Imperial College
London

Stochastic Seismic Waveform Inversion
Using Generative Adversarial Networks
As A Geological Prior

Lukas Mosser, Olivier Dubrule, Martin J. Blunt

PESGB/EAGE ML Workshop 2018

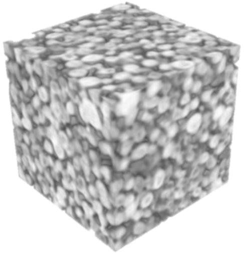
lukas.mosser15@imperial.ac.uk, o.dubrule@imperial.ac.uk, m.blunt@imperial.ac.uk

A Path to Inversion with GANs

Generative Adversarial Networks

Goal: Fast image generation based on samples

(Year 1: Phys. Rev. E / TIPM)

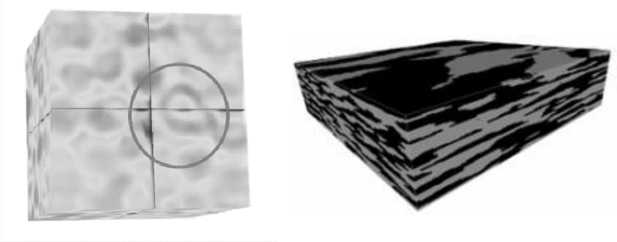


Unconditional Prior

Conditioning of GANs

Goal: Incorporate available data

(Year 1: arXiv:1802.05622)

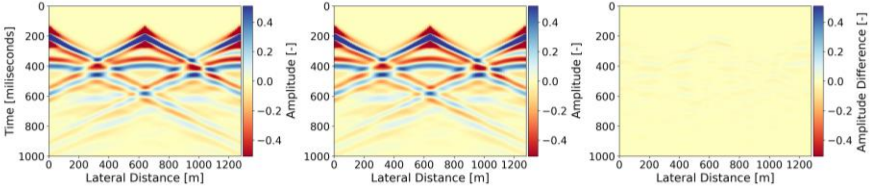


Well Data

GANs for Inverse Problems

Goal: Use GANs as prior for stochastic inversion

(Year 2: arXiv:1806.03720)



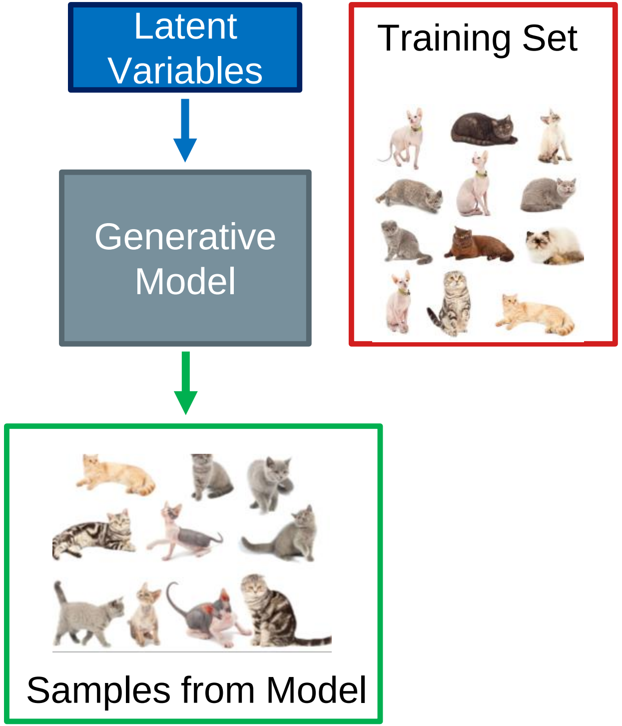
Physics

(Deep) Generative Methods

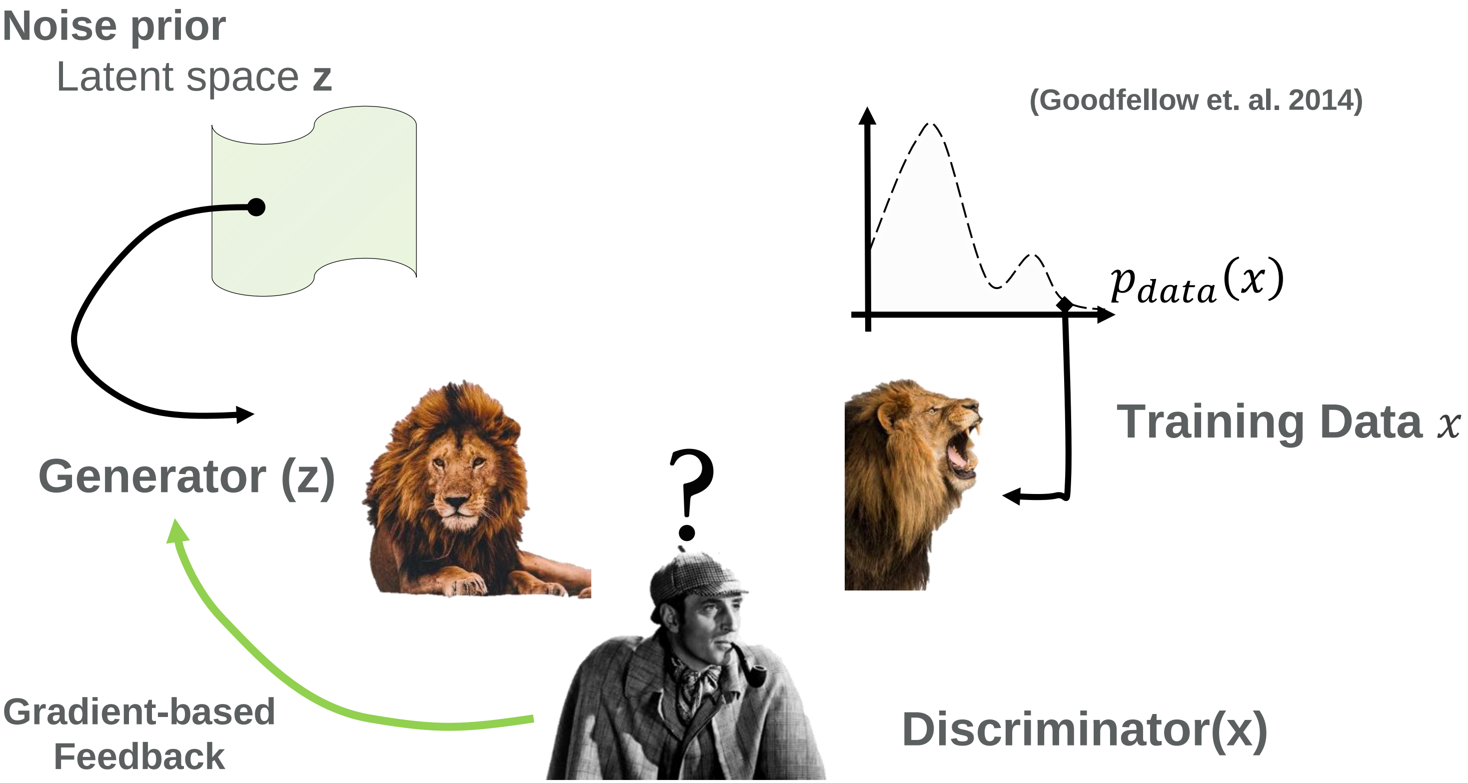
- Task: Draw (new) samples from unknown density given a set of samples

