How to sample from a posterior like you sample from a prior

Goal-Oriented UQ for Inverse Problems via VEDs

AIP 2023, Göttingen

BMA (DTU), Julianne & Matthias Chung (Emory)

Inverse Problems

- Inverse problem formulation

$$\mathbf{b} = F(\mathbf{y}) + \mathbf{e}$$

➡ $\mathbf{b} \in \mathbb{R}^m$ is the measurement

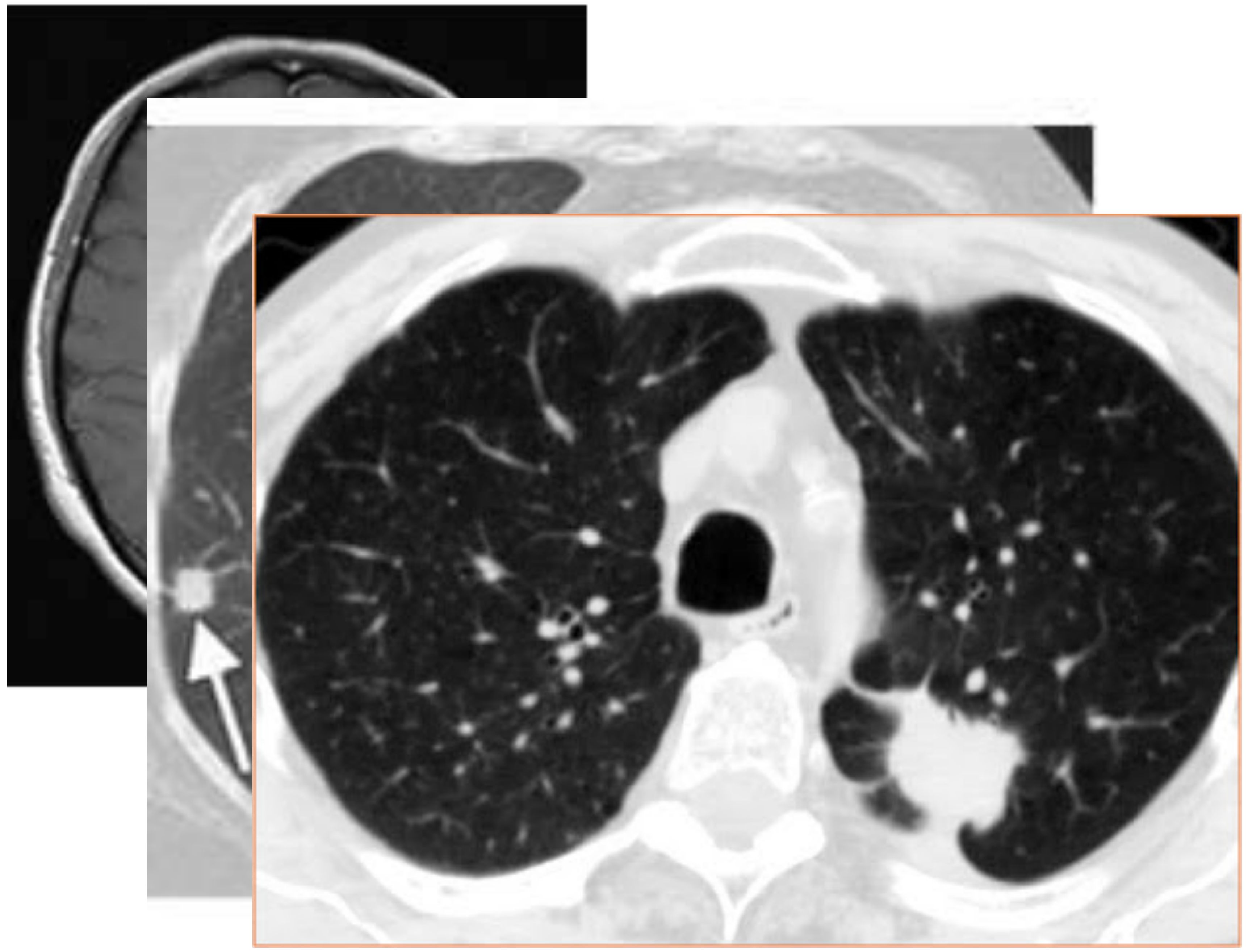
➡ $\mathbf{y} \in \mathbb{R}^n$ is the unknown true parameters

➡ $F : \mathbb{R}^n \to \mathbb{R}^m$ is the parameter to observation map

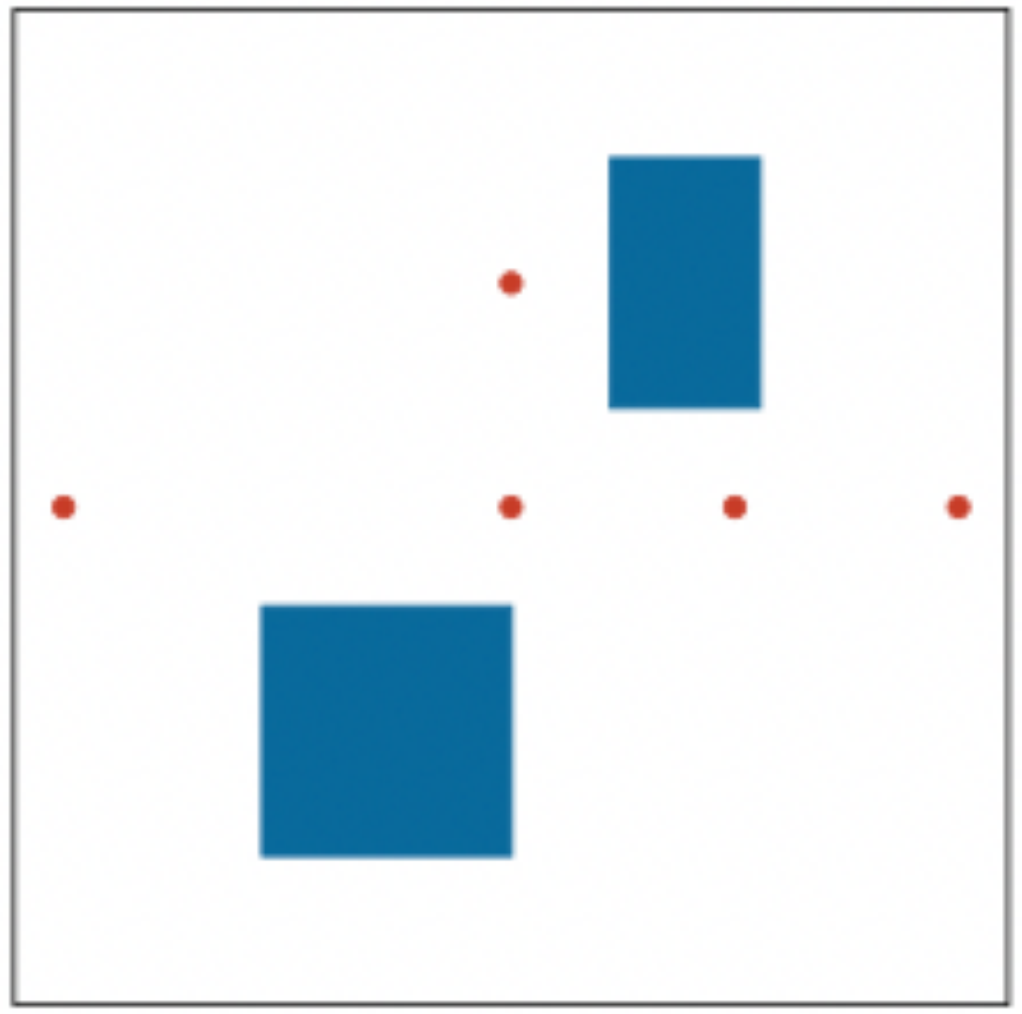
➡ $\mathbf{e} \in \mathbb{R}^m$ is the additive noise

Quantity of Interest in an Inverse Problem

CT:

