

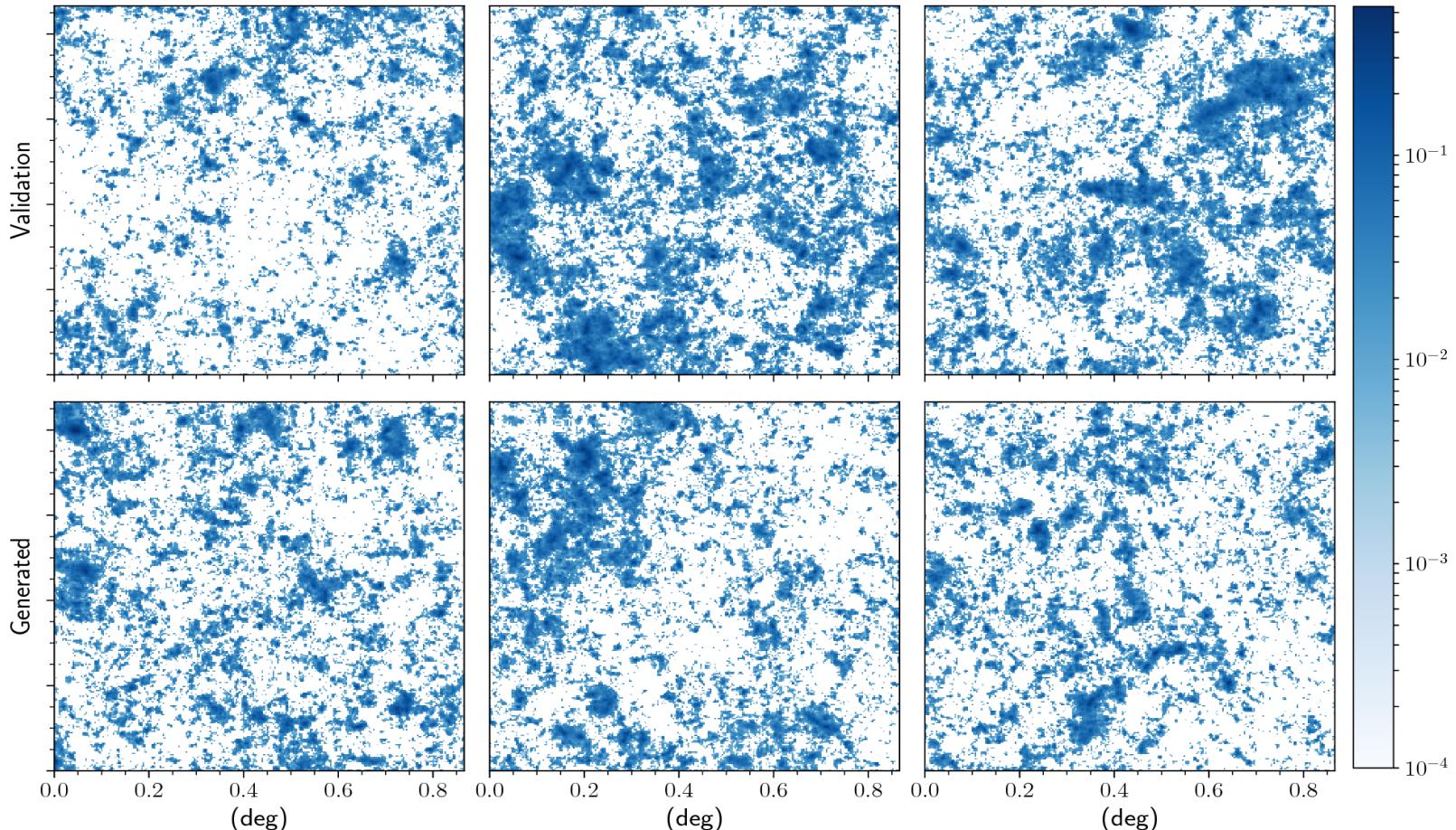
# Towards a Cosmology Emulator using Generative Adversarial Networks

## (Simulation Emulation)

Mustafa Mustafa  
Berkeley Lab.

Heavy Ion Tea, Berkeley Lab.  
06/06/2017

# Cosmo Convergence Maps



Weak lensing convergence maps  $\kappa(v)$  for a  $\Lambda$ CDM 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

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

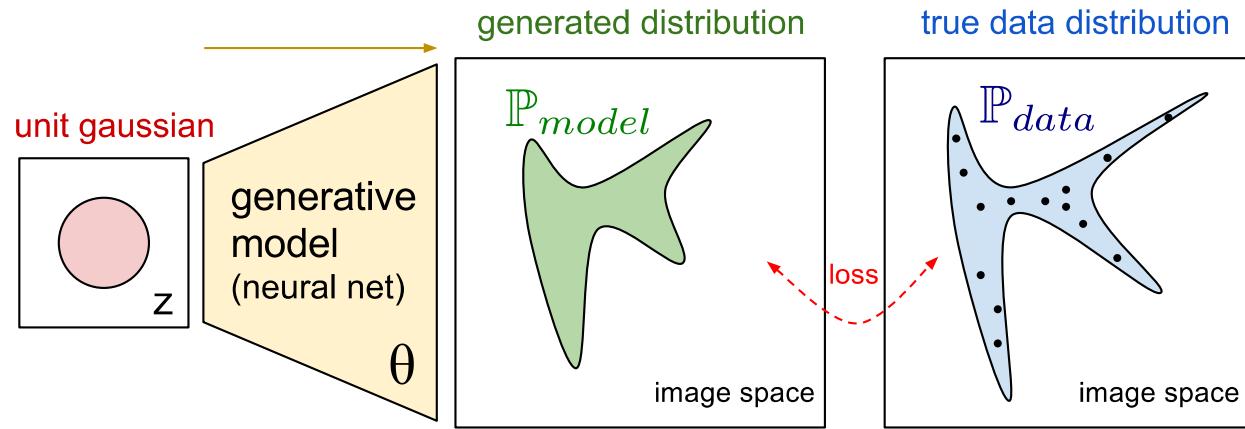
## Generative Models

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  $x_1, x_2, \dots, x_n$  which is assumed to have the same distribution as  $\mathbb{P}_{data}$ .

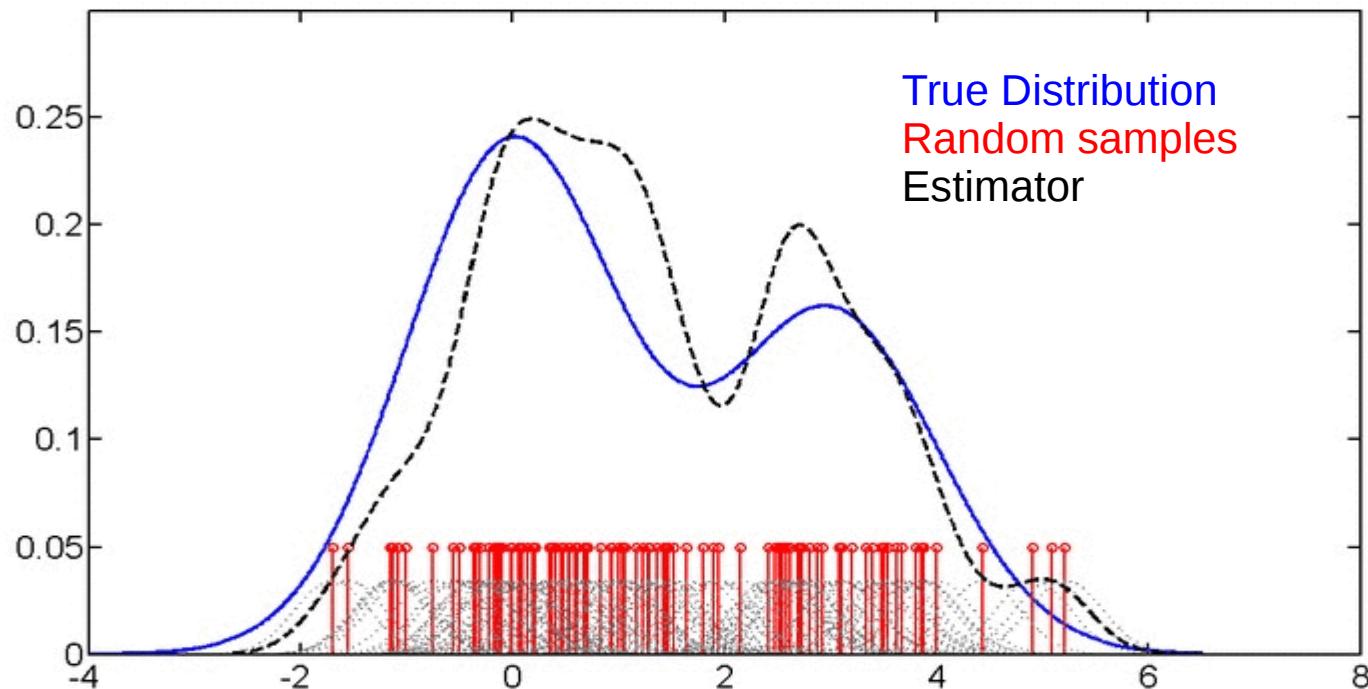
# Generative Models



[blog.openai.com/generative-models](http://blog.openai.com/generative-models)

# Density Estimation

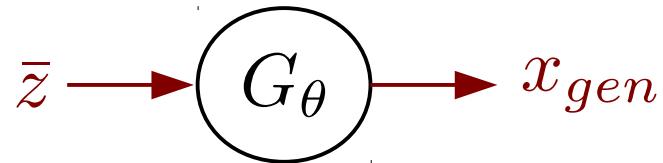
Creative Commons Attribution-Share Alike 3.0 Unported license



