USC University of Southern California

Bayesian Inference in Physics-Driven Problems with Adversarial Priors

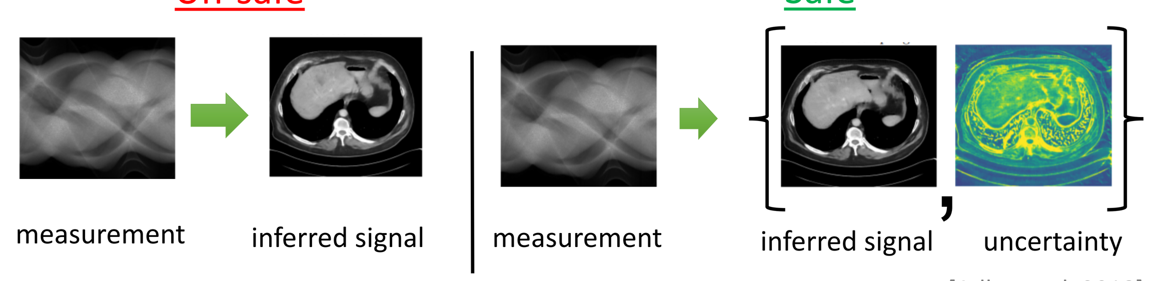
Dhruv V. Patel, Deep Ray, Harisankar Ramaswamy and Assad A. Oberai
{dhruvvpa, deepray, hramaswa, aoberai}@usc.edu

Motivation

Goal: Given $\hat{\boldsymbol{y}} = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{\eta}$, infer $\boldsymbol{p}(\boldsymbol{x}|\hat{\boldsymbol{y}})$.

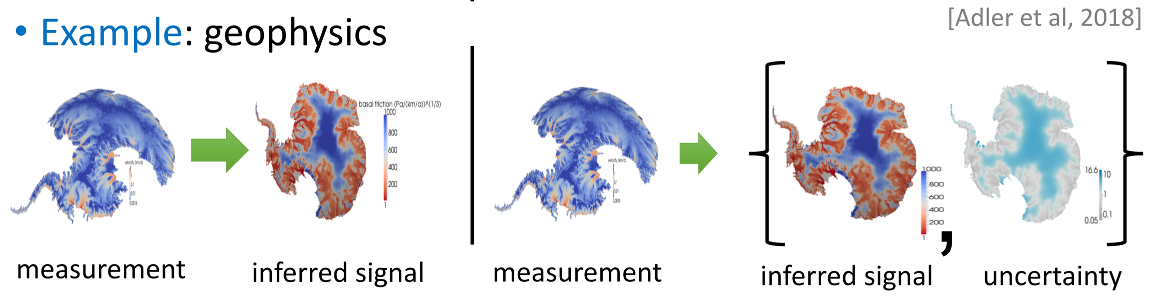
Challenges:
- Measurement may be corrupted by unknown noise.
- Inverse map may not be well-posed.
- Uncertainty in inferred solution critical for applications with high-stake decisions.

Example: medical imaging



[Adler et al, 2018]

- Example: geophysics



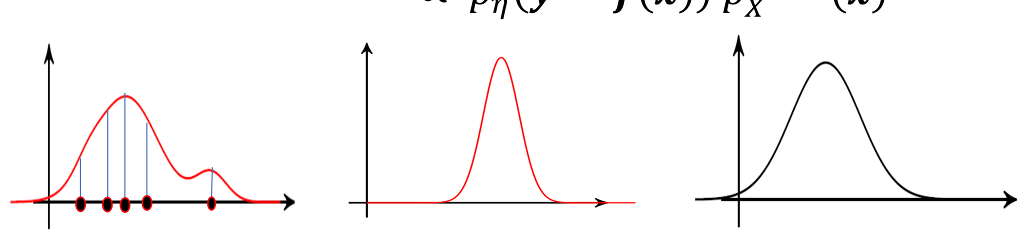
[Isaac et al, 2015]

Bayesian inference

- A principled approach to account for uncertainty in an inverse problem.
- Gives probability distribution over inferred field given some measurement.

