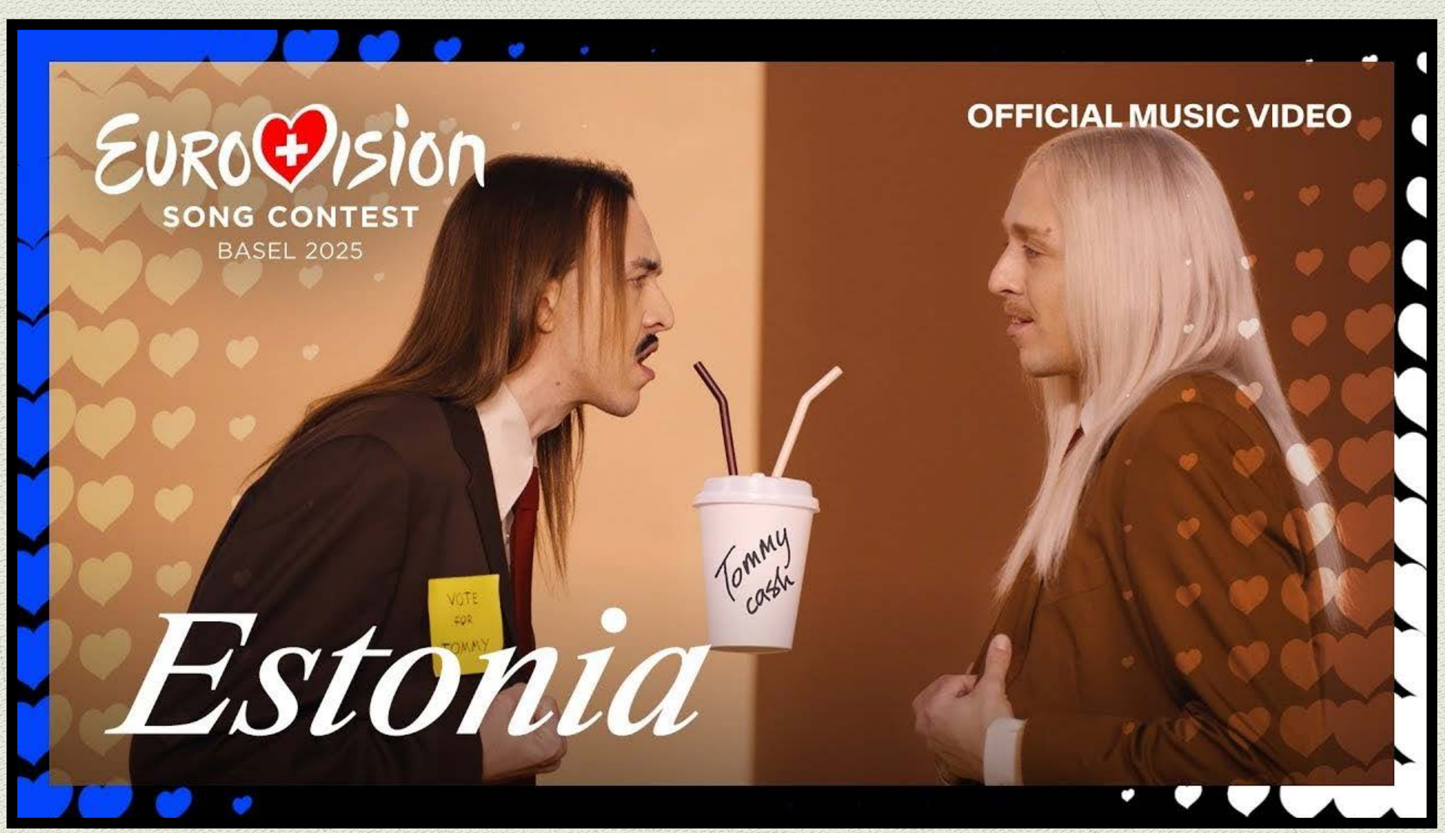
Neural network reconstructions of large-scale structures



Punyakoti Ganeshaiah Veena (Punya)

