Can neural networks emulate physics?

Towards a Cosmology Emulator using Generative Adversarial Networks

Mustafa Mustafa
Data Analytics Services, NERSC
Berkeley Lab.

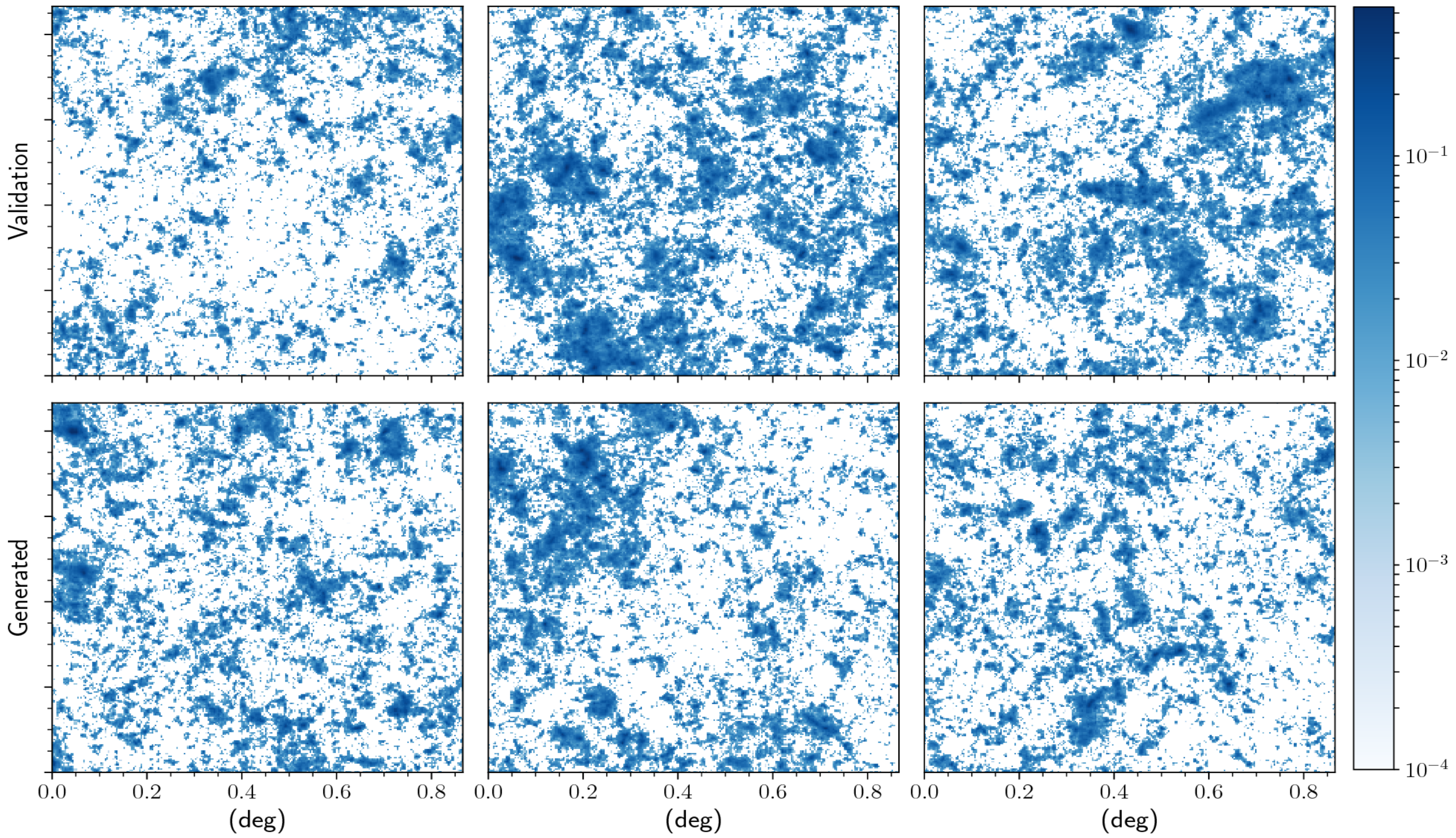
AI@SLAC, SLAC, Stanford University
10/03/2017

The inverse problem of cosmology

The sky surveys collected by observatory experiments pose an inverse problem:

given images of the sky and the “standard model” of cosmology (ΛCDM), can we extract the cosmological parameters of our universe?

Cosmo Convergence Maps



Weak lensing convergence maps κ(**ν**) for a ΛCDM cosmological model.

Generative Models

The central problem of generative models is that given a data distribution $\mathbb{P}_{data}$ can one devise a generator G such that the generated distribution

$$\mathbb{P}_{model} = \mathbb{P}_{data} \, ?$$

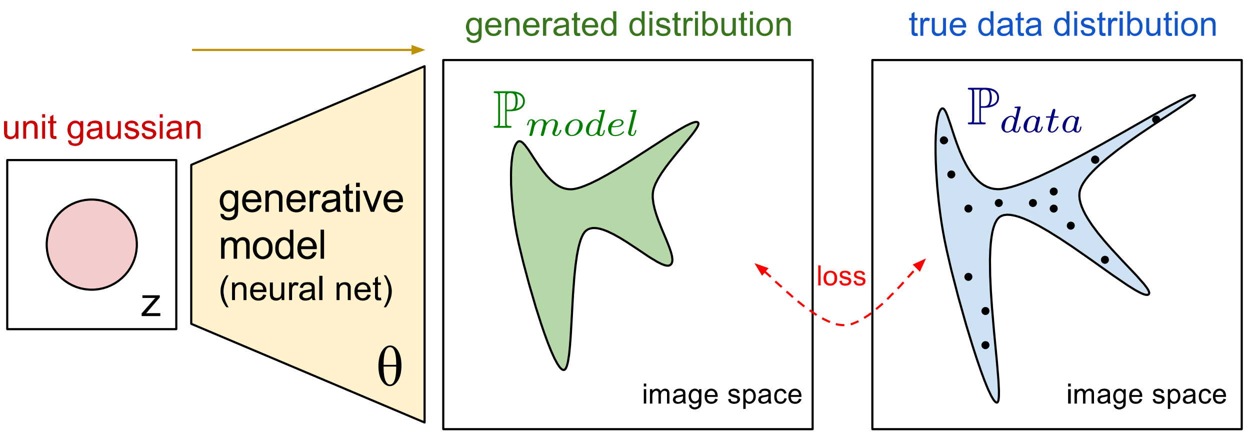
Generative Models

The central problem of generative models is that given a data distribution $\mathbb{P}_{data}$ can one devise a generator G such that the generated distribution

$$\mathbb{P}_{model} = \mathbb{P}_{data} ?$$

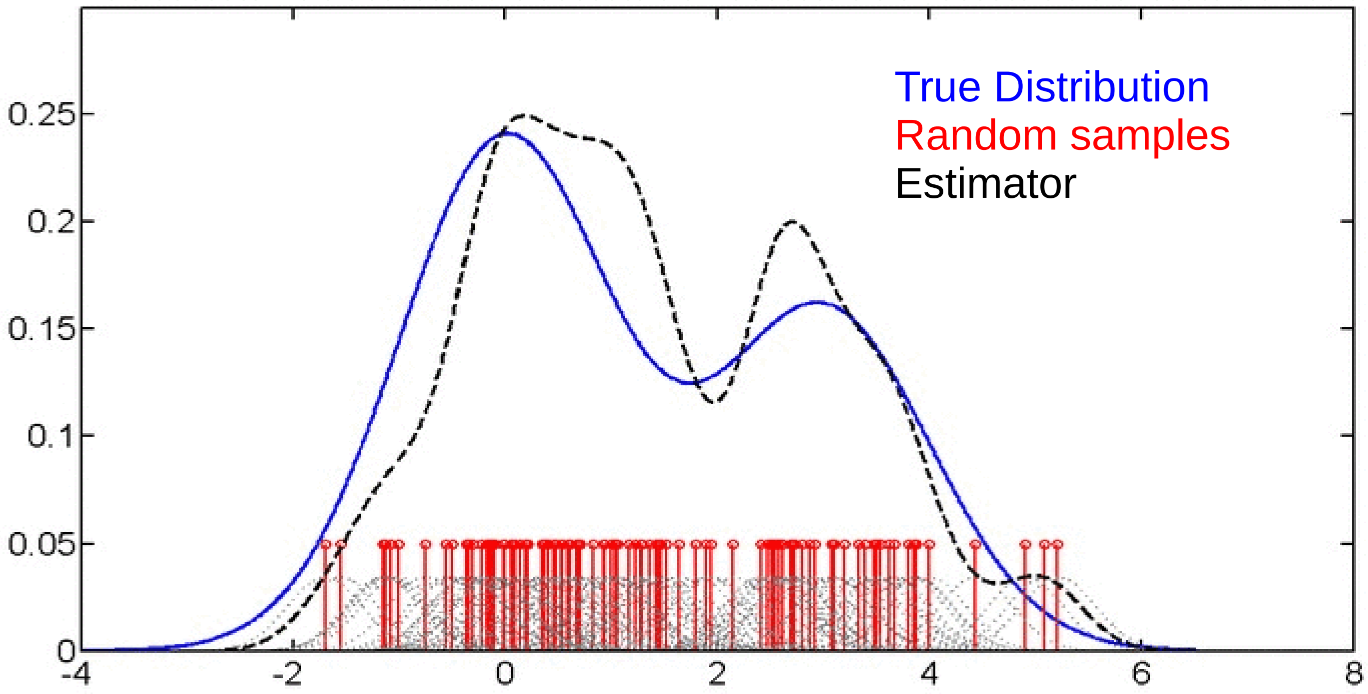
Our information about $\mathbb{P}_{data}$ comes from an independent and identically distributed sample $x_1, x_2, \ldots, x_n$ which is assumed to have the same distribution as $\mathbb{P}_{data}$.