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

# Generative Adversarial Networks

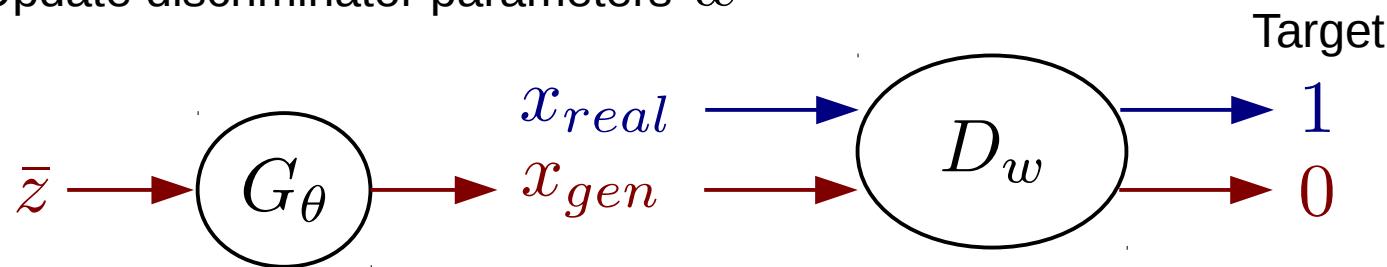
GANs, Goodfellow et al.arXiv:1406.2661



# Generative Adversarial Networks

GANs, Goodfellow et al.arXiv:1406.2661

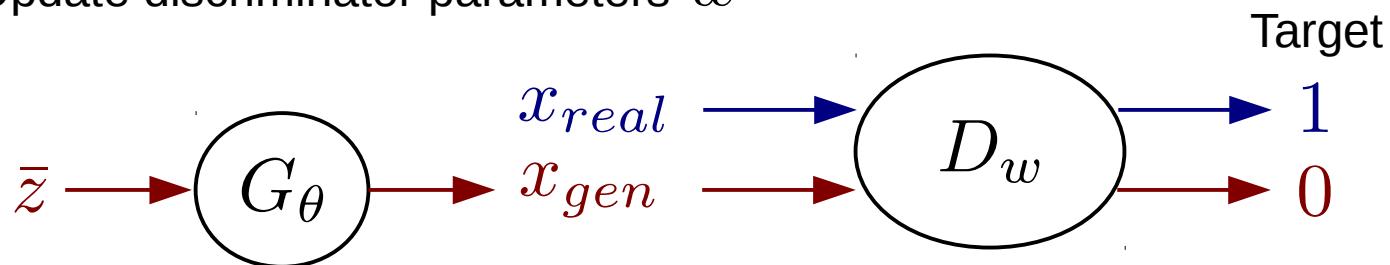
Update discriminator parameters  $w$



# Generative Adversarial Networks

GANs, Goodfellow et al.arXiv:1406.2661

Update discriminator parameters  $w$



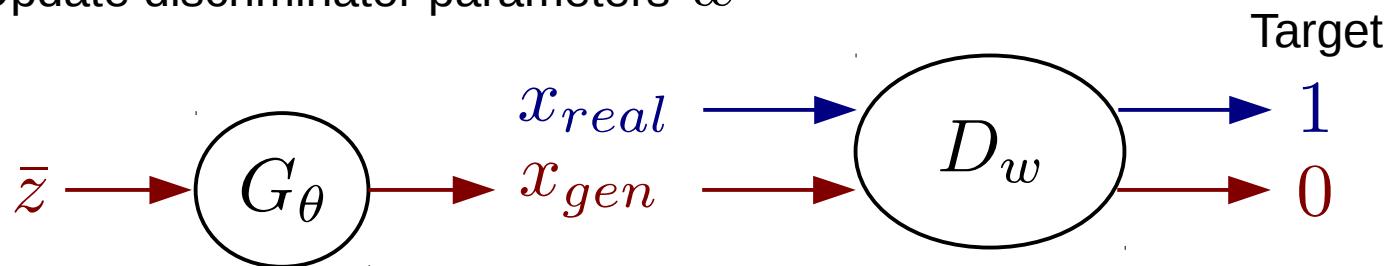
Update generator parameters  $\theta$



# Generative Adversarial Networks

GANs, Goodfellow et al.arXiv:1406.2661

Update discriminator parameters  $w$



Update generator parameters  $\theta$



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

$$G_\theta : \bar{z} \rightarrow x \in \mathbb{R}^{256 \times 256}$$

## Generative Adversarial Networks – Loss function

Minimax game formulation (saturating):

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

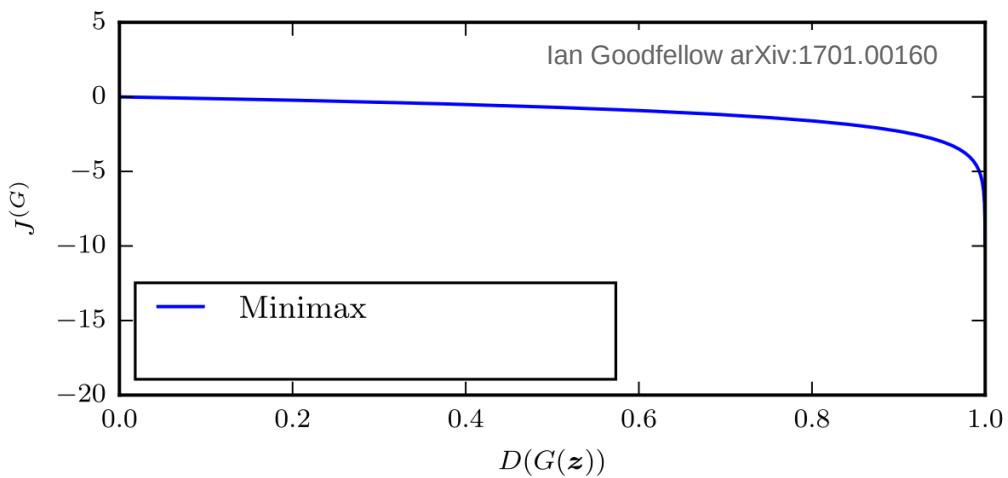
$$J^{(G)} = -J^{(D)}$$

# 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

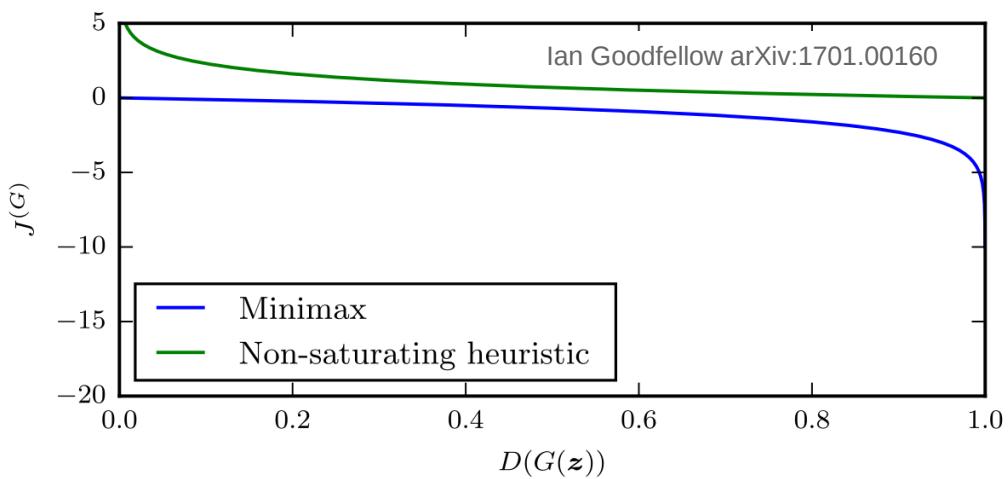
Minimax game formulation (saturating):

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

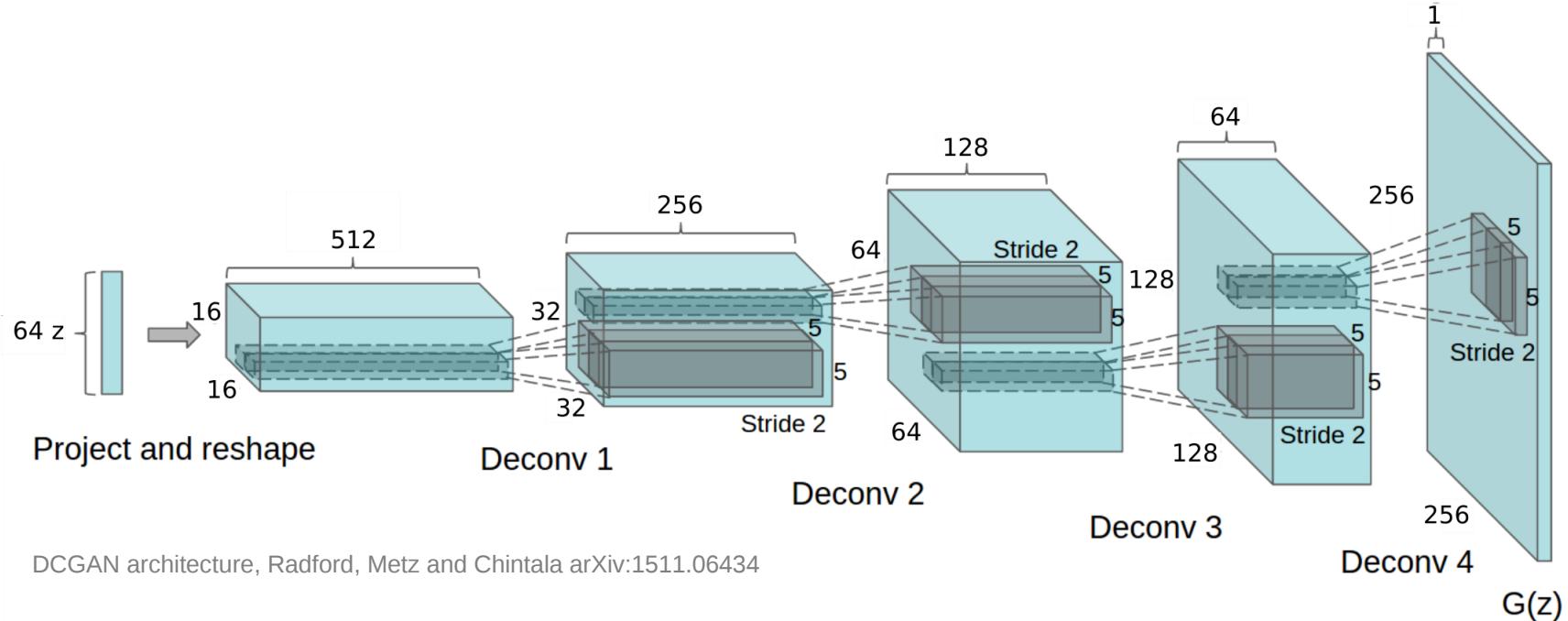
$$J^{(G)} = -J^{(D)}$$

Heuristic loss function (non-saturating):