University of Genoa, Italy

work done with R.Lilow, A.Nusser, E. Branchini and E. Maragliano

Mapping the skies

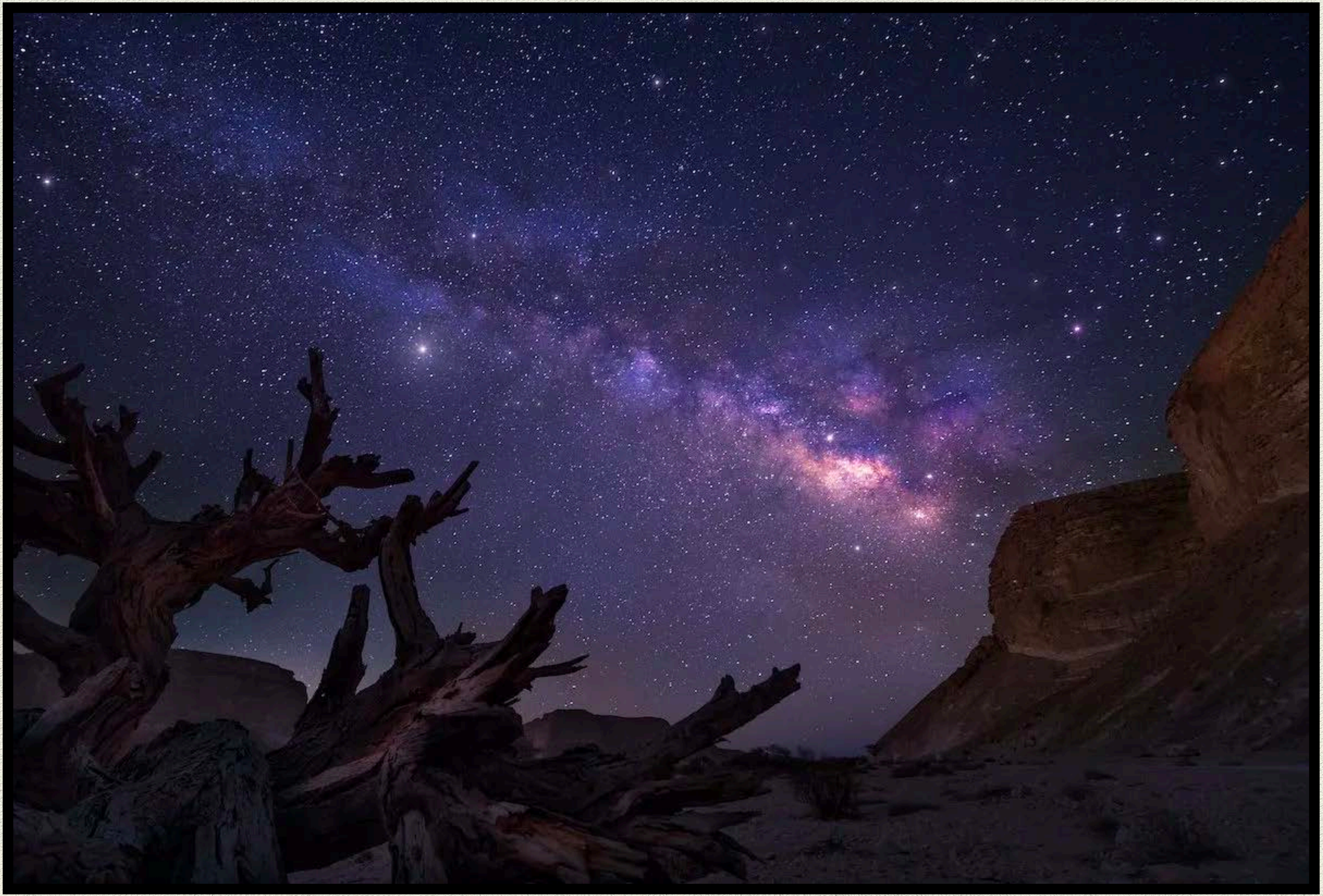


Image credit: Credit: ShutterstockCopyright: Copyright (c) 2015 Cristian Kirshbom/Shutterstock

Mapping the skies

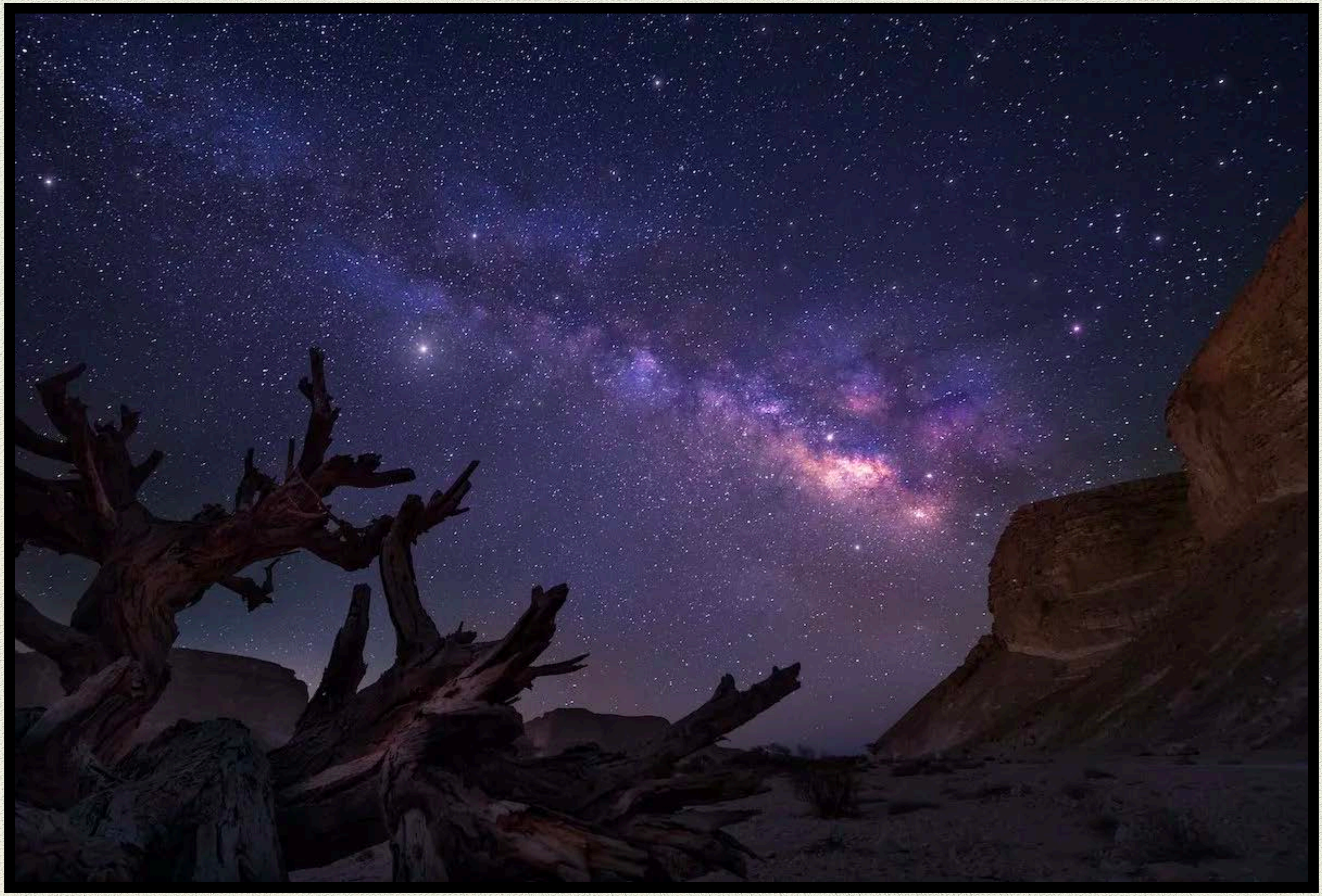
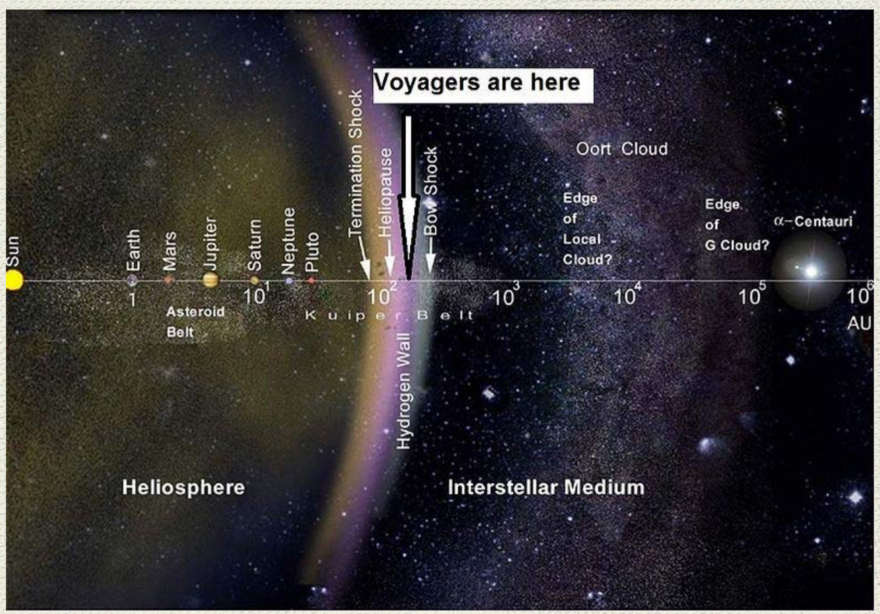