Generative Models



blog.openai.com/generative-models

Density Estimation



Achieving a high fidelity generation scheme amounts to the construction of a density estimator of the training data.

Generative Adversarial Networks

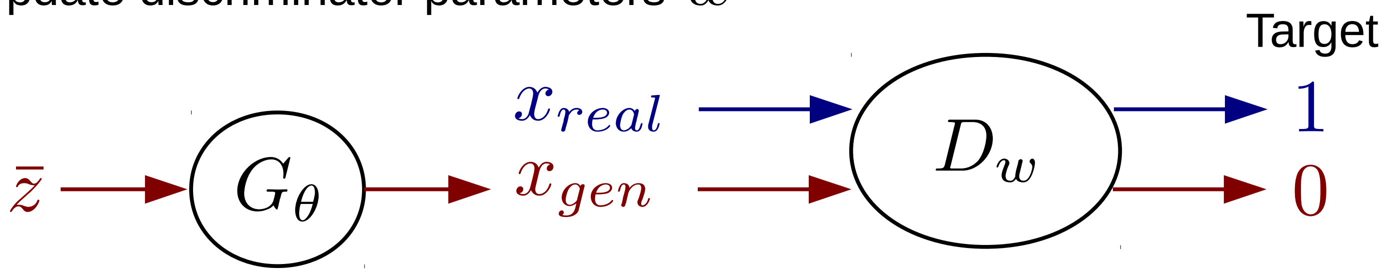
GANs, Goodfellow et al.arXiv:1406.2661

$$\bar{z} \longrightarrow G_\theta \longrightarrow x_{gen}$$

Generative Adversarial Networks

GANs, Goodfellow et al.arXiv:1406.2661

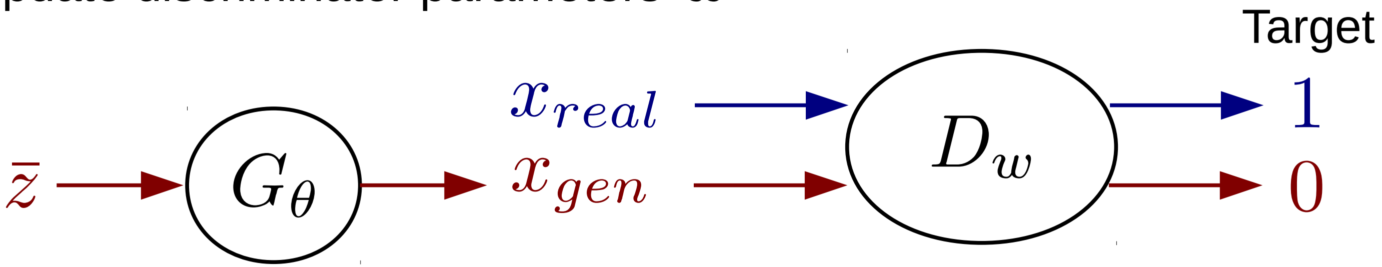
Update discriminator parameters w



Generative Adversarial Networks

GANs, Goodfellow et al.arXiv:1406.2661

Update discriminator parameters w

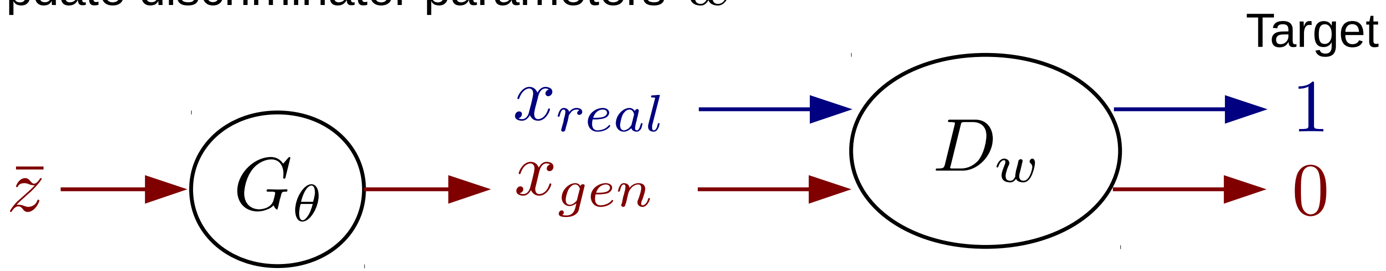


Update generator parameters θ

Generative Adversarial Networks

GANs, Goodfellow et al.arXiv:1406.2661

Update discriminator parameters w



Update generator parameters θ

$$\bar{z} \sim [\mathcal{N}_0(0,1), \ldots, \mathcal{N}_{63}(0,1)]$$

$$G_\theta : \bar{z} \rightarrow x \,\epsilon\, \mathbb{R}^{256\times256}$$

Generative Adversarial Networks – Loss function

Minimax game formulation (saturating):

$$J^{(D)} = -\frac{1}{2}\mathbb{E}_{x \sim \mathbb{P}_{data}} \log D(x) - \frac{1}{2}\mathbb{E}_{z \sim p_z} \log(1 - D((G(z)))$$

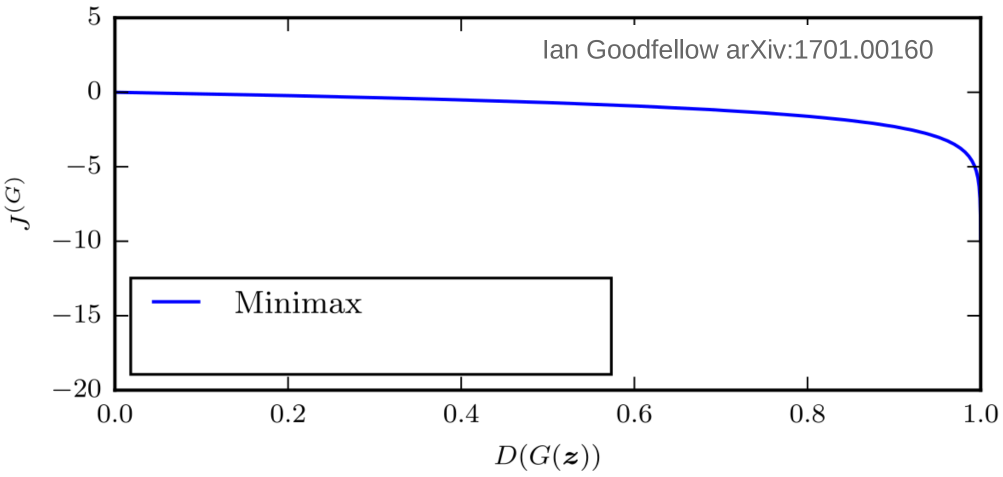
$$J^{(G)} = -J^{(D)}$$

Generative Adversarial Networks – Loss function

Minimax game formulation (saturating):

$$J^{(D)} = -\frac{1}{2}\mathbb{E}_{x\sim\mathbb{P}_{data}} \log D(x) - \frac{1}{2}\mathbb{E}_{z\sim p_z} \log(1 - D((G(z)))$$

$$J^{(G)} = -J^{(D)}$$



Generative Adversarial Networks – Loss function

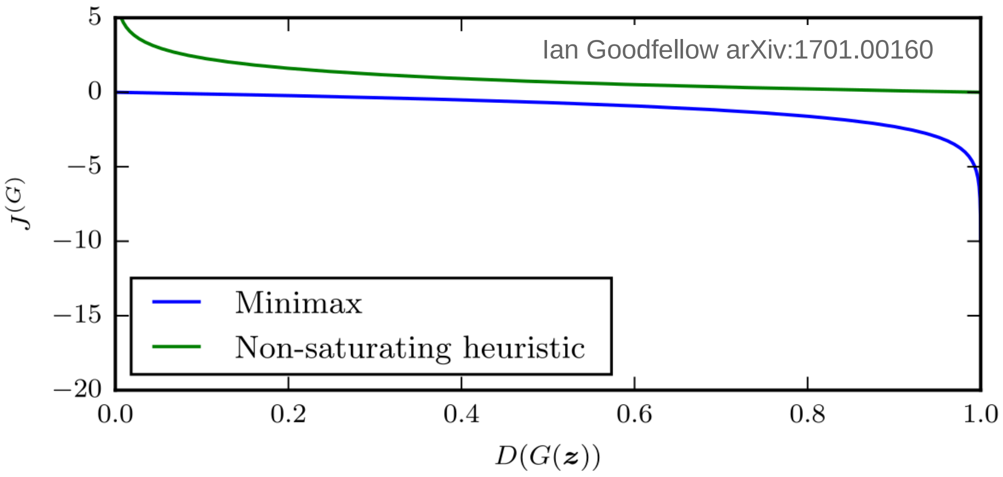
Minimax game formulation (saturating):

$$J^{(D)} = -\frac{1}{2}\mathbb{E}_{x\sim\mathbb{P}_{data}} \log D(x) - \frac{1}{2}\mathbb{E}_{z\sim p_z} \log(1 - D((G(z)))$$

