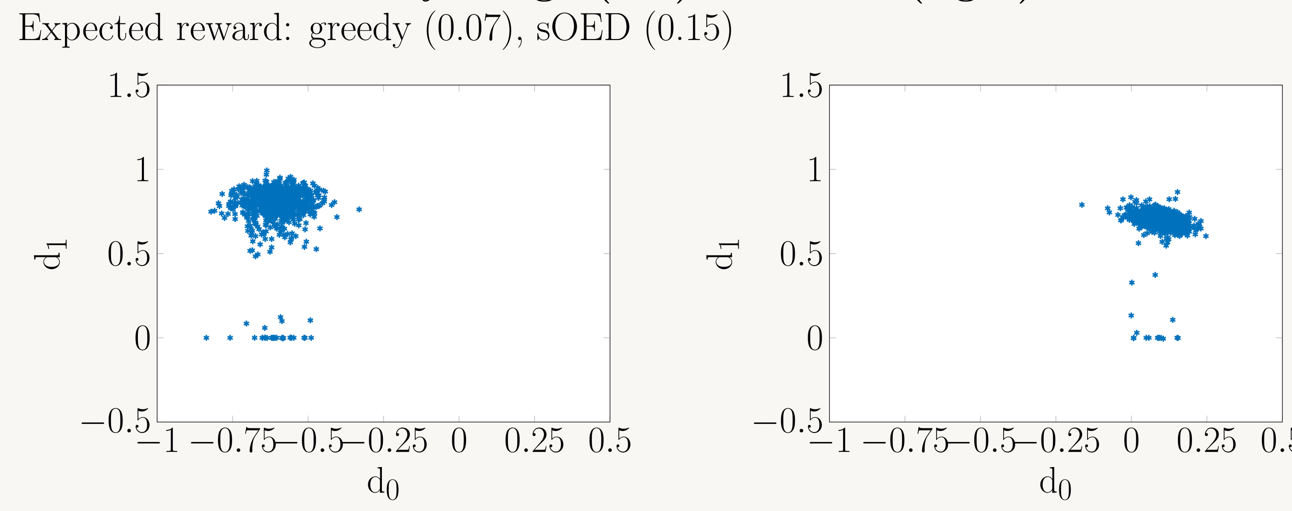
$$y_k = \frac{s}{\sqrt{2\pi}\sqrt{(1.2 + 4Dt)}} \exp\left(-\frac{\|\theta + d_w(t) - z_{k+1}\|^2}{2(1.2 + 4Dt)}\right) + \epsilon_k$$

- 2 experiments
- $\theta \sim \mathcal{N}(0, 2^2)$ starting location: 5.5
- Strong wind blows to the right after first experiment
- Quadratic movement penalty
- Sample evolution trajectory of physical state and plume (left) and belief state (right):



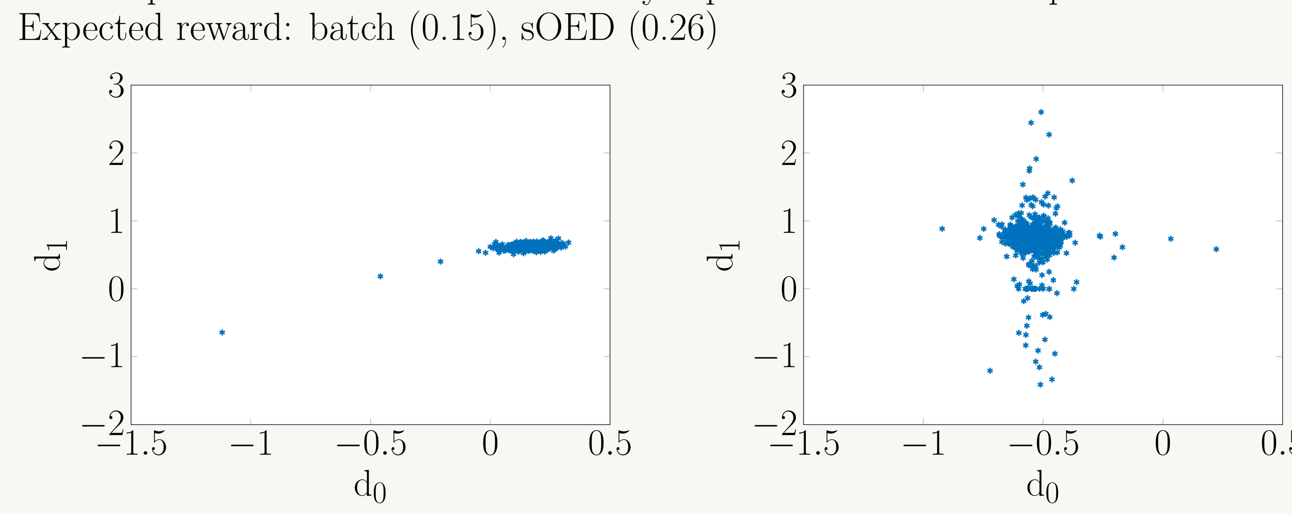
Greedy design (left) vs. sOED (right)