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



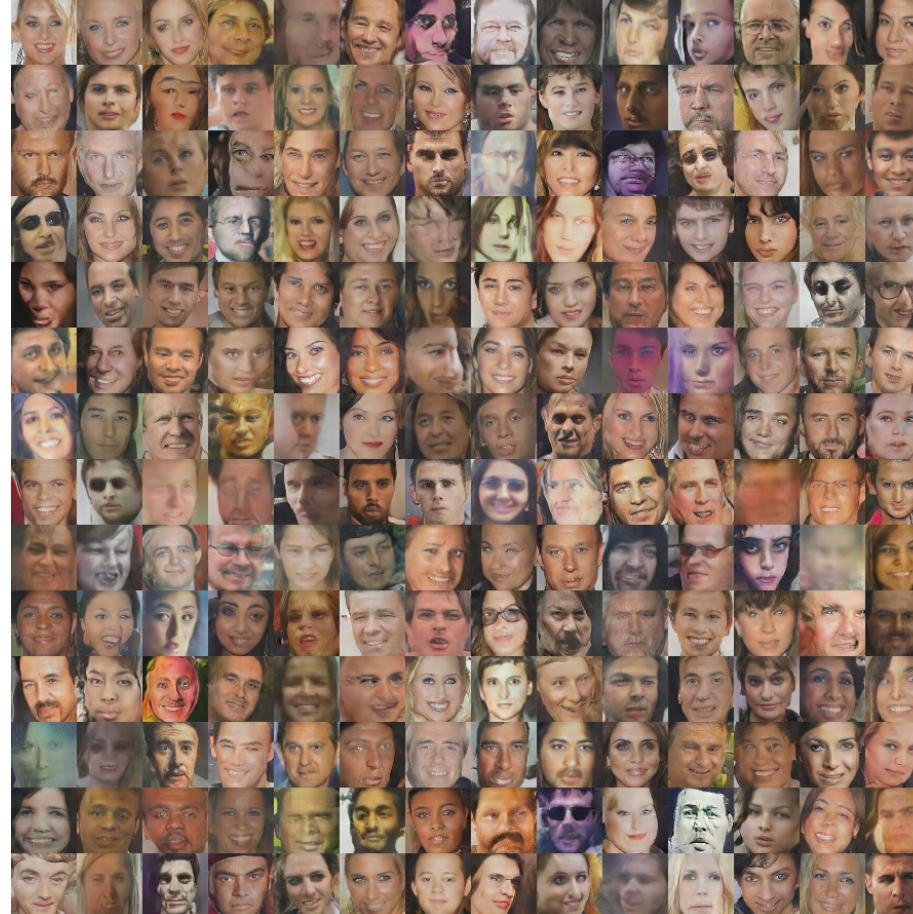
# Deep Convolutional Generative Adversarial Networks (DCGAN)



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)

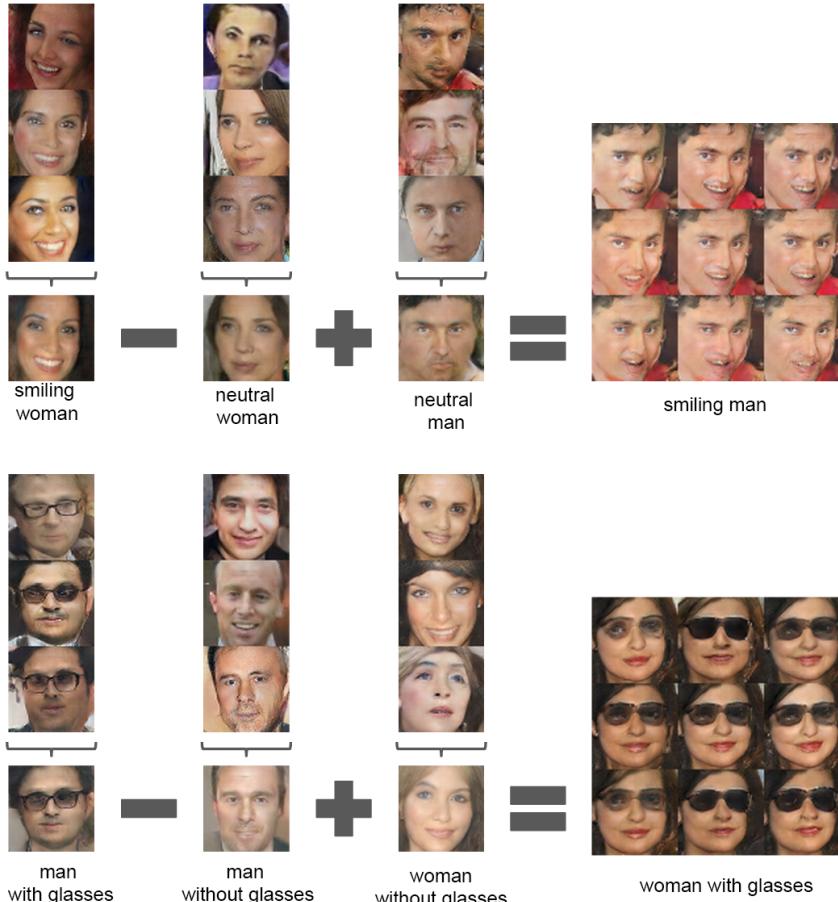
DCGAN architecture, Radford, Metz and Chintala arXiv:1511.06434



Interpolation in the latent space. Rotations are linear!

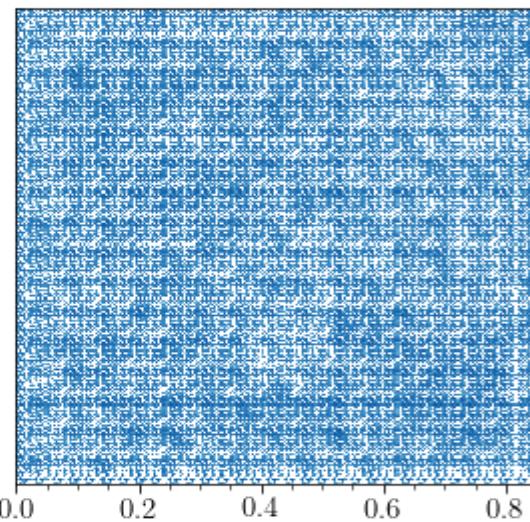
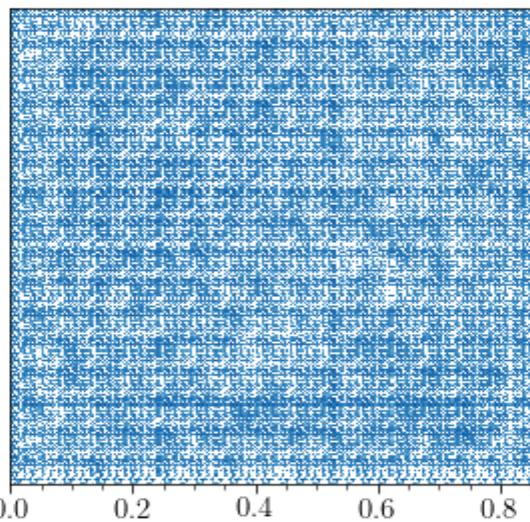
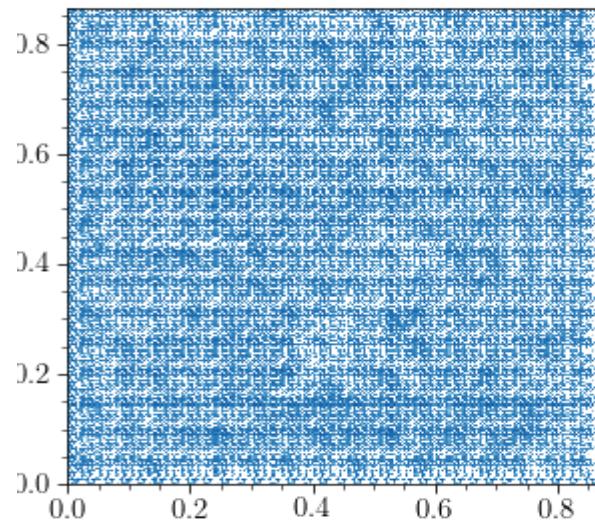
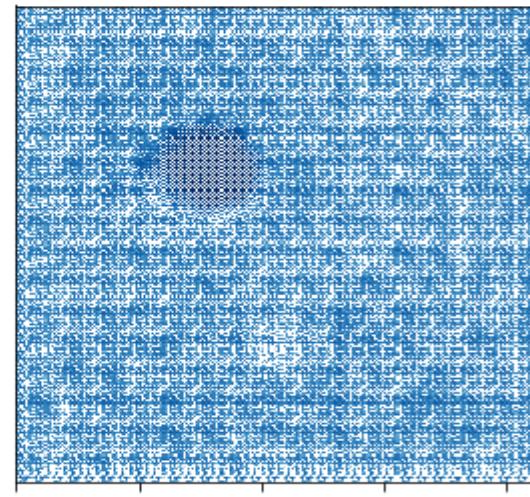
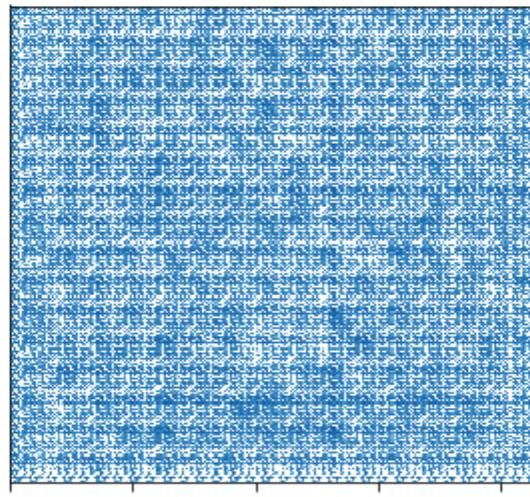
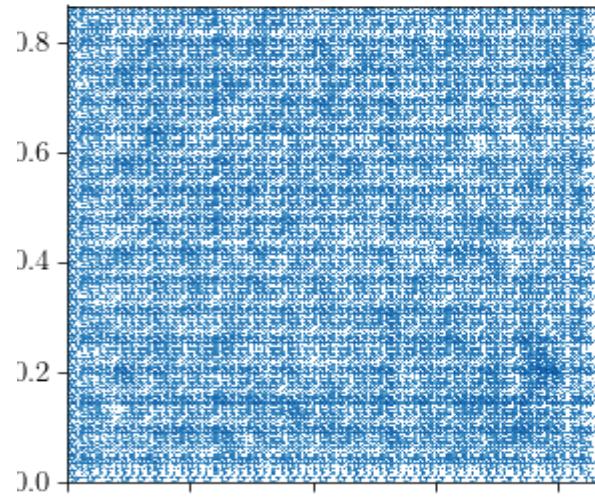
# Deep Convolutional Generative Adversarial Networks (DCGAN)