Image credit: Credit: ShutterstockCopyright: Copyright (c) 2015 Cristian Kirshbom/Shutterstock

Distances

Solar system-nearest star

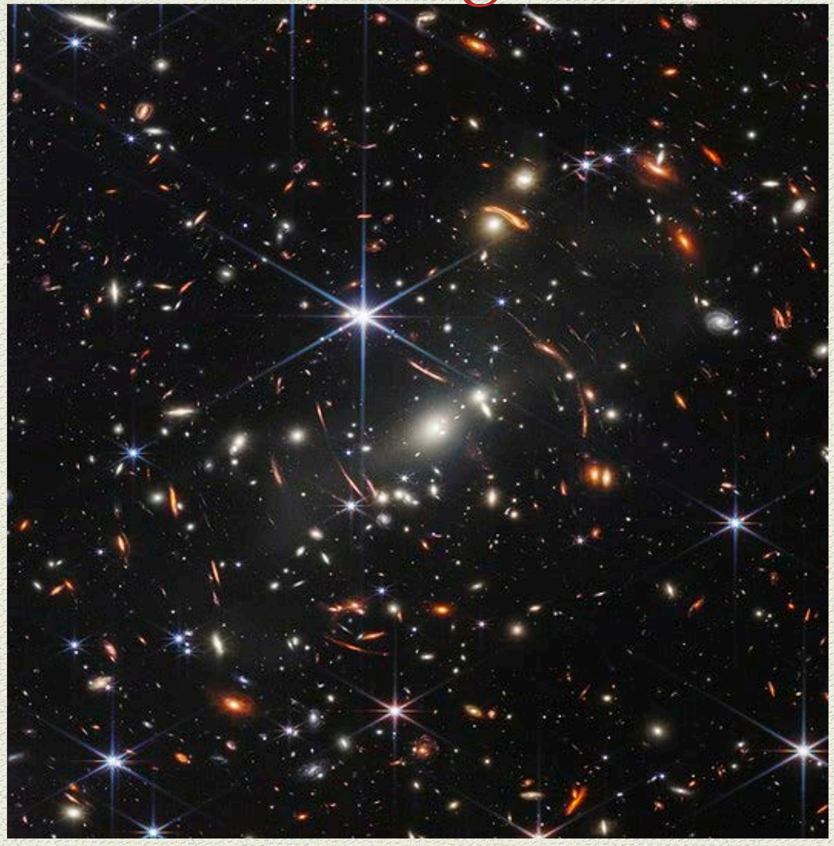


5 pc

Galaxy

15 kpc

Cluster of galaxies

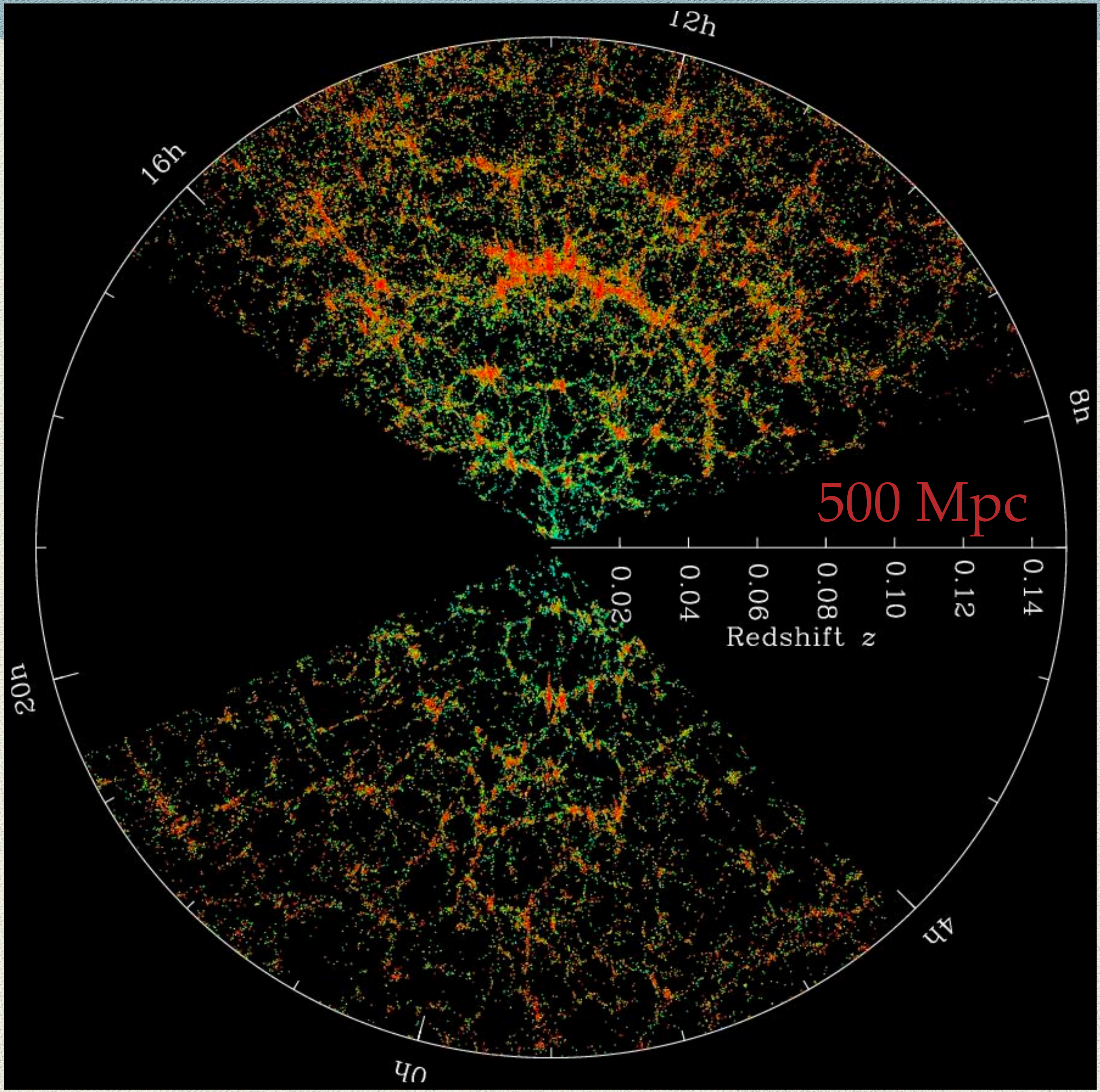


30 Mpc

$$1 \text{ kpc} = 10^3 \text{ pc}$$

