Bayesian inference with Generative Adversarial Network (GAN) Priors

USC University of Southern California

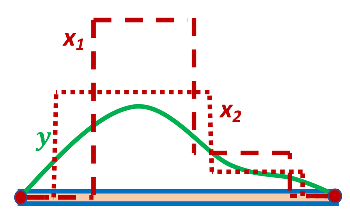
Dhruv Patel and Assad A Oberai

Computation and Data-Driven Discovery group

Bayesian Inference

Anatomy of an inverse problem

$f(x) + n = y$



Forward problem
- Well-posed
  - solution exists
  - solution is unique
  - stable w.r.t. perturbation
- Causal
- Local

Inverse problem
- Ill-posed
  - not meeting well-posedness requirement
  - sparse observations
- Non-causal
- Global

How to tackle ill-posedness?

1.Classical (regularization) approach:
Formulate as an optimization problem.

$$\min_{x} \frac{1}{2}\|f(x) - \hat{y}\|_w^2 + \|x - x_{ref}\|_R^2$$

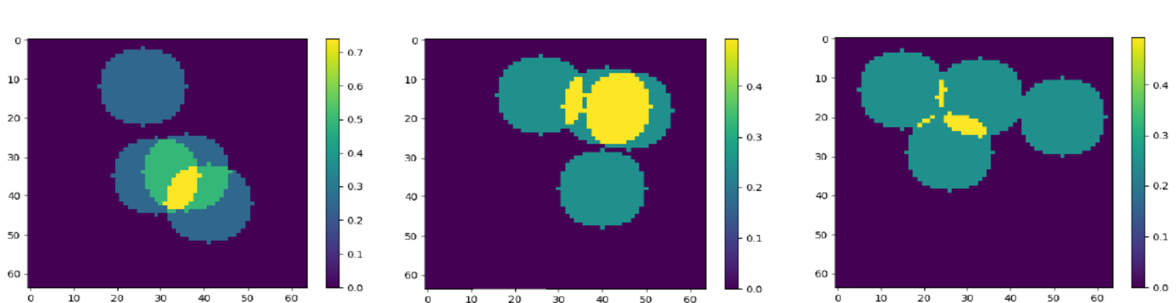
2.Bayesian approach:
Statistical framework to characterize distribution of parameters given some noisy version of measurement.
Posterior distribution

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

Challenge I : Priors
- Finding a quantitative description of informative and feasible priors.
- Typical priors

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

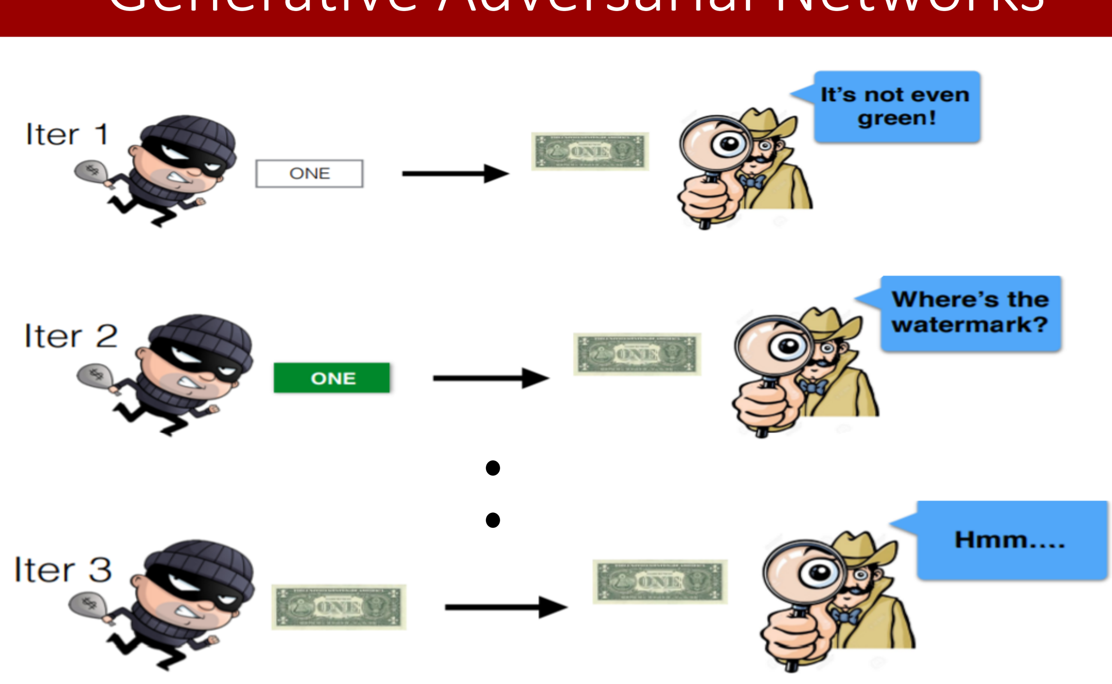
- However, what if the prior knowledge may be something like…