Posterior distribution

$$p_X^{post}(\boldsymbol{x}|\hat{\boldsymbol{y}}) = \frac{1}{Q} p^{like}(\hat{\boldsymbol{y}}|\boldsymbol{x})\, p_X^{prior}(\boldsymbol{x})$$
$$\propto p_\eta(\hat{\boldsymbol{y}} - \boldsymbol{f}(\boldsymbol{x}))\, p_X^{prior}(\boldsymbol{x})$$



Challenge I : Priors

Finding a quantitative description of informative and feasible priors.

Typical priors..

$$p^{prior}(\boldsymbol{x}) = exp\left(-\frac{1}{\sigma^2}\|\boldsymbol{x}\|^2\right)$$

However, what if..
- the prior knowledge is more complex and difficult to characterize analytically.
- not enough domain knowledge is available to construct informative priors.

Challenge II : Sampling
- Inferred signal is high dimensional (10^3-10^7).
- Difficult to sample from high dimensional posterior space using sampling-based methods like MCMC.
- An efficient sampler is difficult to design in high-dimension.

GAN as Prior

Key idea: *Use the distribution learned by GAN as a surrogate for prior distribution and reformulate the inference problem in the low-dimensional latent space of the GAN.*

$$\underset{\boldsymbol{x}\sim p_X^{post}}{\mathbb{E}}[m(\boldsymbol{x})] = \frac{1}{Q}\underset{\boldsymbol{x}\sim p_X^{prior}}{\mathbb{E}}[m(\boldsymbol{x})p_\eta(\hat{\boldsymbol{y}} - \boldsymbol{f}(\boldsymbol{x}))]$$
$$= \frac{1}{Q}\underset{\boldsymbol{x}\sim p_X^{data}}{\mathbb{E}}[m(\boldsymbol{x})p_\eta(\hat{\boldsymbol{y}} - \boldsymbol{f}(\boldsymbol{x}))]$$
$$= \frac{1}{Q}\underset{\boldsymbol{z}\sim p_Z}{\mathbb{E}}[m(\boldsymbol{g}(\boldsymbol{z}))p_\eta(\hat{\boldsymbol{y}} - \boldsymbol{f}(\boldsymbol{g}(\boldsymbol{z})))]$$
$$= \frac{1}{Q}\underset{\boldsymbol{z}\sim p_Z^{post}}{\mathbb{E}}[m(\boldsymbol{g}(\boldsymbol{z}))]$$

where,

$$p_Z^{post} \propto p_\eta\left(\hat{\boldsymbol{y}} - \boldsymbol{f}(\boldsymbol{g}(\boldsymbol{z}))\right)p_Z(\boldsymbol{z})$$

Steps:
1. Learn the prior distribution: train a GAN using samples from data distribution $p_X^{data}(\boldsymbol{x})$.
2. Characterize the posterior distribution: for a given measurement $\hat{\boldsymbol{y}}$, evaluate any statistic of interest $\underset{\boldsymbol{x}\sim p_X^{post}}{\mathbb{E}}[m(\boldsymbol{x})]$.

MCMC Sampling: $p_Z^{MCMC}(\boldsymbol{z}|\hat{\boldsymbol{y}}) \approx p_Z^{post}(\boldsymbol{z}|\hat{\boldsymbol{y}})$.

Evaluate any point estimate $s(\boldsymbol{x})$ as,

$$s(\boldsymbol{x}) \approx \frac{1}{N}\sum_{n=1}^{N} s(\boldsymbol{g}(\boldsymbol{z})), \qquad \boldsymbol{z} \sim p_Z^{MCMC}(\boldsymbol{z}|\hat{\boldsymbol{y}})$$

We use Hamiltonian Monte Carlo (HMC) with burn-in period of 0.5

Convergence of GAN to prior

Assumptions:
- The GAN discriminator $\boldsymbol{d}(\boldsymbol{x};\boldsymbol{\phi})$ is smooth.
- Derivates of $\boldsymbol{d}$ with respect to its weights $\boldsymbol{\phi}$ form a dense subset of $C_b(\Omega_x)$.
- Number of training samples N and number of weights N_ϕ are large enough.