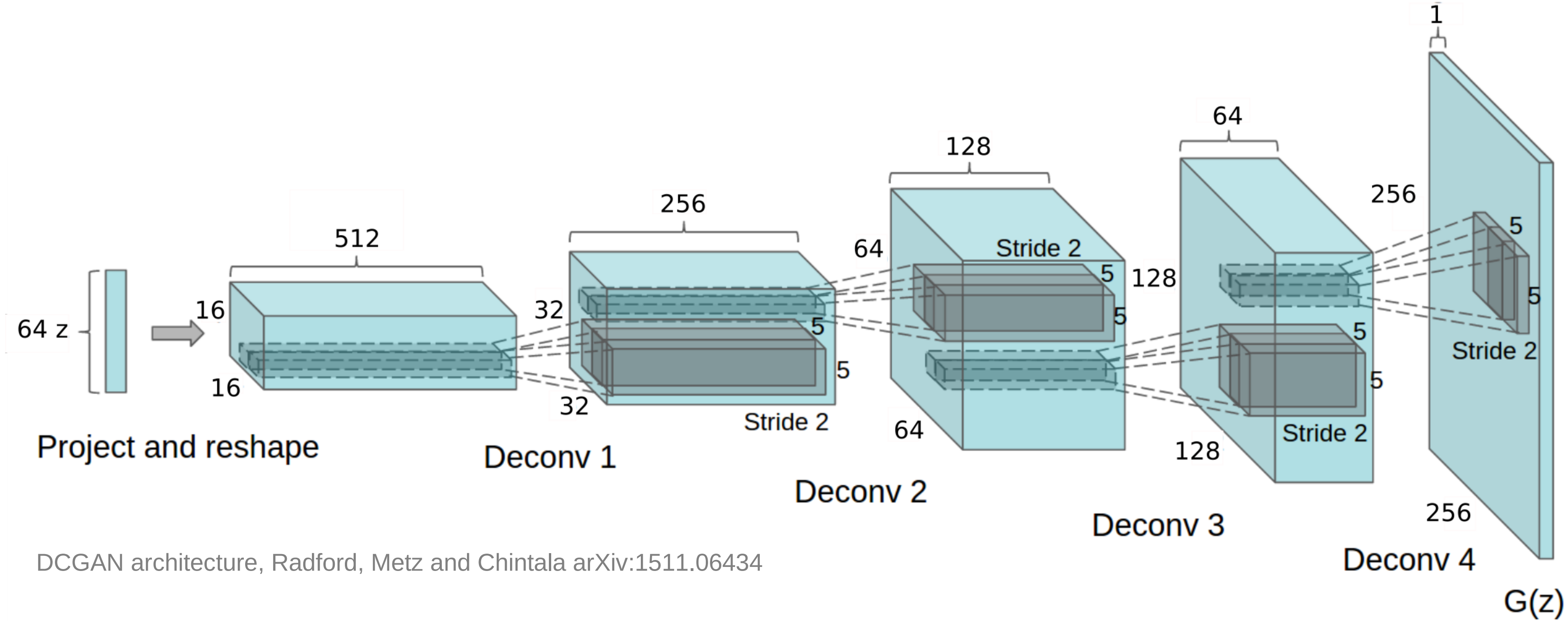
$$J^{(G)} = -J^{(D)}$$

Heuristic loss function (non-saturating):

$$J^{(G)} = -\frac{1}{2}\mathbb{E}_{z\sim p_z} \log D(G(z))$$



Deep Convolutional Generative Adversarial Networks (DCGAN)



DCGAN architecture, Radford, Metz and Chintala arXiv:1511.06434

Deep Convolutional Generative Adversarial Networks (DCGAN)

DCGAN architecture, Radford, Metz and Chintala arXiv:1511.06434

DCGAN generated celebrity face images

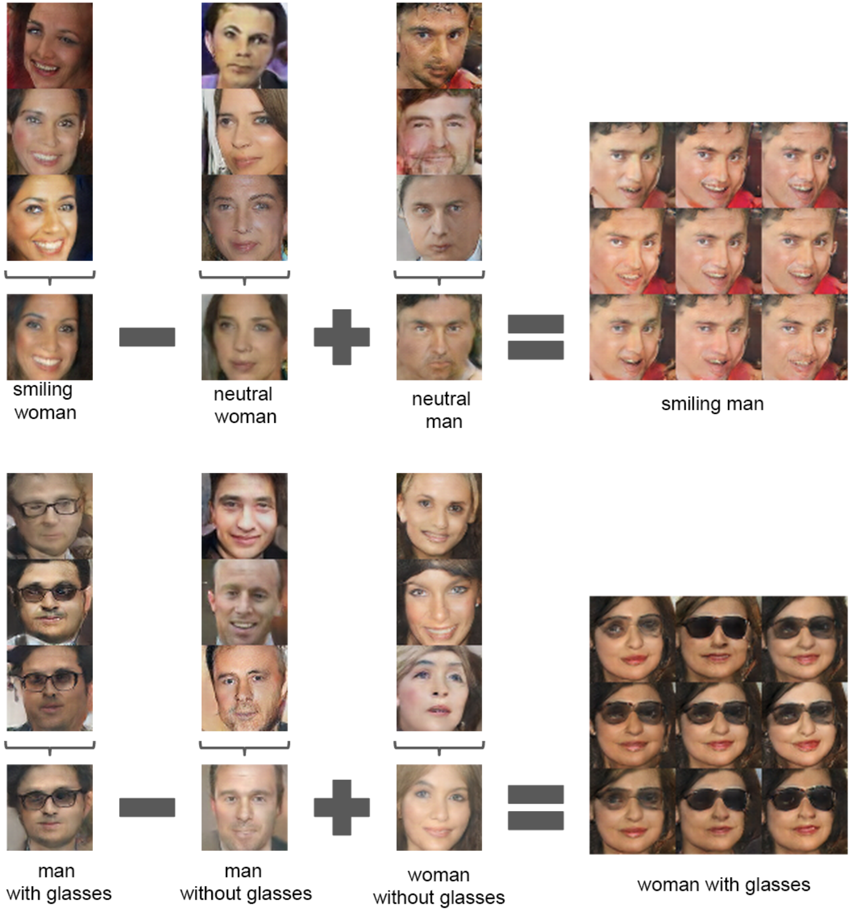
Deep Convolutional Generative Adversarial Networks (DCGAN)

DCGAN architecture, Radford, Metz and Chintala arXiv:1511.06434

Interpolation in the latent space. Rotations are linear!

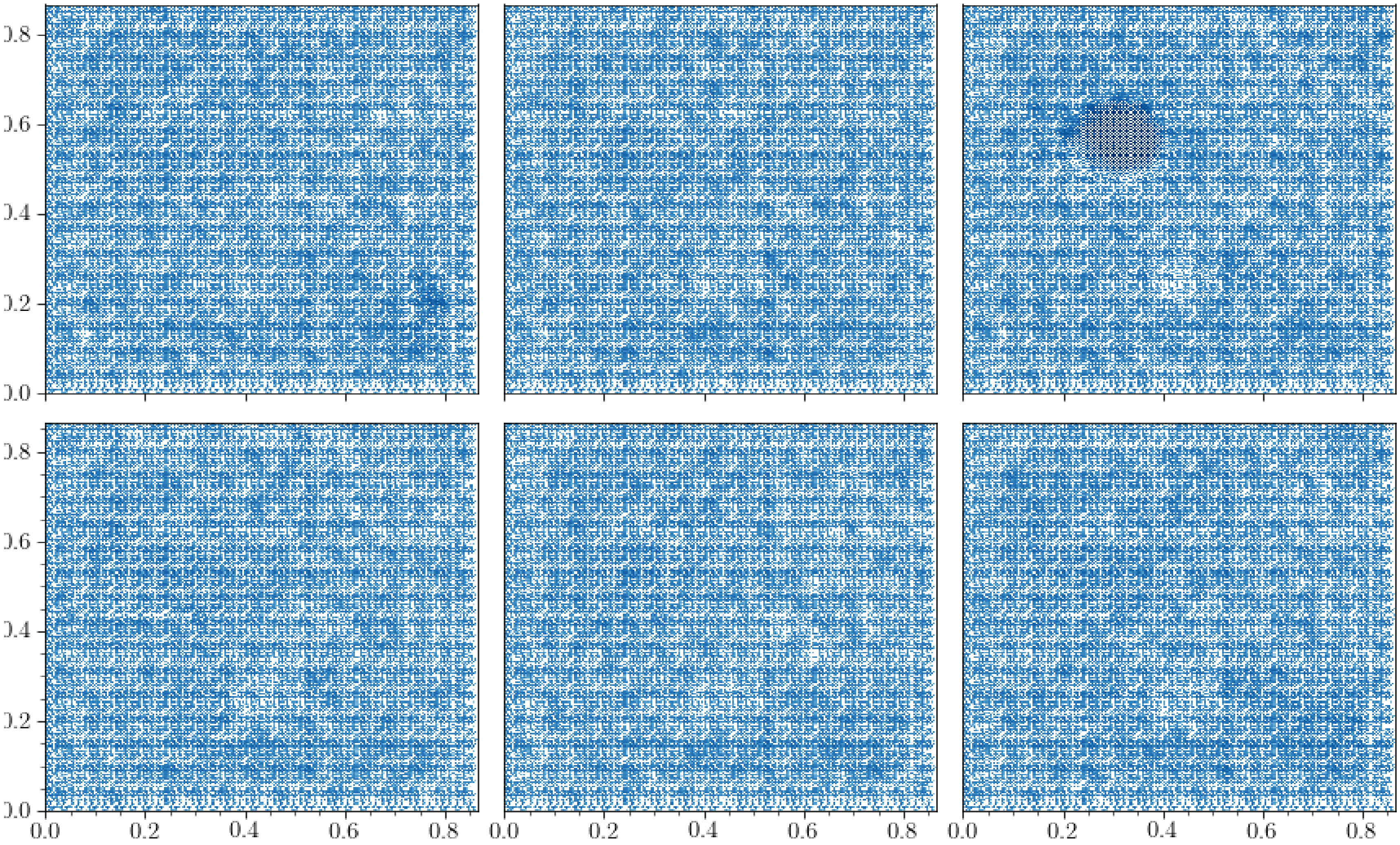
Deep Convolutional Generative Adversarial Networks (DCGAN)