$$1 \text{ Mpc} = 10^3 \text{ kpc}$$

Universe around us.



Redshift space distortions

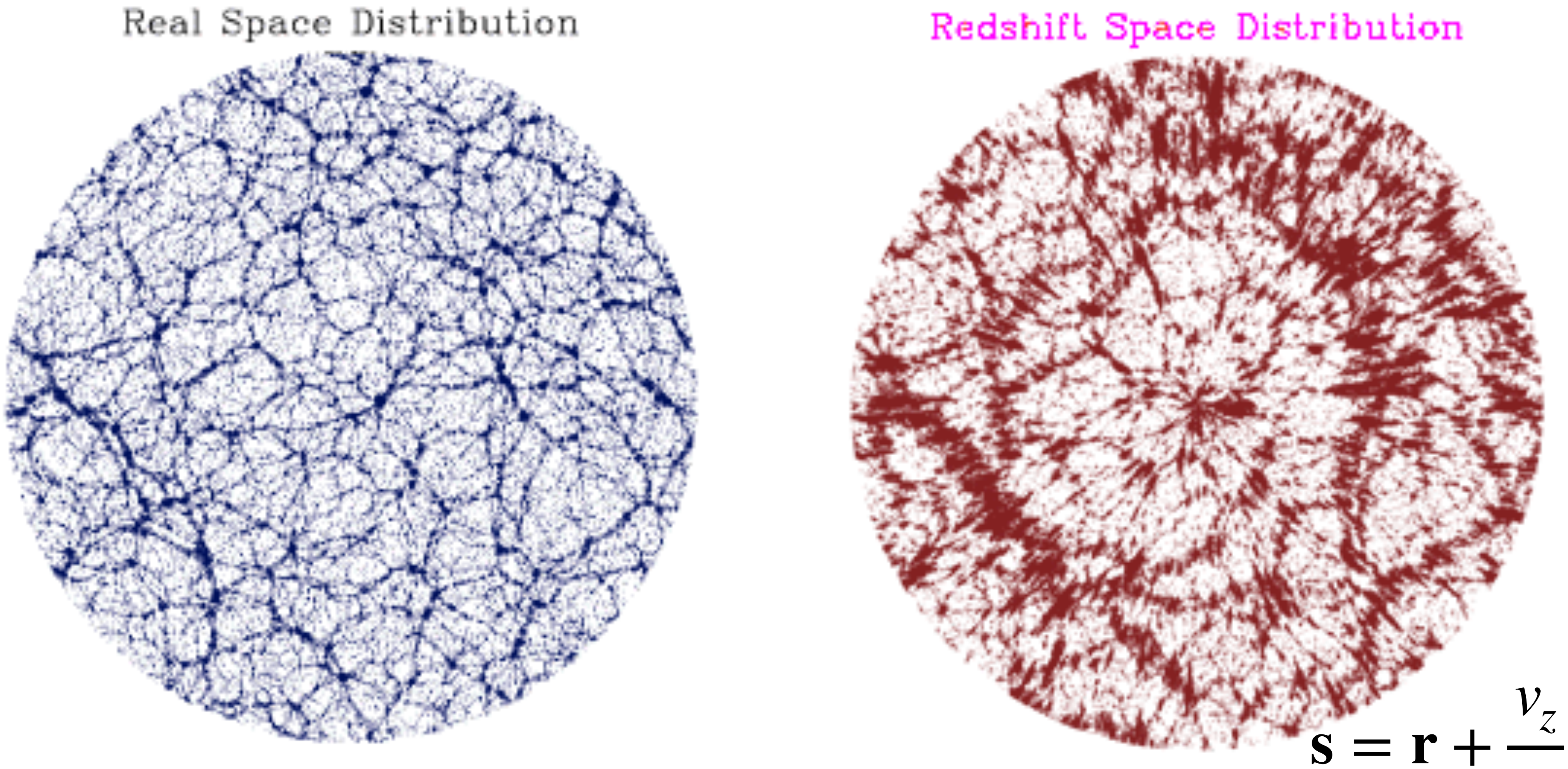
- Peculiar velocity: velocity of galaxy away from the Hubble flow:

$$v_{pec} = v_{observed} - v_{Hubble}$$

- Distribution of galaxies in real space v/s redshift space.