DCGAN architecture, Radford, Metz and Chintala arXiv:1511.06434



Arithmetics in the latent space

# Generative Adversarial Networks



# Evaluation of Generative Models

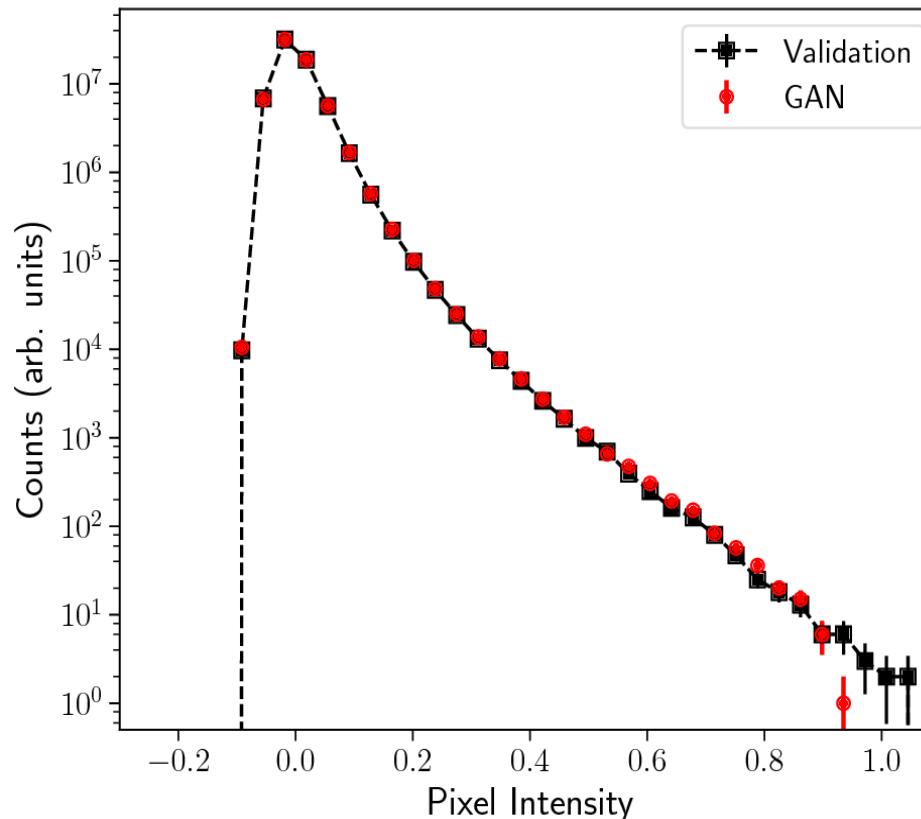
How close is  $\mathbb{P}_{model}$  to  $\mathbb{P}_{data}$ ?

## 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.

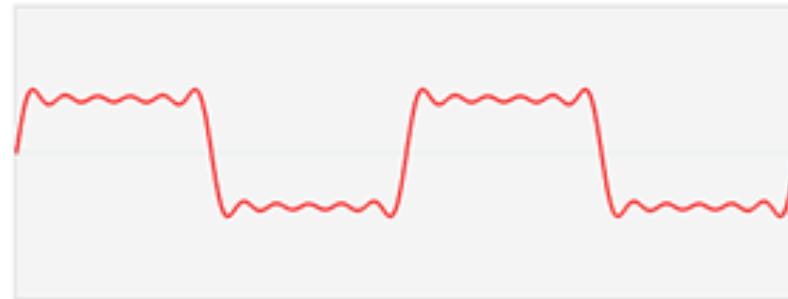
# Convergence Maps First Order Statistics



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

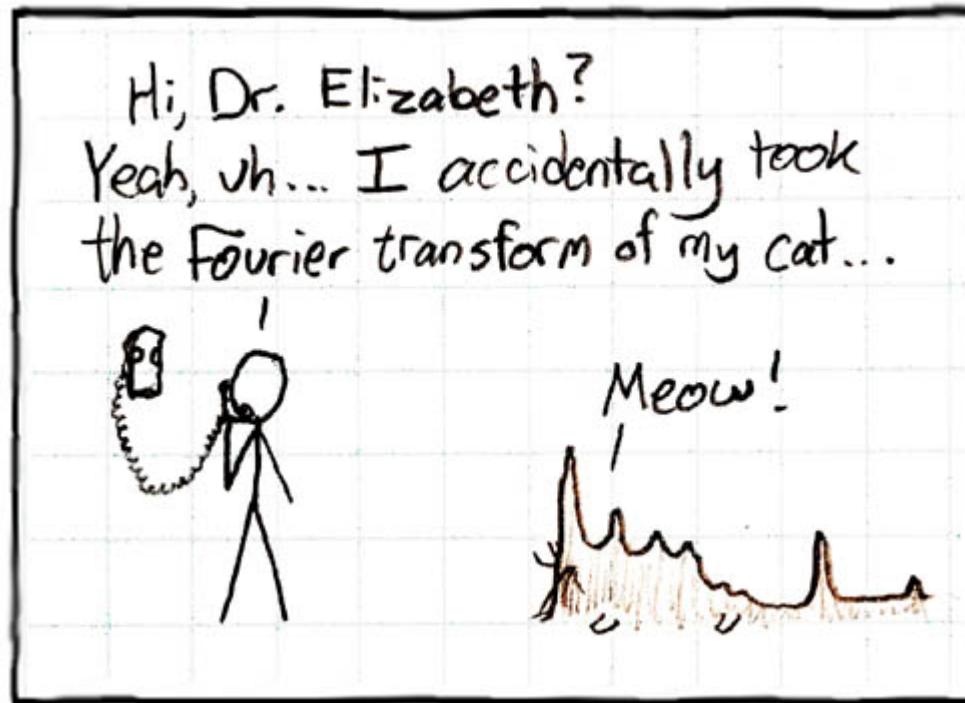
# Fourier Spectral Analysis

[quora.com/Whats-the-use-of-Fast-Fourier-Transform](https://www.quora.com/Whats-the-use-of-Fast-Fourier-Transform)