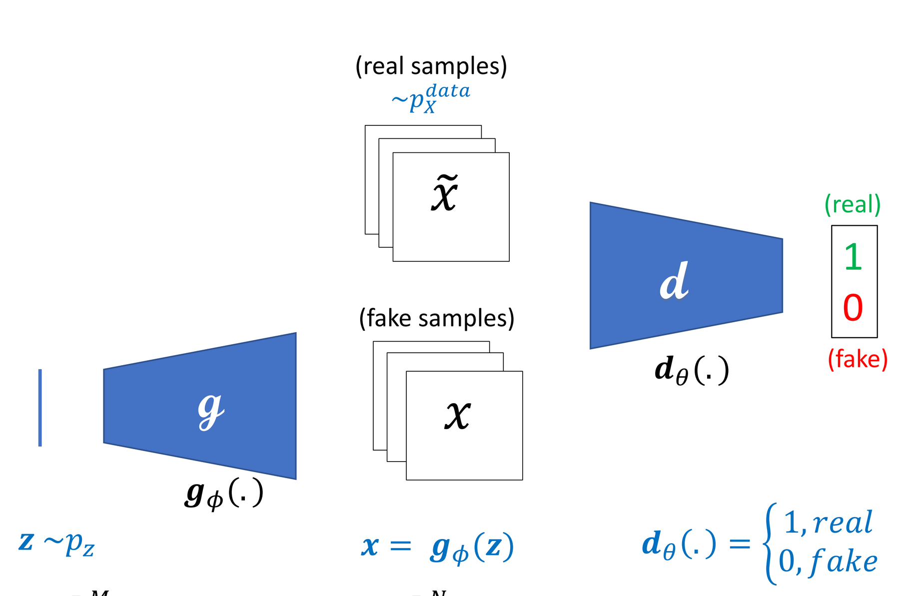
Challenge II : Sampling
- Typical physics-driven inverse problem contain large no. of parameters (10^4-10^7).
- MCMC approximation of posterior requires many samples from posterior, where each sample involves the solution of PDE.
- An efficient sampler is difficult to design for high-dimensional parameter space.

Central Idea
- Use deep generative model (GAN)s as *priors* in Bayesian inference by learning the parameter distribution directly from data.
- Demonstrate GANs as a tool to reduce the dimensionality of parameter space for *efficient posterior sampling*.
- Use the quantified uncertainty information for *optimal design of experiments* leading to efficient parameter inference.

Generative Adversarial Networks



GAN, the two-player min-max game



➢Weights of g and d are obtained by solving min-max problem:

$$\phi^*, \theta^* = \arg\min_{\phi} \arg\max_{\theta} \left[\mathop{\mathbb{E}}_{x \sim p_X^{data}}[\log d_\theta(x)] + \mathop{\mathbb{E}}_{z \sim p_Z}[1 - \log d_\theta(g_\phi(z))] \right]$$

➢For a network with infinite capacity and adequate training time:

$$\mathop{\mathbb{E}}_{x \sim p_X^{data}}[m(x)] = \mathop{\mathbb{E}}_{x \sim p_X^{gen}}[m(x)] = \mathop{\mathbb{E}}_{z \sim p_Z}[m(g(z))]$$