$$\mathbf{s} = \mathbf{r} + \frac{v_z}{H}\hat{\mathbf{z}}$$

Redshift space distortions

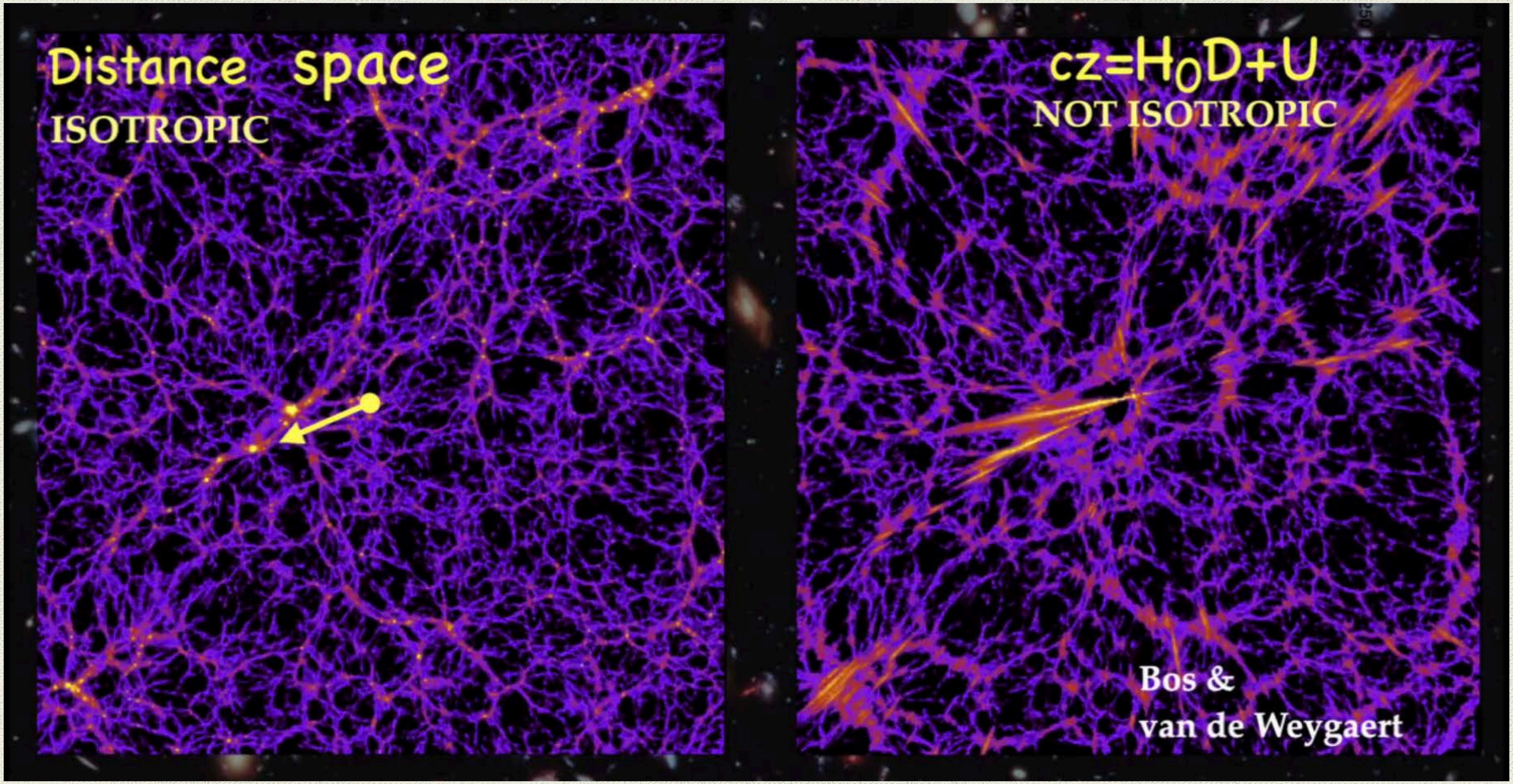


$$\mathbf{s} = \mathbf{r} + \frac{v_z}{H}\,\hat{\mathbf{z}}$$

Image from: feldman.ku.edu

- **Small-scale non-linear distortions: Fingers of God: Elongated along the line of sight**
- **Large-scale linear distortions: Kaiser effect: overdensity squished along the line of sight**

