SAND2022-15501C

Sandia National Laboratories

Leveraging Physics-based Surrogates for Efficient Density Estimation of Sparse Observable Data on Low-dimensional Manifolds

Tian Yu Yen and Tim Wildey

November 5, 2022

Talk for SIAM TXLA Section Meeting 2022

MS 26: Scientific Deep Learning

Acknowledgements: *This material is based upon work supported by the U.S. Department of Energy, Office of Science, ASCR, Early Career Research Program.*

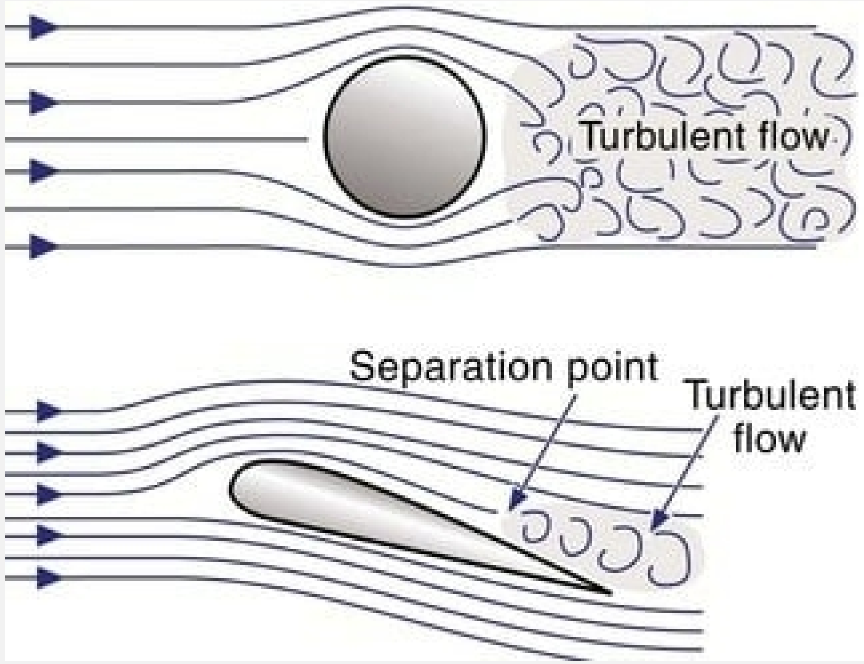
Outline

1. **The Context:** Stochastic Inverse Problems (Permeability Fields)
2. **The Method:** Data-consistent Inversion (DCI)
3. **The Challenge:** Tackling Density Estimation in High-dimensions
4. **Some Reflections:** Future Work and Analysis

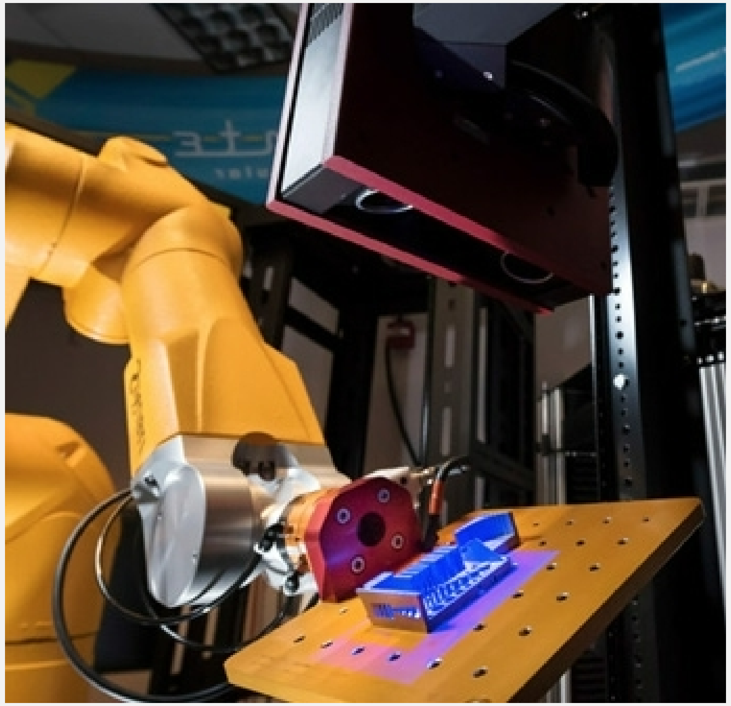
The Context

Stochastic Inversion with Permeability Fields

Solving stochastic inverse problems has many important applications.