Main Problem: How to find the generative model?

- Generative Adversarial Networks (GAN)
  - Two competing Neural Networks
- Variational Autoencoders (VAE)
  - Bayesian Graphical Model of data distribution
- Autoregression (Pixel-CNN)
  - Conditional Distribution on every sample
- Many More ...



Generative Adversarial Networks – Toy Example



Noise prior

Latent space **z**

(Goodfellow et. al. 2014)

$p_{data}(x)$

Training Data x

Generator (z)

Gradient-based Feedback

Discriminator(x)

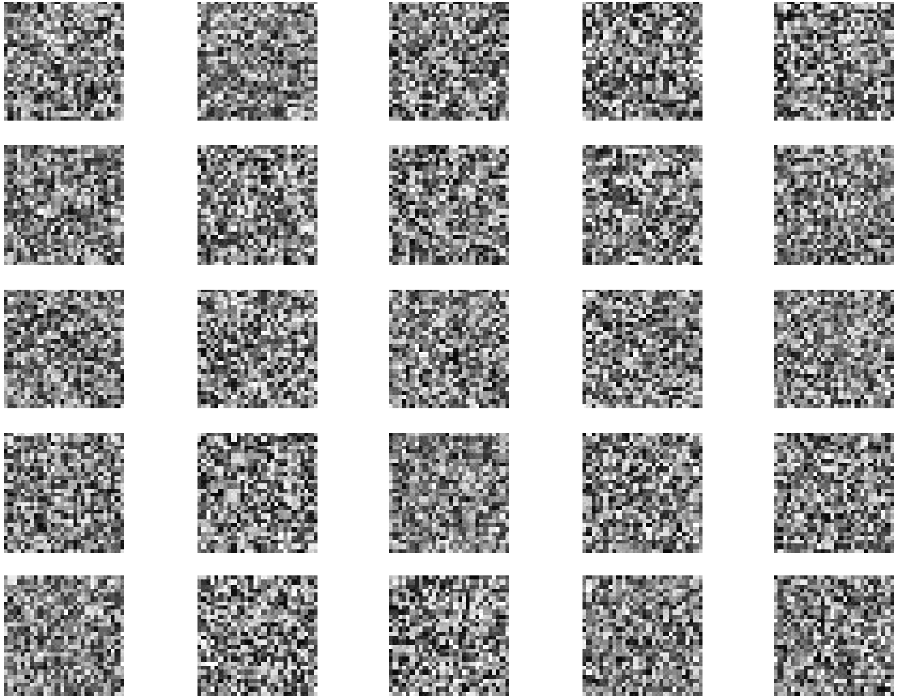
Generative Adversarial Networks – Training

- **Requirements:**
  - **Training Set of data**
  - **Generator – creates samples G(z)** $\mathbf{z} \sim \mathcal{N}(0,1)^{d\times 1\times 1\times 1}$ $G_\theta : \mathbf{z} \rightarrow \mathbb{R}^{1\times 64\times 64\times 64}$
  - **Discriminator – evaluates samples** $D_\omega : \mathbb{R}^{1\times 64\times 64\times 64} \rightarrow [0,1]$
  - **Cost function:** $\min_{\theta} \max_{\omega} \{\mathbb{E}_{\mathbf{x}\sim p_{data}}[log\ D_\omega(\mathbf{x})] + \mathbb{E}_{\mathbf{x}\sim p_{\mathbf{z}}}[log(1 - D_\omega(G_\theta(\mathbf{z})))]\}$
- **GAN training – two step procedure in supervised way**
  - **Discriminator training step – Generator fixed**
    - **Train on real data samples**
    - **Train on fake samples**
  - **Generator training step – Discriminator fixed**
    - **Push generator towards “real” images**

GAN Training Example - MNIST

Training Images

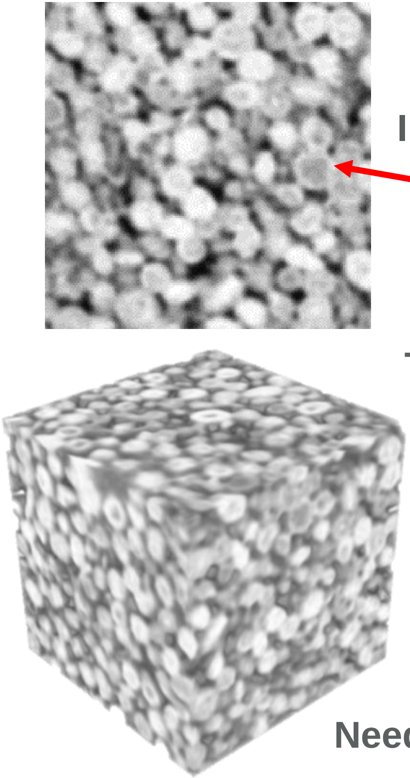
Generative Model (GAN)