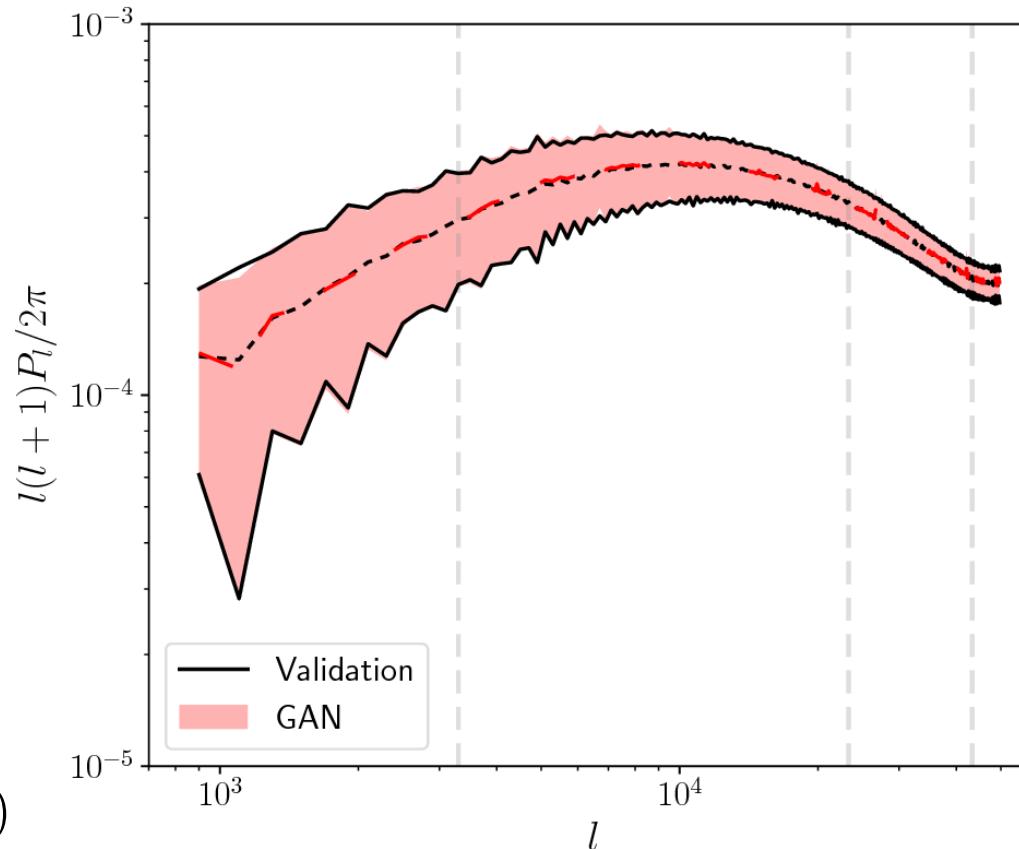
# Fourier Spectral Analysis

<https://xkcd.com/26/>



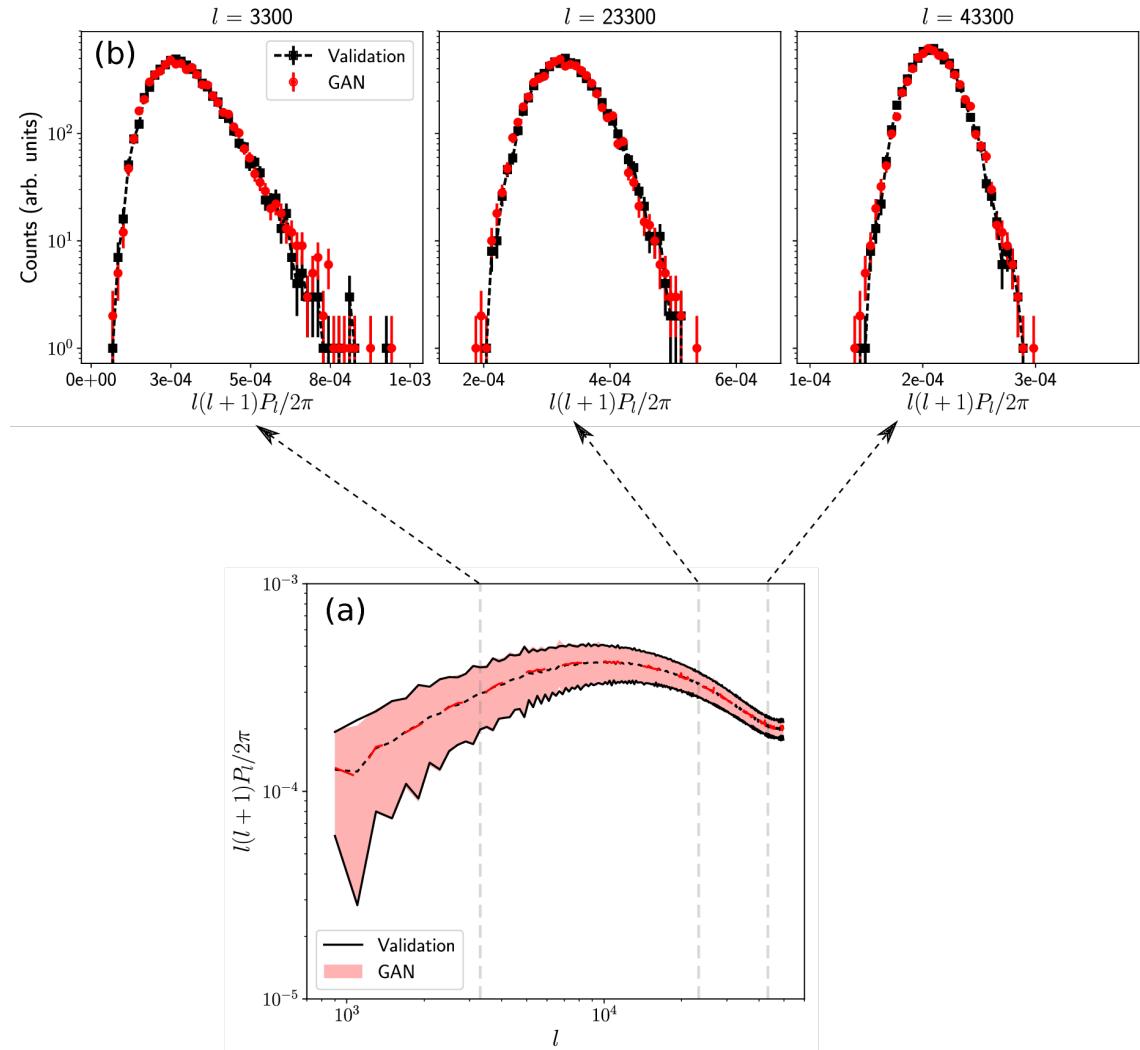
# Fourier Spectral Analysis: Power Spectrum