In studies of hypersonic flight, quantifying aleatoric uncertainties in turbulent flows



In additive manufacturing, quantifying variability between component parts due to manufacturing

Example Stochastic Inverse Problem: Fluid Flow and Permeability Fields

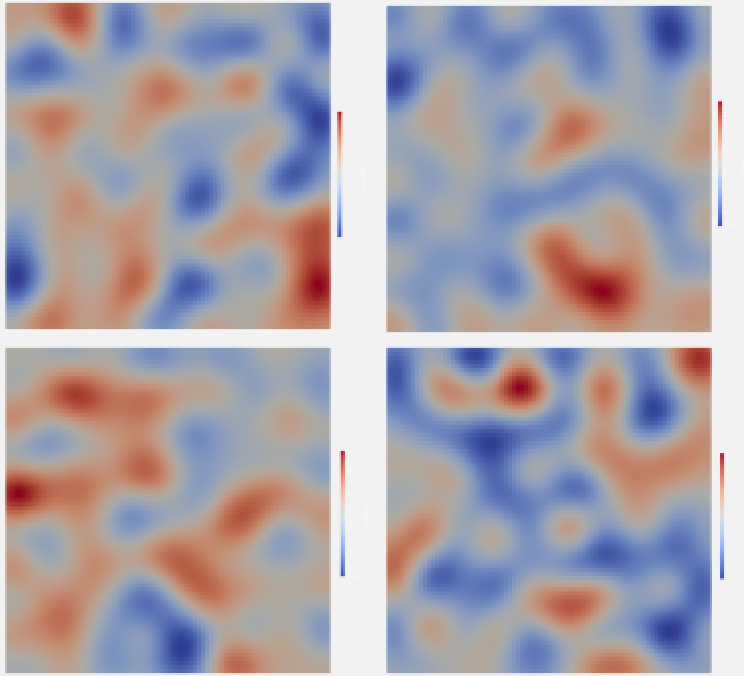
Suppose we want to solve a Poisson equation in 2D using a mixed method.

$$\mathbf{u} = -\mathbf{K}\nabla p,$$
$$\nabla \cdot \mathbf{u} = f$$

- i.e., a Darcy fluid flow problem through porous media with permeability field K.
- The permeability is modeled as a *random field*...

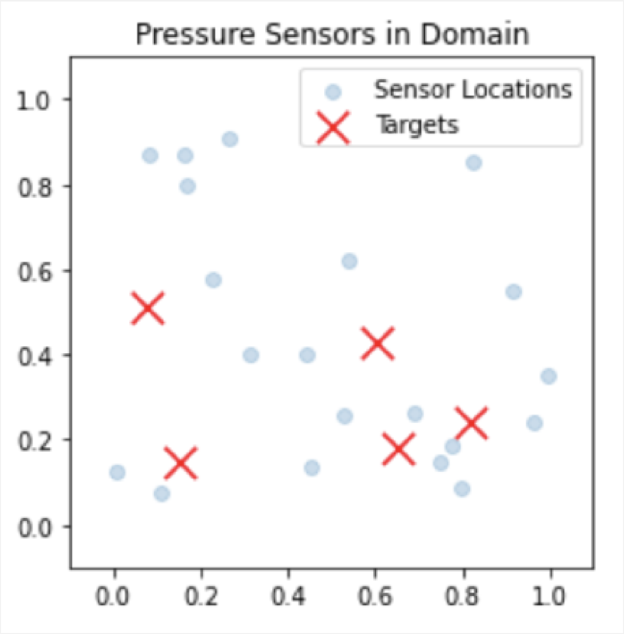
Example Stochastic Inverse Problem: Fluid Flow and Permeability Fields

Karhunen-Loeve (KL) Expansion used to model the permeability field.



Computational Model

Inverse Problem



Goal: Update parameters of KL expansion for better predictions.


The Method
Using Data-consistent Inversion to Solve Stochastic Inverse Problems

Data-consistent Inversion: What is the method?