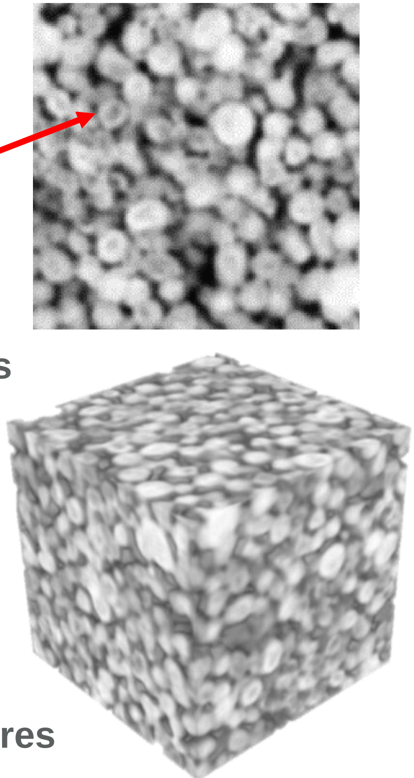
Credit: @eriklindernoren

Unconditional Simulation – Pore Scale

Ketton Training Image



GAN generated sample



Intergranular Porosity
Moldic Features
Micro-Porosity

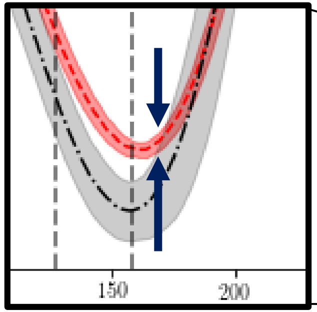
Training Time: 8 hours
Generation: 5 sec.

High visual quality
Needs quantitative measures

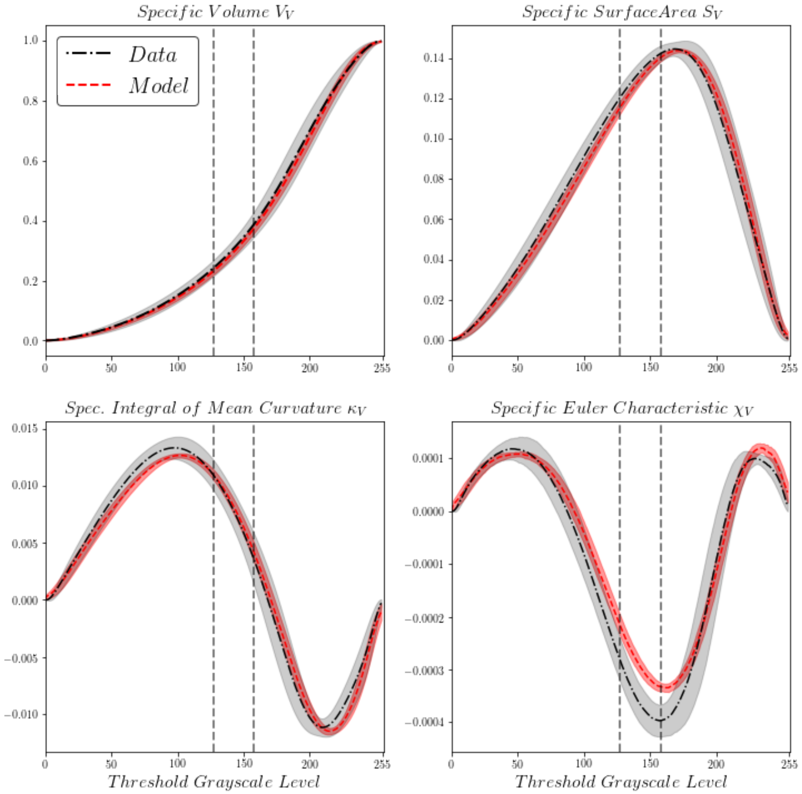
Morphological Properties of Generated Images