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

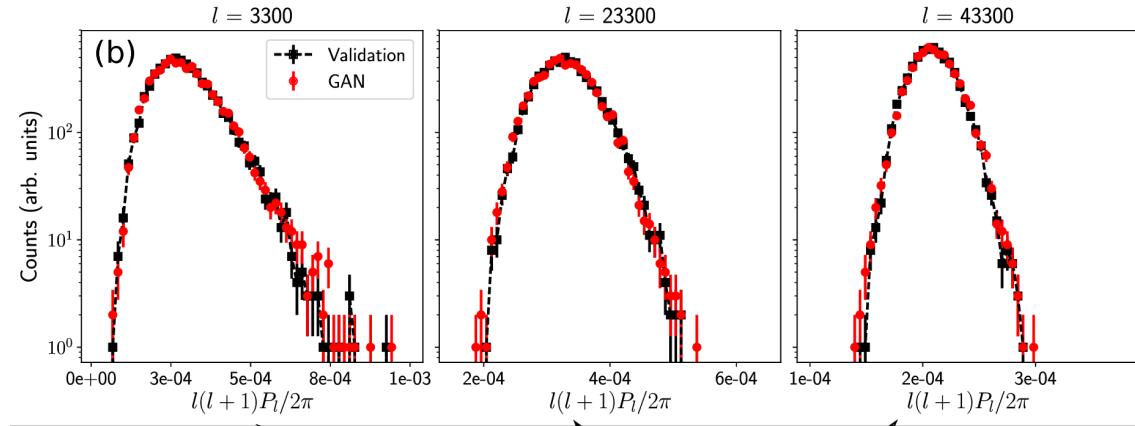


Bands are  $\mu(l) \pm \sigma(l)$

# Fourier Spectral Analysis: Power Spectrum

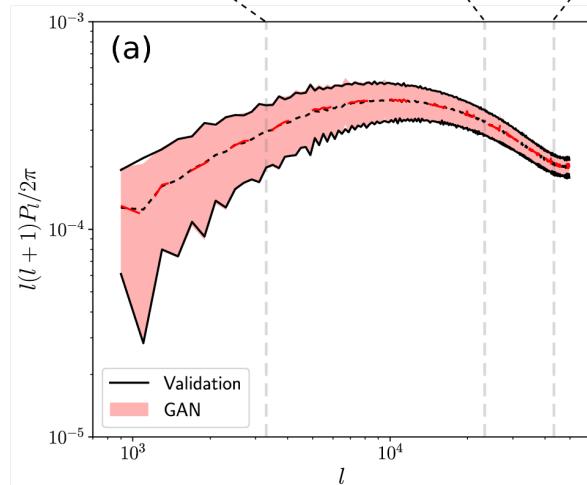


# Fourier Spectral Analysis: Power Spectrum



Kolmogorov-Smirnov

# moments	p-value
242	> 0.995
6	> 0.93

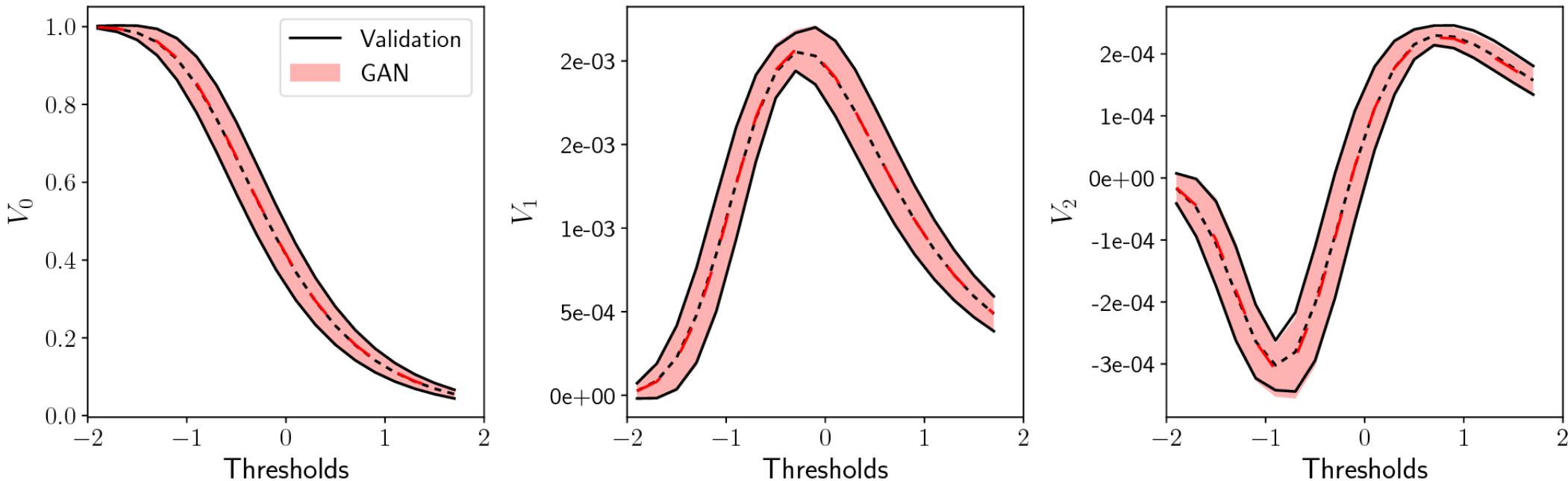


## Non-Gaussian Corrections

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

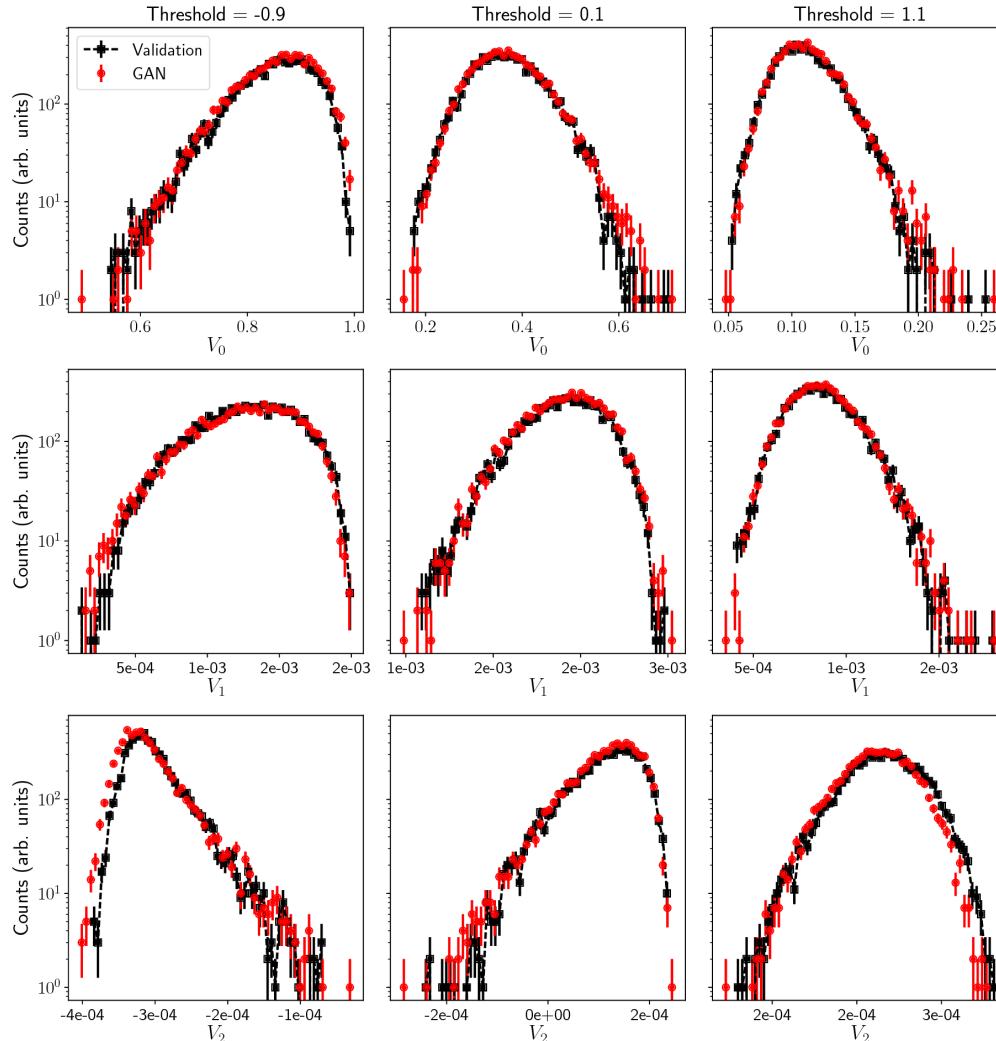
# Non-Gaussian Corrections

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

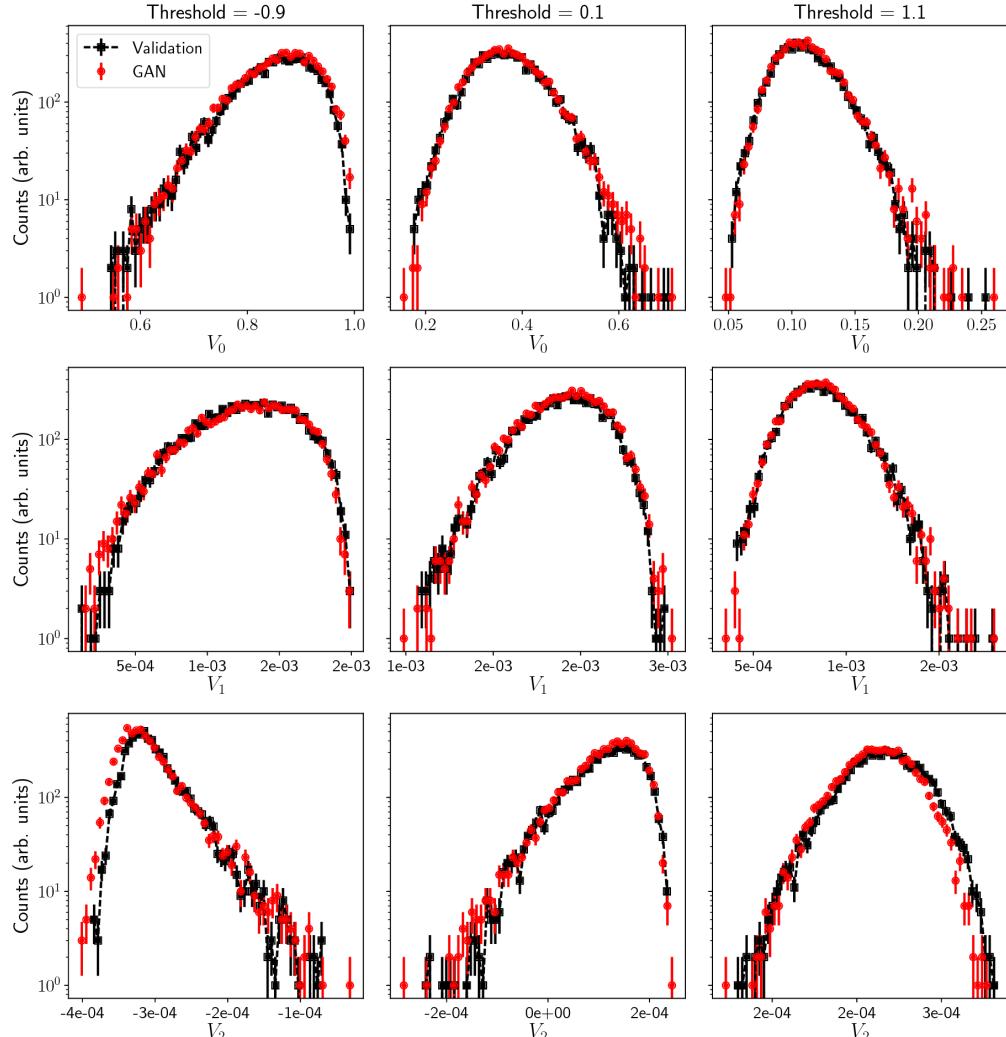


The three Minkowski Functionals are sensitive to the higher order correlations.