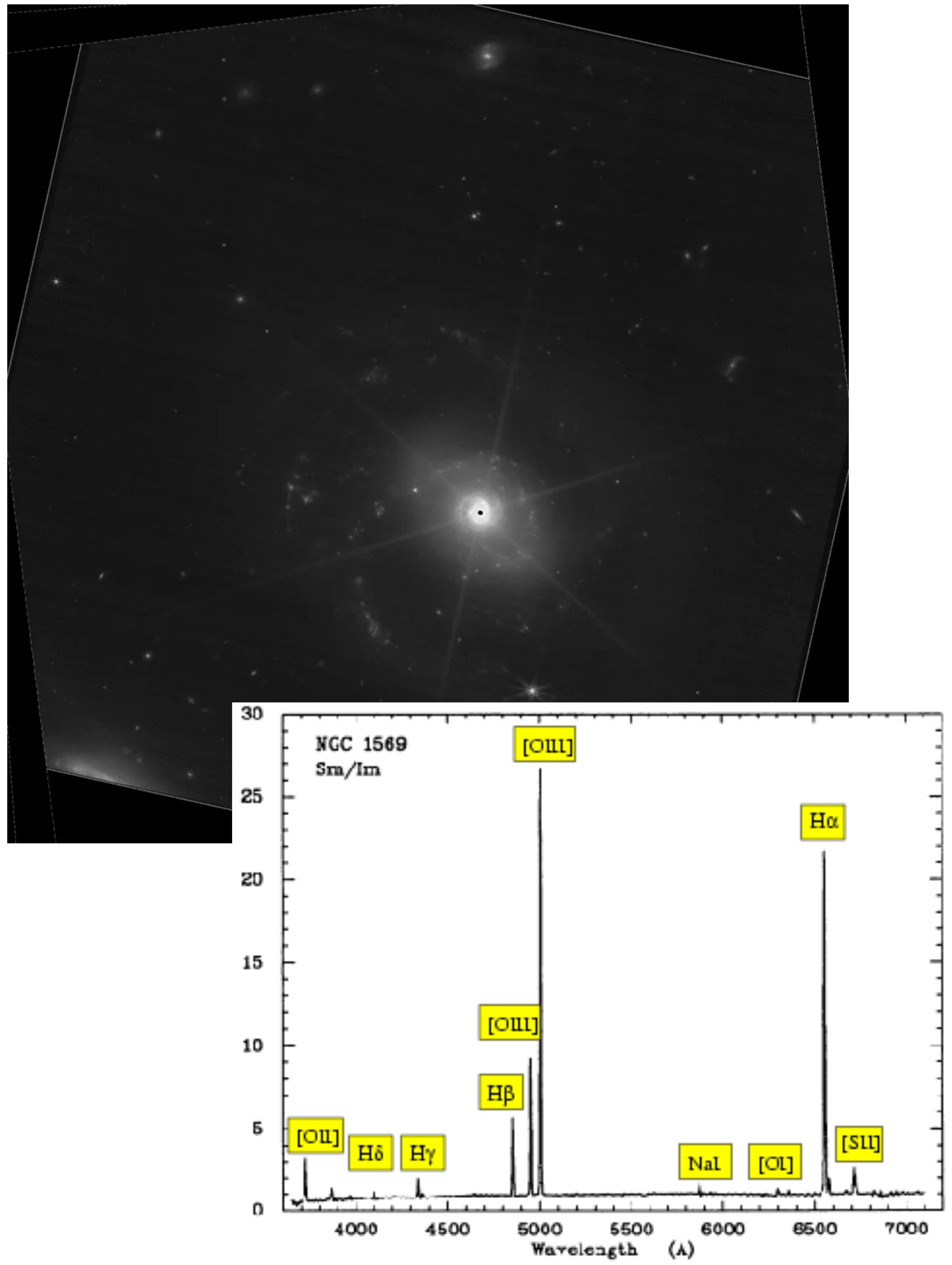


optimal design:



A. Attia, et al. (2018)

de-blurring:



Goal-Oriented Inverse Problems (goIP)

- Quantity of interest (according to the current literature):
  - maximum, minimum, average, integral, …
  - Not even well-defined, e.g. expert's opinion
- In many applications we are only interested in QoI

$$\mathbf{b} = F(\mathbf{y}) + \mathbf{e}$$

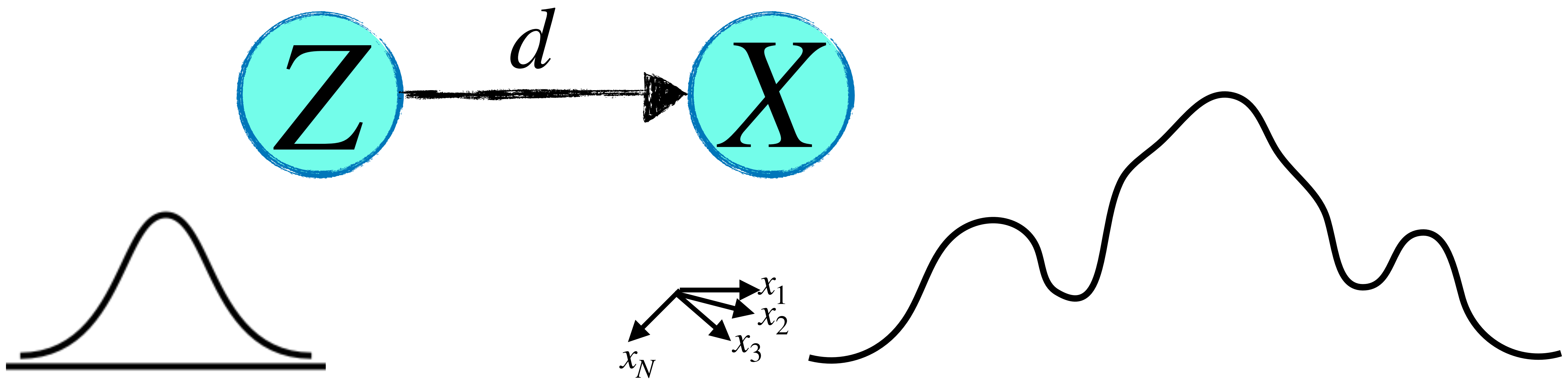
$$\mathbf{x} = G(\mathbf{y})$$

➡ $G : \mathbb{R}^n \to \mathbb{R}^q$ prediction operator

➡ $\mathbf{x}$ is low dimensional, dim$(\mathbf{x}) \ll$ dim$(\mathbf{y})$

Latent Variables

- Suppose that we we find a hidden random variable Z that describes X:



MS51 2:

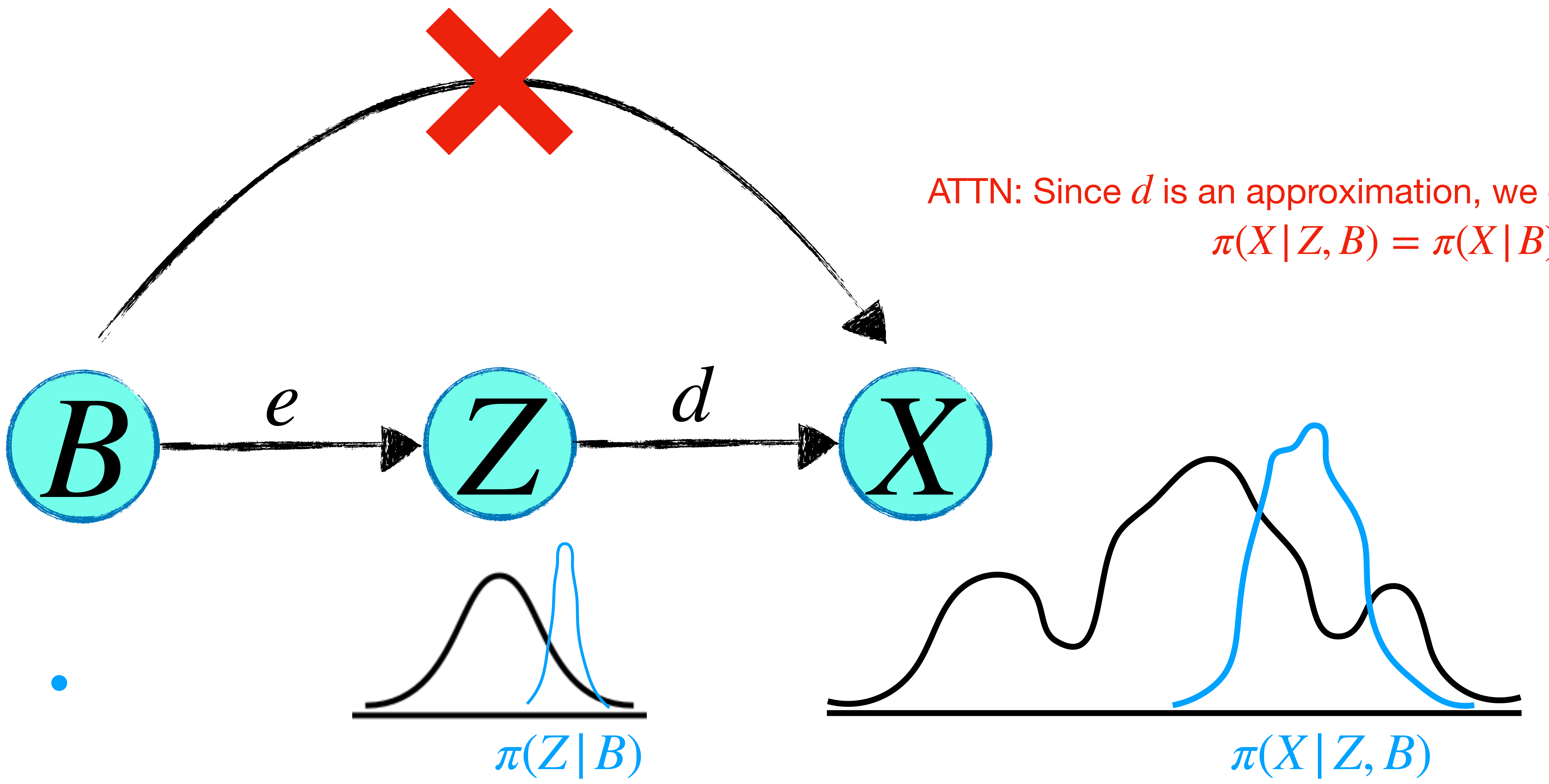
VG1.108

17:00

$$\pi(X|Z) = \delta$$

Latent variables and Inverse Problems

- Now suppose that we have an observation $\mathbf{b}$



ATTN: Since d is an approximation, we cannot ensure that
$\pi(X|Z,B) = \pi(X|B)$

Bayesian Assumption

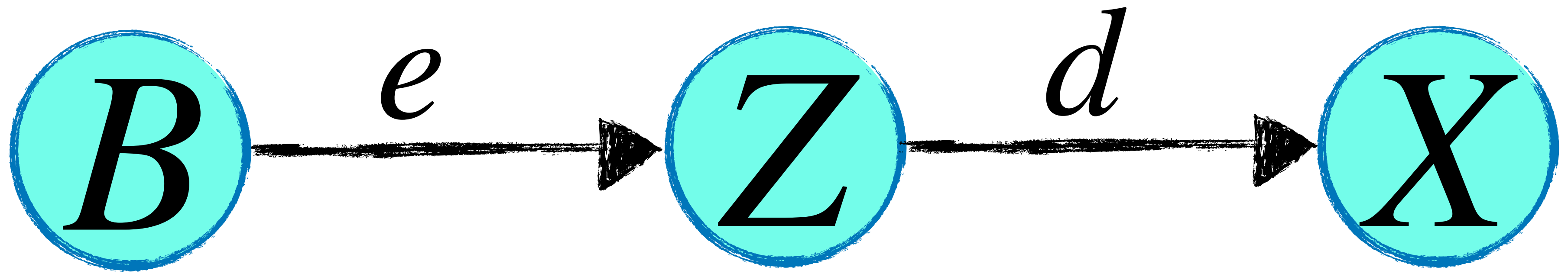
- We assume that Z parameterizes X such that X and Z are indistinguishable to B, i.e.,