Characteristic Functionals captured by GAN model

Largest Error (~20%)
-> Euler Characteristic



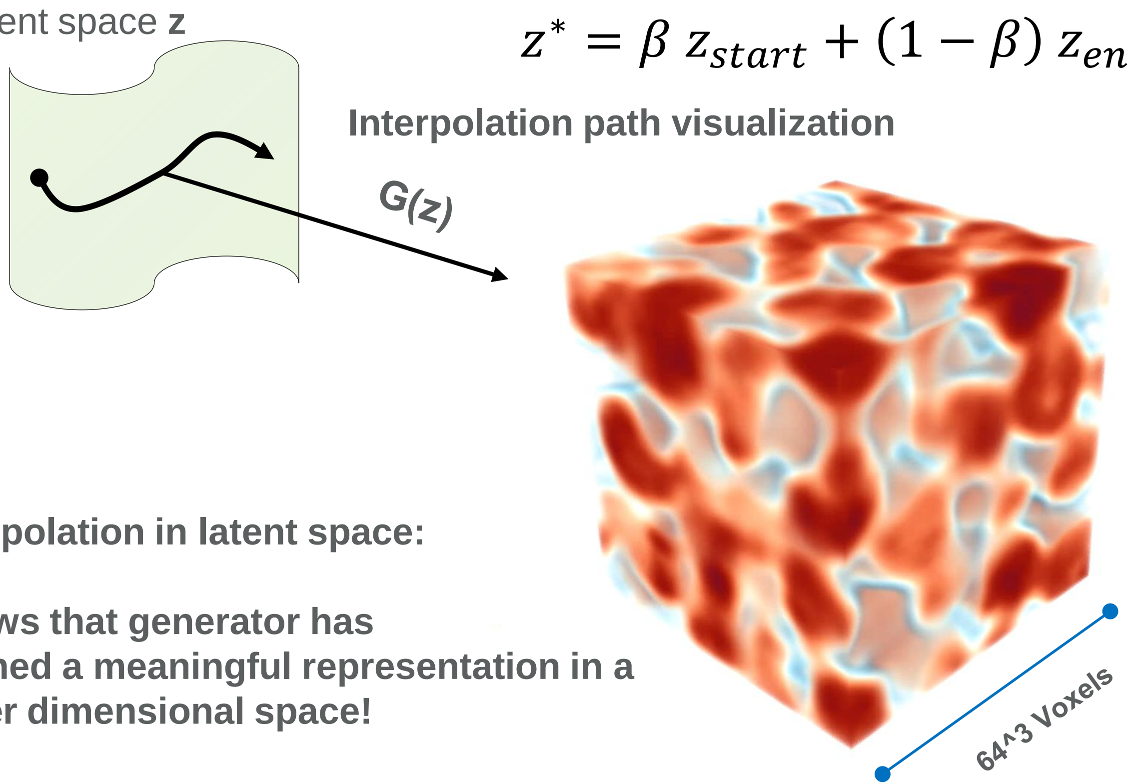
Smaller variance of GAN images



Latent Space Interpolation – Image Parameterisation

$$z^* = \beta\, z_{start} + (1-\beta)\, z_{end}, \beta \in [0,1]$$

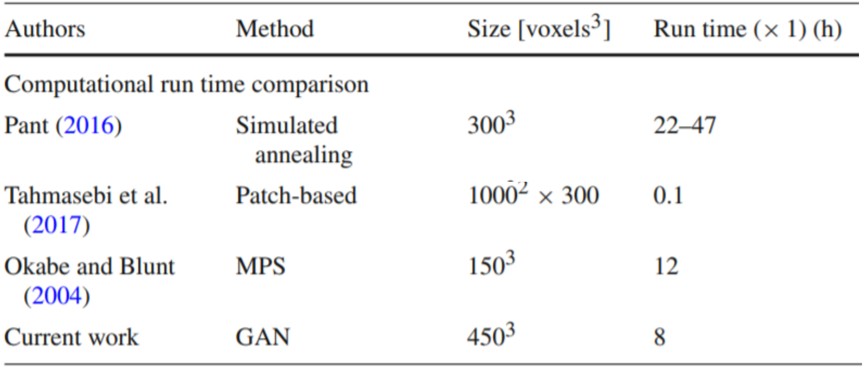
Interpolation path visualization



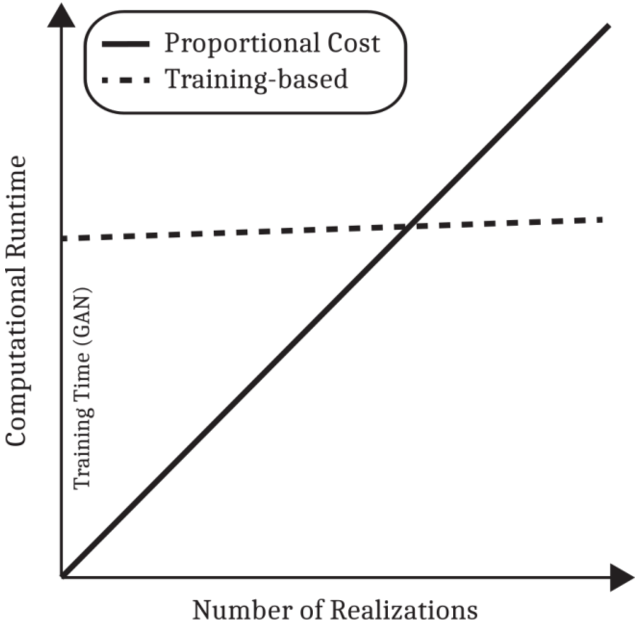
Interpolation in latent space:

Shows that generator has learned a meaningful representation in a lower dimensional space!

Computational Effort



Authors	Method	Size [voxels^3]	Run time (× 1) (h)
Computational run time comparison			
Pant (2016)	Simulated annealing	300^3	22–47
Tahmasebi et al. (2017)	Patch-based	$1000^2 \times 300$	0.1
Okabe and Blunt (2004)	MPS	150^3	12
Current work	GAN	450^3	8



Main Computation cost training:
Amortizes with number of samples due to low per sample cost / runtime

Image Inpainting (Yeh et al. 2016)

Task: Restore missing details given a corrupted / masked image $M \cdot \tilde{x}$

Use a generative model G(z) to find missing details, conditional to given information.