# Non-Gaussian Corrections



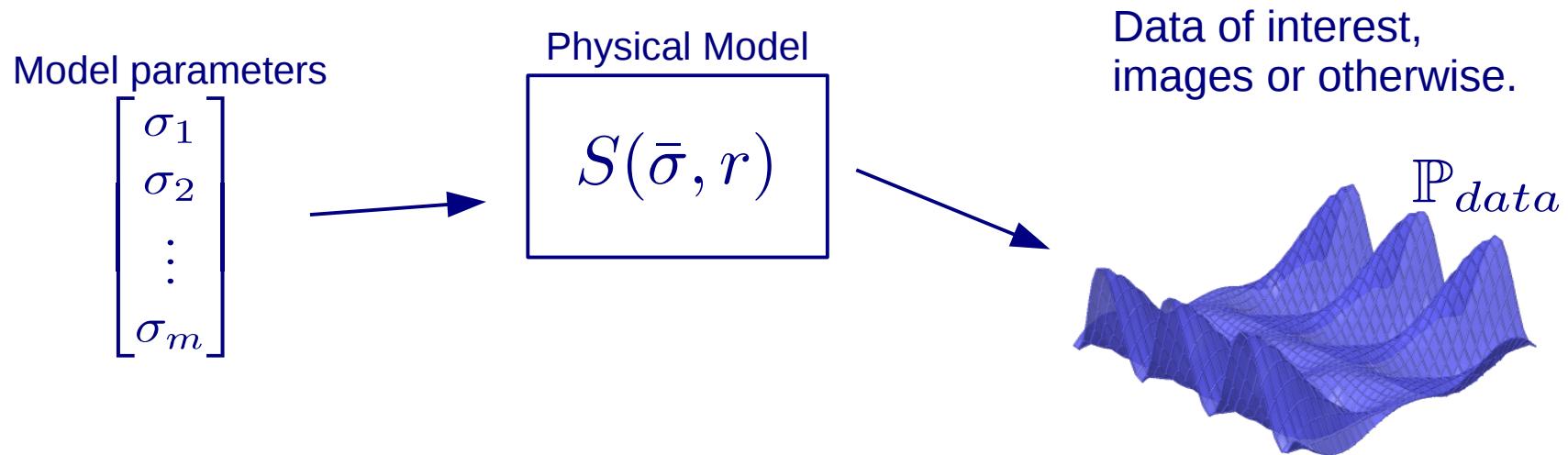
# Non-Gaussian Corrections



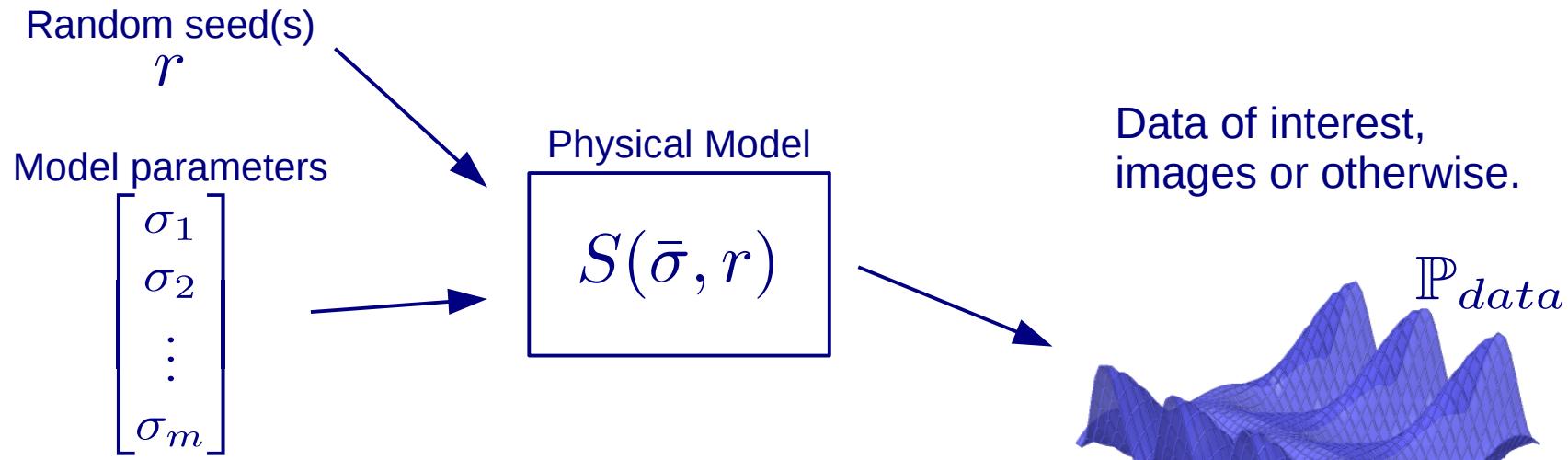
## Kolmogorov-Smirnov

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

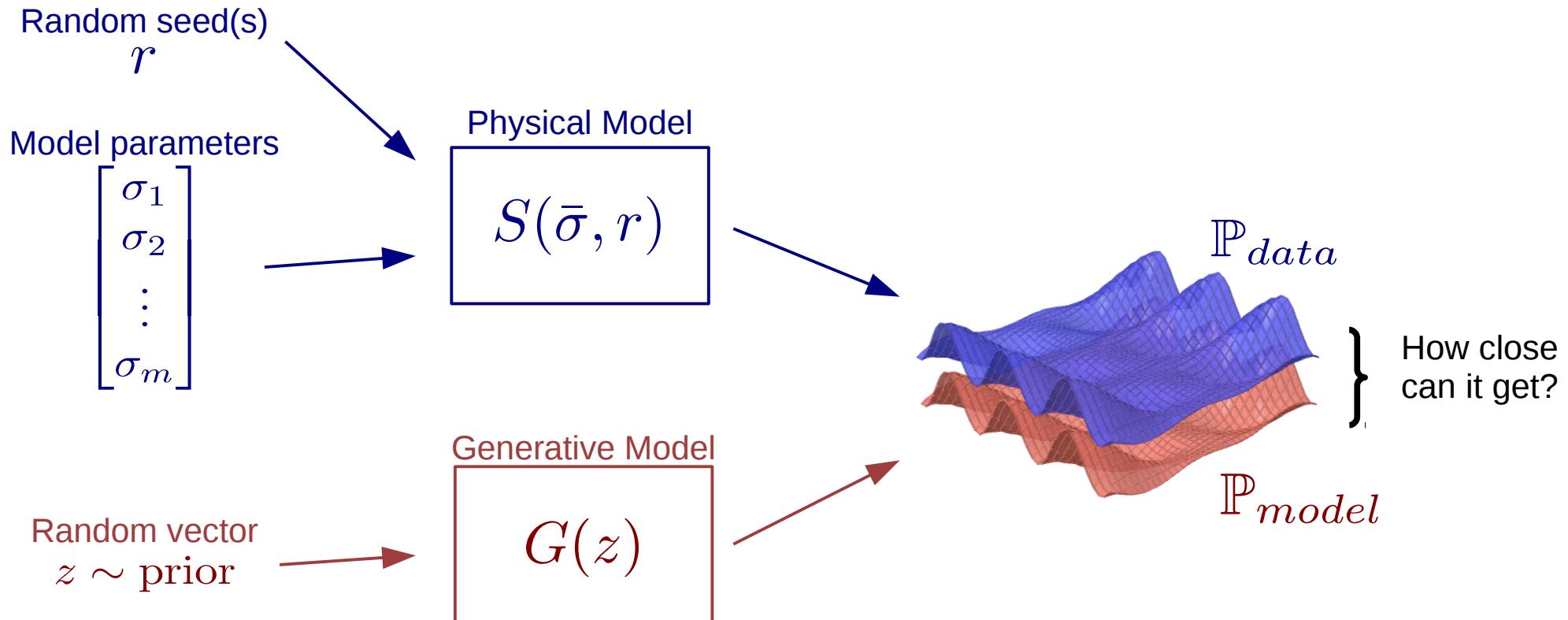
# Generative Models for Emulating Scientific Data



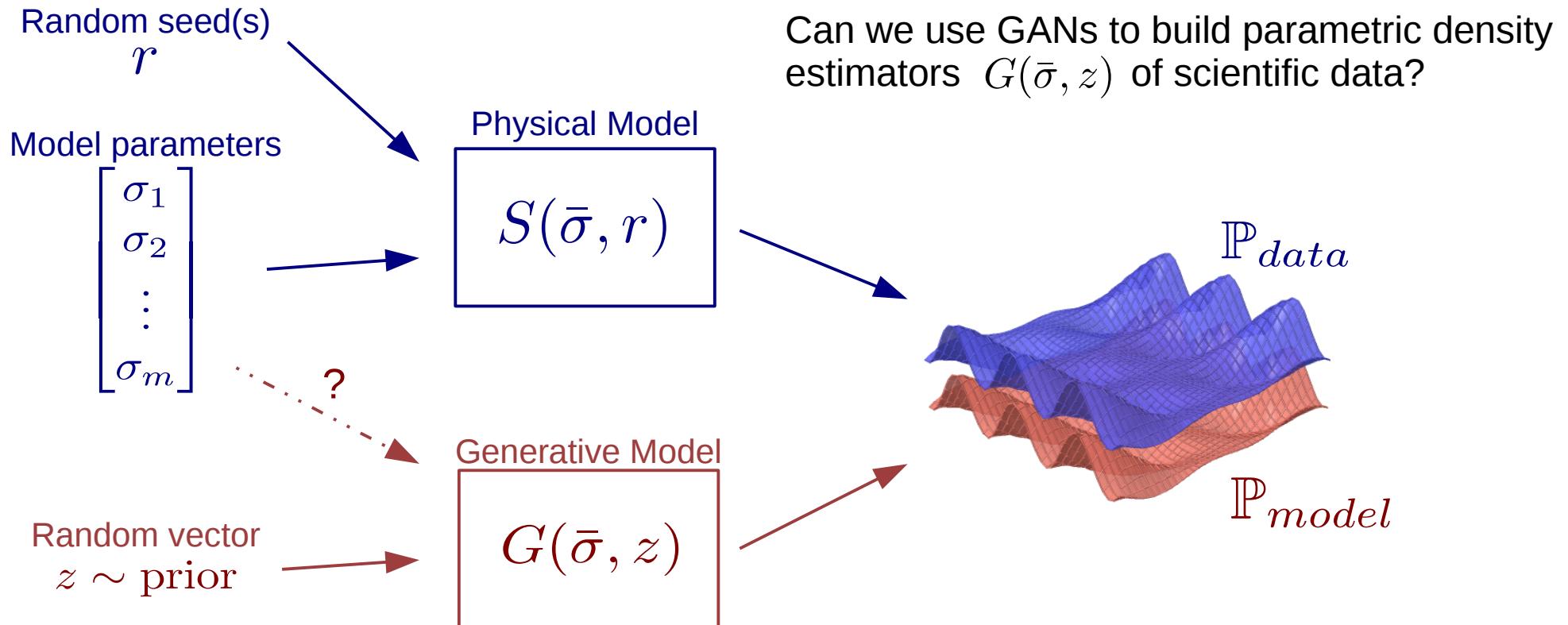
# Generative Models for Emulating Scientific Data



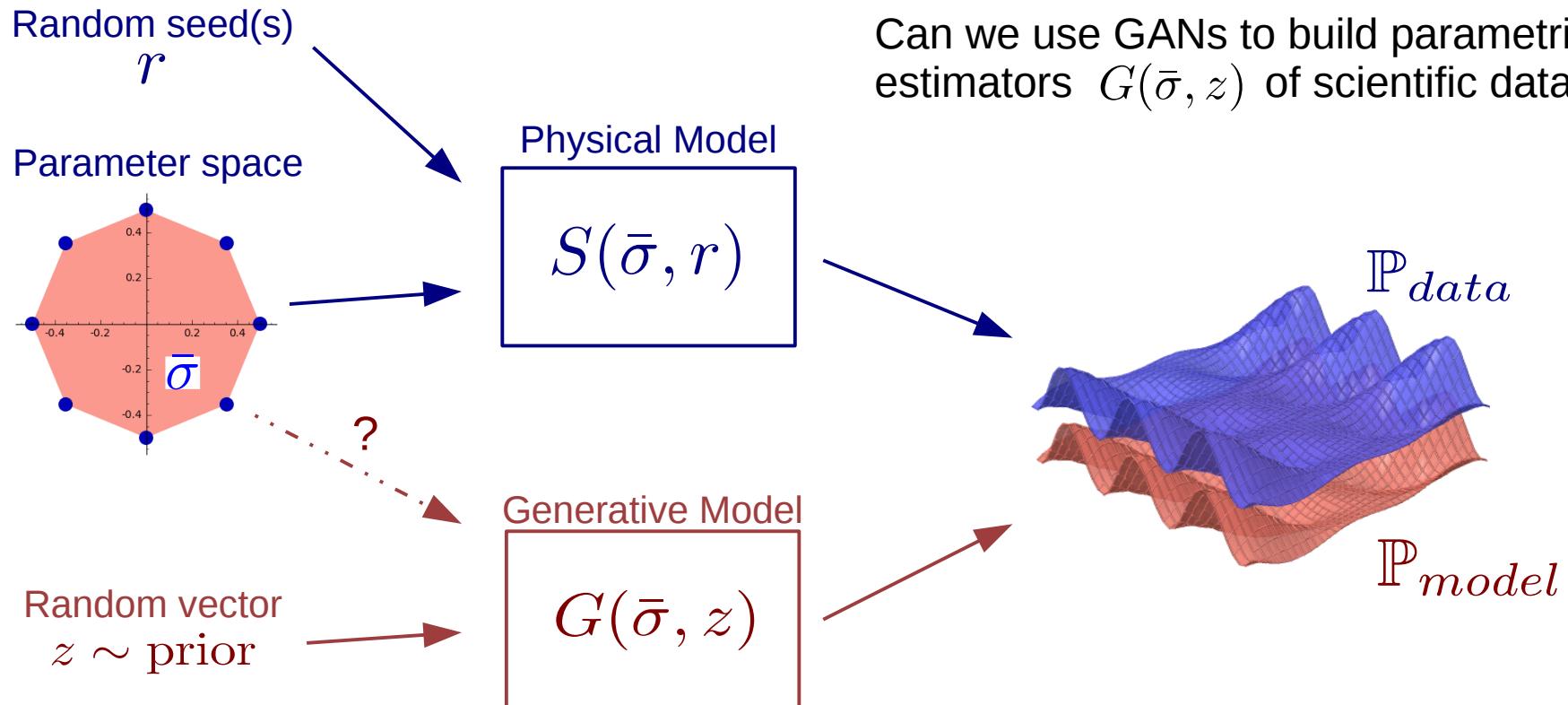
# Generative Models for Emulating Scientific Data