$$\pi(B \mid X) = \pi(B \mid X, Z) = \pi(B \mid Z)$$

Bayesian Assumption and the Latent Variable

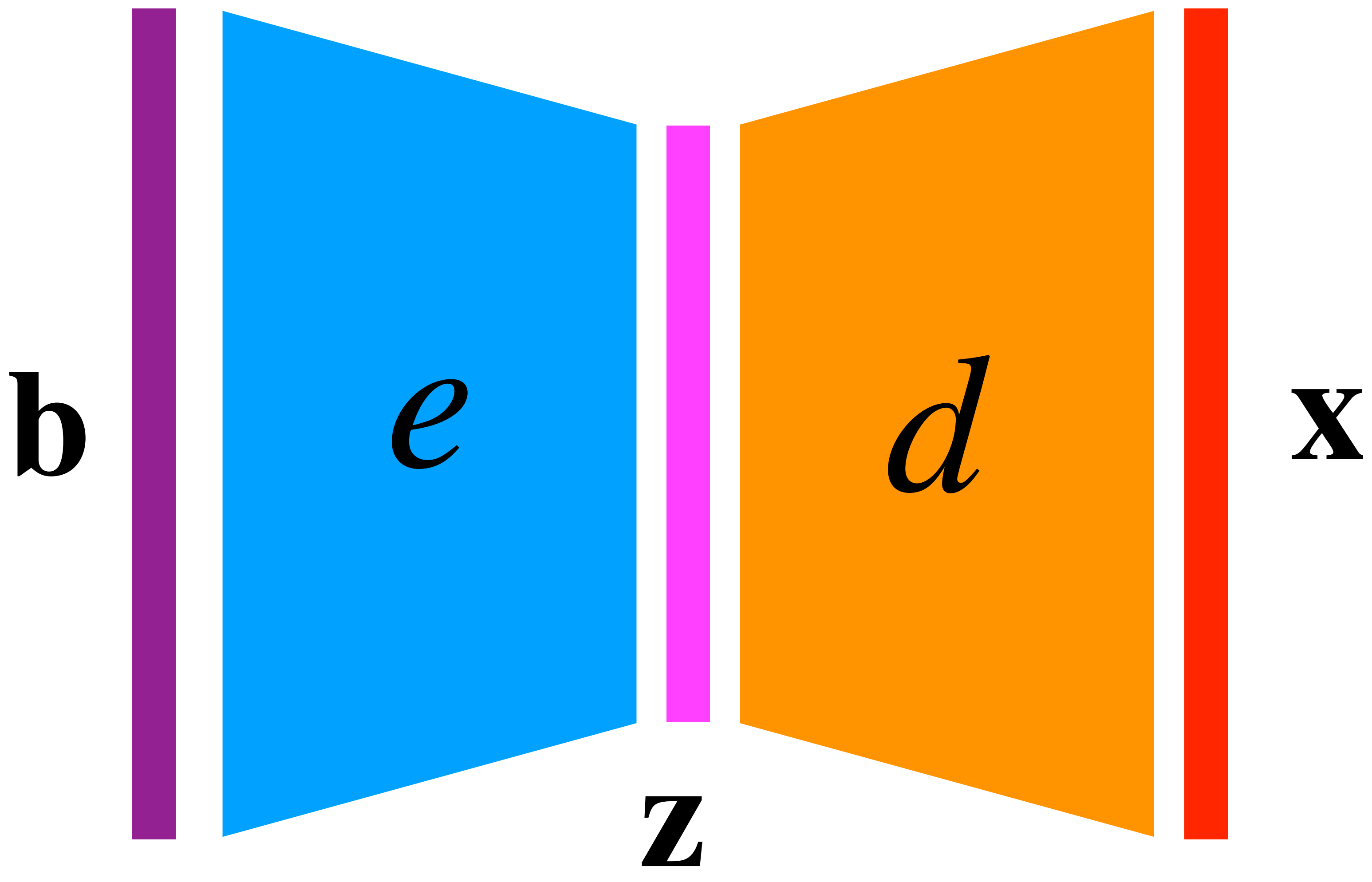
- Proposition:

$$\pi(X|B,Z) = \frac{\pi(Z|X)}{\pi(Z|B)}\underbrace{\frac{\pi(Z|X)}{\pi(Z|B)}}_{\text{parameterization bias}}\cancel{\pi(X|B)} \qquad \pi(X|B)$$



Variation Encoder-Decoders

- Encoder-Decoder networks:



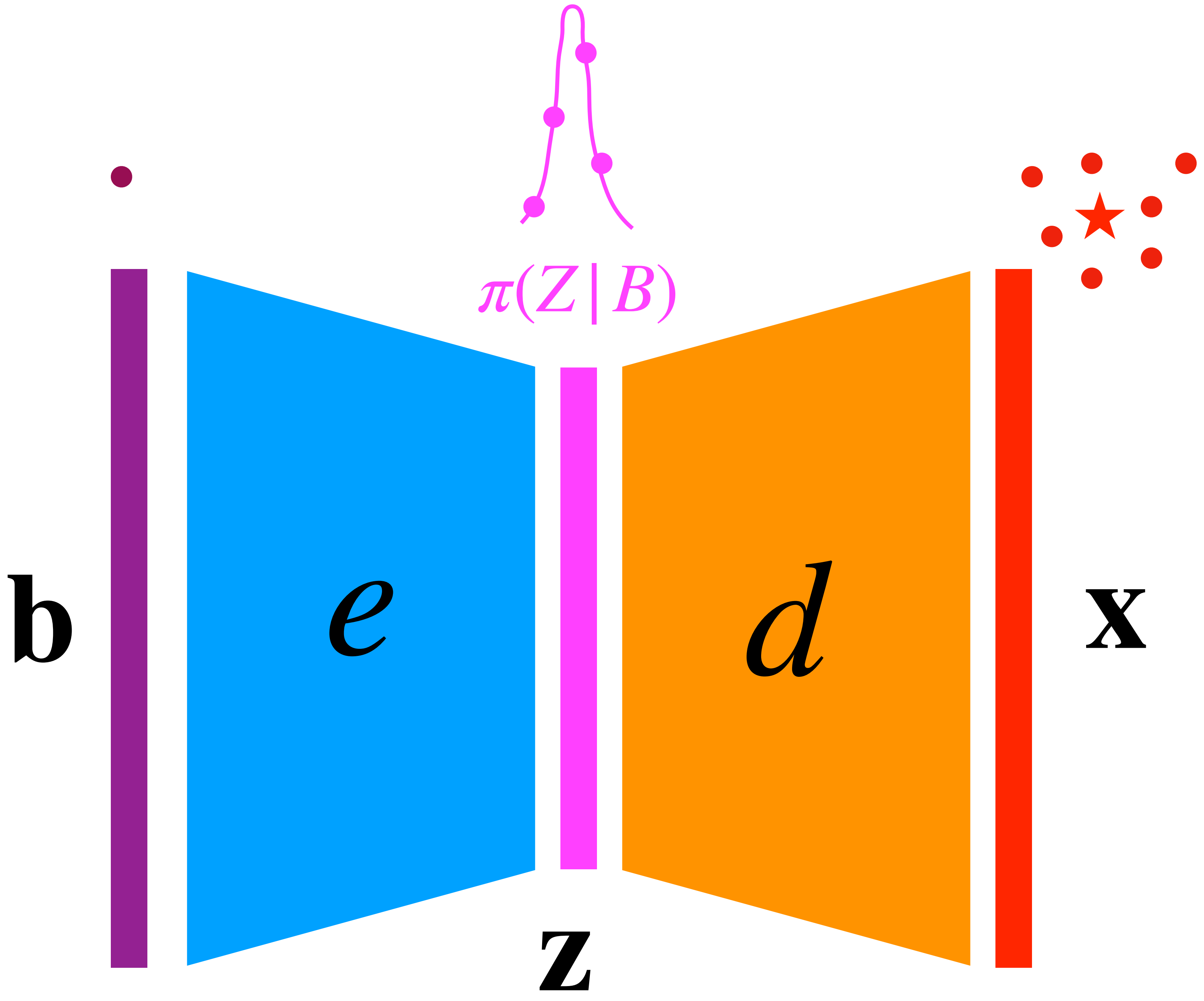
- Loss function:

$$\mathscr{L}(\mathbf{b}, \mathbf{x} \,|\, e, d) = D_{\mathrm{KL}}(\pi_d(Z \,|\, X) \,||\, \pi_e(Z \,|\, B))$$

Simplification of the Loss

- Proposition (simplified):

$$\mathscr{L}(\mathbf{b}, \mathbf{x} \,|\, e, d) = D_{\mathsf{KL}}(\pi(Z \,|\, B) \,||\, \mathscr{N}(0,1)) + \mathbb{E}\ \lambda \|\mathbf{x} - \bar{\mathbf{x}}\|_2^2$$

