Mapping Dark Matter to Sunyaev-Zel'dovich with Neural Networks

1. Convolutional Neural Network (CNN) approach
2. DeepSet approach

Leander Thiele (Princeton)

Motivation

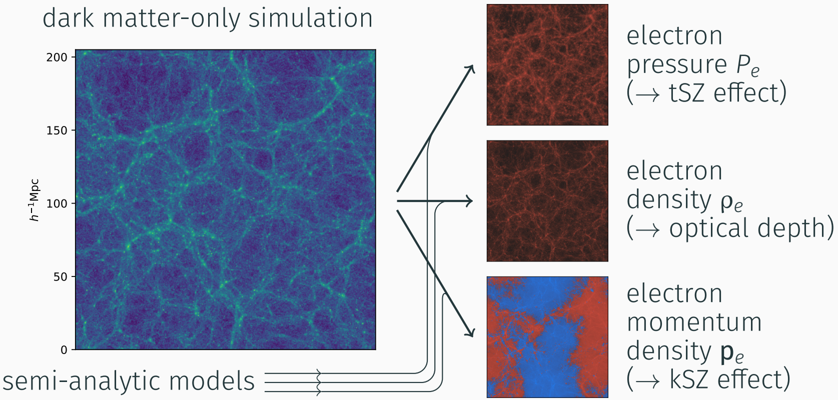
- goal: predict baryonic fields from gravity-only (N-body) simulations
- simplification: astrophysical processes are more local than gravity, baryons trace DM distribution fairly well $\rightarrow$ local machine-learning approach
- 1st use: rapidly generate vast amounts of data, to:
  - model summary statistics and their distribution
  - model cross-correlations (e.g. WL-tSZ)
  - perform likelihood-free inference (at summary-statistic- or field-level)
- 2nd use: interpret and learn something about the connection between astrophysics and cosmology
- focus in this talk: Sunyaev-Zel'dovich effects

Teaching neural networks to generate Fast Sunyaev Zel'dovich Maps

Leander Thiele, Francisco Villaescusa-Navarro, David N. Spergel, Dylan Nelson, Annalisa Pillepich

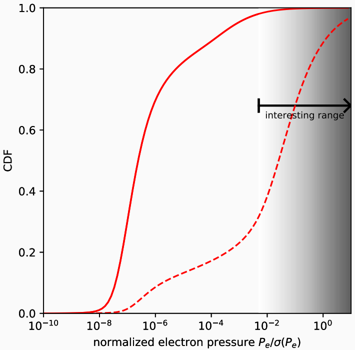
2007.07267

Idea



- simulation data from IllustrisTNG300 and zoom-ins
- work directly with 3-dimensional field
- only $z = 0$ so far

Sparsity



Few interesting voxels

$\rightarrow$ biased training samples:

zoom-ins for tSZ,

mass biases o/wise

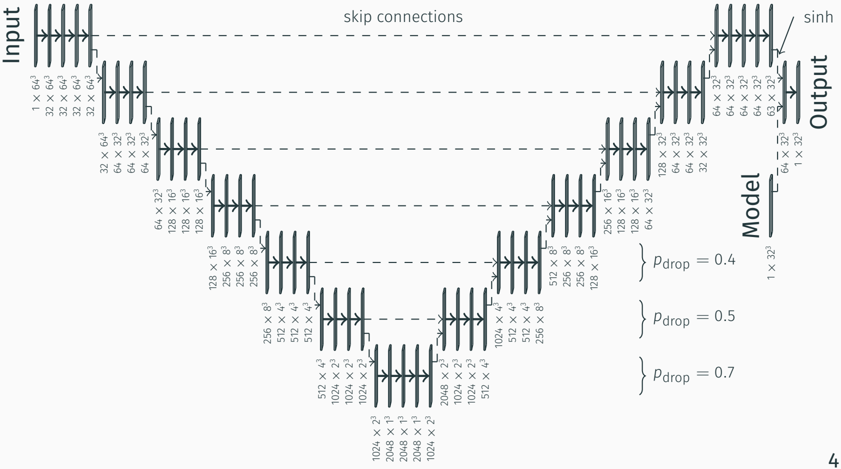
Tailed distributions

$\rightarrow$ input transformation

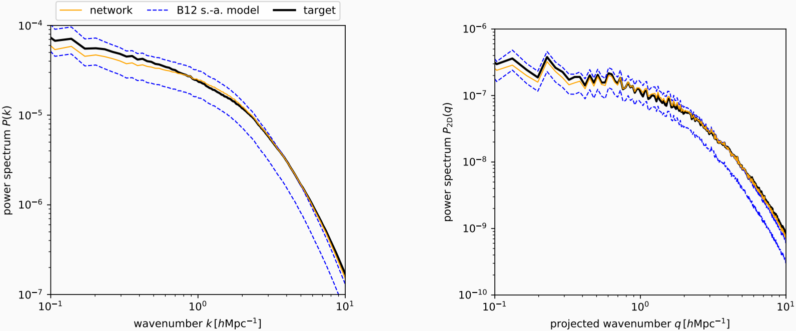
$\rightarrow$ epoch-dependent loss function