DCGAN architecture, Radford, Metz and Chintala arXiv:1511.06434



Arithmetics in the latent space

Generative Adversarial Networks



Mustafa Mustafa (2017)

Evaluation of Generative Models

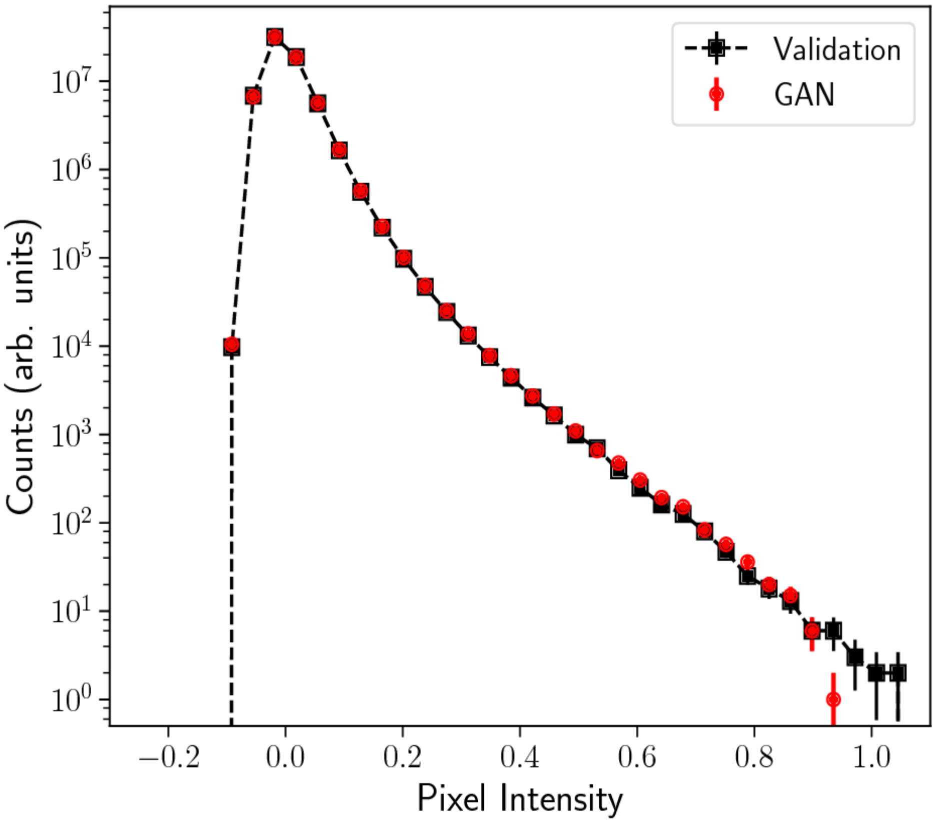
How close is $\mathbb{P}_{model}$ to $\mathbb{P}_{data}$?

Evaluation of Generative Models

How close is $\mathbb{P}_{model}$ to $\mathbb{P}_{data}$?

We think that when it comes to practical applications of generative models, such as in the case of emulating scientific data, the criterion to evaluate generative models is to study their *ability to reproduce the characteristic statistics* which we can measure on the original dataset.