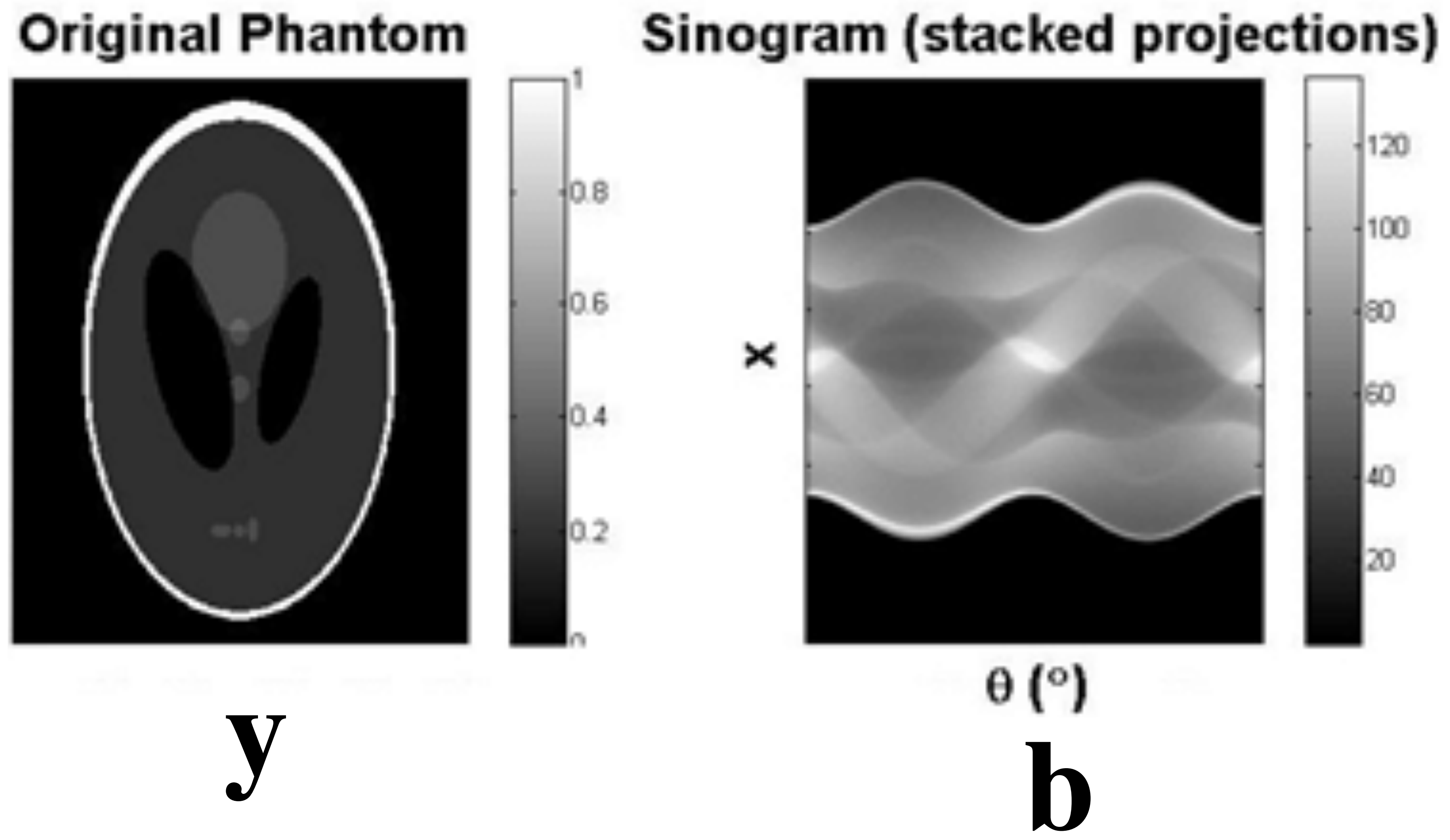


Example

Deterministic X-ray CT

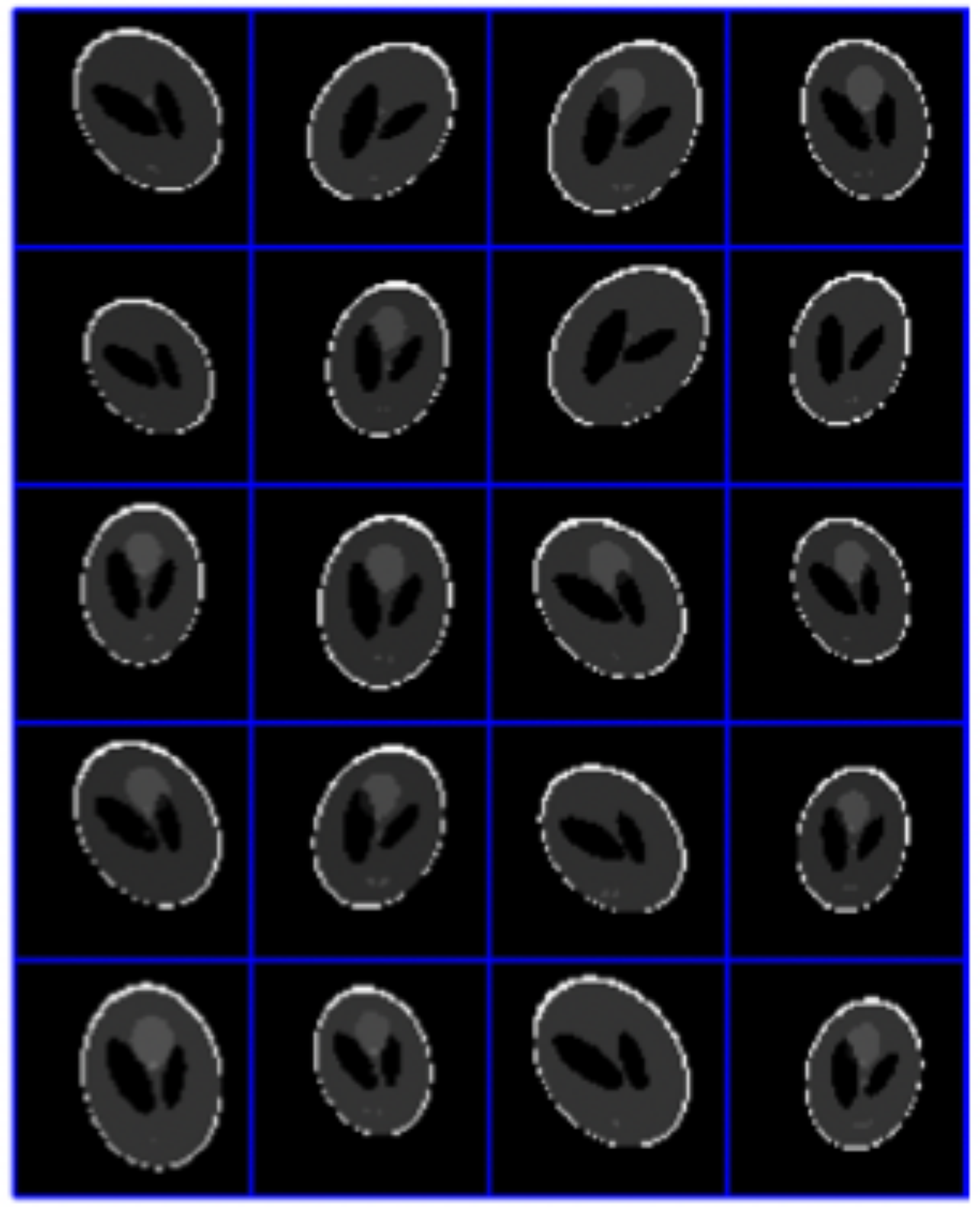
- Edge preserving reconstruction:

$$\mathbf{y}(x) = \text{argmin} ||F(\mathbf{y}) - \mathbf{b}||_2^2 + x||D\mathbf{y}||_1$$

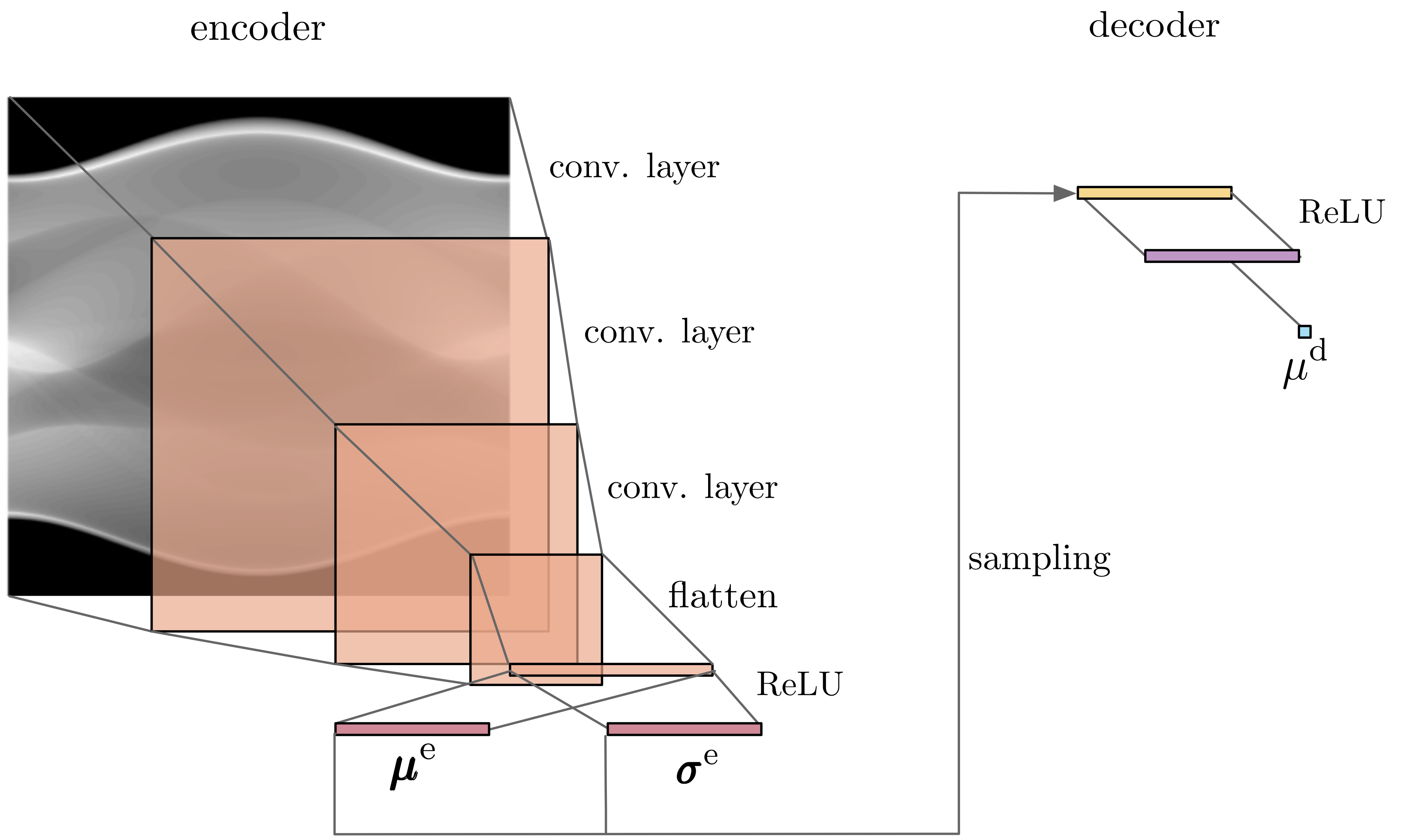


BMA, M&J Chung [2021]: Learning regularization parameters of inverse problems via deep neural networks