A measure-theoretic approach…

$E(r) = 1$

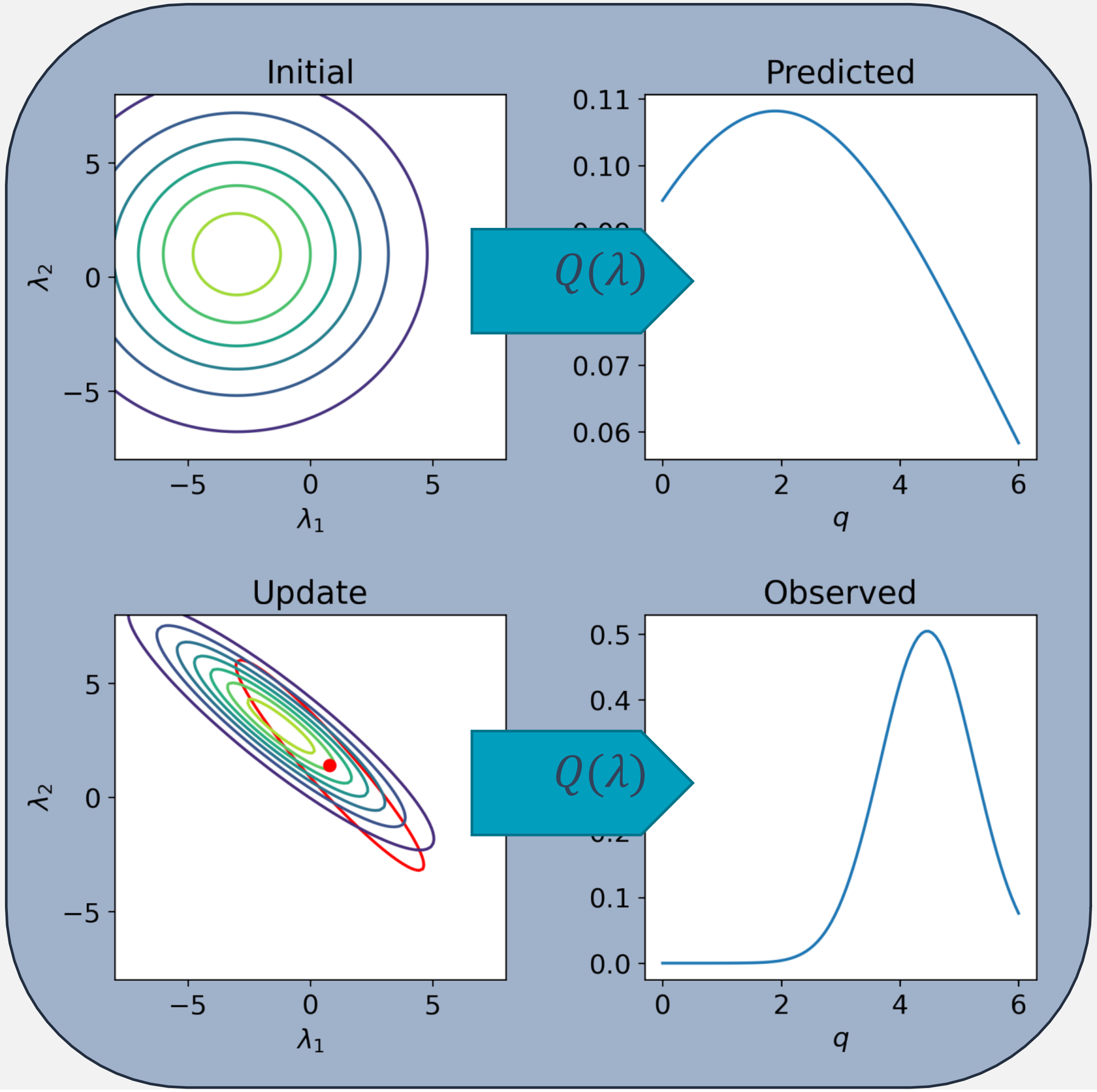
$$\pi_{update}(\lambda) = \pi_{init}(\lambda)\frac{\pi_{obs}(Q(\lambda))}{\pi_{predict}(Q(\lambda))}$$