$\rightarrow$ semi-analytic models

Tune hyperparameters & network architecture on electron pressure, then apply to density & momentum. Spatial problem with translational symmetry → convolutional net (CNN).



Results: electron pressure (tSZ)



→ projection improves network-fiducial agreement

DeepSets applied to Clusters: Machine learning the Lagrangian way

Leander Thiele, Miles Cranmer, William Coulton,
Shirley Ho, David N. Spergel

work in progress!

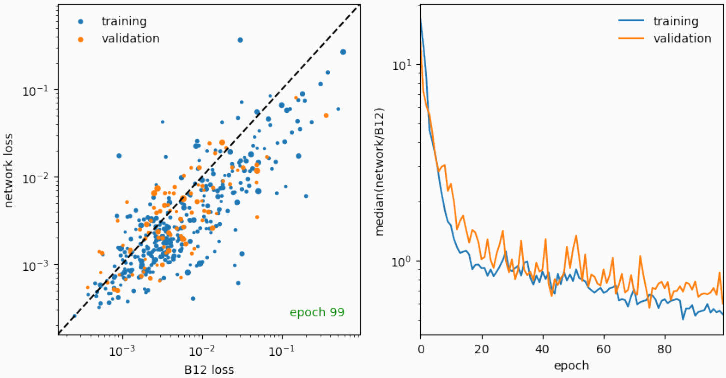
Problems with previous approach, and what they teach us

- ended up concentrating training on massive halos $\rightarrow$ let us focus on those for now!
- if we only want to work with halos, translational symmetry is broken (finding halo centers is a mostly solved problem)
- with CNNs we spend a lot of resources on boring regions because we need to cover large scales but still require decent resolution because scales are coupled
- interpretability is rather poor
- maybe CNNs are not the best approach!
- (not related to CNN architecture) there is stochasticity in the baryonic fields which we should try to model

Idea

- Given Dark Matter particles within $\mathcal{O}(1)R_{\text{halo}}$. Can we work directly with this simulation representation?
- This is a set of features (no ordering) $\rightarrow$ rotation-equivariant DeepSet.
- Intuition: the simulation representation should be ideal to overcome the sparsity problem.
- Incorporate stochasticity using a conditional VAE architecture.
- architecture components (modular=interpretable):
    - spherically symmetric approximation
    - miscentering correction
    - deformations
    - local environment ($\sim$ 100 kpc)
    - halo-scale features
    - probabilistic

Preliminary results



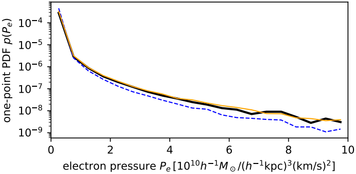
- B12 = Battaglia+2012 GNFW with best fit parameters
- overfitting a constant problem (only working with IllustrisTNG300)
- this plot does not include the entire network architecture!

Conclusions

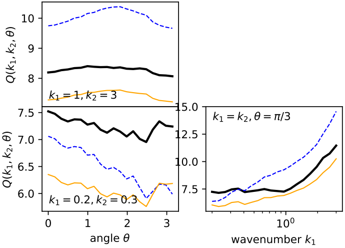
- CNN produces good results on select summary statistics
- but sparsity and coupling between scales make us believe CNNs may not be ideal architecture
- stochasticity is likely a subleading correction but should also be accounted for if the distribution of summary statistics is desired
- DeepSet approach appears more natural, stay tuned!
- future work:
    - combine DeepSet architecture with CNN, test performance on summary statistics
    - interpret the trained DeepSet architecture
    - $z > 0$, lightcones
    - multifield