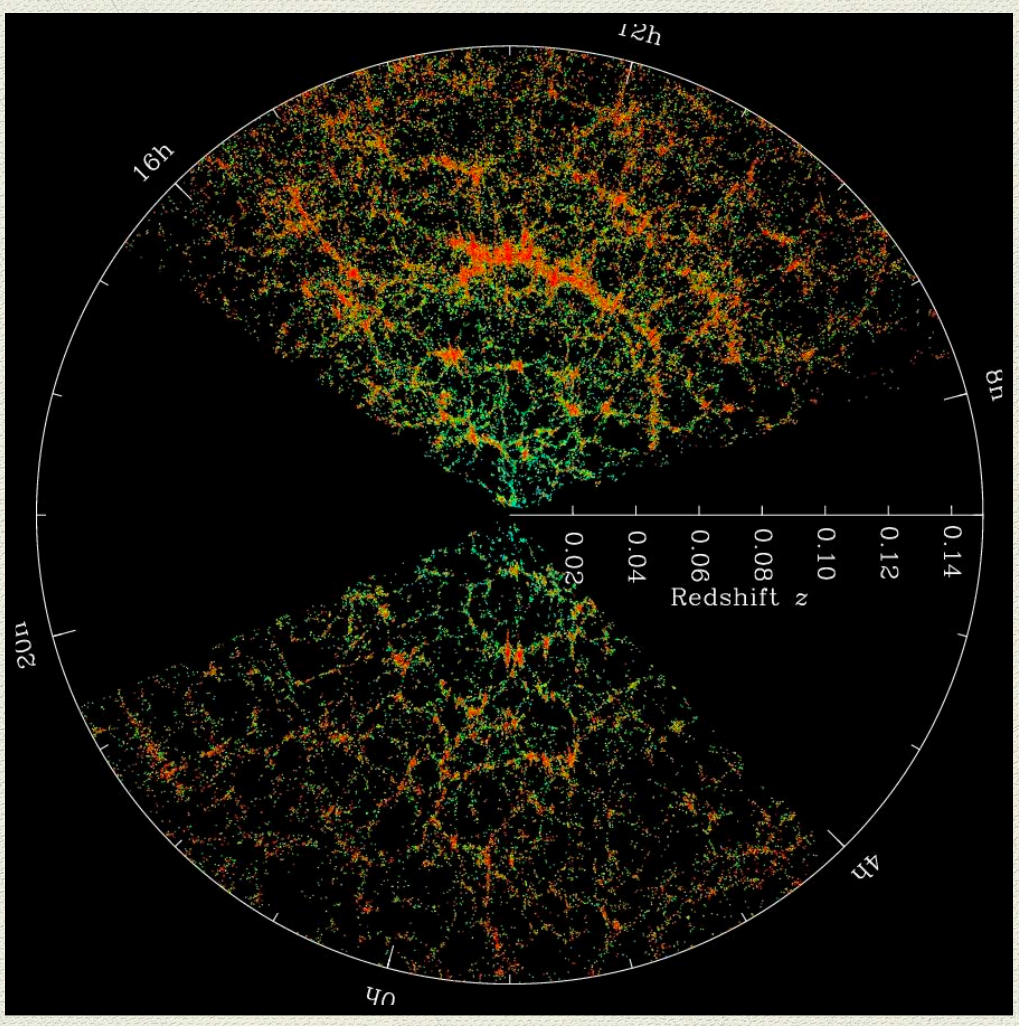
Redshift space distortions



Slide credit: A. Nusser

Noisy, missing and incomplete data

- Discrete sampling: shot noise
- Biased tracers: model of bias
- Redshift space distortions - structures are elongated along the line-of-sight.
- Gaps in the data - eg. galaxies in the ZoA are obscured by star, dust and gas, survey selection functions



True mapping the Universe

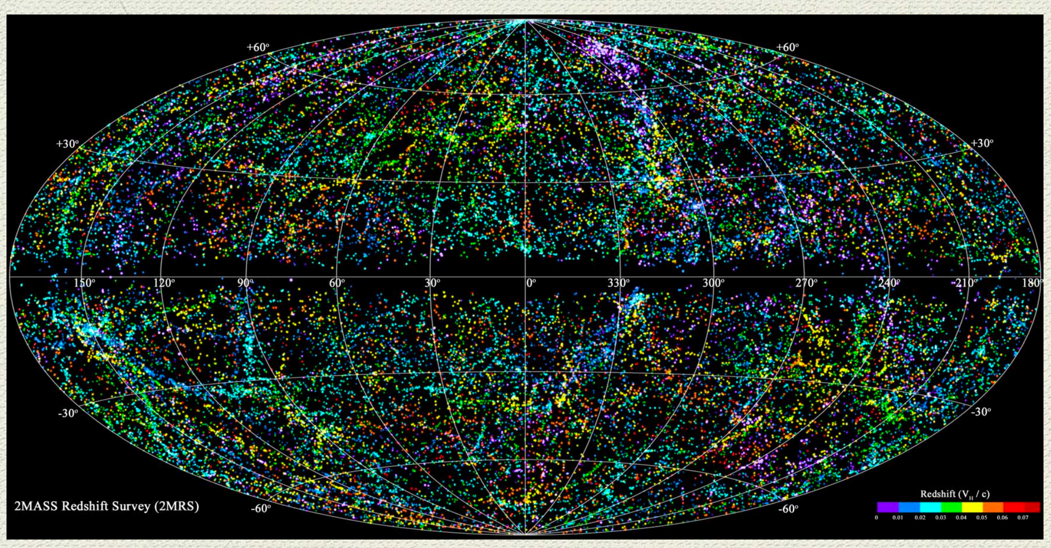
- Infer the true matter density and flows in 3D

- Compare inferred and observed velocity fields ==> how galaxies populate dm haloes, gravity

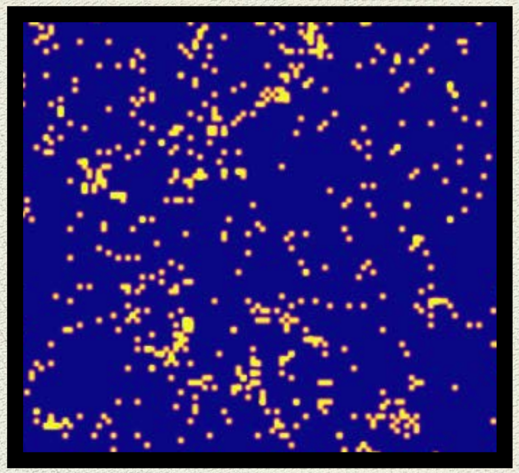
- Constraints on the cosmological parameters - *for the cosmology that we train on.*

$$-\frac{1}{H}\vec{\nabla}_r.\vec{v}_{lin} = f\delta \quad f \approx \Omega_m^{0.55}$$