GAN as Priors

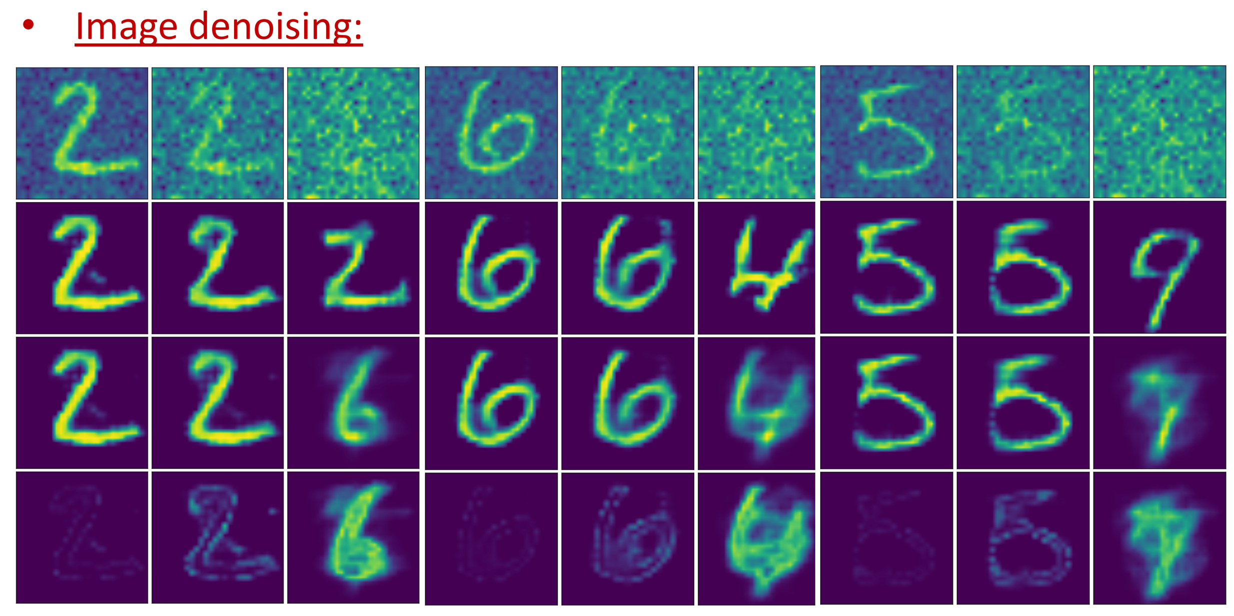
$$\mathop{\mathbb{E}}_{x \sim p_X^{post}}[m(x)] = \frac{1}{\mathbb{Z}} \mathop{\mathbb{E}}_{x \sim p_X^{prior}}[m(x) p_\eta(\hat{y} - f(x))]$$
$$= \frac{1}{\mathbb{Z}} \mathop{\mathbb{E}}_{x \sim p_X^{data}}[m(x) p_\eta(\hat{y} - f(x))]$$
$$= \frac{1}{\mathbb{Z}} \mathop{\mathbb{E}}_{z \sim p_Z}[m(g(z))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