Assumption: *Predictability assumption*

- Given initial assumptions about λ, model $Q(\lambda)$ can predict the data

Idea of Method: *Update initial assumptions by*

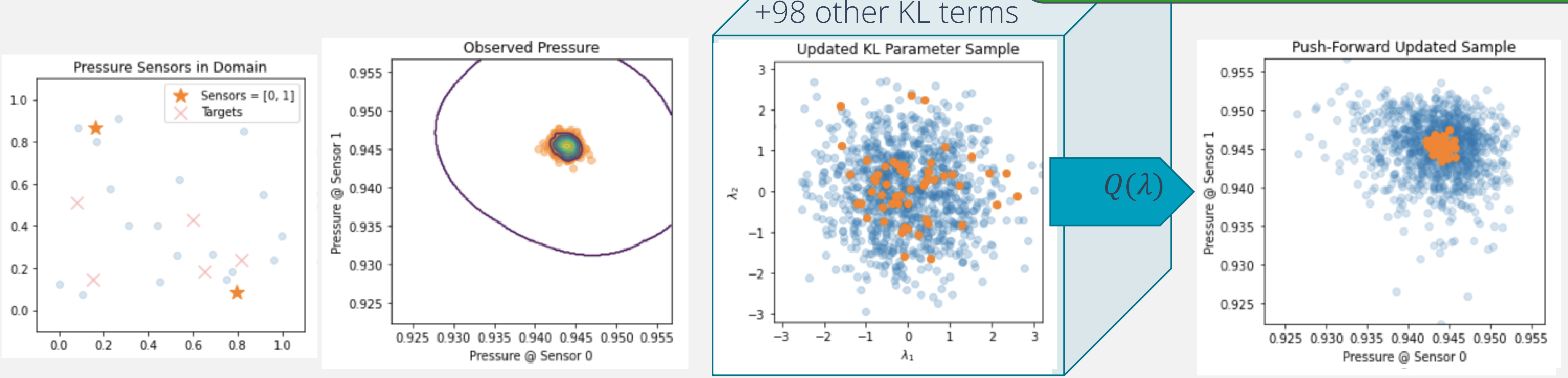
- Re-weighting initial with ratio of predicted (push-forward) to observed density



Data-consistent Inversion: A Consistent Solution

How does it work?

$$\pi_{update}(\lambda) = \pi_{init}(\lambda)\frac{\pi_{obs}(Q(\lambda))}{\pi_{predict}(Q(\lambda))}$$

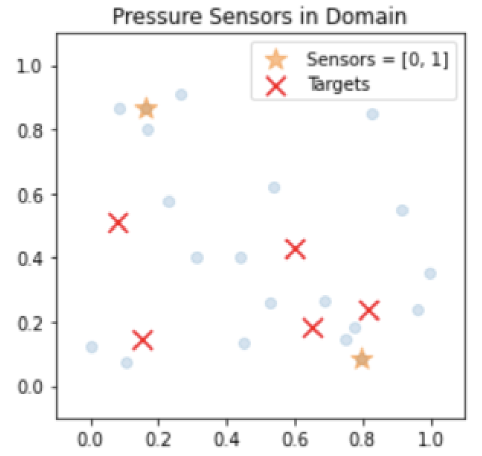


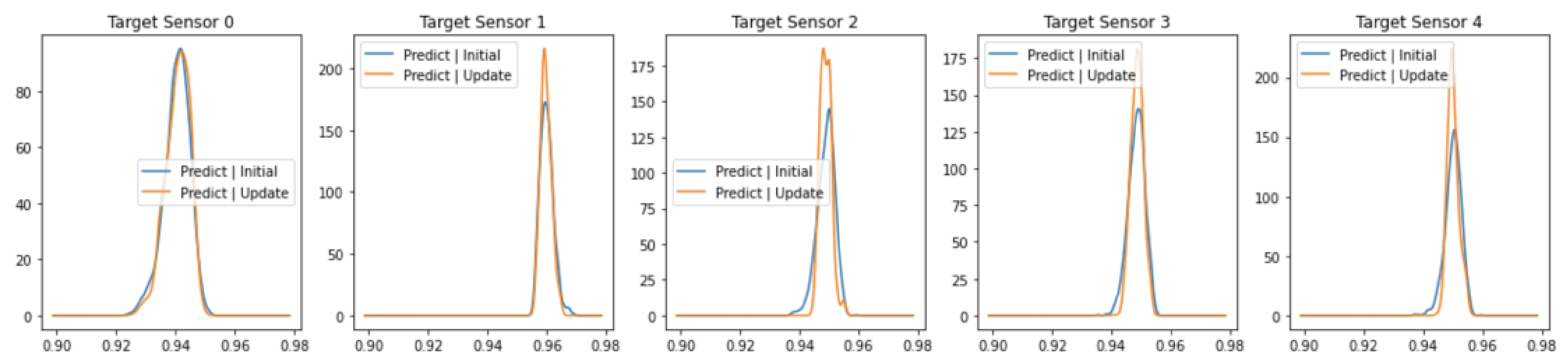
$\pi_{update}(\lambda)$ is a *consistent* solution to the stochastic inverse problem!