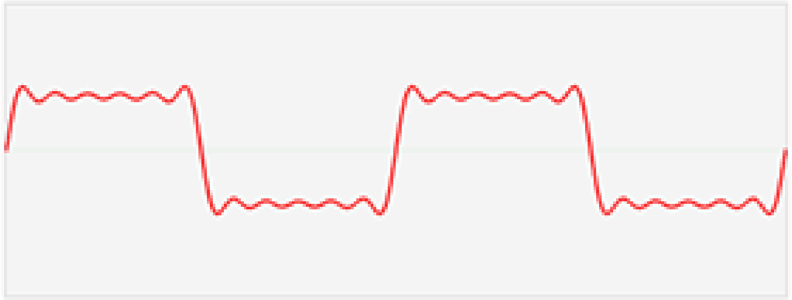
Convergence Maps First Order Statistics



Kolmogorov-Smirnov two tailed test yields p-value >0.999

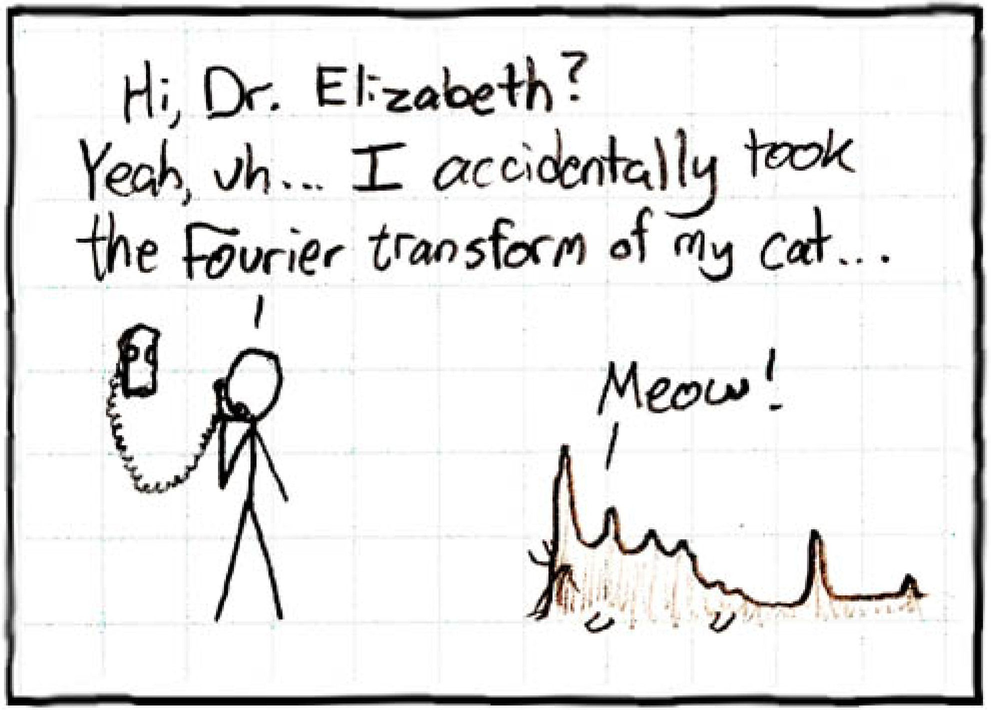
Fourier Spectral Analysis

quora.com/Whats-the-use-of-Fast-Fourier-Transform



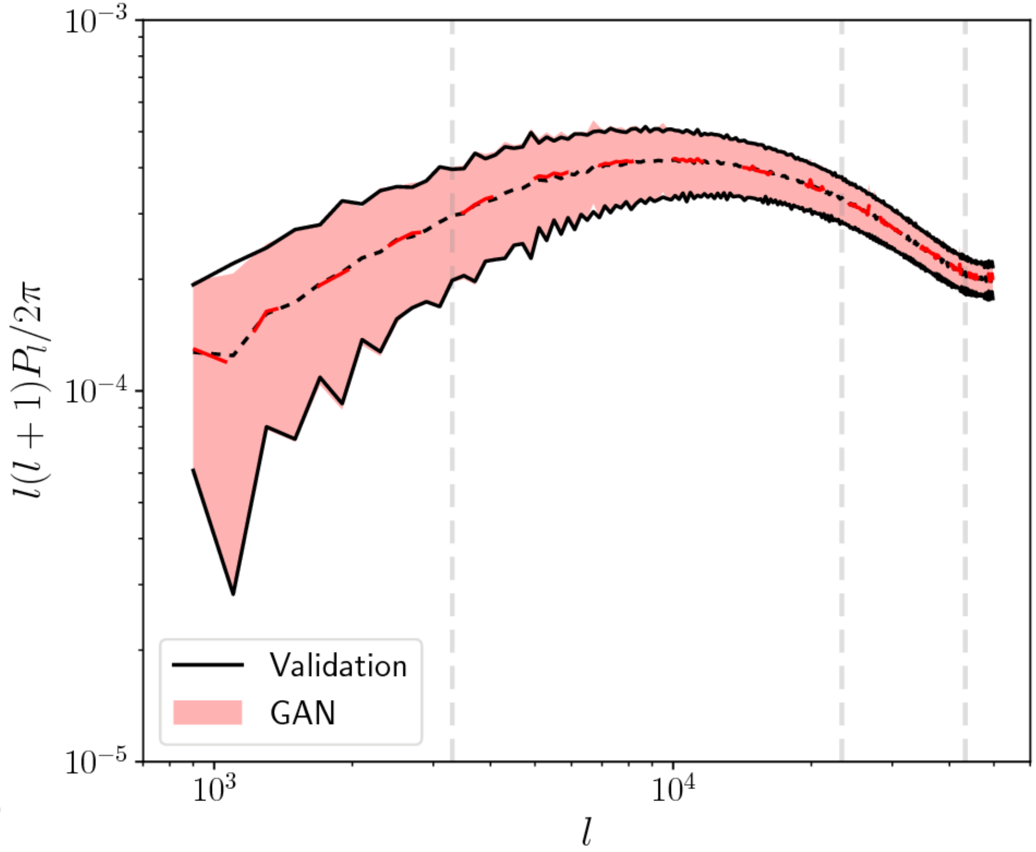
Fourier Spectral Analysis

https://xkcd.com/26/



Fourier Spectral Analysis: Power Spectrum

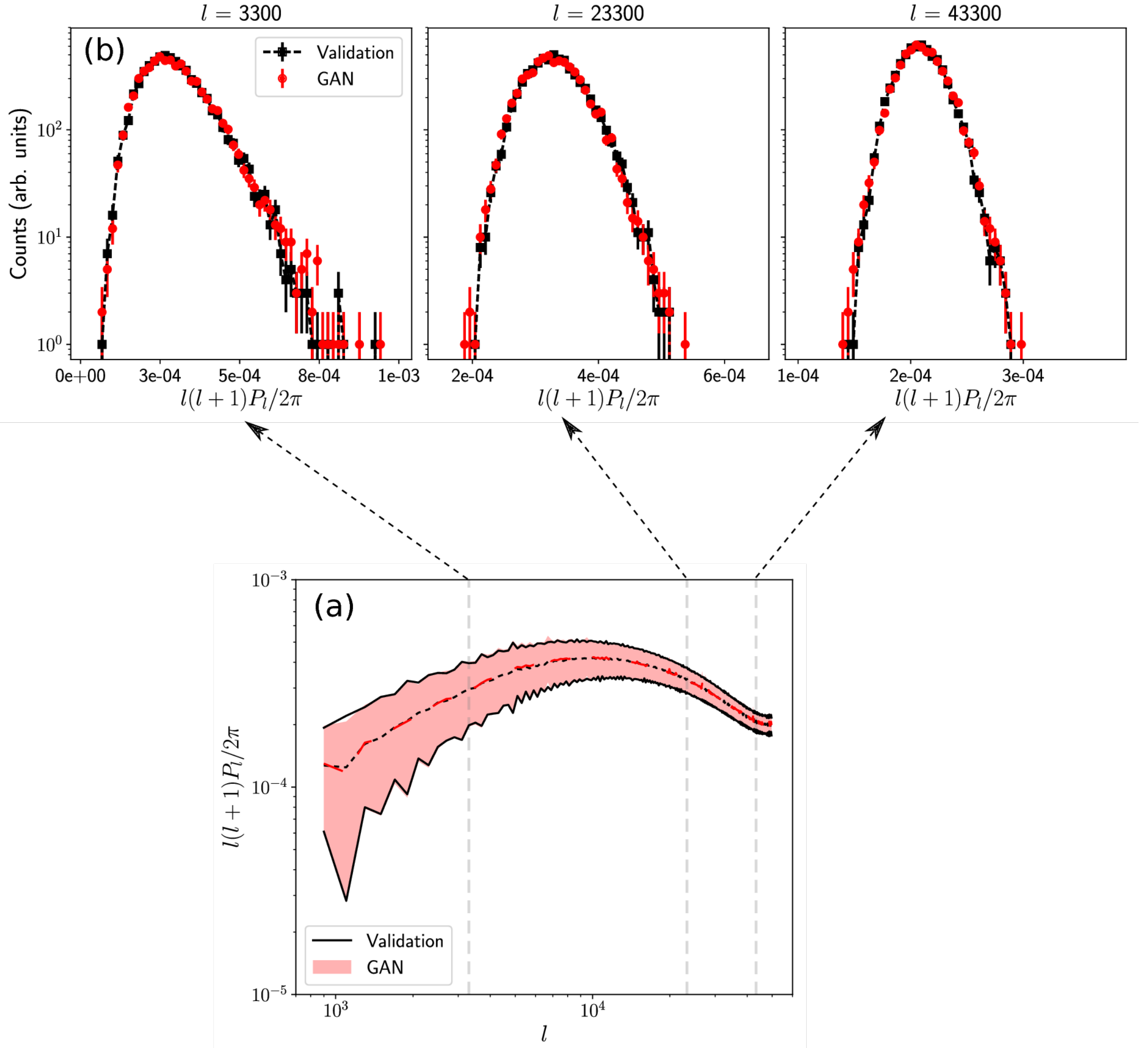
$$\langle \tilde{\kappa}(l)\tilde{\kappa}^*(l')\rangle = (2\pi)^2\delta_D(l-l')P_\kappa(l)$$



Bands are $\mu(l) \pm \sigma(l)$

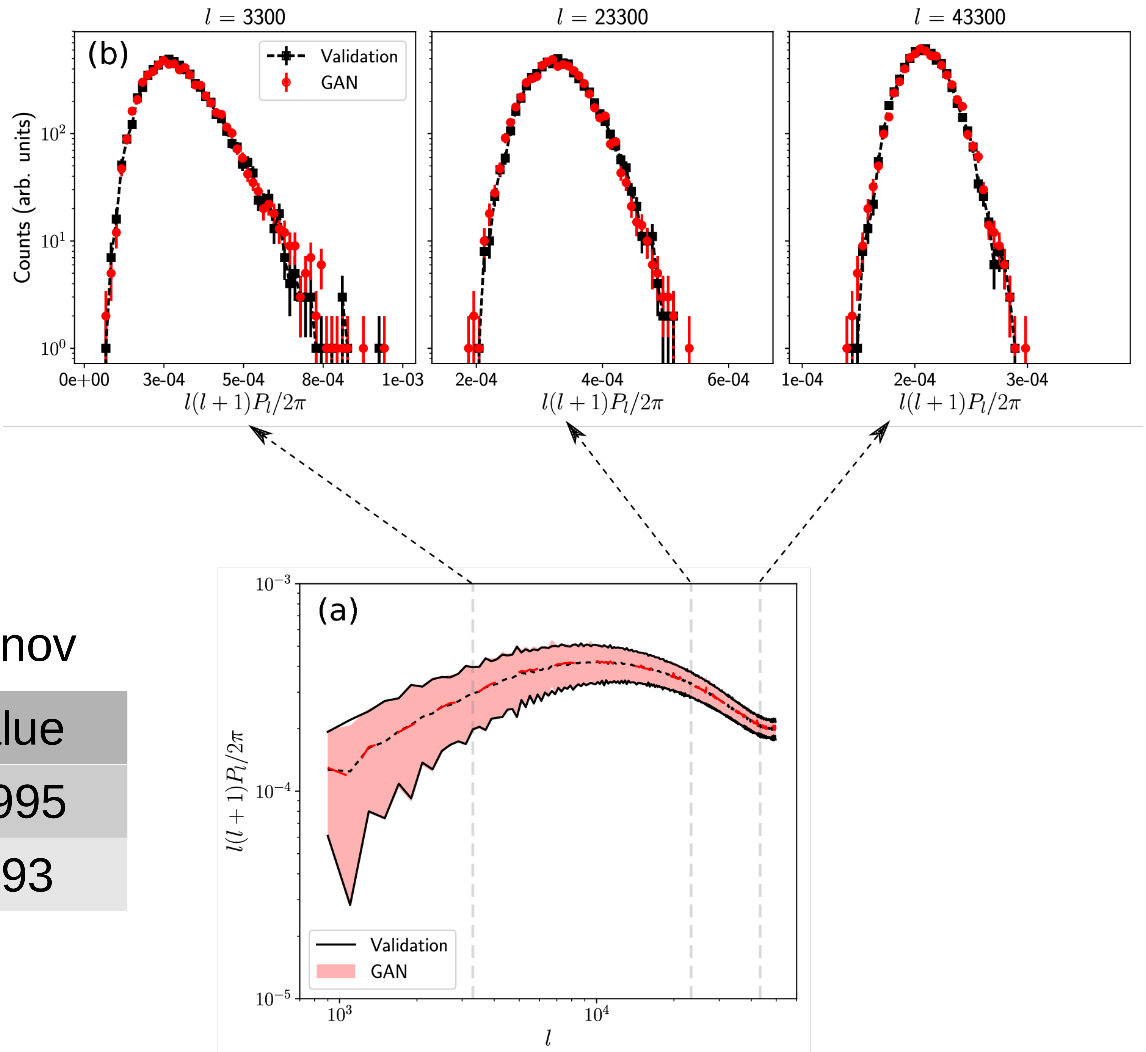
Mustafa Mustafa (2017)