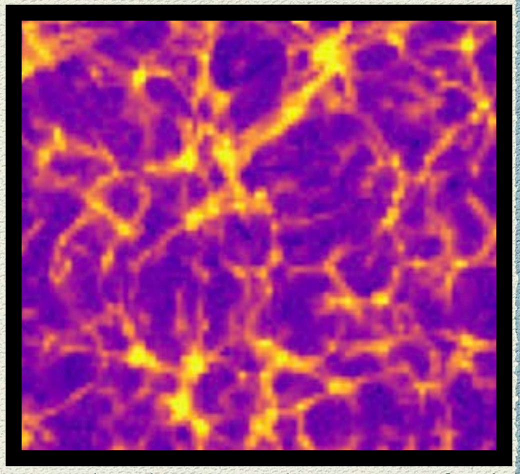
- Baryon Acoustic Oscillations



2MRS



galaxy distribution

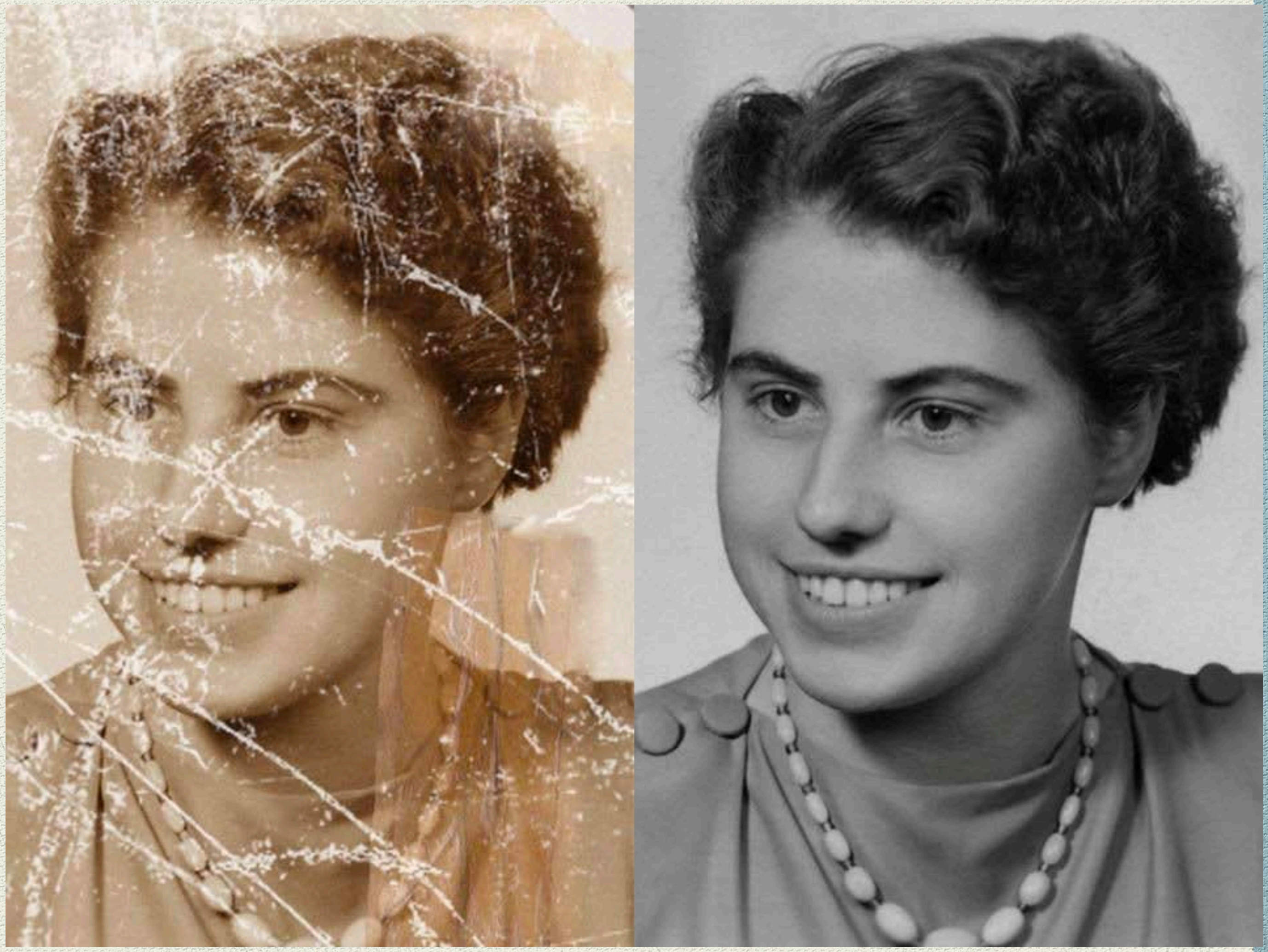


underlying density field

Can we remove **distortions, fill the gaps,** and **de-bias** using **neural nets**?