Training VEDs for the X-ray CT



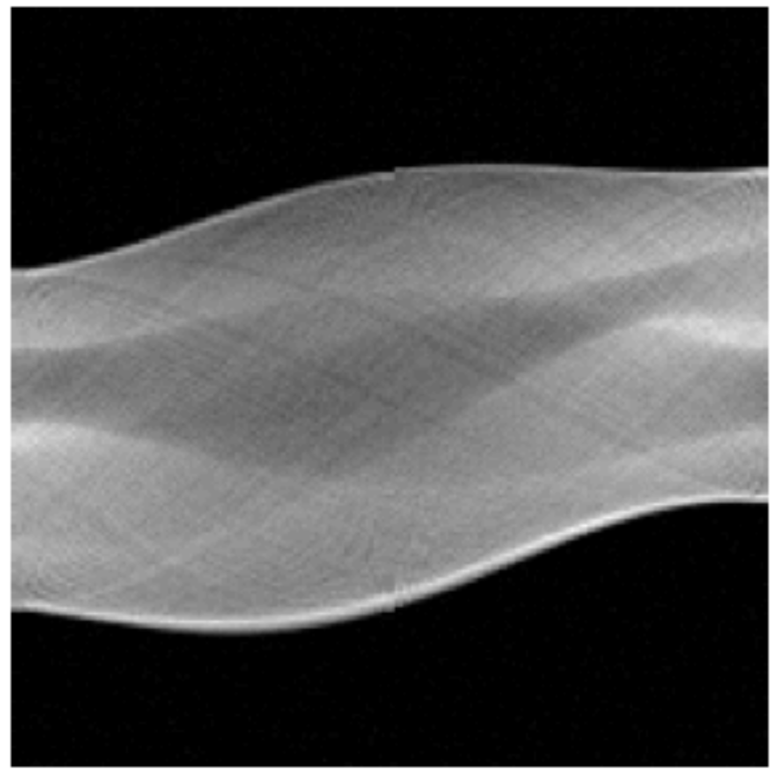
- Training data: Randomized Shepp-Logan phantoms. [M. Chung]
- x is obtained through a bi-level optimization problem.
- 2×10^4 data points
- Fixed forward problem.
- Training over 10^4 epochs.