Fourier Spectral Analysis: Power Spectrum



Mustafa Mustafa (2017)

Fourier Spectral Analysis: Power Spectrum



Kolmogorov-Smirnov

# moments	p-value
242	> 0.995
6	> 0.93

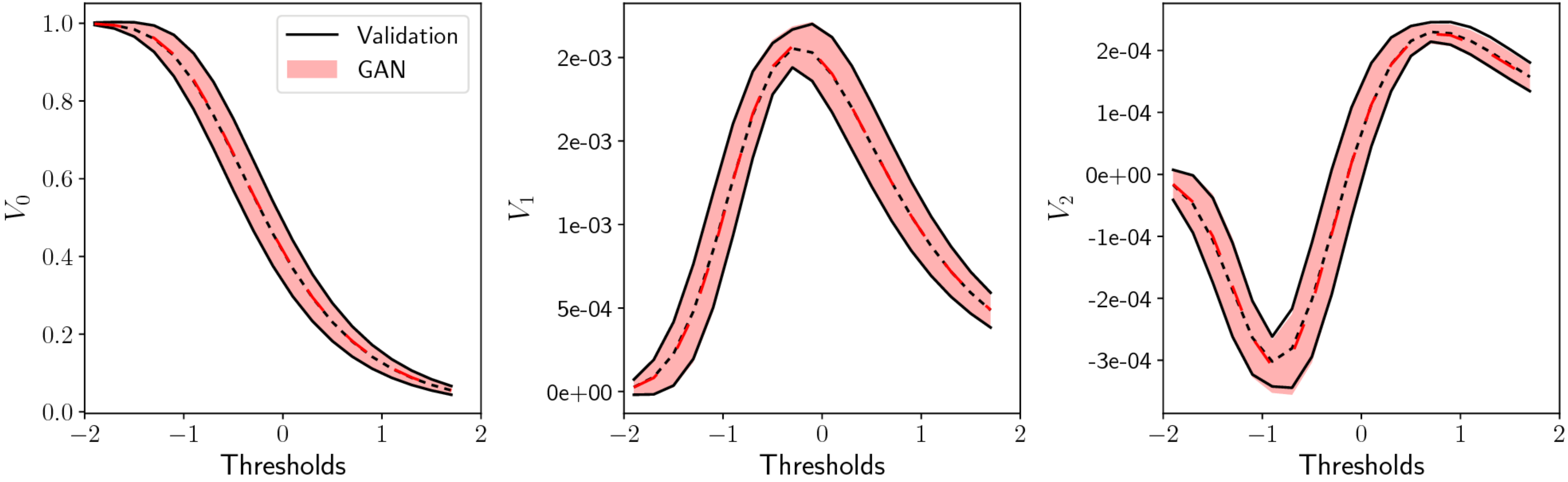
Mustafa Mustafa (2017)

Non-Gaussian Corrections

The power spectrum captures the Gaussian structures in the images. However, gravity produces non-Gaussian structures

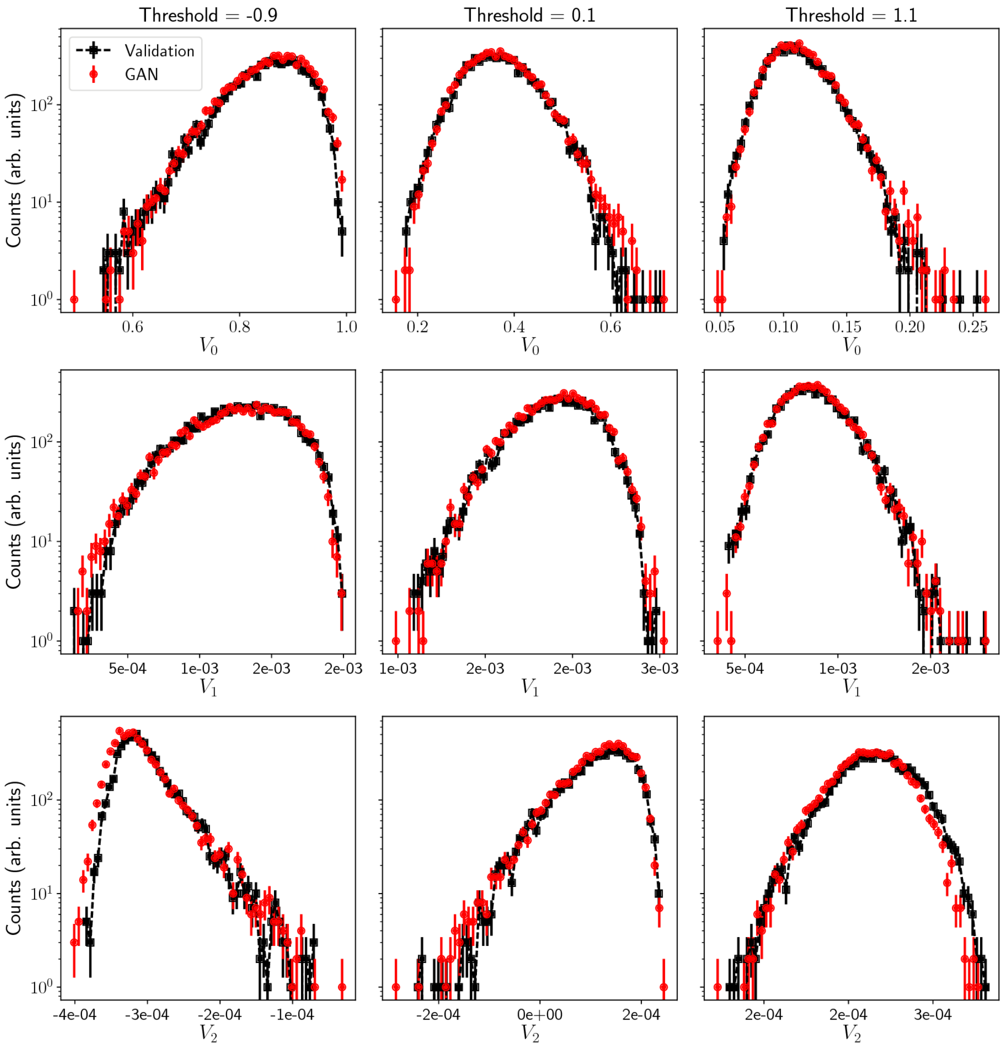
Non-Gaussian Corrections

The power spectrum captures the Gaussian structures in the images. However, gravity produces non-Gaussian structures

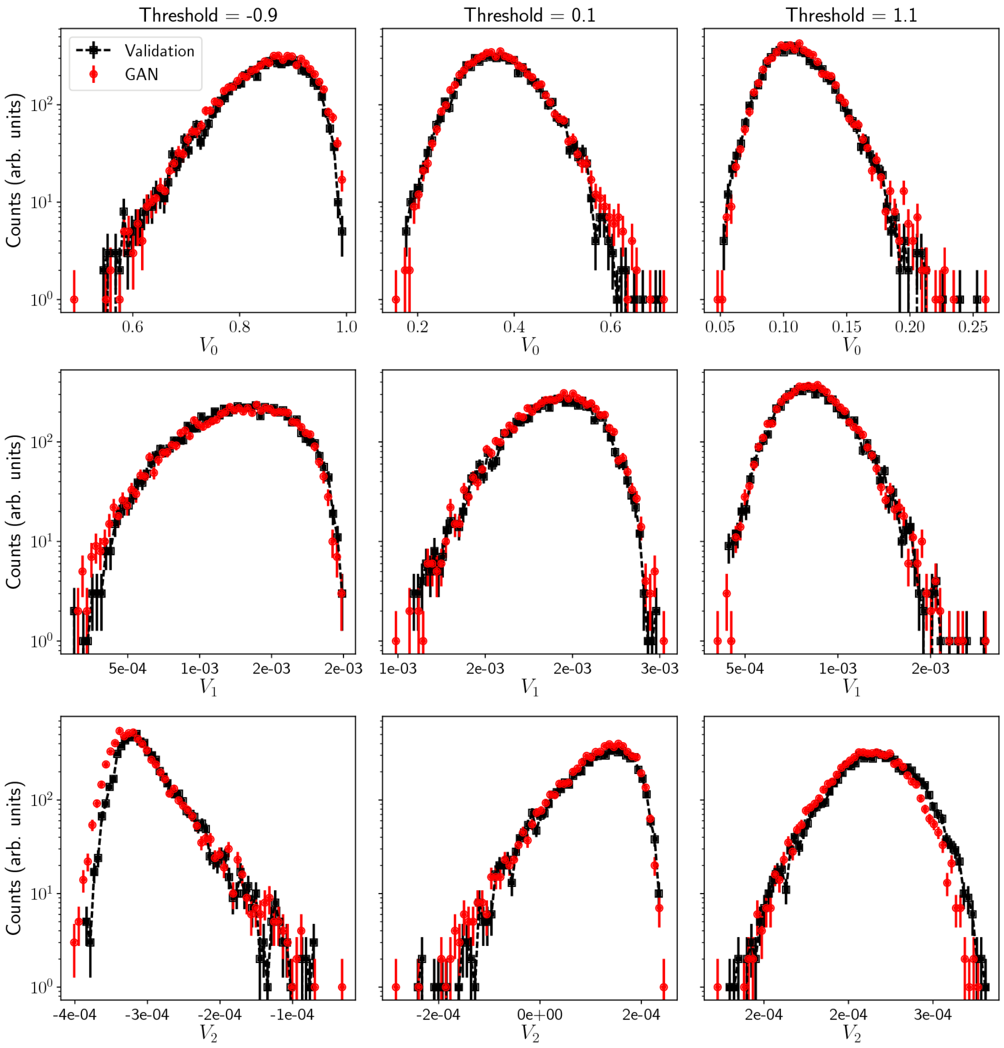


The three Minkowski Functionals are sensitive to the higher order correlations.