Expected reward: greedy (0.07), sOED (0.15)



Batch design (left) vs. sOED (right)

A more precise instrument available only if posterior variance drops < 3
Expected reward: batch (0.15), sOED (0.26)



Results: Source Inversion in 2D Space

A more challenging problem:

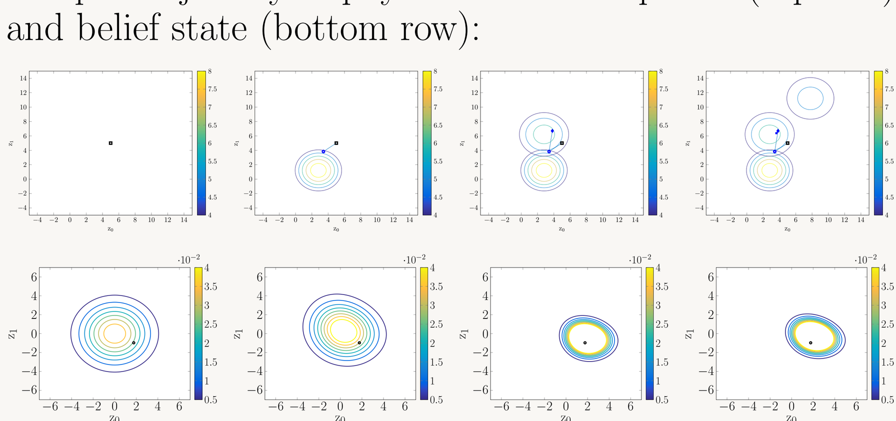
- 3 experiments
- 2D parameter and physical spaces

$$\theta \sim \mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}2^2 & 0\\ 0 & 2^2\end{bmatrix}\right)$$

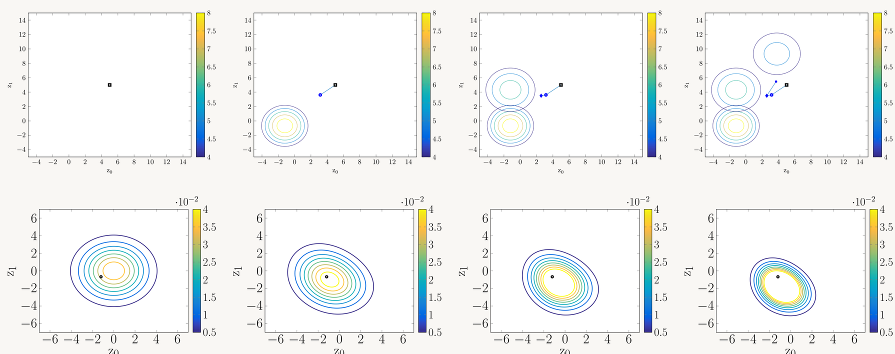
- Starting location = [5, 5]
- 15 dimensional joint distribution map
- Variable wind:
  blows north after experiment 1
  north-east after experiment 2

$\Rightarrow$ **Takeaway:** policy inherently adaptive, selects design based on what occurred so far!