Contextual Loss: $L_{content} = \lambda||M \cdot G(z) - M \cdot \tilde{x}||_2$

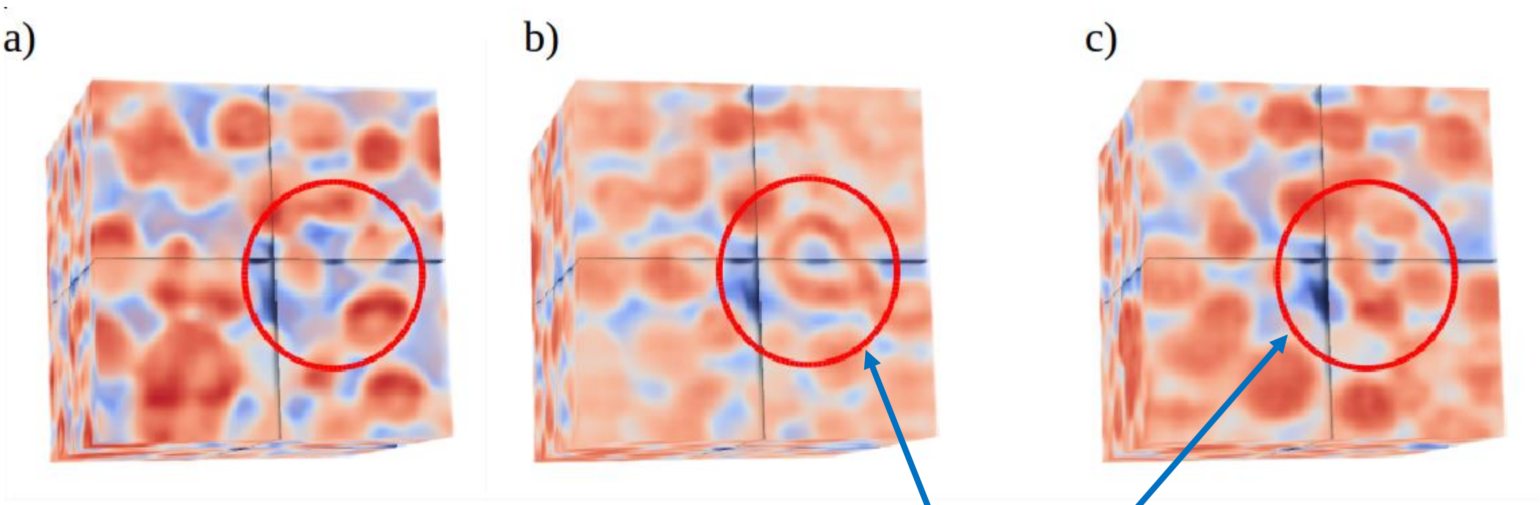
Perceptual Loss: $L_{perc} = log(1 - D(G(z)))$

Optimize loss by modifying latent vector z

$M \cdot \tilde{x}$ **Human Artist** $L_2\ Loss$ $L_{content} + L_{perc}$

Credit: Kyle Kastner

Conditioning – Pore Scale Example



Conditioning Data
Ground Truth Volume

Stochastic Sample 1
Conditioned to Data

Stochastic Sample 2
Conditioned to Data

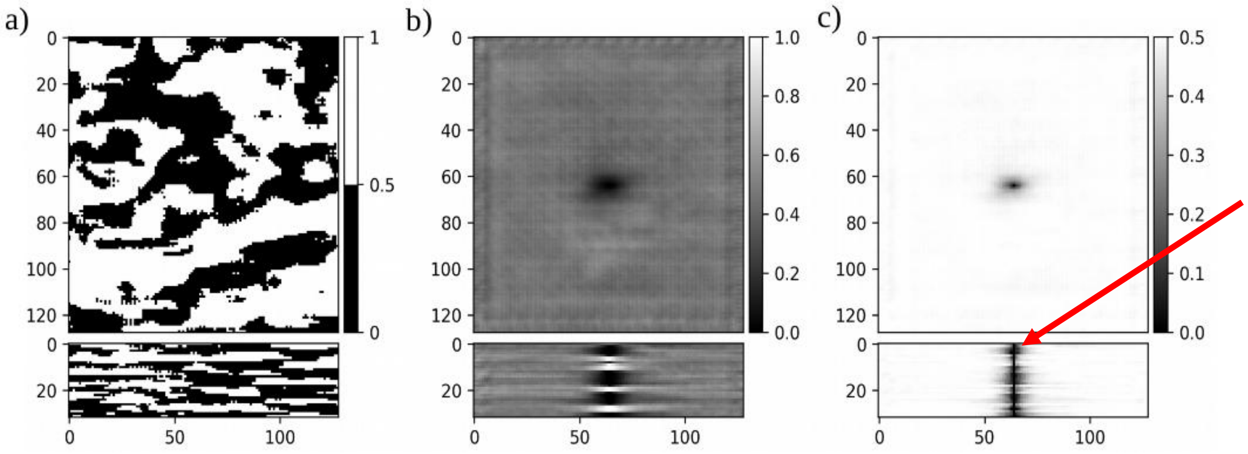
Same 2D conditioning data leads to varied realizations in 3D

Conditioning – Reservoir Scale Example

Maules Creek Training Image (Credit G. Mariethoz)

Pre-trained 3D-Generative Adversarial Network

Condition to single well (1D conditioning) from ground truth data:



Single Realization

Mean (N=1000)

Standard Dev. (N=1000)

