Creating simulations of our universe using GANs

Juan Jose Zamudio, Atakan Okan, Asena Derin Cengiz, Seda Bilaloglu

Team: Jaas, *Advisor: Shirley Ho, Flatiron Institute*

Objective & Data

- The objective is to use GANs to learn the Hydrogen distribution in universe simulations (that took 14 million GPU hours) to generate new simulations using GANs resulting in 10^6 decrease in computational resources needed and time.
- The data consists in cosmological simulations from IllustrisTNG. It consists in cube of size 2048x2048x2048, each side representing 75 Mpc/h (244,600,000 light years).
- We take samples with replacement of size 64x64x64 as inputs. Each cell represents the HI/total mass in the position (x,y,z).

Methods

- Two summary statistics were calculated in addition to 2D and 3D plots to gauge how similar the generated samples are.

3D Power Spectrum

- Power spectrum measures the density contrast of the real and generated universes by calculating the difference between the local density and the mean density as a function of scale.

Models

Wasserstein GAN with Gradient Penalty

- The WGAN seeks to minimize Earth Mover distance between real and generated samples distributions , gradient penalty is implemented to ensure the Lipschitz constraint is met.

$$L = \mathbf{E}_{z\sim P_z}[D(G(z))] - \mathbf{E}_{x\sim P_r}[D(x)] + \lambda \mathbf{E}_{\hat{x}\sim P_{\hat{x}}}[(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2]$$

Wasserstein GAN Architecture

- WGAN Architecture is using DCGAN architecture adapted to 3D data with 3D Convolutions and Convolution Transpose operations. Also, with the addition of gradient penalty, all Batch Normalization operations in the discriminator were removed.

Maximum Mean Discrepancy GAN

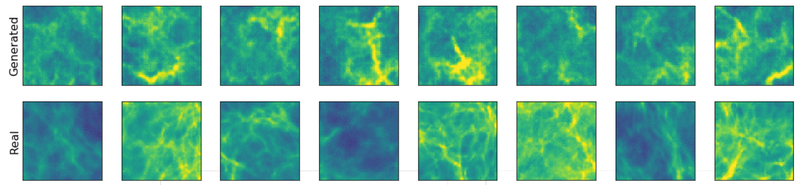
- The MMDGAN seeks to minimize the distances between kernelized embeddings of the real and generated samples. Only the maximum kernelized embedding distance is taken into account and the kernel bandwidth hyperparemeter search is done to accommodate IllustrisTNG data. Weight clipping is added to improve stability of the training process.

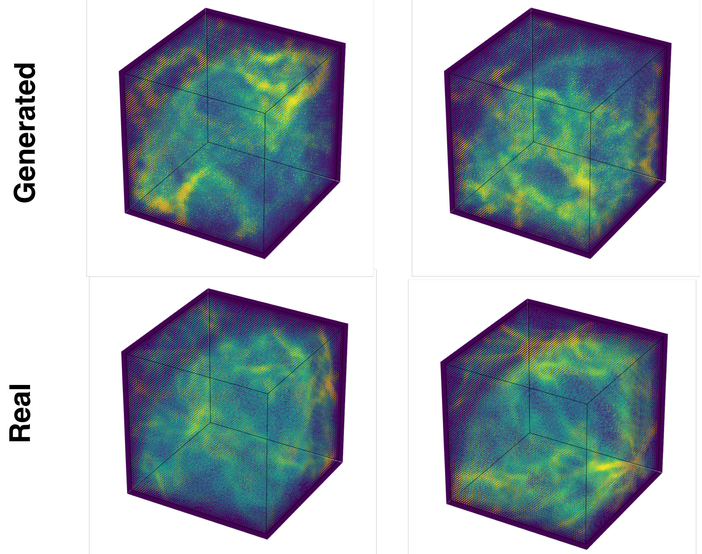
$$MMD(\mathbb{P}, \mathbb{Q}; \mathcal{H}) = \sup_{f\in\mathcal{H}, \|f\|_{\mathcal{H}}\leq 1} \mathbb{E}_{\mathbb{P}} f(X) - \mathbb{E}_{\mathbb{Q}} f(Y)$$

MMDGAN Architecture

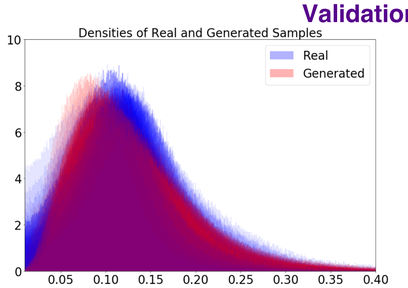
- MMDGAN Architecture is using DCGAN architecture adapted to 3D data. In addition, Upsample-Convolution layers were used in the Generator to alleviate the “checkerboard” problem. MMDGAN Discriminator also employs a decoder to learn from autoencoding real and generated inputs.