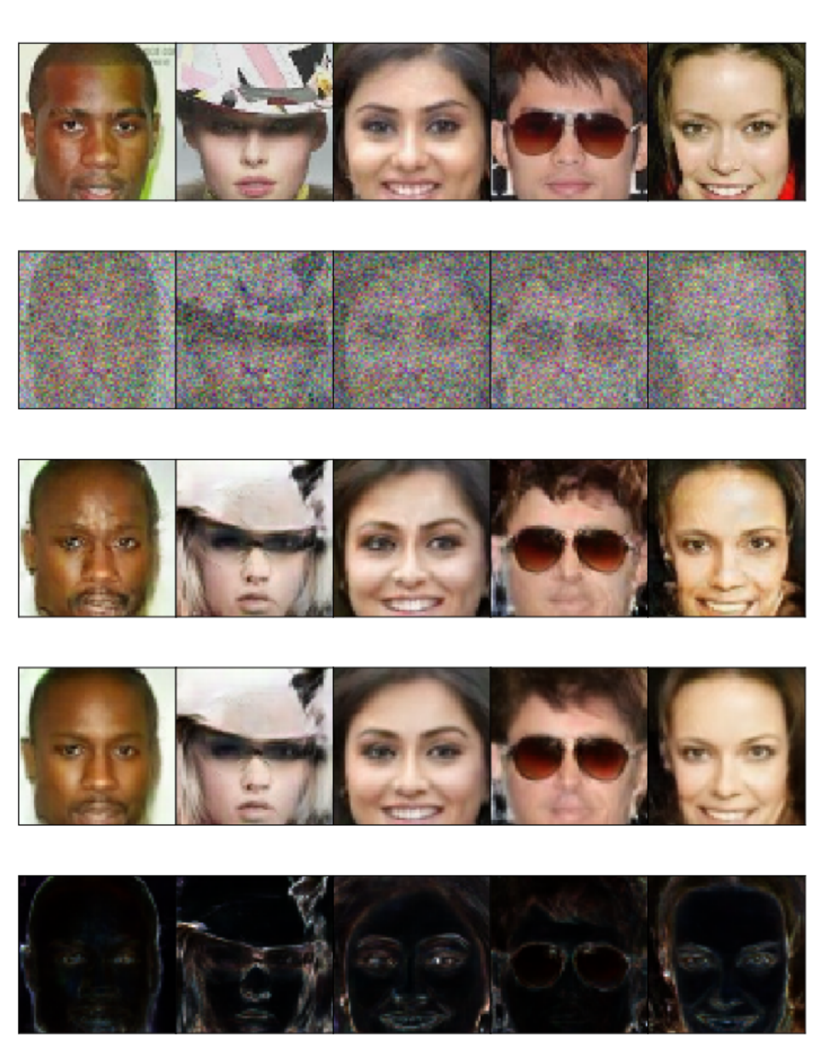
where,

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

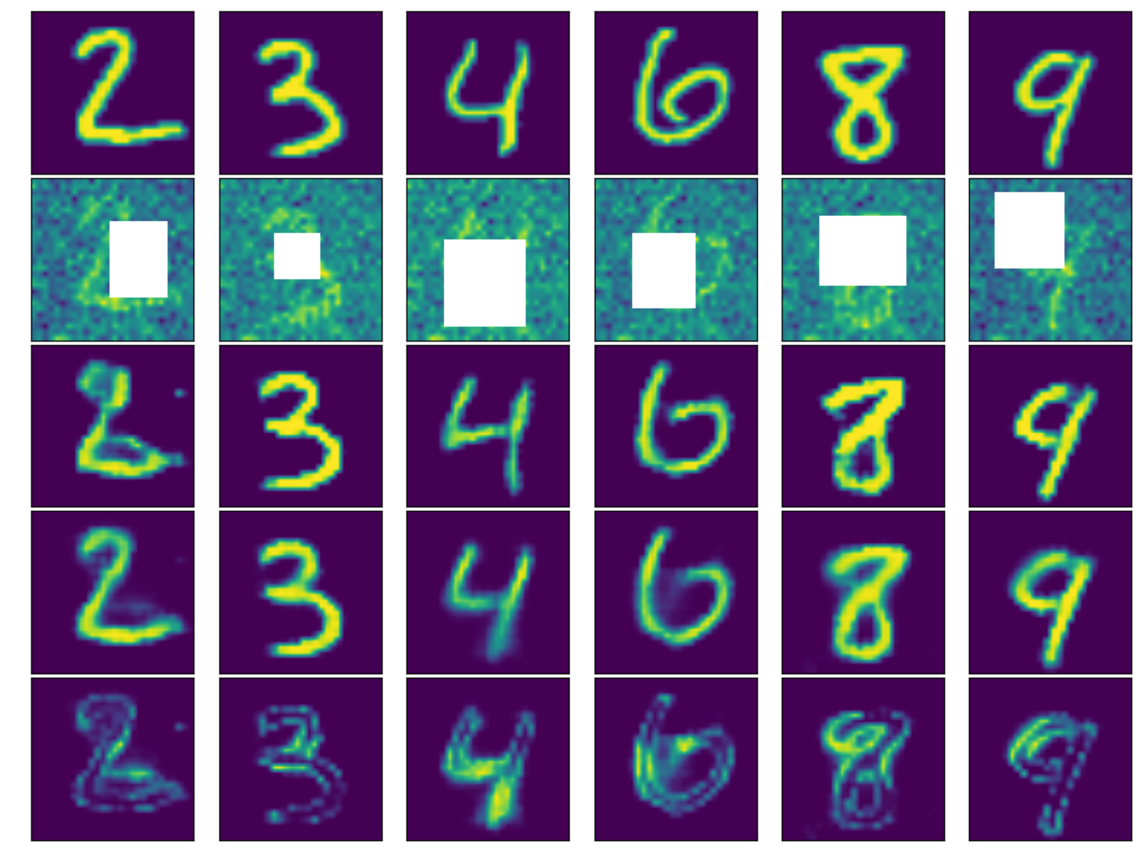
Experimental Results

- Image denoising:

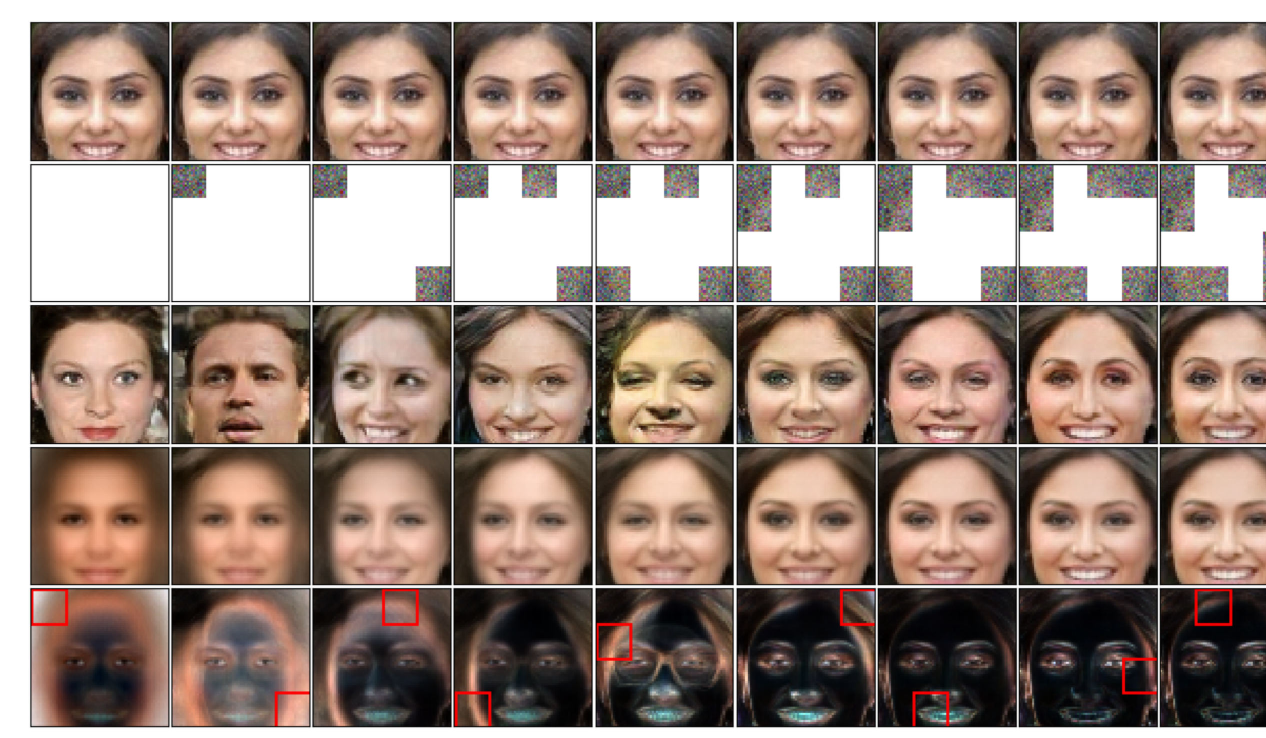




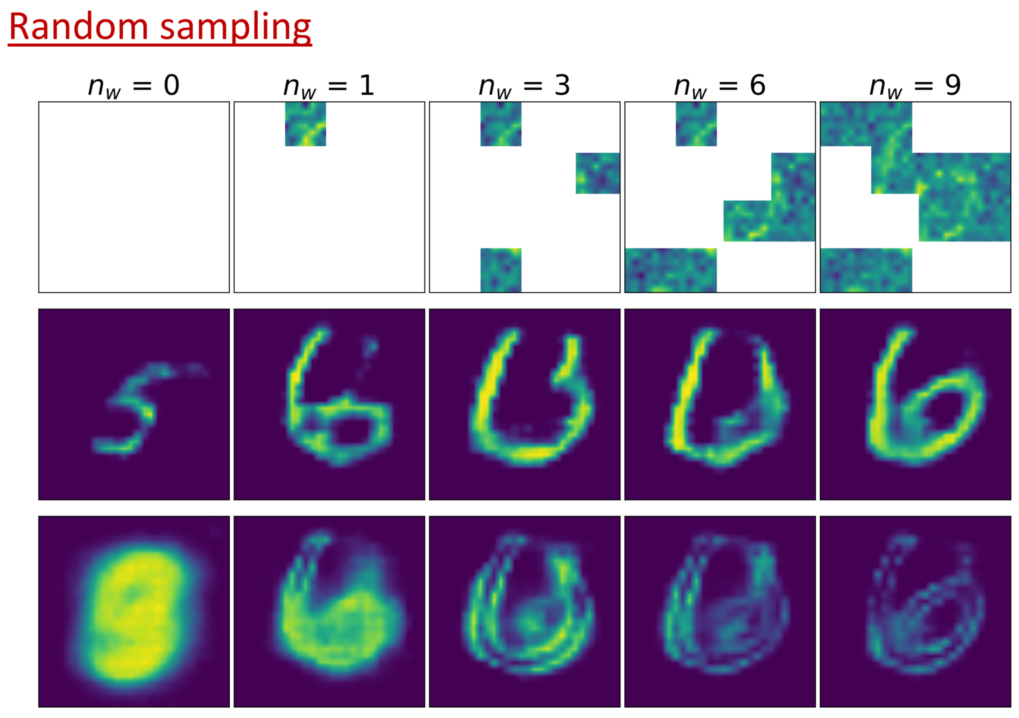
- Image inpainting + denoising:



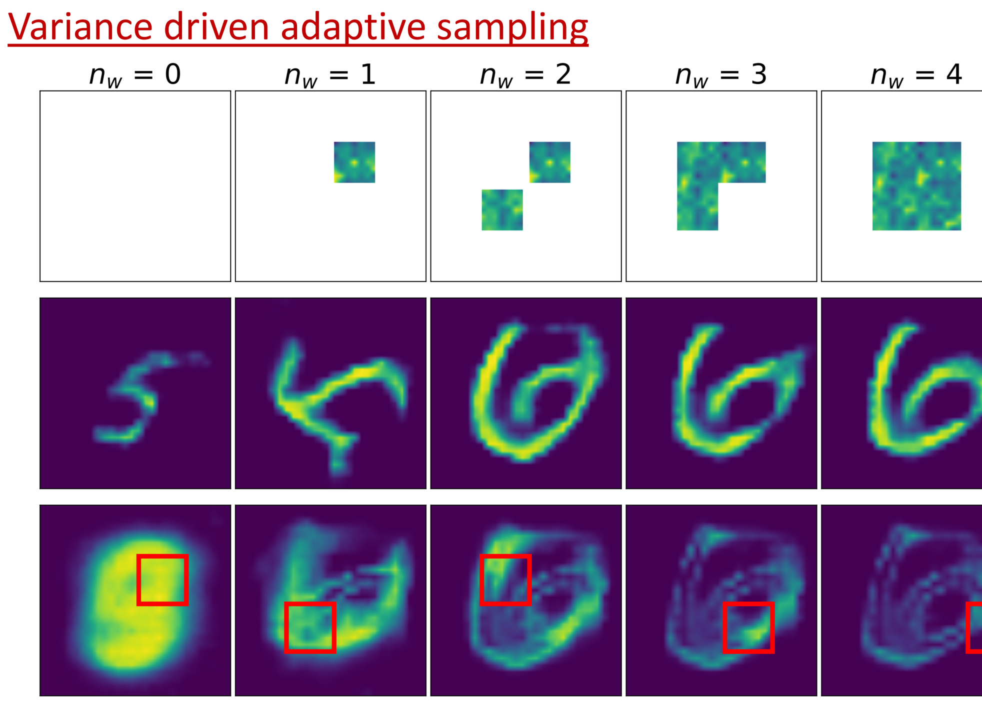
- Optimal experimental design / Active learning