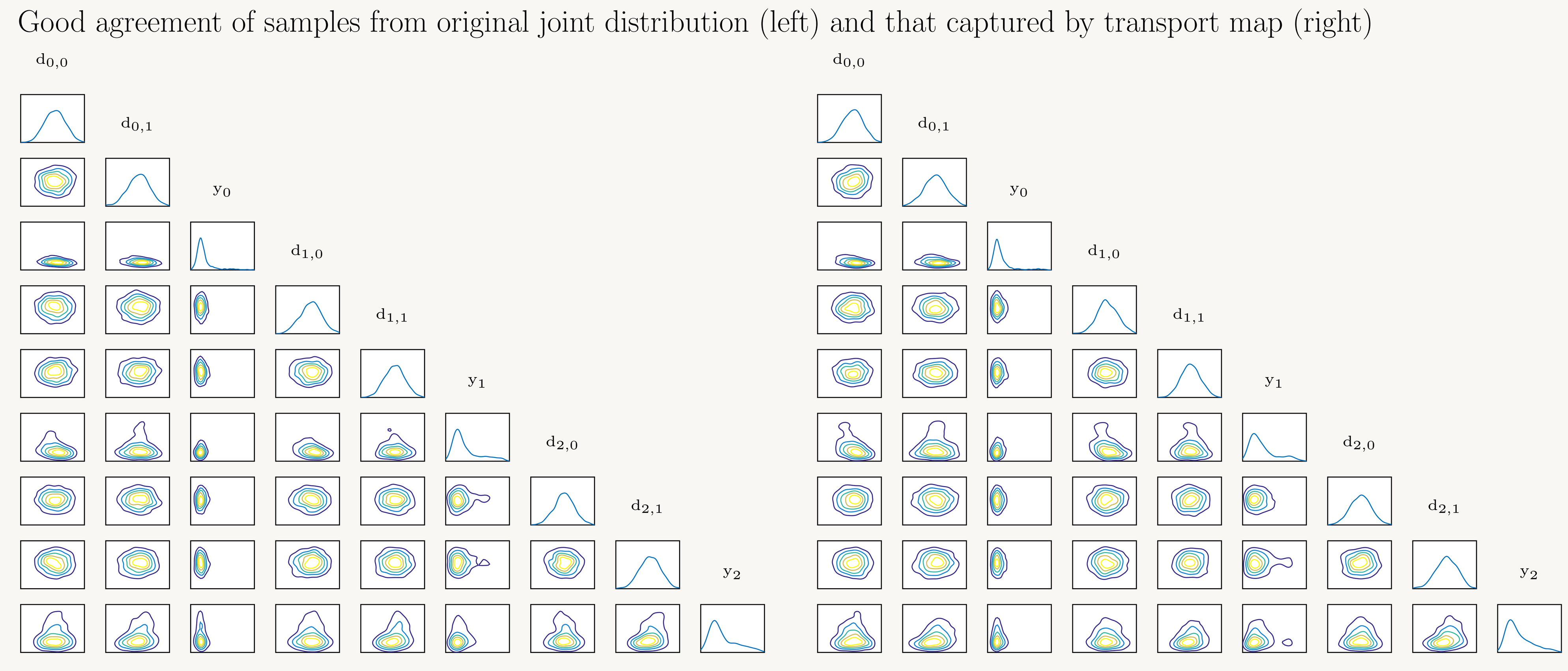
Sample trajectory of physical state and plume (top row) and belief state (bottom row):



Another sample trajectory:



Good agreement of samples from original joint distribution (left) and that captured by transport map (right)



Support for this work was provided through the Air Force Office of Scientific Research (AFOSR) Computational Mathematics Program and the Defense Advanced Research Projects Agency (DARPA) program on Enabling Quantification of Uncertainty in Physical Systems (EQUiPS). Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525.