Non-Gaussian Corrections



Non-Gaussian Corrections



Kolmogorv-Smirnov

# thresholds	p-value
34	>0.999
16	>0.97
6	>0.6
1	0.32

Mustafa Mustafa (2017)

Generative Models for Emulating Scientific Data

Model parameters

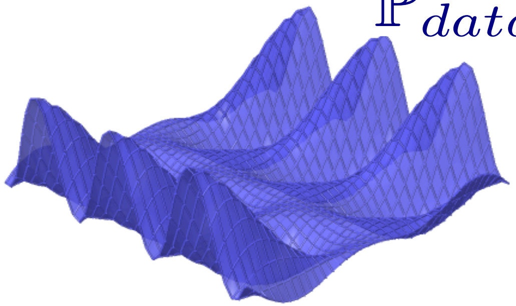
$$\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_m \end{bmatrix}$$

Physical Model

$$\boxed{S(\bar{\sigma}, r)}$$

Data of interest,
images or otherwise.

$\mathbb{P}_{data}$



Generative Models for Emulating Scientific Data

Random seed(s)
r

Model parameters

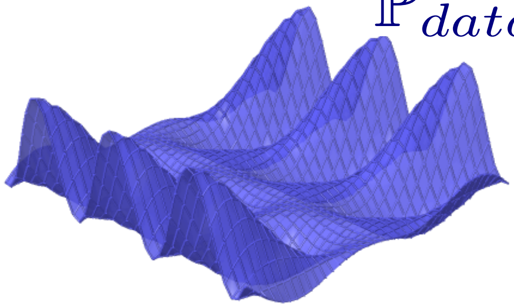
$$\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_m \end{bmatrix}$$

Physical Model

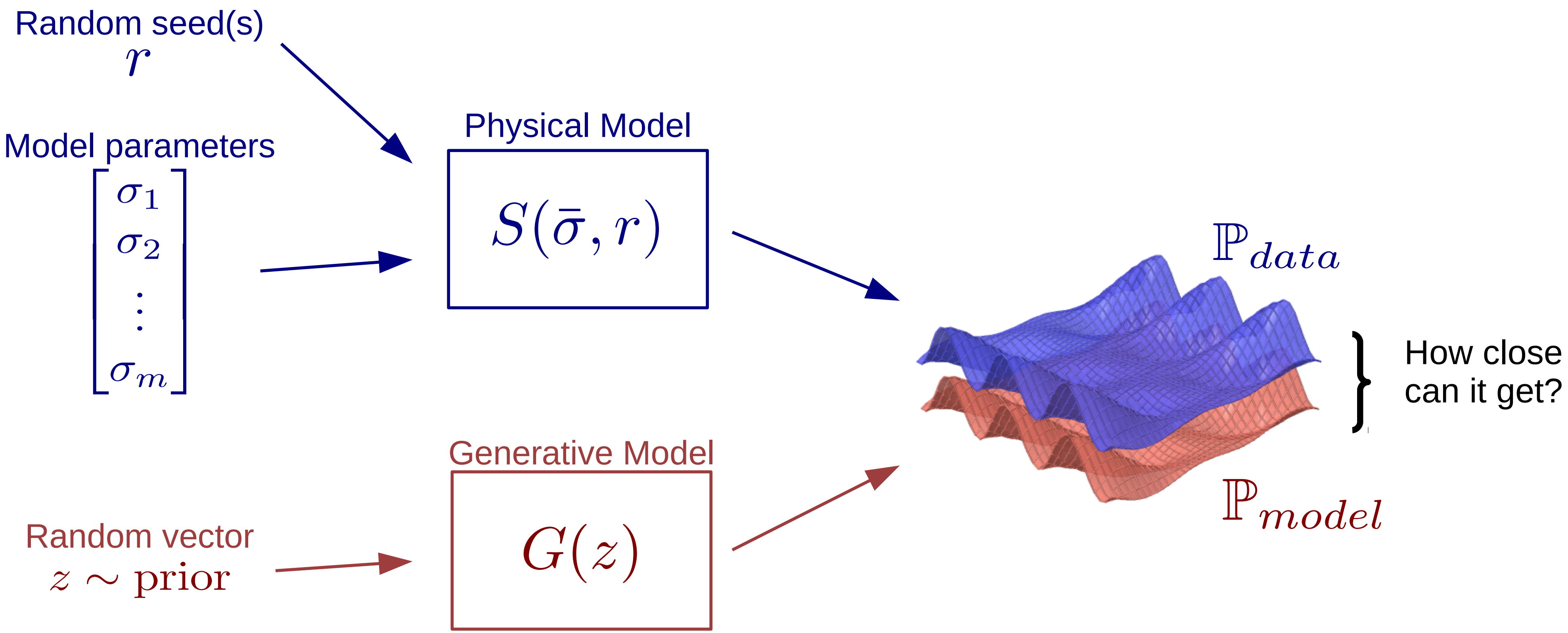
$$\boxed{S(\bar{\sigma}, r)}$$

Data of interest,
images or otherwise.

$\mathbb{P}_{data}$



Generative Models for Emulating Scientific Data



Generative Models for Emulating Scientific Data

Can we use GANs to build parametric density estimators $G(\bar{\sigma}, z)$ of scientific data?

Random seed(s)
r

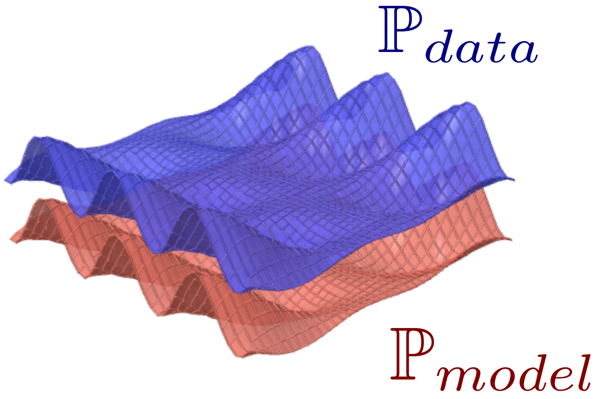
Model parameters
$$\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_m \end{bmatrix}$$

Physical Model
$$\boxed{S(\bar{\sigma}, r)}$$

?

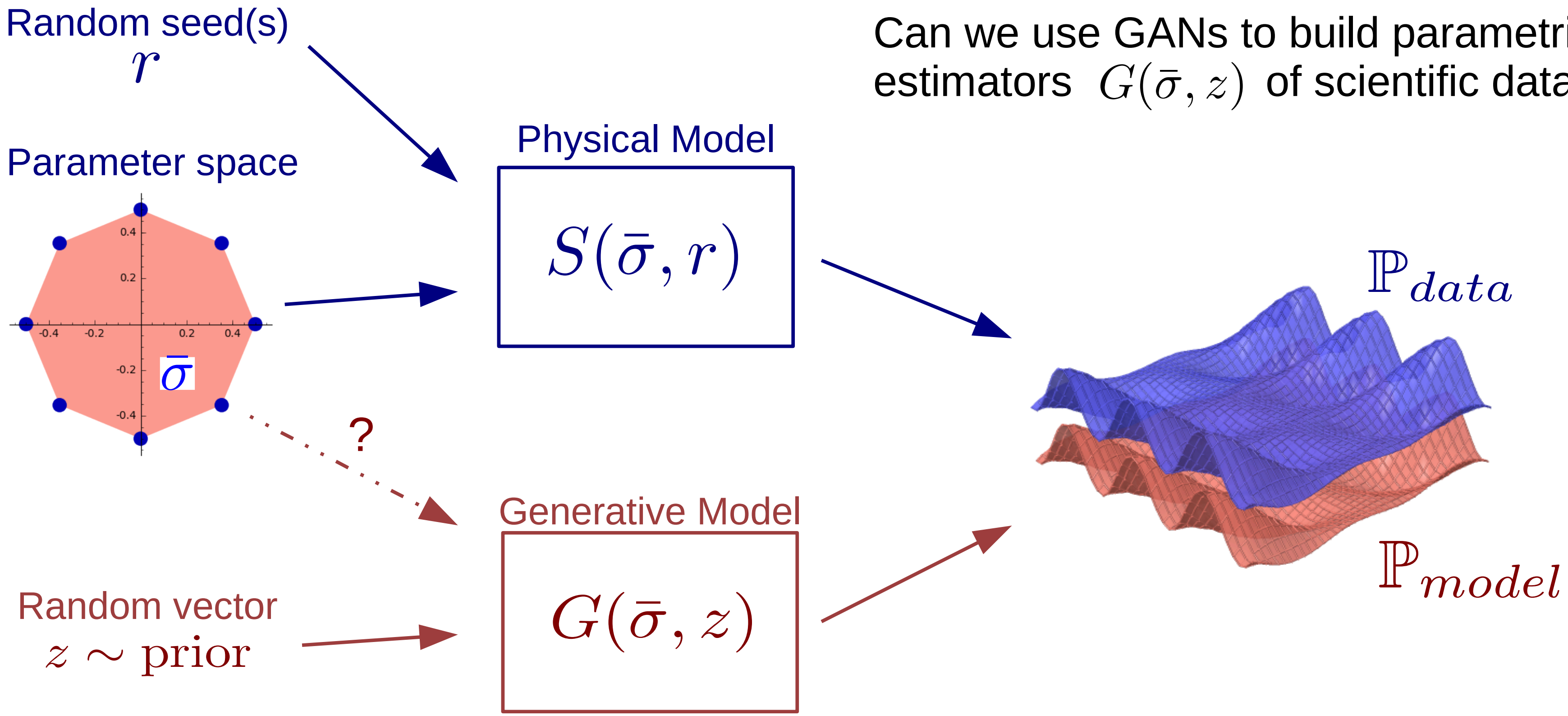
Generative Model
$$\boxed{G(\bar{\sigma}, z)}$$