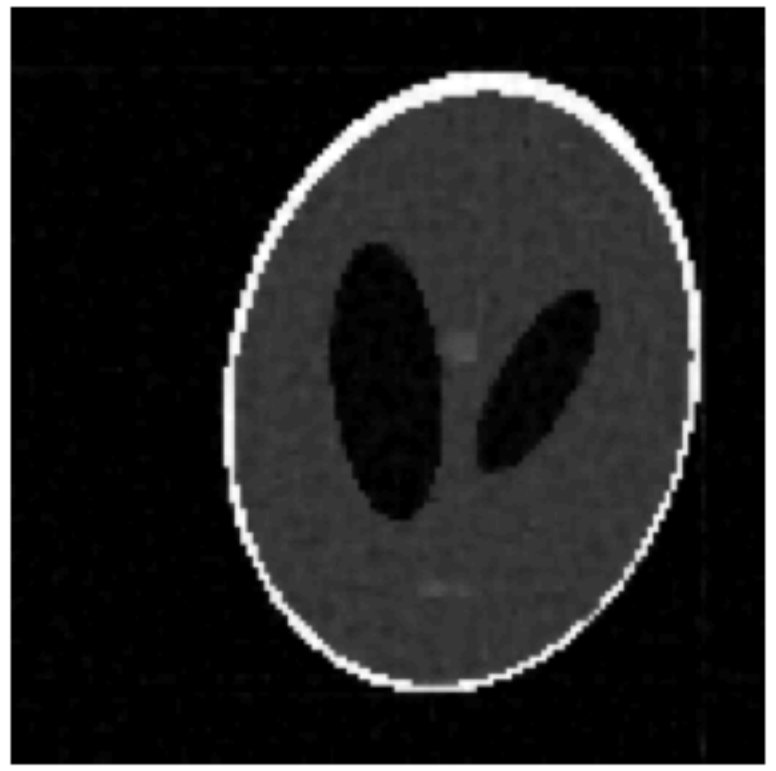
Results

Uncertain view angles

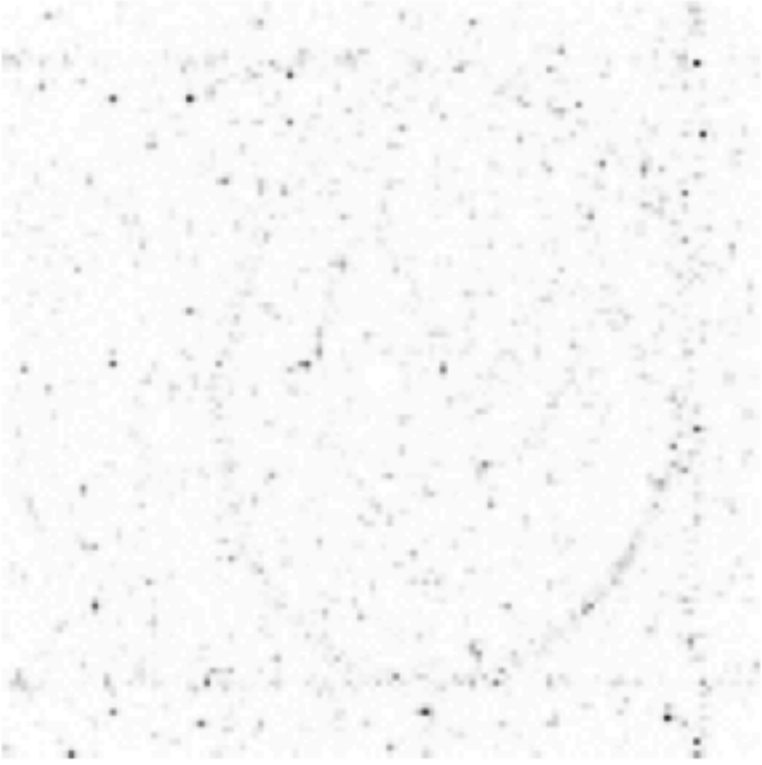
true image $\mathbf{y}_{\mathrm{true}}$



sinogram $\mathbf{b}$



mean reconstruction

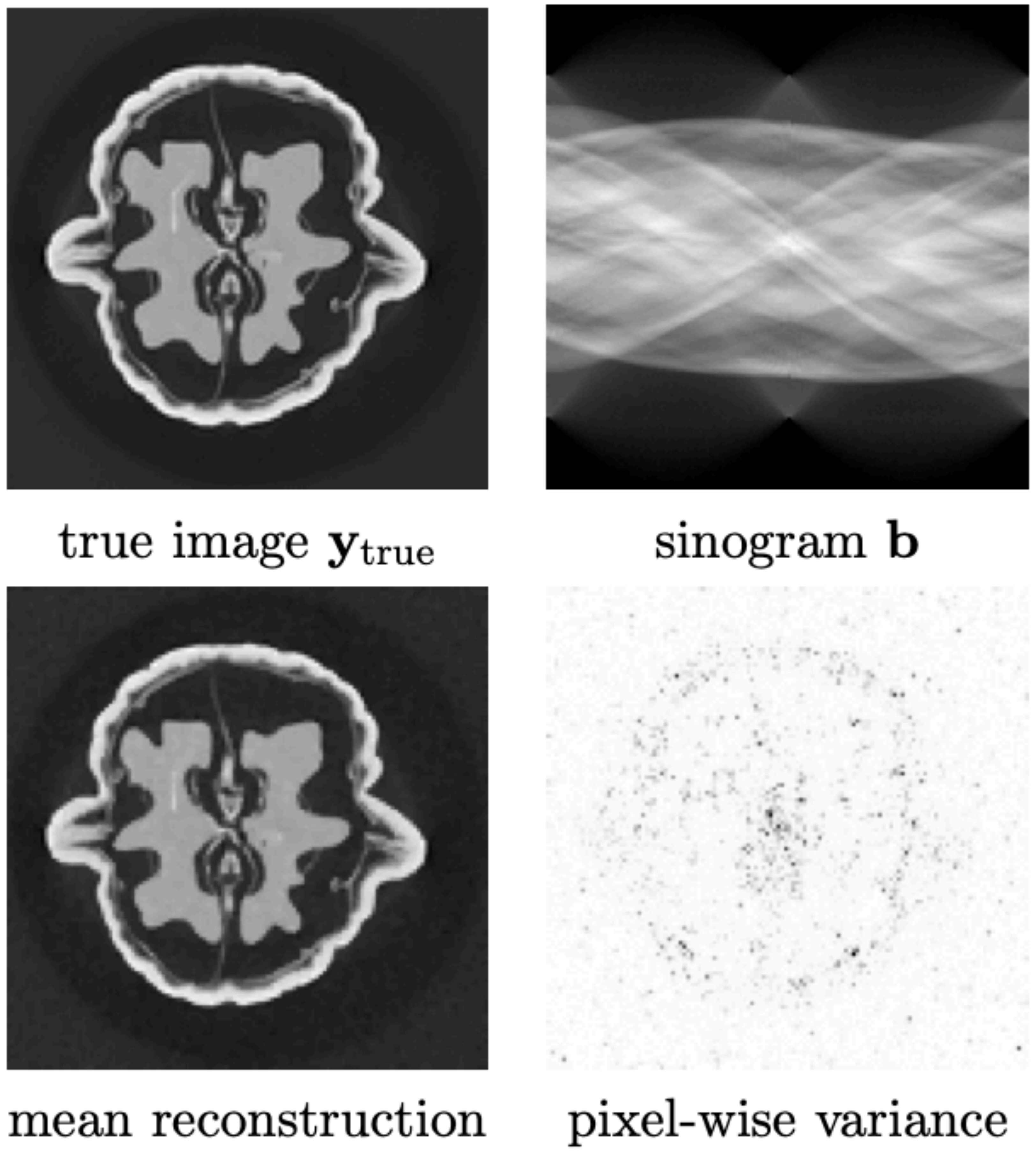


pixel-wise variance

Resutls

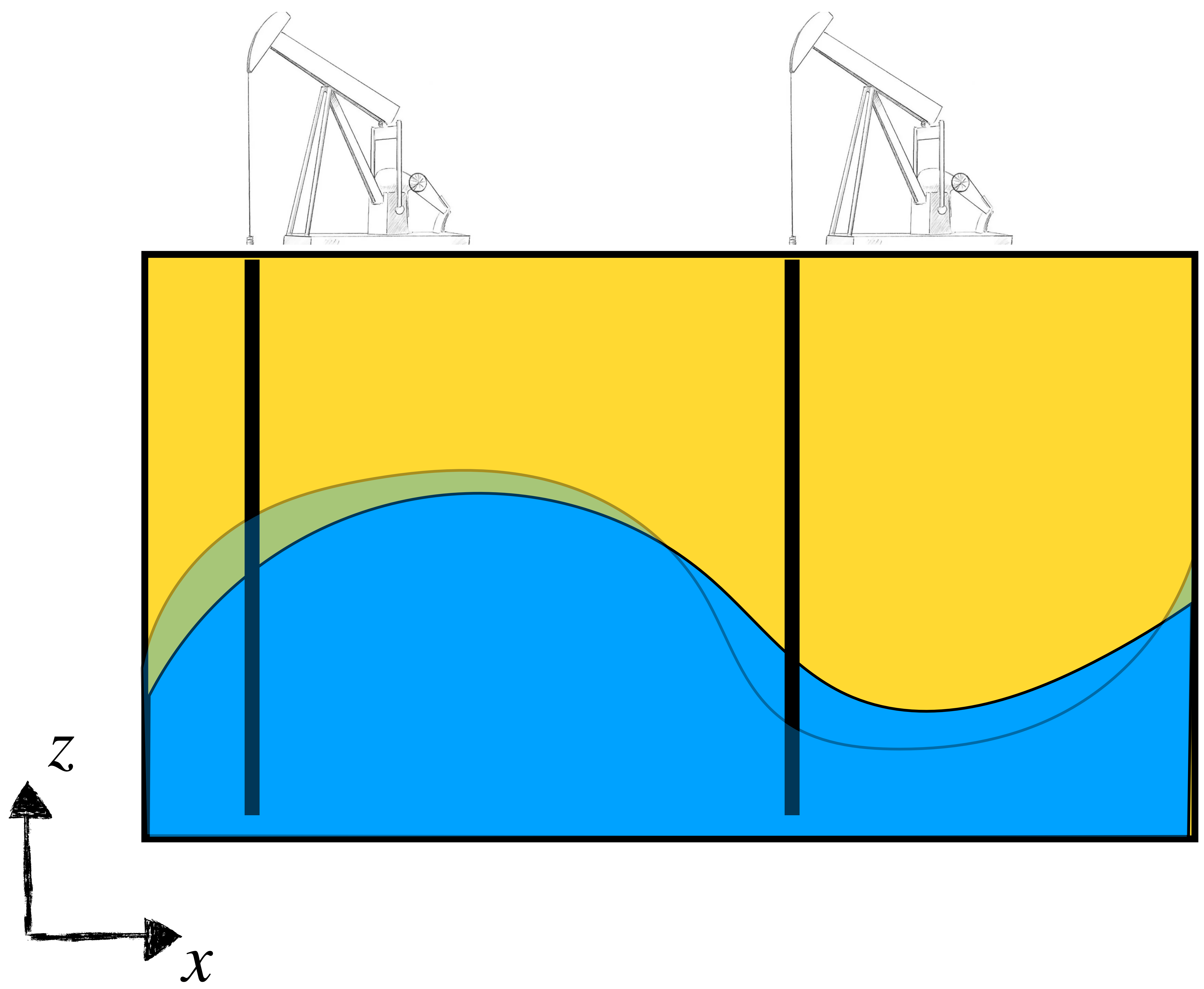
Out-of-prior sample

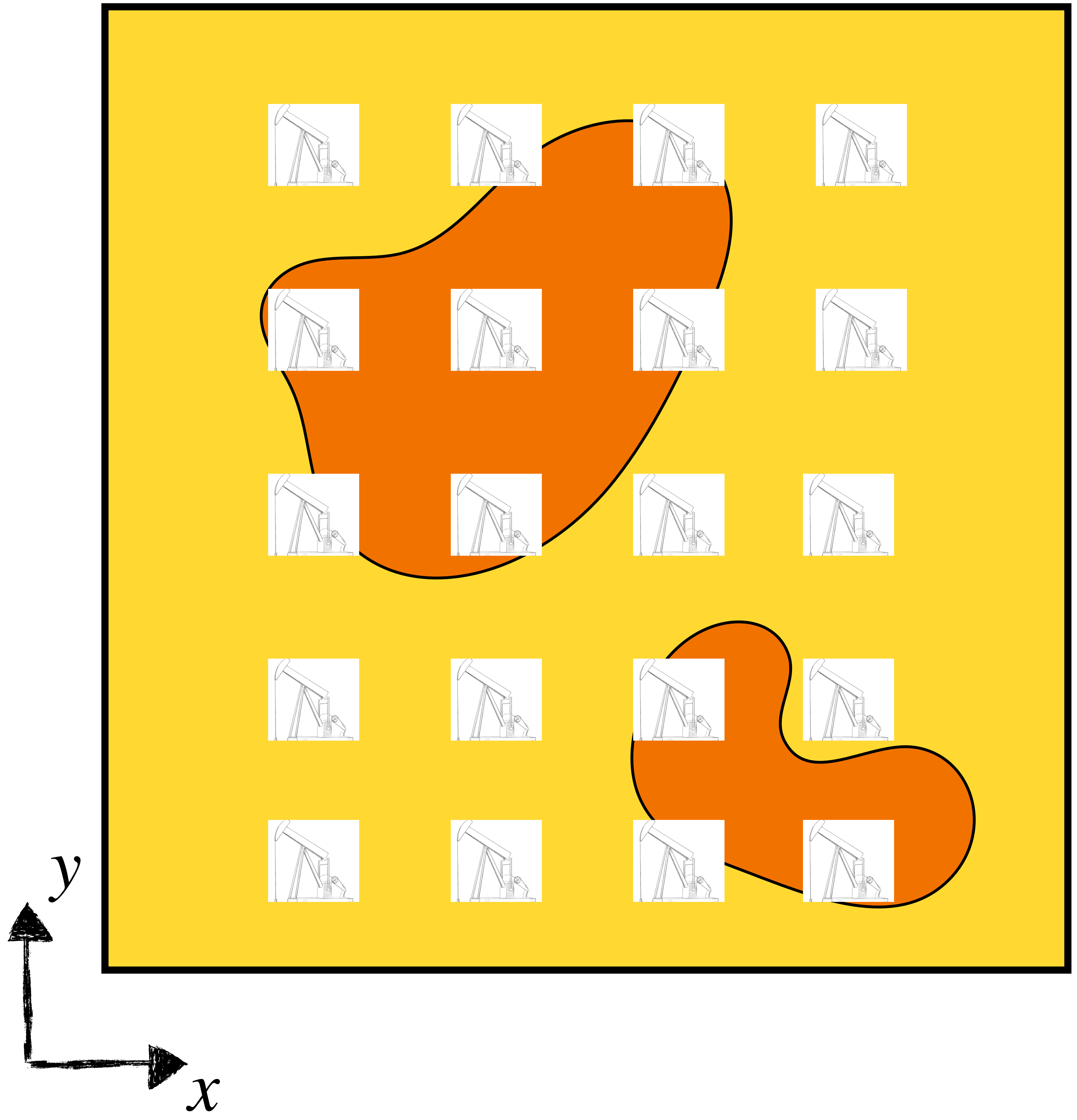
- The Walnut phantom: [FIPS, 2015]



Example

Hydraulic Tomography





Hydraulic Tomography problem

Mathematical model

- Mathematical model for the Hydraulic head

$$\nabla \cdot (\kappa \nabla h) = q_i \delta(x_i^{\text{well}})$$

and

$$\kappa \nabla h(x) \cdot n = 0 \qquad x \text{ on top}$$

$$h = 0 \qquad x \text{ not on top}$$

Hydraulic Tomography problem

Prior model - levelset prior

- We assume there is an underlying Gaussian random function:

$$X \sim (0,C) \ , \quad C = (\sigma I + \Delta)^2$$

  where

- The Porosity parameter κ is then piecewise constant with

$$\kappa = \frac{c^+}{2}(1 + \text{sign}(X)) + \frac{c^-}{2}(1 - \text{sign}(X))$$

Hydraulic Tomography problem
Goal

- We can expand X in the basis of Eigen vectors of C:

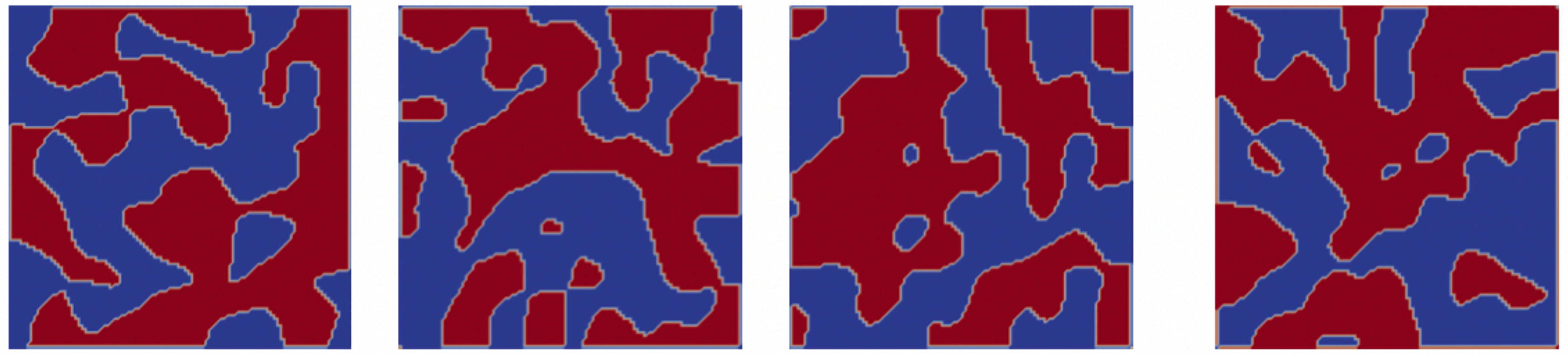
$$X = \sum_{i=1}^{\infty} x_i \sqrt{\lambda_i} e_i$$

- The goal is to recover the first q coefficients, i.e. $x_1, \ldots, x_q$

Hydraulic Tomography problem

Training VED network

- We collect 10^4 data $\{\mathbf{b}, \mathbf{x}\}$ from the prior

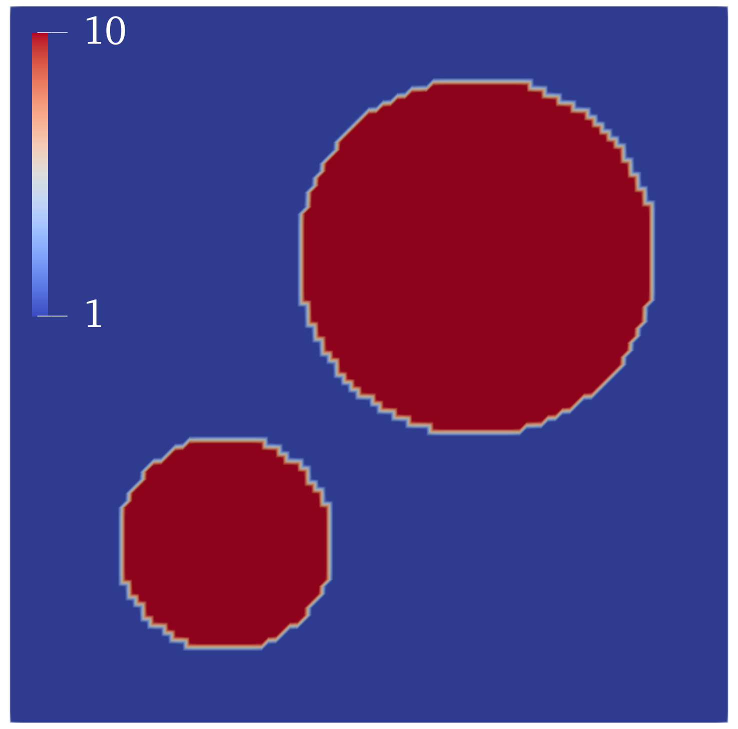


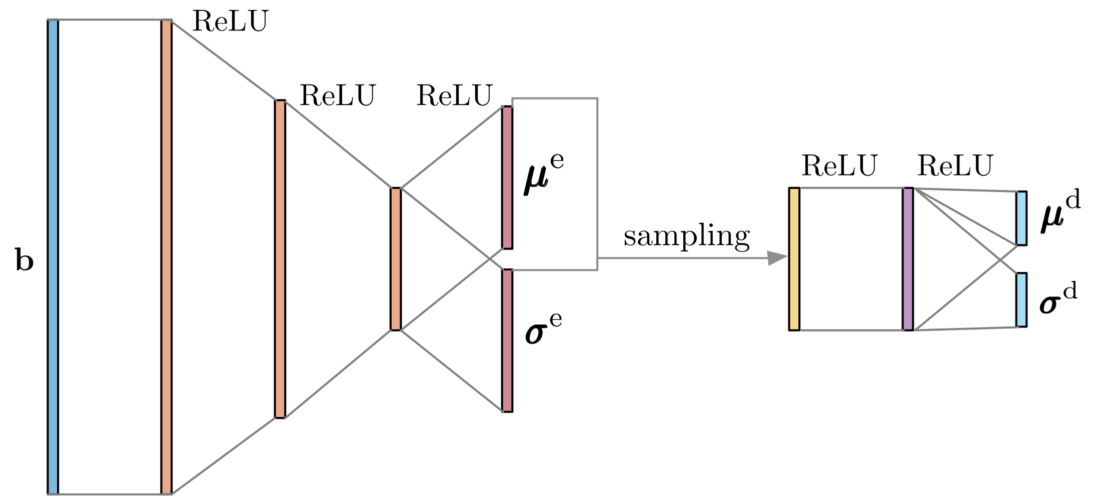
(a) samples from prior distribution visualized on the conductivity field

- We have a fully-connected feed-forward (3 hidden layers) encoder and decoder (1 hidden layer)

- True conductivity is out of prior

- Comparing with MCMC with 10^6 samples.

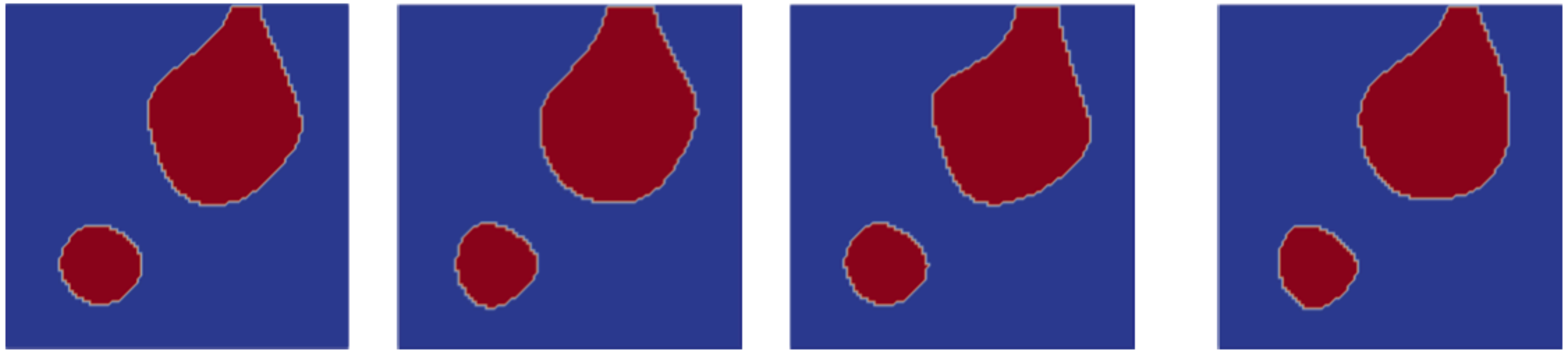




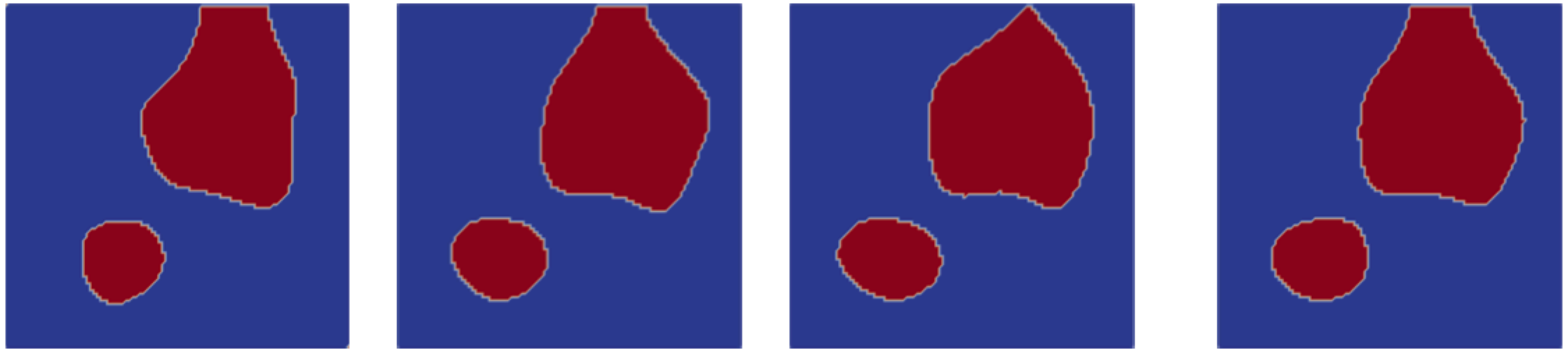
Results

Comparing to pCN-MCMC

- Samples from the posterior:



(b) samples from the VED posterior predictive visualized on the conductivity field