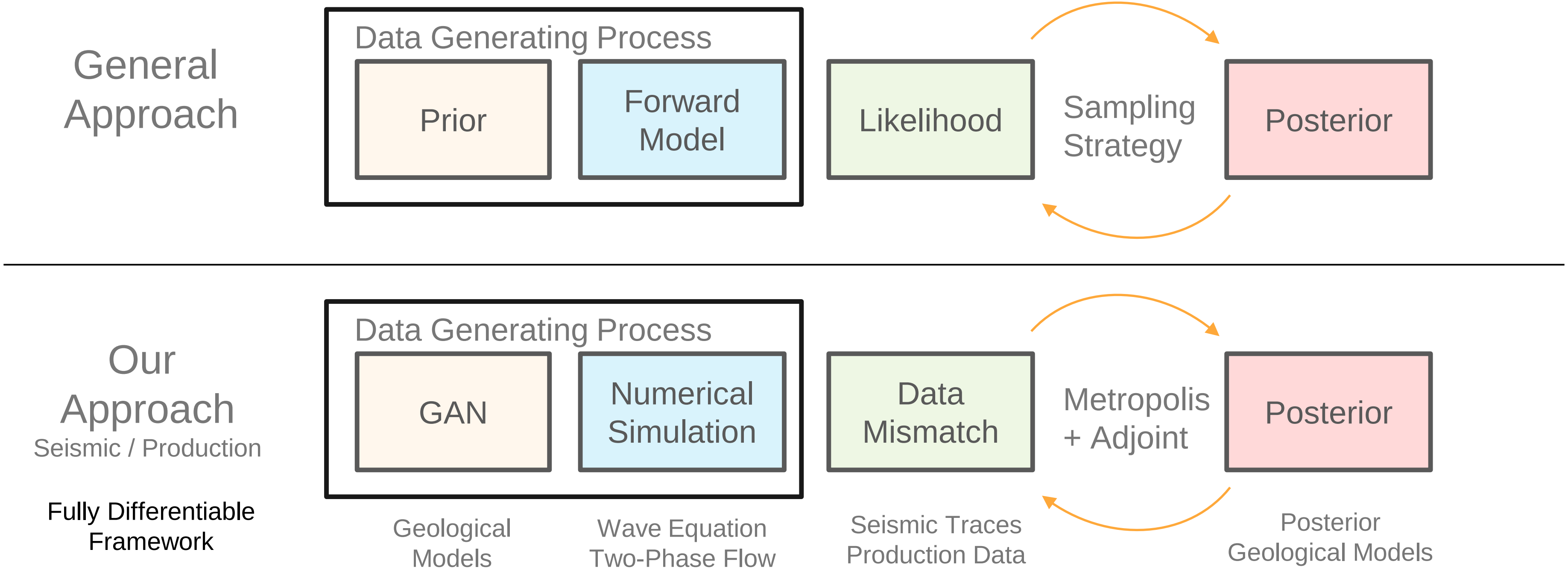
No Variance at Well

GANs for Inverse Problems

Goal: Use GANs as prior for stochastic inversion

Bayes' Rule

$$p(\mathbf{z}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{z})p(\mathbf{z})}{p(\mathbf{d})} \propto p(\mathbf{d}|\mathbf{z})p(\mathbf{z})$$



General Approach

Data Generating Process

Prior

Forward Model

Likelihood

Sampling Strategy

Posterior

Our Approach

Seismic / Production

Fully Differentiable Framework

Data Generating Process

GAN

Numerical Simulation

Data Mismatch

Metropolis + Adjoint

Posterior

Geological Models

Wave Equation
Two-Phase Flow

Seismic Traces
Production Data

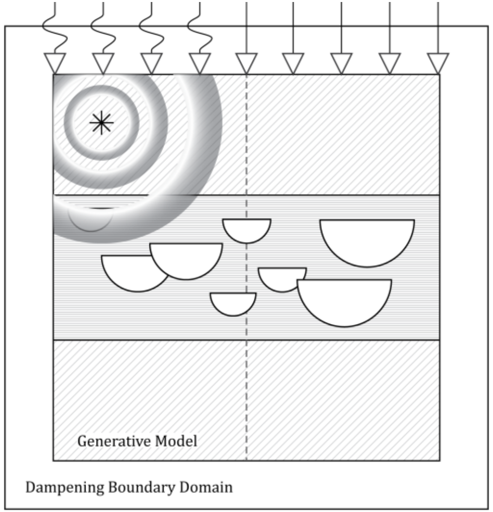
Posterior
Geological Models

Stochastic Inversion with GAN priors

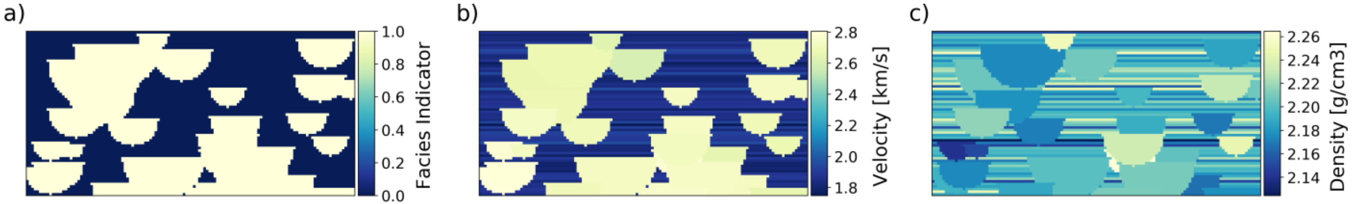
- Prior represented by GAN:
  Pre-train on geological models of river channels
  ~ 5000 training images, synthetic object-based

- GAN maps from latent-space to image space of geological models

- GAN outputs 3 channels:
  - Facies Probability (0 – Shale, 1- Sand)
  - Acoustic p-wave velocity
  - Rock density



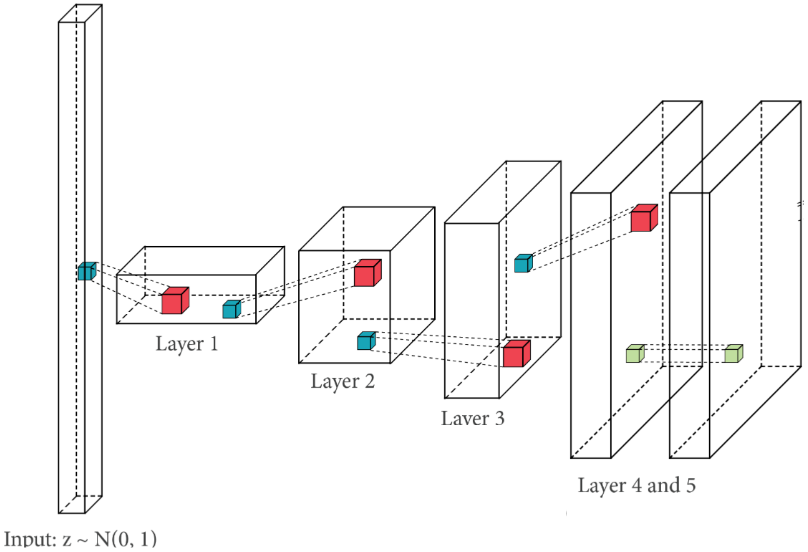
Computational Domain



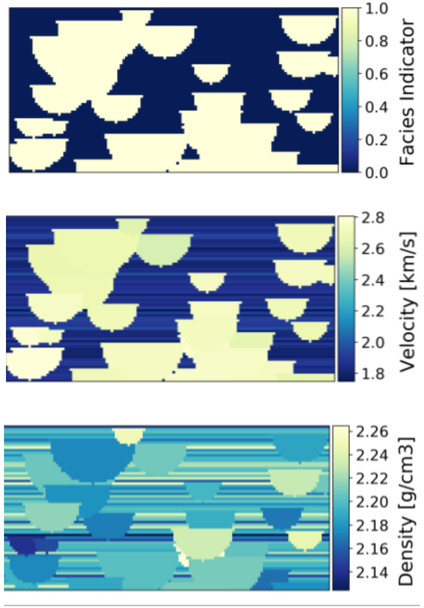
Example geological model and ground truth model for synthetic simulations