Backup electron pressure (tSZ) I

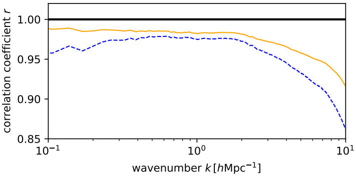
PDF



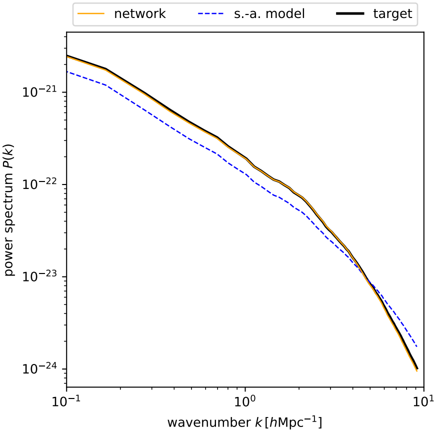
bispectrum



correlation

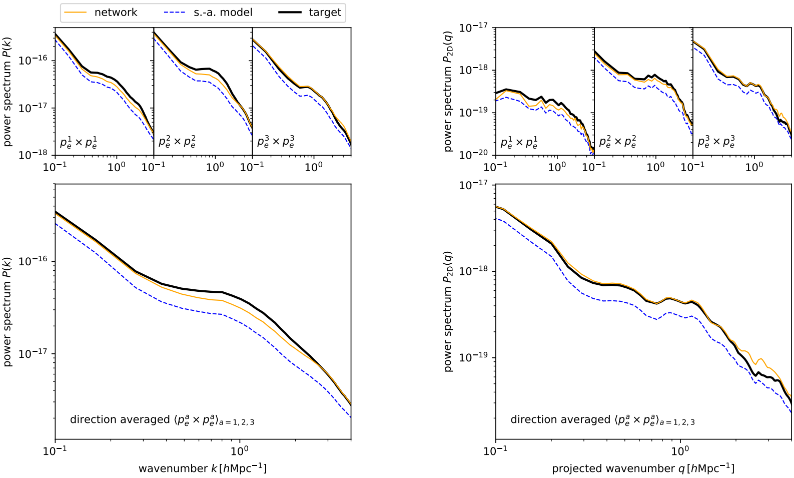


Backup electron density (optical depth)



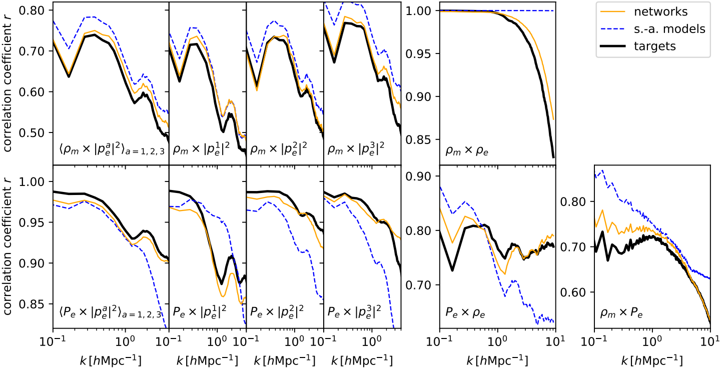
$\rightarrow \rho_e$ easier target than P_e: $P_e \sim \rho_e T_e$

Backup electron momentum density (kSZ)



→ sub-optimal network architecture for vectors

Backup cross-correlations



$\rightarrow$ model quality is important

Backup DeepSet Architecture

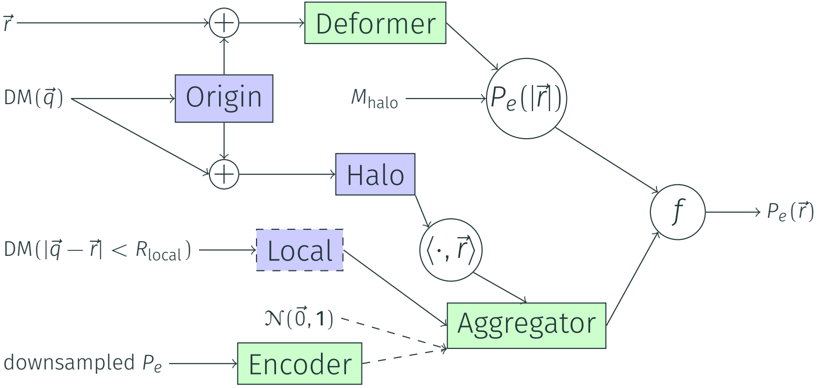
$\vec{q}$ = DM position

$\vec{r}$ = baryon position

= scalar DeepSet

= vector DeepSet

= MLP



(scalars and vectors describing halo passed at various points)