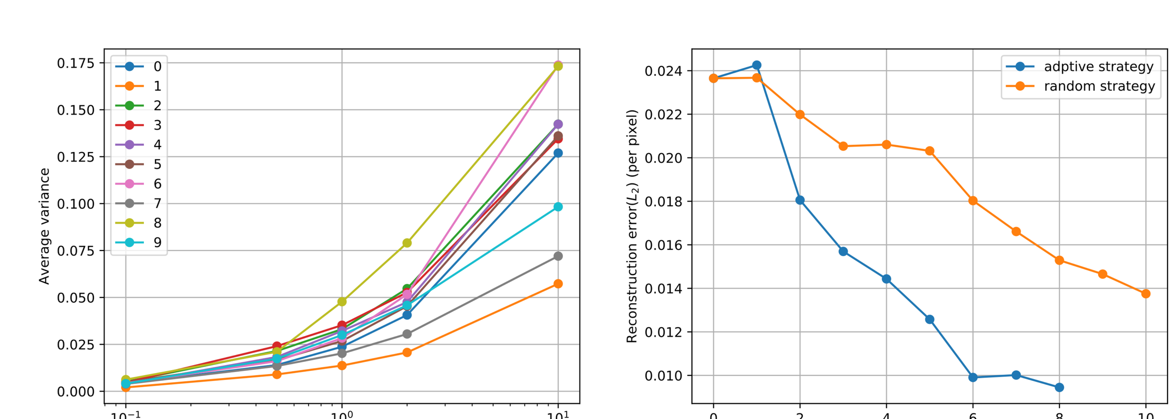


- Random sampling

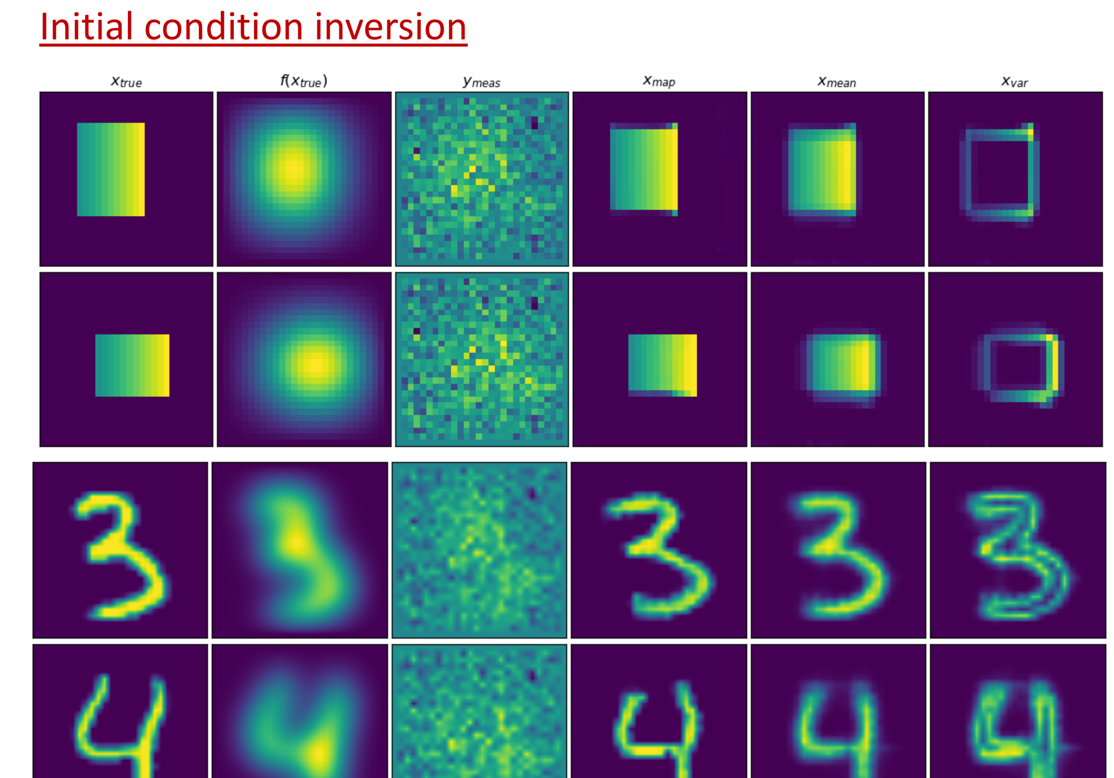


- Variance driven adaptive sampling





- Initial condition inversion



Reference

Dhruv Patel, Assad Oberai, **Bayesian Inference with Generative Adversarial Network Priors** *arxiv :1907.09987[stat.ML].*