Network Architecture - 2D Convolutional Network

Represent G(z) and D(x) as deep neural networks:



Generator Network Architecture



Channel 1

Channel 2

Channel 3

Discriminator: Wasserstein Critic / Discriminator

Posterior Sampling Strategy – Gradient Descent

- Recall Bayes' Rule:

Goal: Find posterior of latent variables controlling GAN

$$p(\mathbf{z}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{z})p(\mathbf{z})}{p(\mathbf{d})} \propto p(\mathbf{d}|\mathbf{z})p(\mathbf{z})$$

- Perform Gradient Descent on mismatch by changing latent vector **z**

$$\mathbf{z}_{t+1} = \underbrace{\mathbf{z}_t}_{\text{Latent Vector}} + \overbrace{\epsilon_1}^{\text{Step Size}} \underbrace{\frac{\partial \|S(G_\theta(\mathbf{z}_t)) - d^{obs}\|_2}{\partial G_\theta(\mathbf{z}_t)}}_{\text{Adjoint-State Method}} \overbrace{\frac{\partial G_\theta(\mathbf{z}_t)}{\partial \mathbf{z}_t}}^{\text{NN - Backpropagation}}$$

- Choose random starting latent vectors **z(t=0)** and minimize mismatch

- Works, but can lead to low diversity. Formalisation -> MALA sampling (Nguyen et al, 2017)

Posterior Sampling Strategy - MALA

- Recall Bayes' Rule: Goal: Find posterior of latent variables controlling GAN

$$p(\mathbf{z}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{z})p(\mathbf{z})}{p(\mathbf{d})} \propto p(\mathbf{d}|\mathbf{z})p(\mathbf{z})$$

- Metropolis – Adjusted - Langevin Algorithm

$$\mathbf{z}_{t+1} = \underbrace{(1-\lambda)\mathbf{z}_t}_{\text{Weight-Decay}} + \overbrace{\epsilon_1}^{\text{Step Size}} \underbrace{\frac{\partial\|S(G_\theta(\mathbf{z}_t)) - d^{obs}\|_2}{\partial G_\theta(\mathbf{z}_t)}}_{\text{Adjoint-State Method}} \overbrace{\frac{\partial G_\theta(\mathbf{z}_t)}{\partial \mathbf{z}_t}}^{\text{NN - Backpropagation}} + \underbrace{\mathcal{N}(0, \epsilon_2)}_{\text{Noise-Perturbation}}$$

- Perform MALA by gradient descent with annealing step-size and adding noise

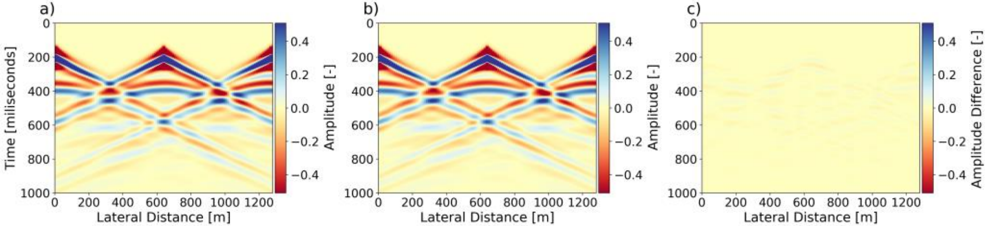
- Additional well data included by cross-entropy term on facies probability or L2-Norm for continuous properties.

Numerical Results – Stochastic Inversion

PyTorch

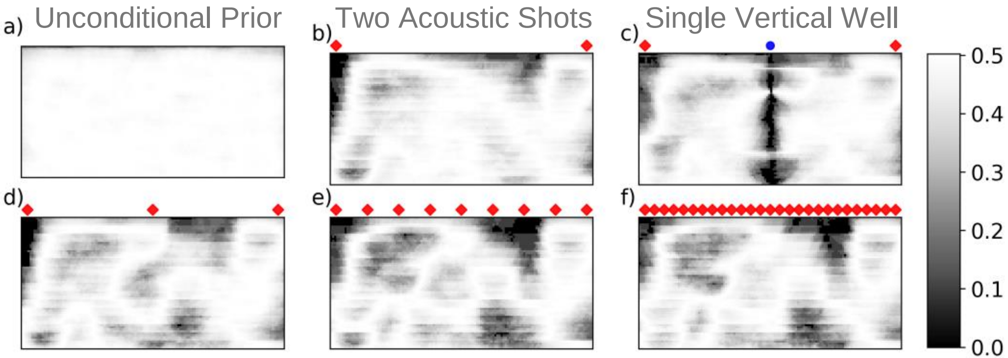
Devito

- Data Mismatch in Seismic Domain < 5%:



Ground Truth Posterior Sample Data Mismatch

- Perform posterior sampling (N=100) for increasing shot number (◆)