Information Gain at Sensors of Interest

Average Information Gain:

- (KL Divergence) = 0.083
- …More information gain with more sensors?





Data-consistent Inversion: Benefits and Drawbacks

Other benefits:

- Generally requires less model evaluations than hier. bayes
- Provides sanity check of predictability assumption ($E(r) = 1$)
- Density estimation in data space rather than parameter space

Some drawbacks:

- Density estimation in data space difficult when $\mathbf{dim}(D)$ is large

The Approach

Utilizing Nonlinear Dimension Reduction (e.g., Manifold Learning)

To Enable Density Estimation for Data-consistent Inversion

Is the data high-dimensional?

The manifold hypothesis states that the dimension of "high-dimensional" data is only superficially large…

- Data lie on a low dimensional manifold embedded in data space $\boldsymbol{D}$

In many cases of interest, a reasonable assumption!

- Multiple measurements made on physical systems likely to have structured correlation determined by physics laws…

Manifold Hypothesis: Consequences for DCI

Suppose there exists a manifold described by $z \in \mathbb{R}^m$, $m \ll \dim(D)$,

Let $f: Z \rightarrow D$ with $f(z) = q$,

$f(z)$ is <u>injective</u> (*man. hyp.*) $\Rightarrow$ $\pi_D(q) = \pi_Z\big(f^{-1}(q)\big) \cdot \det|J^T J|^{-1/2}$

$$\pi^{update}(\lambda) = \pi^{init}(\lambda) \cdot \frac{\pi_D^{obs}(Q(\lambda))}{\pi_D^{pred}(Q(\lambda))} = \pi^{init}(\lambda) \cdot \frac{\pi_Z^{obs}(f^{-1} \circ Q(\lambda))}{\pi_Z^{pred}(f^{-1} \circ Q(\lambda))}$$

Can we find a transformation of $f^{-1}: D \rightarrow Z$?

Manifold Hypothesis: Observations about f^{-1}

$$\pi^{update}(\lambda) = \pi^{init}(\lambda) \cdot \frac{\pi_Z^{obs}(f^{-1} \circ Q(\lambda))}{\pi_Z^{pred}(f^{-1} \circ Q(\lambda))}$$

Goal: find a transformation $f^{-1}: D \rightarrow Z$…

- $\dim Z \ll \dim D$
- Density estimation in Z is easier…
- Leverage *predicted* samples to learn manifold (n obs. can be small!)
- Computation of determinant-Jacobian of f^{-1} is not necessary!

Lots of Options: (*dimension reduction + density estimation*)

- Linear PCA + KDE
- Isomap (nonlinear) + Normalizing Flows

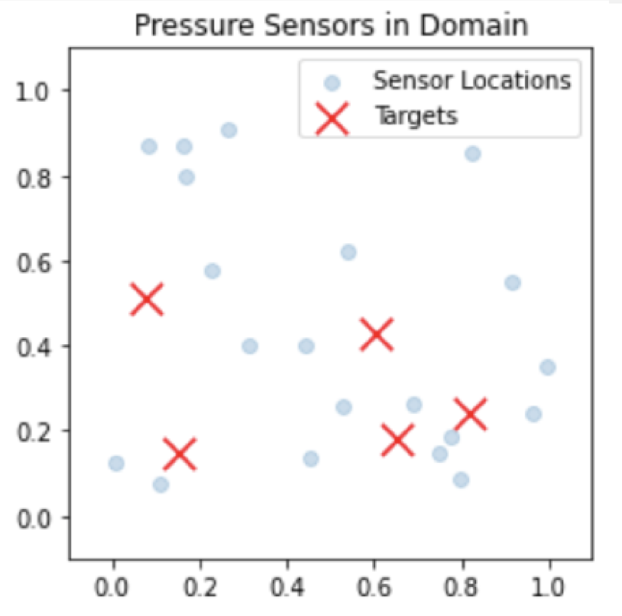
Fluid Flow and Permeability Field Example