Convergence results: *given any* $m \in C_b(\Omega_x)$,

$$\underset{\boldsymbol{z}\sim p_Z}{\mathbb{E}}\left[\left(\underset{\boldsymbol{z}\sim p_X}{\mathbb{E}}[m(\boldsymbol{x})] - \frac{1}{N}\sum_{n=1}^{N} m\left(\boldsymbol{g}(\boldsymbol{z}^{(n)})\right)\right)^2\right] < C\left(N^{-1} + (N_\phi)^{-2\alpha}\right)$$

where C, α are positive constants depending on m.

Results

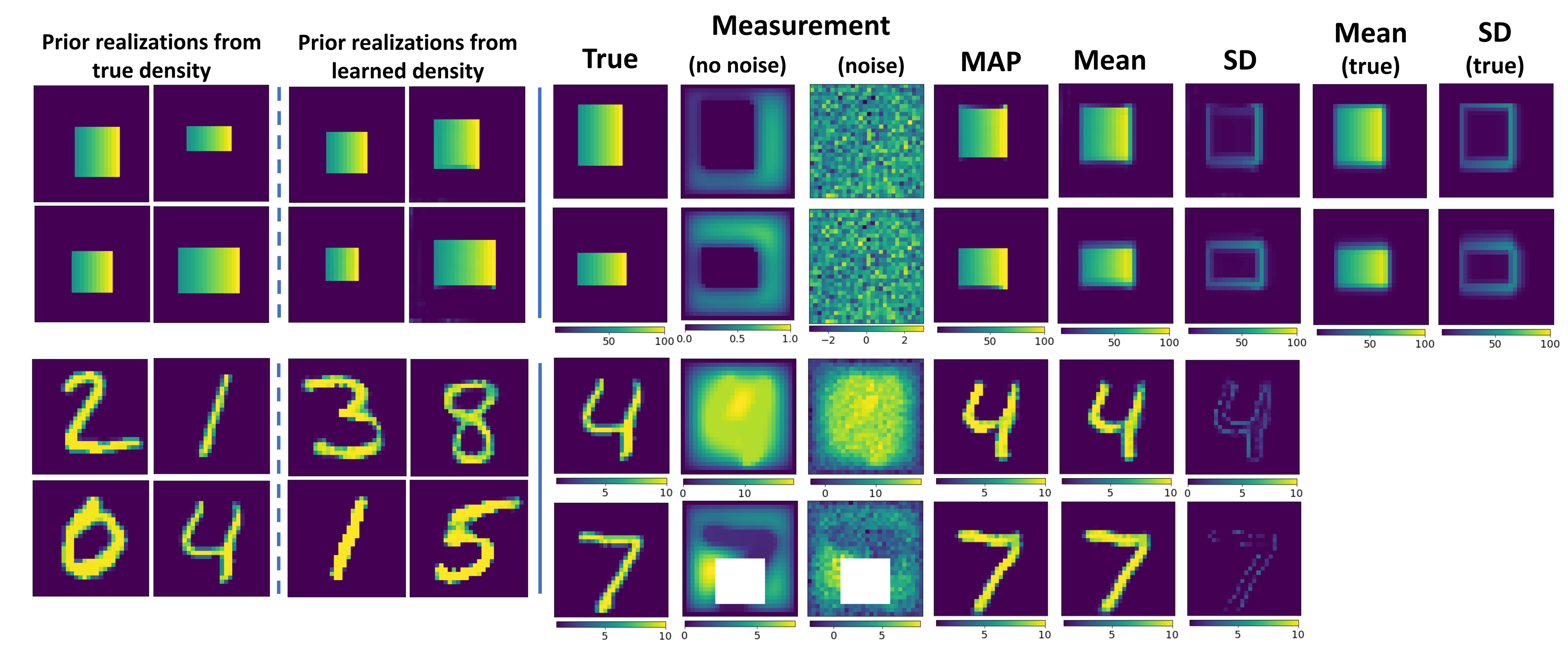
Inverse heat conduction:

$$-\nabla.(\boldsymbol{x}(s)\nabla \boldsymbol{y}(s))\big|_\Omega = \boldsymbol{b}(s)$$
$$\boldsymbol{y}(s)|_{\partial\Omega} = 0$$