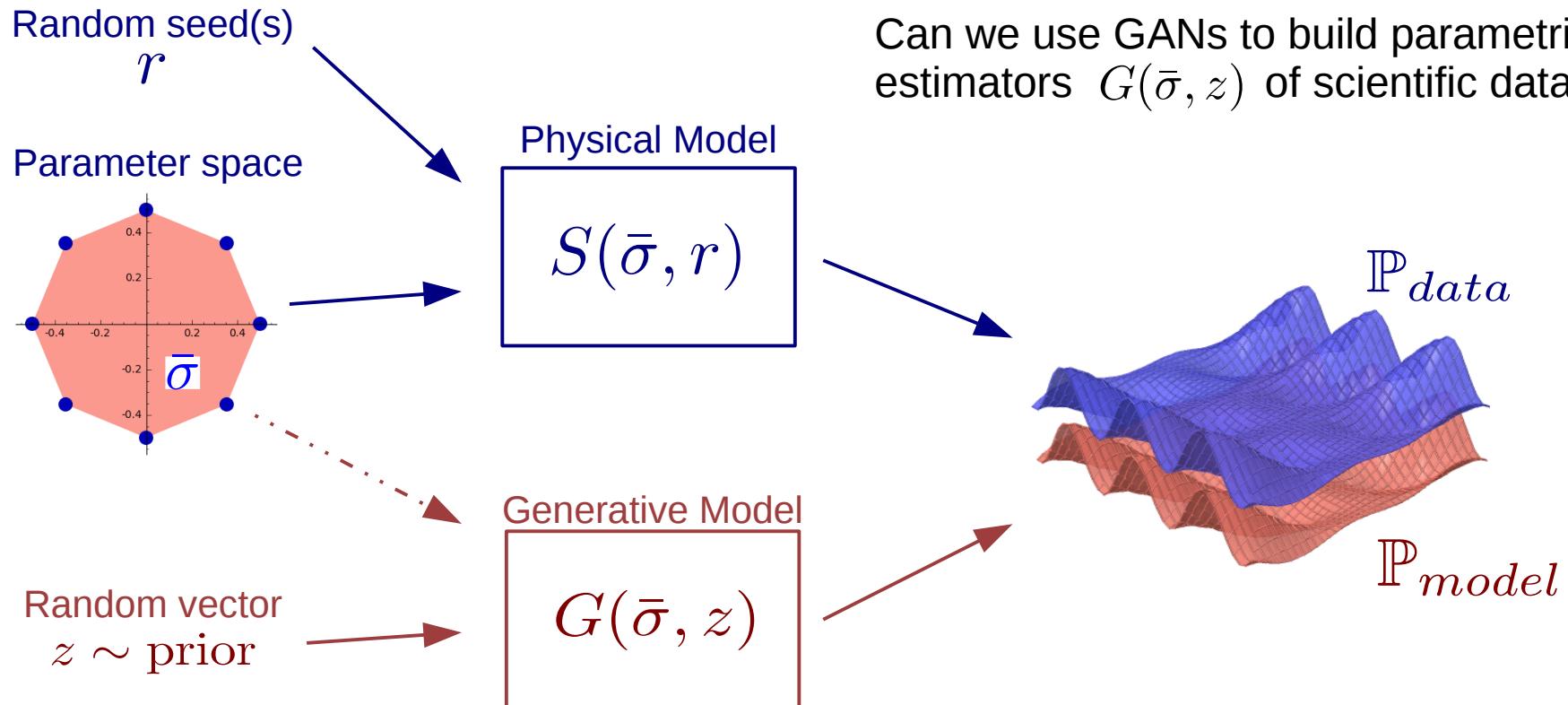
# Generative Models for Emulating Scientific Data



# Generative Models for Emulating Scientific Data



# Generative Models for Emulating Scientific Data

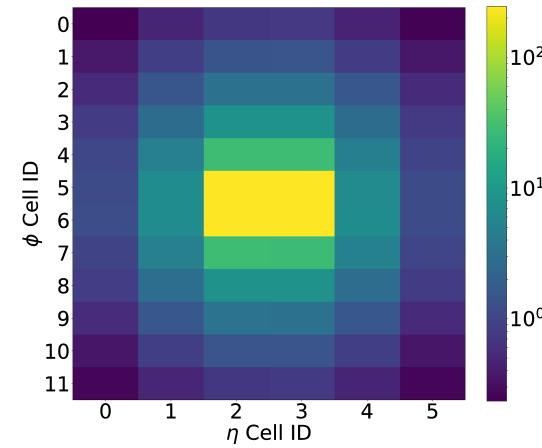
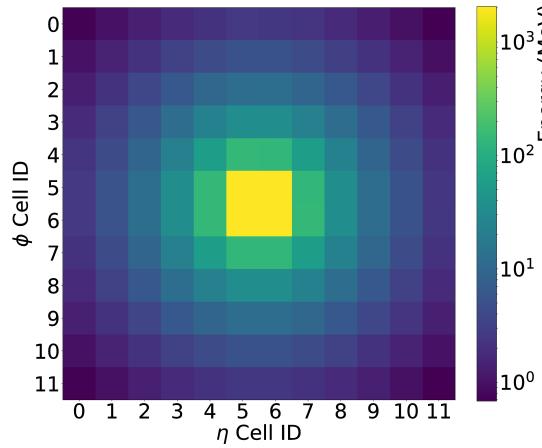
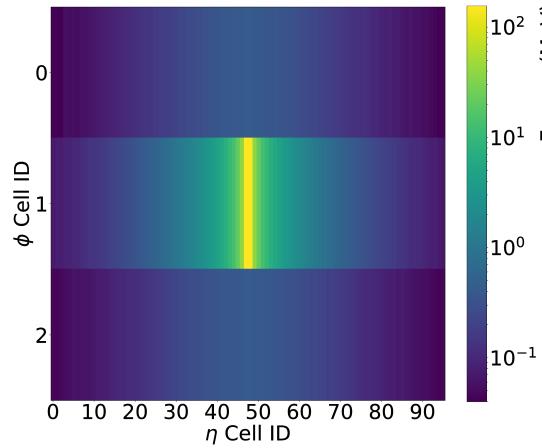


Can we use GANs to build parametric density estimators  $G(\bar{\sigma}, z)$  of scientific data?

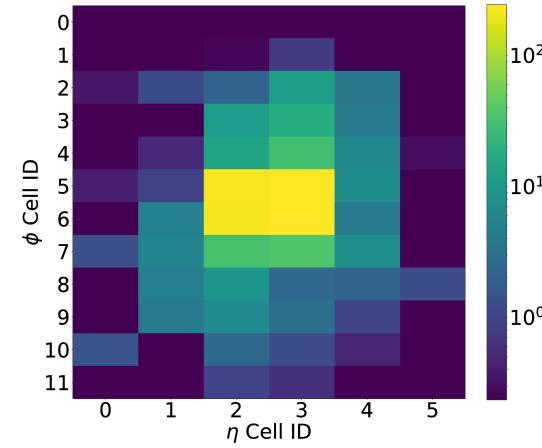
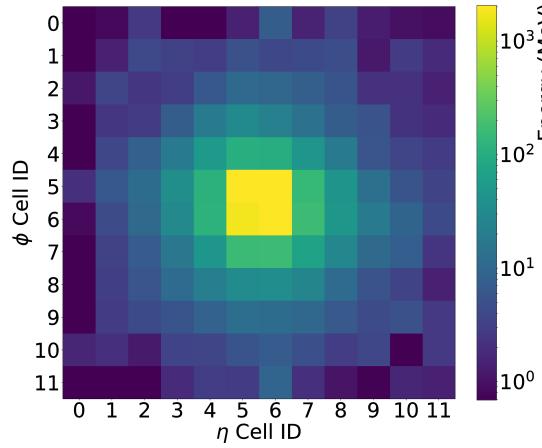
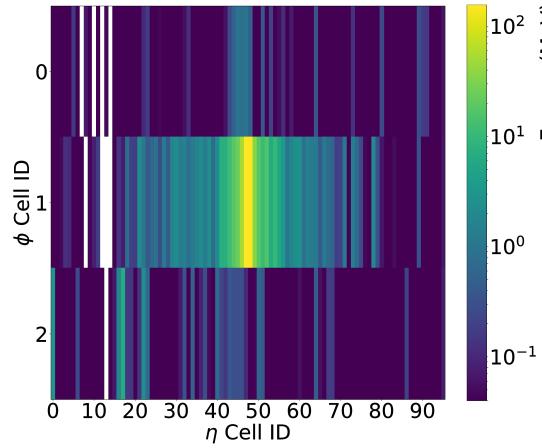
Such generators would likely exclude regions in parameter space where the physical model  $S(\bar{\sigma}, r)$  exhibits critical behavior.

# CaloGAN: Simulating 3D Calorimeter Showers using GANs

GEANT



CaloGAN



Paganini, de Oliveira and Nachman arXiv:1705.02355

## Summary and Outlook

- We have shown with statistical confidence that GANs can emulate  $\Lambda$ CDM 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  $\Lambda$ CDM 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, will be posted to arXiv tonight!

## Summary and Outlook

- We have shown with statistical confidence that GANs can emulate  $\Lambda$ CDM 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, will be posted to arXiv tonight!
- 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