Use the observed pressure at all 20 sensors!

- KDEs perform poorly…
- $\bar{r} = 0.000012$

$E(r) = 1$

$$\pi_{update}(\lambda) = \pi_{init}(\lambda)\frac{\pi_{obs}(Q(\lambda))}{\pi_{predict}(Q(\lambda))}$$



General Idea with Linear PCA + KDE

1. Sample initial KL parameters
2. Compute predicted sensor data
3. Perform Linear PCA on predicted data
4. Transform observed data to PC-space
5. Compute KDEs on both observed and predicted PCA data
6. Apply DCI to obtain solution

$$\pi^{update}(\lambda) = \pi^{init}(\lambda) \cdot \frac{\hat{\pi}_Z^{obs}\left(f^{-1} \circ Q(\lambda)\right)}{\hat{\pi}_Z^{pred}\left(f^{-1} \circ Q(\lambda)\right)}$$

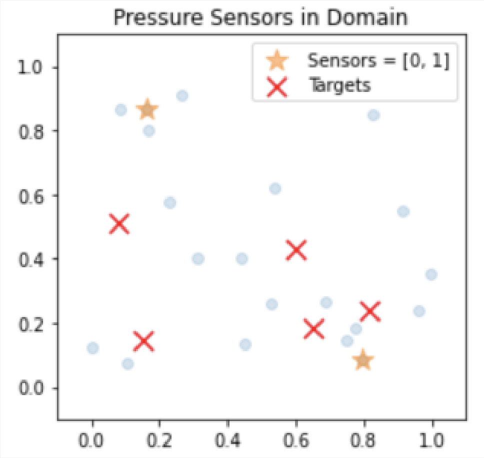
Full Disclosure:

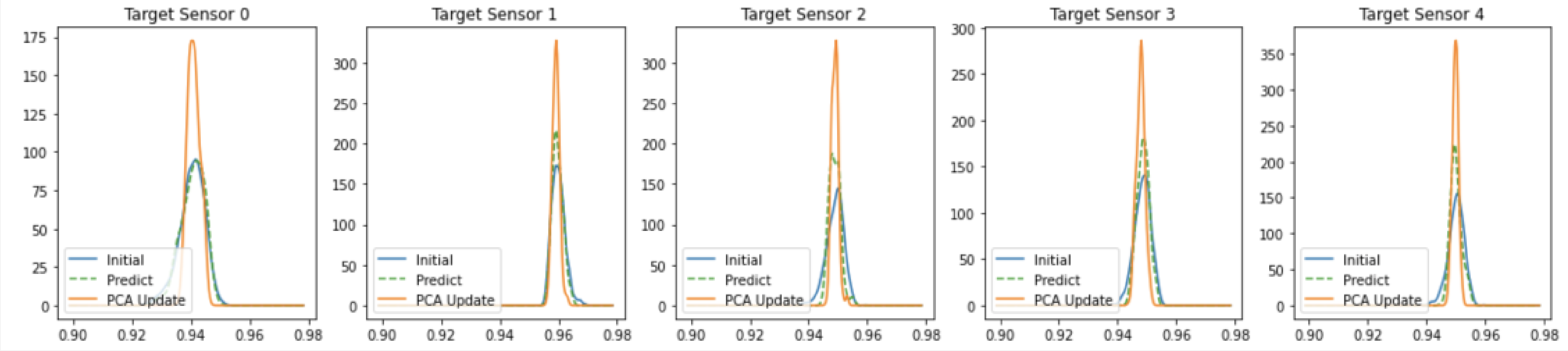
$\bar{r} = 3.074$

Information Gain at Sensors of Interest

Average Information Gain:

- (KL Divergence) = 0.369
- >> first just two sensors (gain = 0.083)

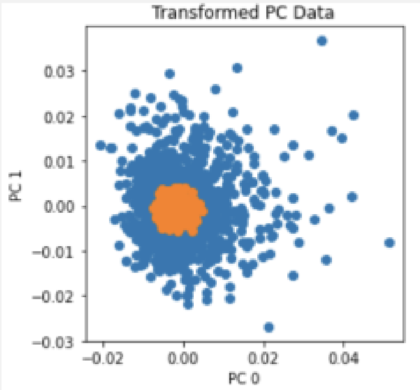




Problems: KDE in Linear PCA space

Data-consistent solution gives good information gain, but…

- *GKDE violates predictability assumption!*
  (though the assump. not violated by data)



Issues:

- 1000 samples insufficient for GKDE with 5 principal components
- Choosing a bandwidth challenging

Isomap + Normalizing Flow

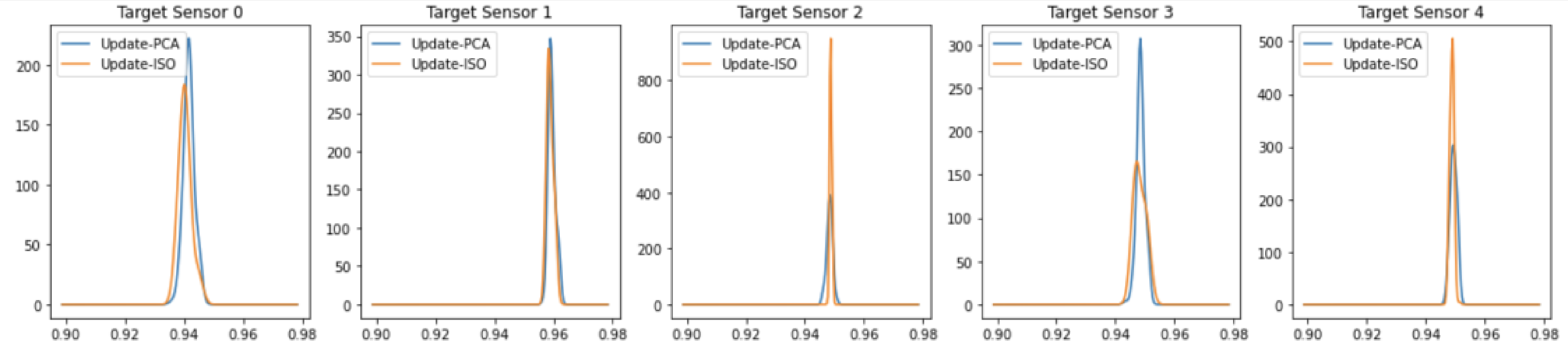
Isomap: nonlinear dimension reduction technique to find low-dim. embedding...

Choose 5 components: explains ~90% of var.

Normalizing Flow: neural network approach to density estimation

Full Disclosure:
$\bar{r} = 0.489$

Results: (KL Divergence) = 0.644



Problems: Isomap + Normalizing Flows

Obtaining reliable density estimates using these ML techniques...

- Requires parameter tuning
- Dependent on the network architecture
- Stochastic optimizers stuck in local minima
- ML techniques require large amounts of samples too

Side Note: PCA + Normalizing Flows
$\bar{r} = 1.001$

Takeaway:

- In theory, any (*dimension reduction* + *density estimation*) can be used in conjunction with data-consistent inversion to find a solution
- In practice, finding a f^{-1} such that the density estimation problem is consistently tractable is difficult
  - Especially, when number of samples is *small!*

Some Reflections

Future Work and Analysis

Conclusions

1. Data-consistent Inversion can *efficiently* solve stochastic inverse problems with high-dimensional data $(\mathbf{dim(D)})$ large)…
   a) *When there exists low-dimensional manifold…*
   b) *When we can find a reasonable manifold (**dimension reduction**)…*
   c) *When we can approximate the density (**density estimation**) on the manifold…*

2. Many new cutting-edge techniques for tackling *b)* and *c)*, which should we choose and when?

3. What is a sufficient sampling size to obtain reliable solutions?