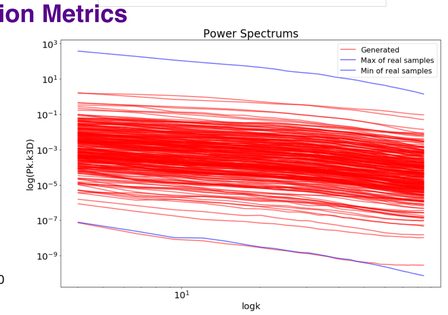
WGAN Results



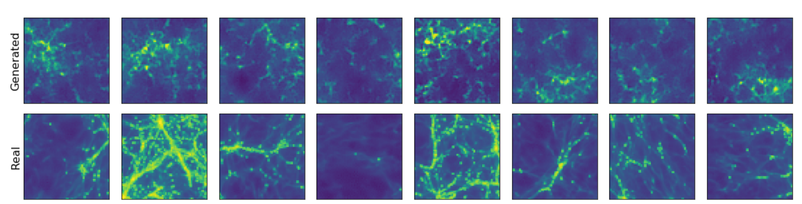


Validation Metrics

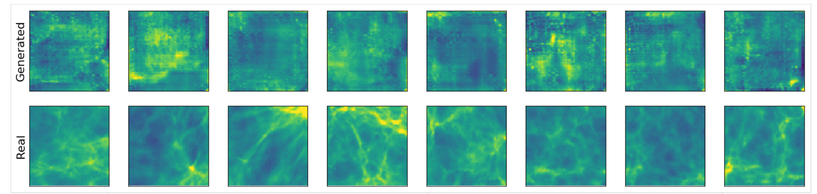


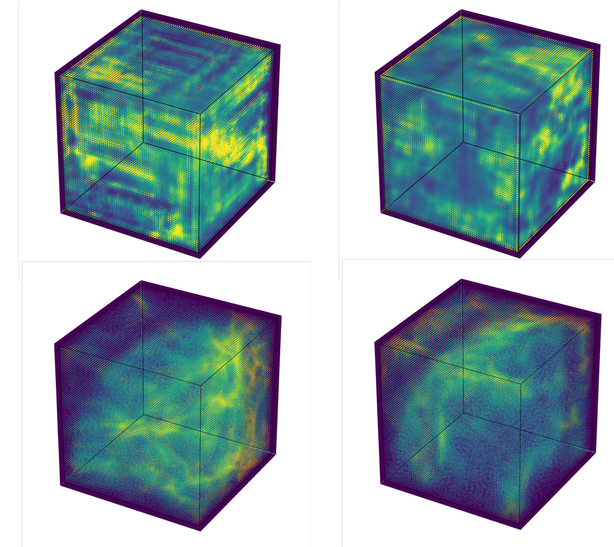


WGAN in 2D

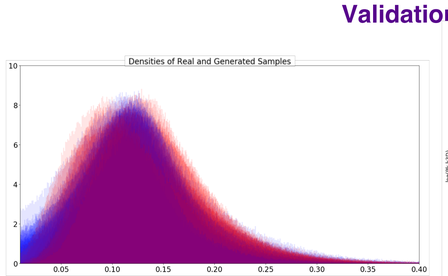


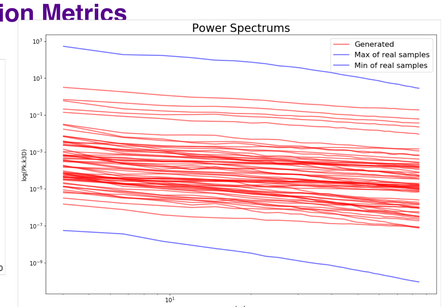
MMD GAN Results





Validation Metrics





Conclusion and Future Work

- We believe that multi-GPU training will aid the stability and the convergence speed of MMDGAN in addition to other forms of penalties and regularization methods.
- For the WGAN-GP, deeper architectures with more channels for the generator would be helpful to model the skewed distribution of our hydrogen data. For its occasional checkboard problem, implementing Upsample-Convolution layers should improve results too.

Acknowledgements

- We are grateful for our wonderful advisors Shirley Ho, Francisco Villaescusa-Navarro and Siyu He at the Flatiron Institute Center for Computational Astrophysics for all their help.