Unconditional Prior

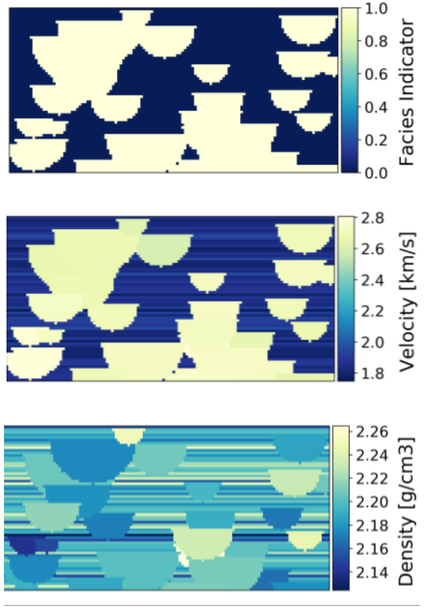
Two Acoustic Shots

Single Vertical Well

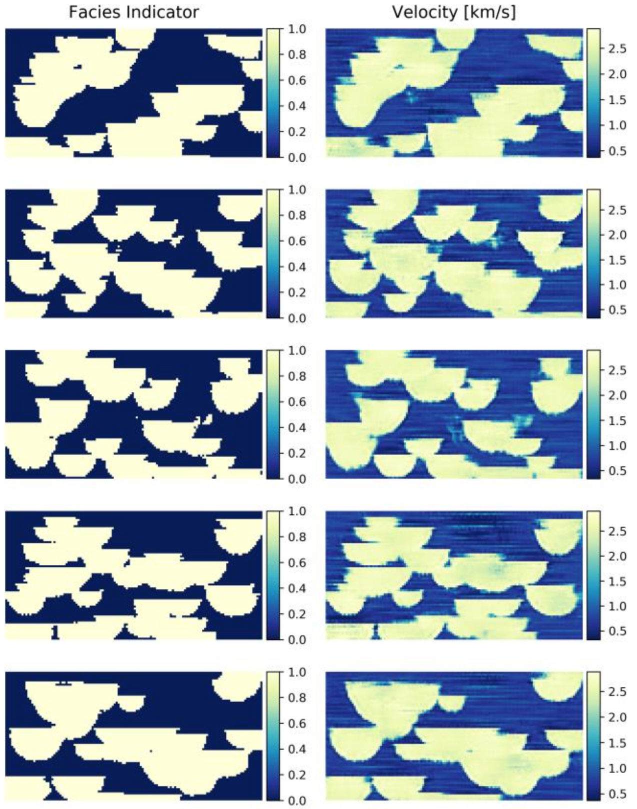
- Higher shot number leads to narrower posterior, well matches > 95% accuracy

Samples – (Seismic) Inversion

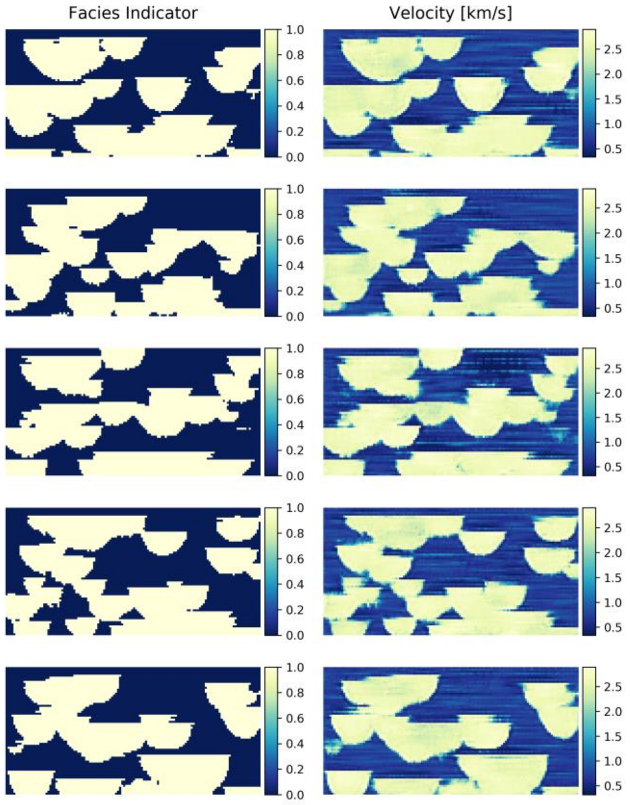
Ground Truth Example



2 Acoustic Sources



27 Acoustic Sources



Conclusions

- GANs can be used as efficient parameterizations of geological models
- Continuous, non-linear and differentiable representations of image distributions
- GANs do not alleviate the need for training images
- Can be challenging to train and quality control – mode collapse, training instabilities
- GANs can be used to represent a solution space for ill-posed inverse problems when combined with a posterior sampling method such as MALA.

Evolution of channels during sampling process

Imperial College London

Thank you! Questions?

References

Reconstruction of three-dimensional porous media using generative adversarial neural networks. ***Physical Review E*, *96*(4), 043309**,
Mosser, L., Dubrule, O., & Blunt, M. J. (2017).

Stochastic reconstruction of an oolitic limestone by generative adversarial networks. ***Transport in Porous Media*, 1-23,**
Mosser, L., Dubrule, O., & Blunt, M. J. (2017).

Conditioning of Generative Adversarial Networks for Pore and Reservoir Scale Models, **80th EAGE Conference**,
Mosser L., Dubrule, O., & Blunt, M. J. (2018).

Stochastic seismic waveform inversion using generative adversarial networks as a geological prior. ***arXiv preprint arXiv:1806.03720***,
Mosser, L., Dubrule, O., & Blunt, M. J. (2018).

Parametric generation of conditional geological realizations using generative neural networks. Chan, S. and Elsheikh, A.H., (2018)
arXiv preprint arXiv:1807.05207

Generating Realistic Geology Conditioned on Physical Measurements with Generative Adversarial Networks. Dupont, E., Zhang, T., Tilke, P., Liang, L. and Bailey, W., (2018) ***arXiv preprint arXiv:1802.03065***