(c) MCMC samples from posterior predictive visualized on the conductivity field

Results

Comparing to pCN-MCMC

- Mean and variance of the posterior

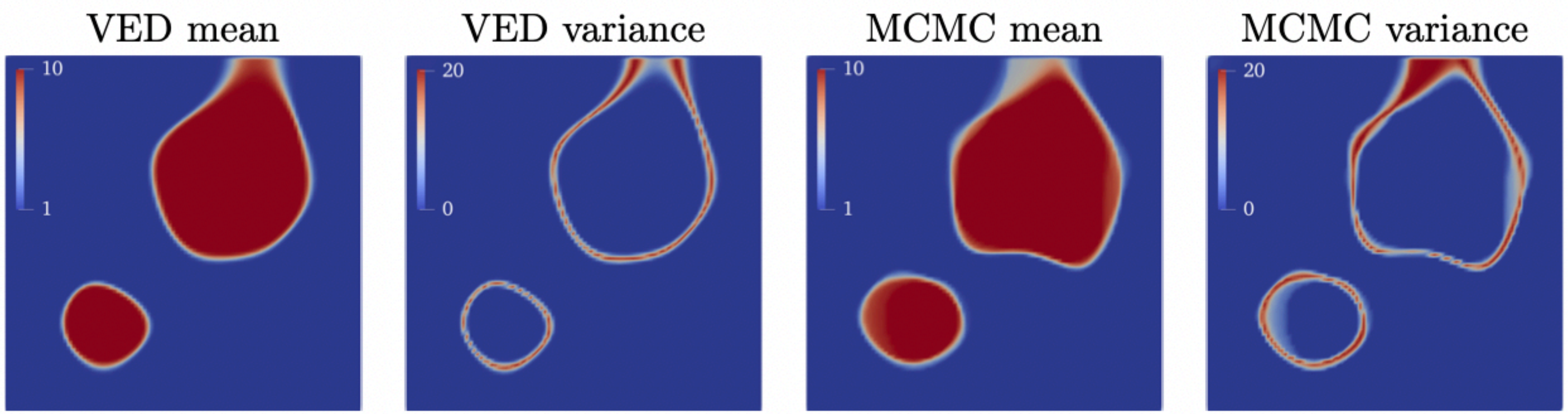
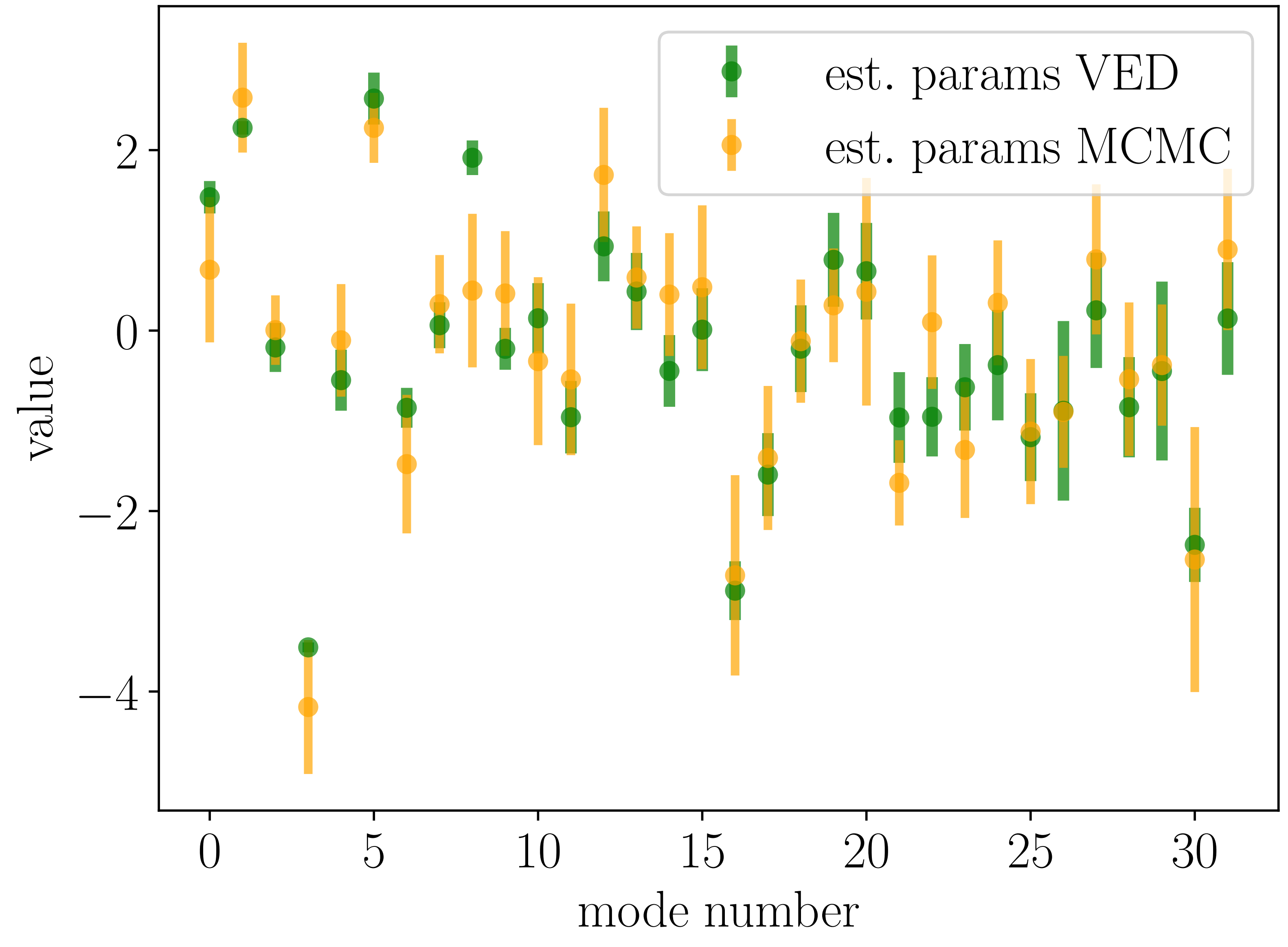


Figure 13: Means and variances of the conductivity fields for samples from the VED posterior predictive (left) and MCMC samples from the posterior predictive (right).

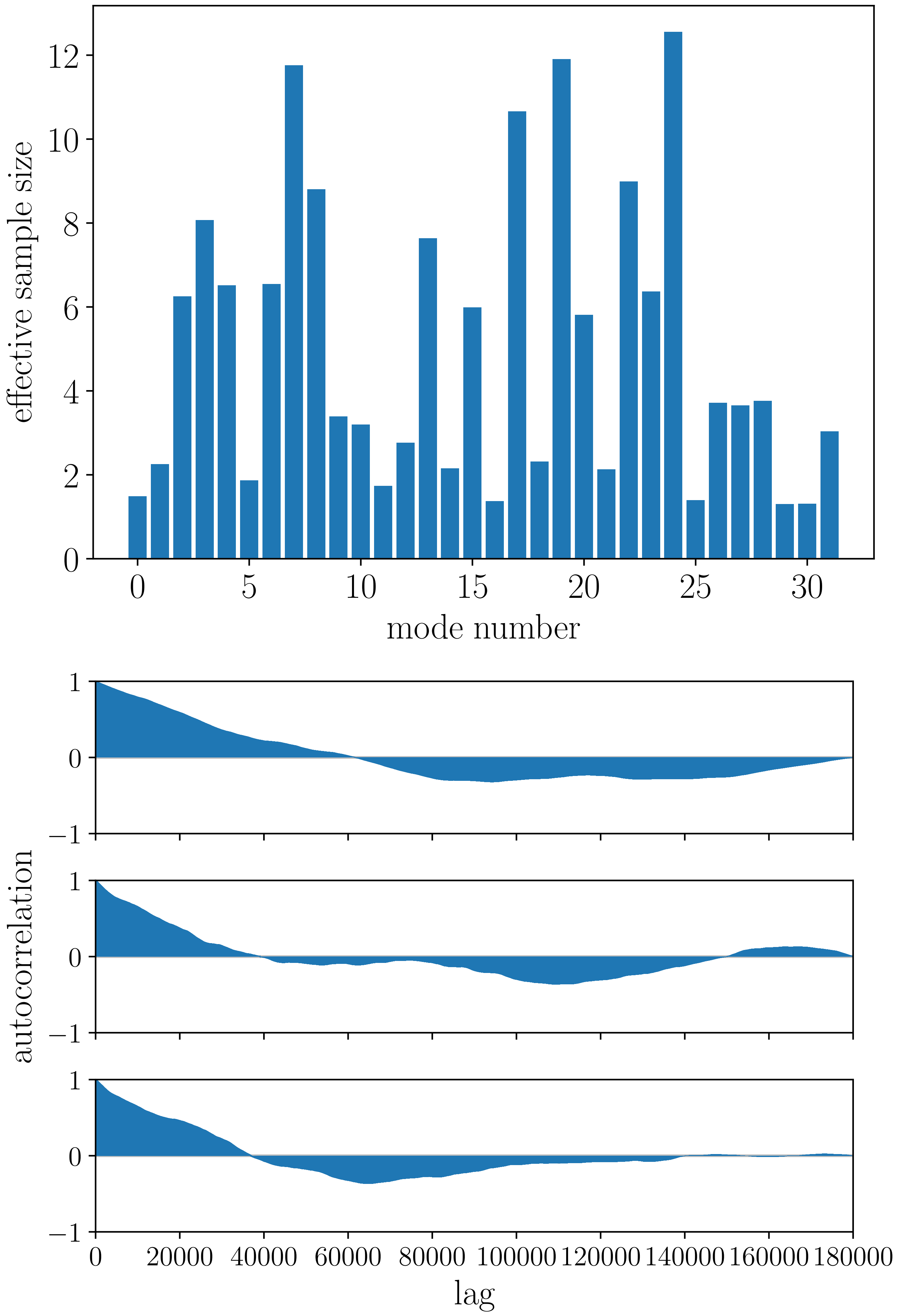
Results

Comparing to MCMC

Comparing coefficients



MCMC diagnostics



	Elapsed time (s) for computing 1000 samples	Elapsed time (s) for computing 10 *independent* samples
MCMC (pCN)	270	2.5×10^5
VED sampling	20	0.02
speed-up	13.5	1.2×10^8

Table 1: CPU times for sampling from the posterior predictive using MCM versus using VED sampling. The speed-up in terms of obtaining independent samples is significant.

Conclusions:

- Efficient tool for UQ for inverse problem.
- We can achieve UQ for deterministic problems using data.
- Can we use sampling data to train a network?
- Unlocking larger dimensions

References:

- Goal-oriented Uncertainty Quantification for Inverse Problems via Variational Encoder-Decoder Networkse. BMA, T&J Chung. (Preprint)
- Learning regularization parameters of inverse problems via deep neural networks. BMA, T&J Chung. Inverse Problems, IOPScience, 2021.