$\boldsymbol{x}$:= thermal conductivity

$\boldsymbol{y}$:= temperature

$\boldsymbol{b}$:= heat source



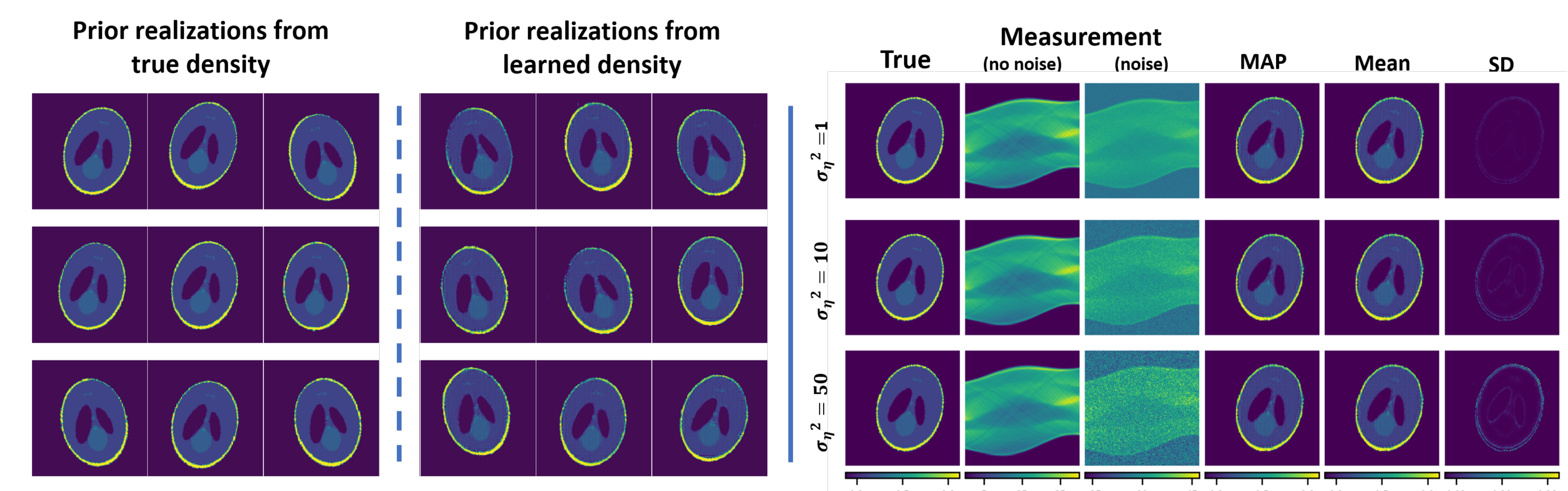
Inverse Radon transform:

$$\boldsymbol{y}_{i,j} = \int_{l(t_i,\psi_j)} \boldsymbol{x}\, \boldsymbol{dl}$$

$\boldsymbol{x}$:= image

$\boldsymbol{y}$:= sinogram

$\boldsymbol{l}_{(t_i,\psi_j)}$:= ray/line

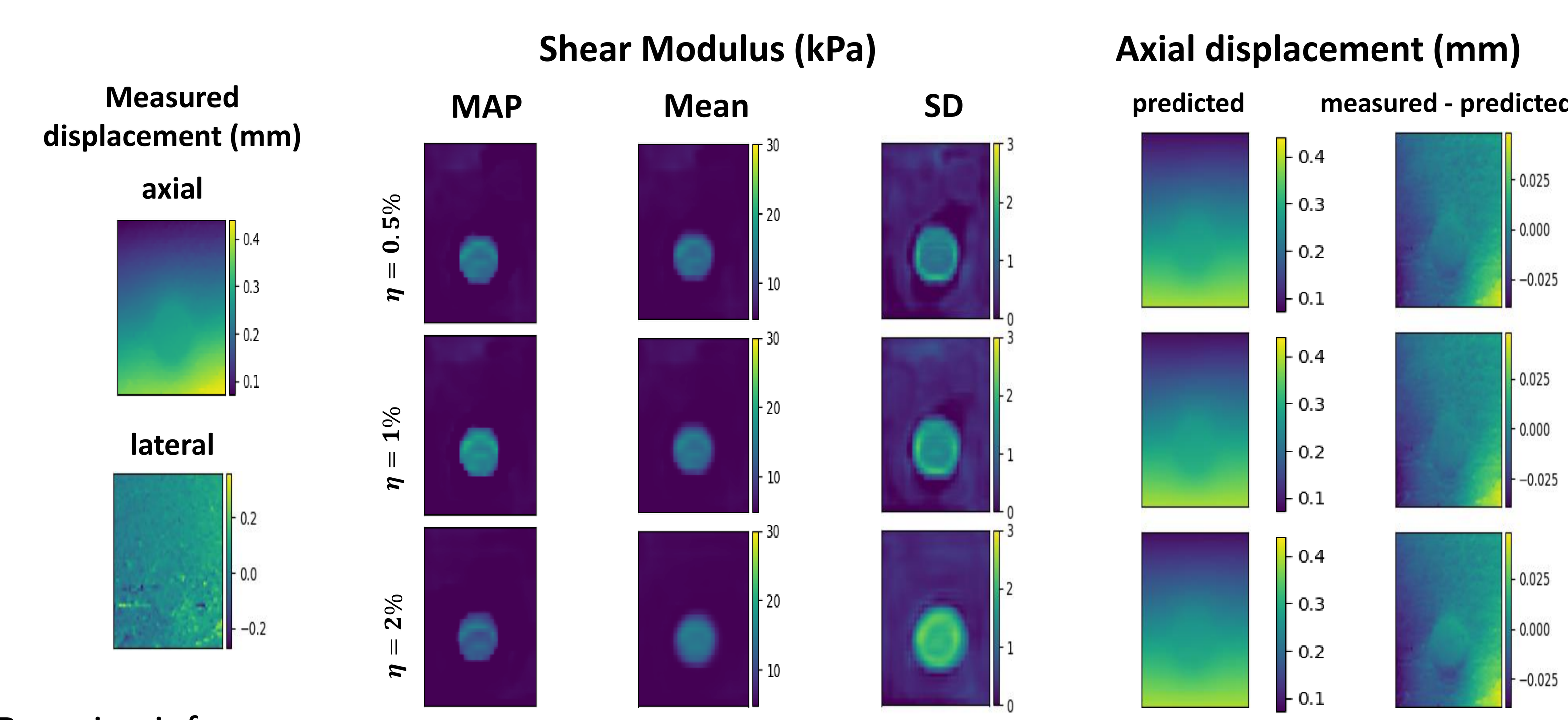


Elasticity imaging: (using experimental data)

$$\nabla.\boldsymbol{\sigma}\big|_\Omega = \boldsymbol{0}$$
$$\boldsymbol{y}\big|_{\Gamma_D} = \boldsymbol{y}_D$$
$$\boldsymbol{\sigma}.\boldsymbol{n}\big|_{\Gamma_N} = \boldsymbol{\tau}$$
$$\boldsymbol{\sigma} = 2\boldsymbol{x}(\nabla^s\boldsymbol{y} + (\nabla.\boldsymbol{y})\boldsymbol{I}) \coloneqq \text{plane-stress}$$

$\boldsymbol{x}$:= shear modulus

$\boldsymbol{y}$:= displacement field



Key takeaways:

- ✓ GANs trained to learn complex priors in Bayesian inference.
- ✓ Significant reduction in dimension for efficient sampling .
- ✓ Convergence estimate of GAN showing decay of estimated statistics with the size of training set and network.
- ✓ Demonstration of method of problems motivated by physics, in the presence of noisy and occluded measurements.

Acknowledgement: The support from ARO grant W911NF2010050 is acknowledged.

Paper: D. Patel, D. Ray, H. Ramaswamy, A. Oberai, "*Bayesian Inference in Physics-Driven Problems with Adversarial Priors*", Deep inverse workshop (NeurIPS 2020): https://openreview.net/pdf?id=P-0ae-EbP8