Random vector
$z \sim \text{prior}$



Generative Models for Emulating Scientific Data

Can we use GANs to build parametric density estimators $G(\bar{\sigma}, z)$ of scientific data?



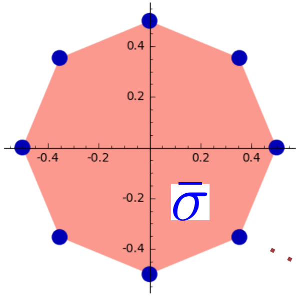
Mustafa Mustafa (2017)

Generative Models for Emulating Scientific Data

Can we use GANs to build parametric density estimators $G(\bar{\sigma}, z)$ of scientific data?

Random seed(s)
r

Parameter space



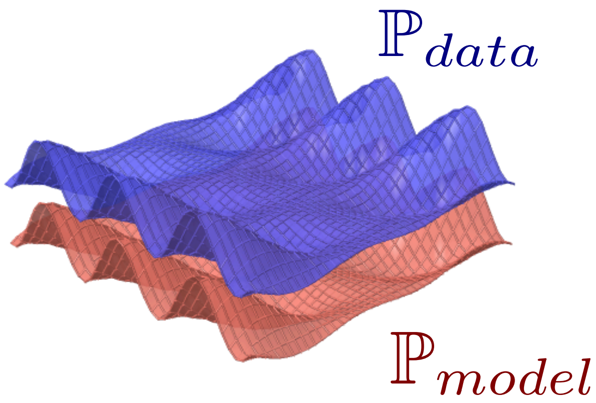
Physical Model

$$S(\bar{\sigma}, r)$$

Generative Model

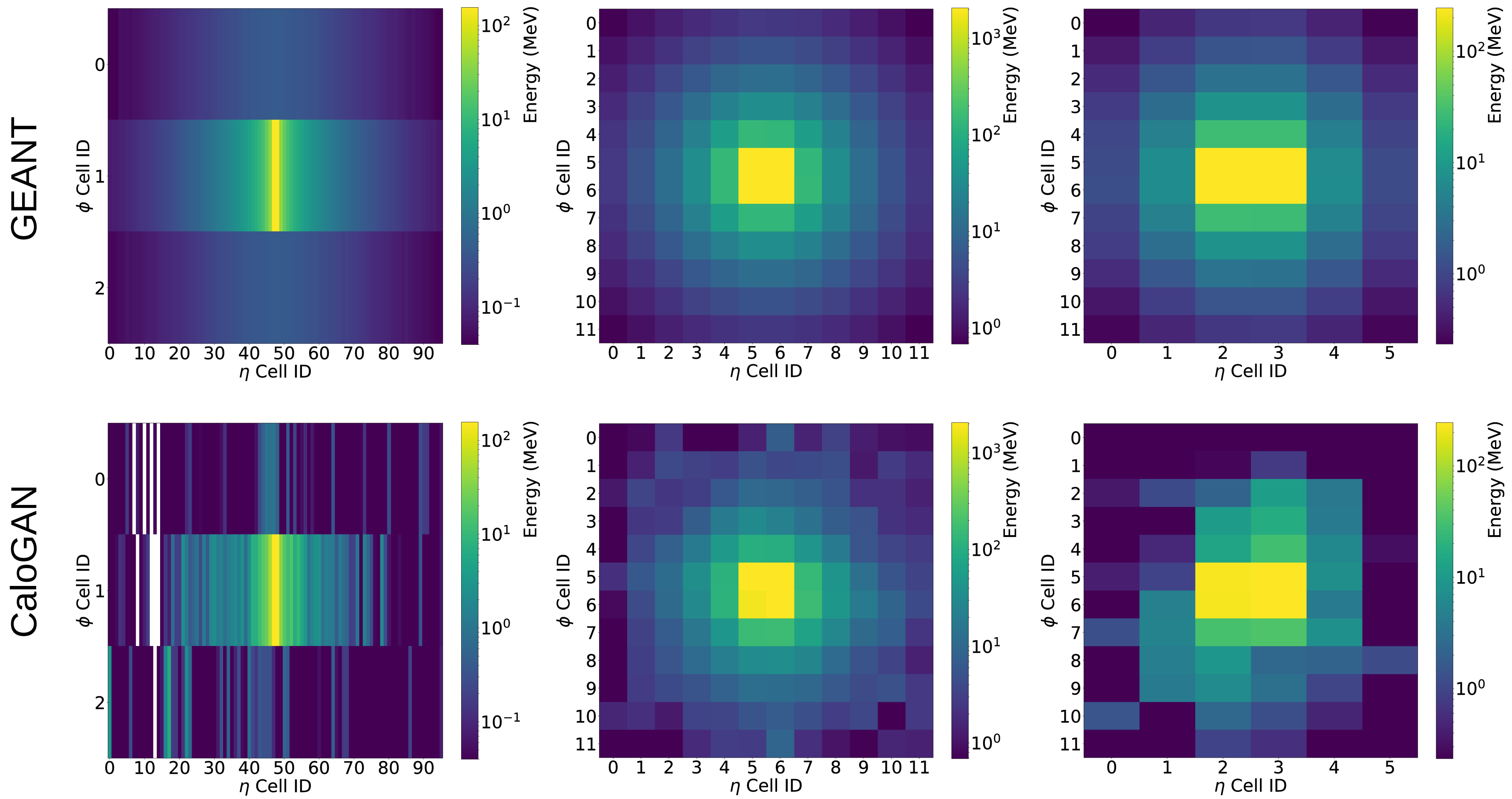
$$G(\bar{\sigma}, z)$$

Random vector
$z \sim \mathrm{prior}$



Such generators would likely exclude regions in parameter space where the physical model $S(\bar{\sigma}, r)$ exhibits critical behavior.

CaloGAN: Simulating 3D Calorimeter Showers using GANs



Paganini, de Oliveira and Nachman arXiv:1705.02355

Summary and Outlook

- We have shown with statistical confidence that GANs can emulate ΛCDM cosmological model convergence maps
  - Fourier spectrum of generated maps match that of a validation dataset
  - Non-Gaussian structures are discovered and emulated by the generator
- Deep generative models have the potential of creating high-fidelity computationally inexpensive emulators of scientific data.

Summary and Outlook

- We have shown with statistical confidence that GANs can emulate ΛCDM cosmological model convergence maps
  - Fourier spectrum of generated maps match that of a validation dataset
  - Non-Gaussian structures are discovered and emulated by the generator

- Deep generative models have the potential of creating high-fidelity computationally inexpensive emulators of scientific data.
- Manuscript: “Creating Virtual Universes using Generative Adversarial Networks”, Mustafa, Bard, Lukic, Al-Rfou, Bhimji, arXiv:1706.02390

Summary and Outlook

- We have shown with statistical confidence that GANs can emulate ΛCDM cosmological model convergence maps
  - Fourier spectrum of generated maps match that of a validation dataset
  - Non-Gaussian structures are discovered and emulated by the generator

- Deep generative models have the potential of creating high-fidelity computationally inexpensive emulators of scientific data.
- Manuscript: "Creating Virtual Universes using Generative Adversarial Networks", Mustafa, Bard, Lukic, Al-Rfou, Bhimji, arXiv:1706.02390

- Future lines of work:
  - investigate the ability of GANs to interpolate in the parameter space of physical models
  - multi dimensional data: 1 & 3D (see CaloGAN), time dependence, "sequential data"
  - using NN interpretation techniques to gain insight in what these networks are learning