Filling gaps and correcting distortions

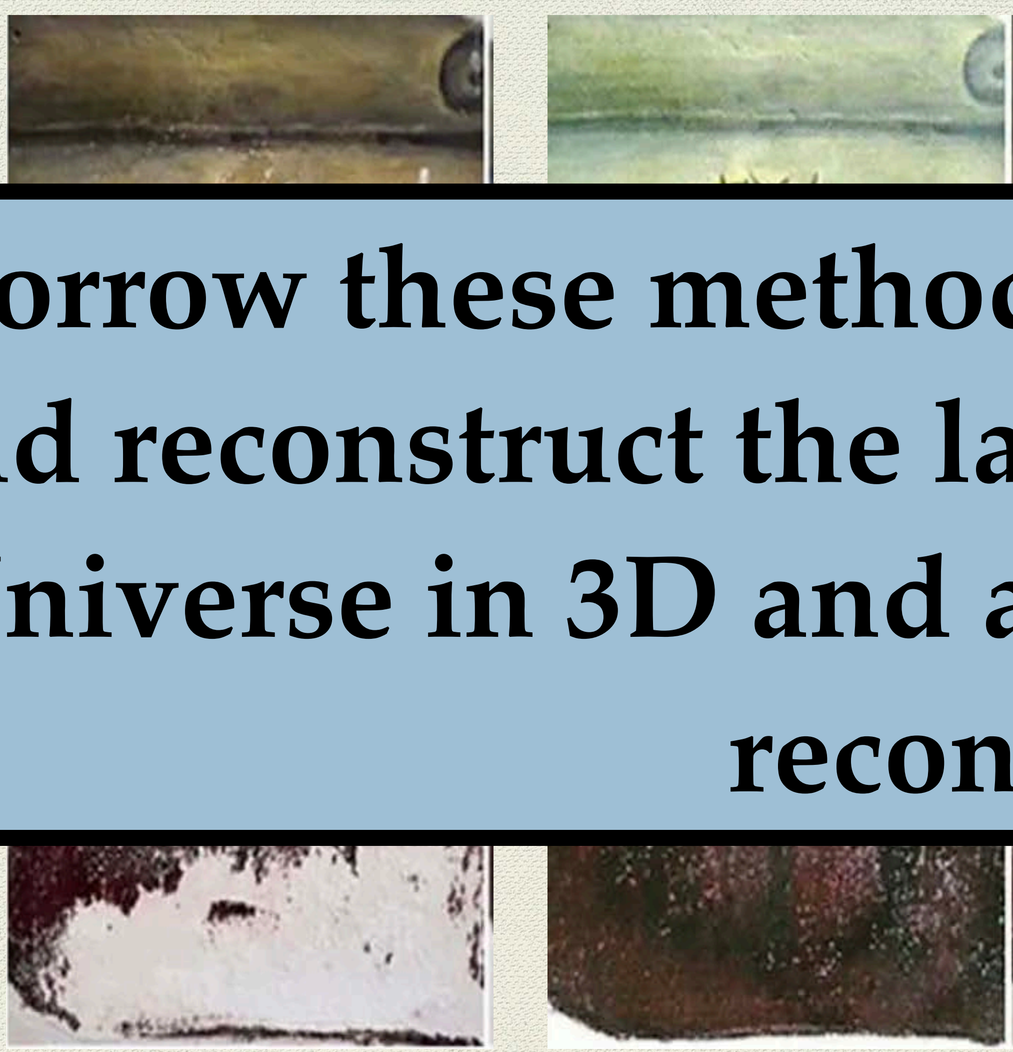
Example of restoration work: Elias Garcia Martinez's work 'Ecce Homo'

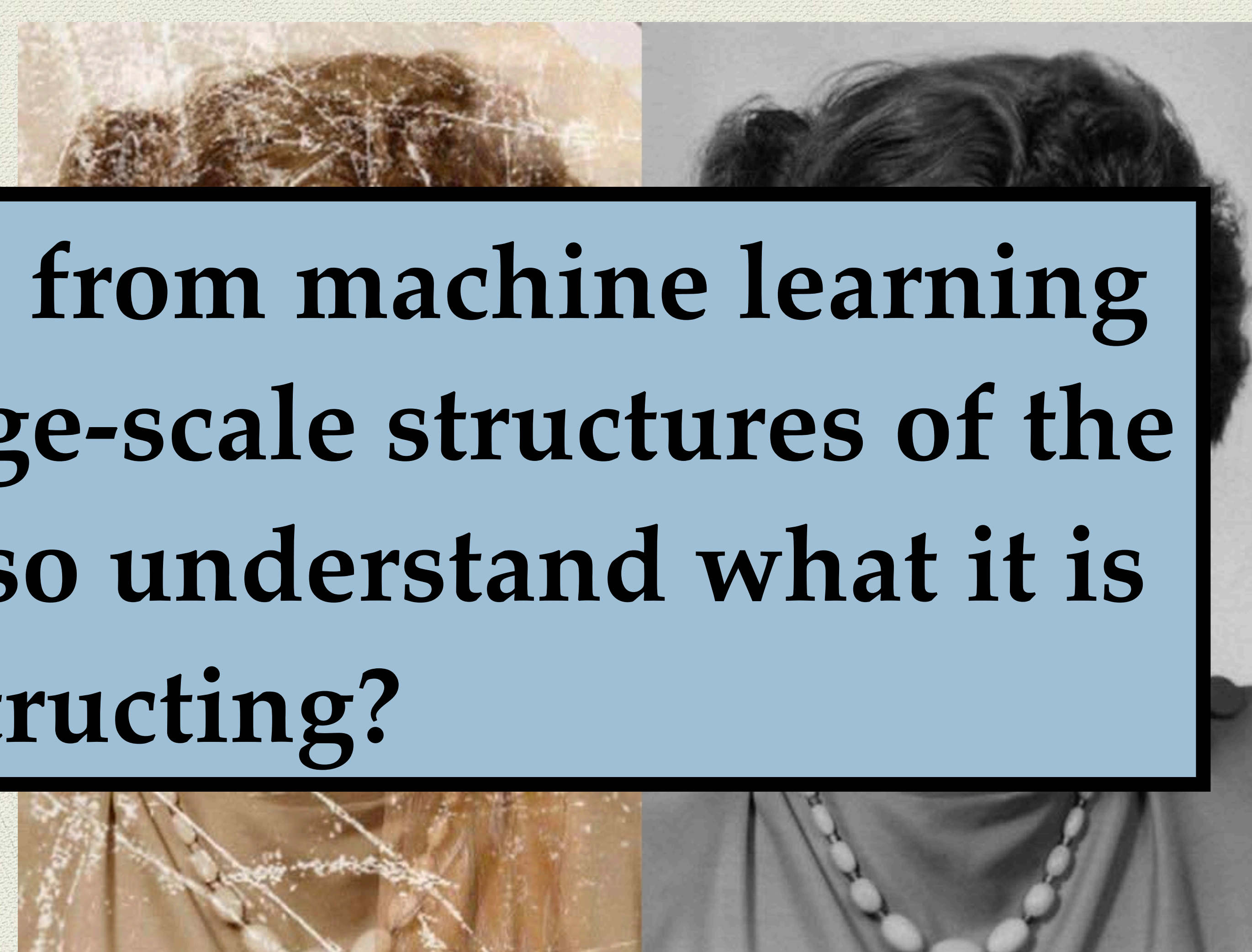


taken from:

Modern techniques - neural networks.

Filling gaps and correcting distortions





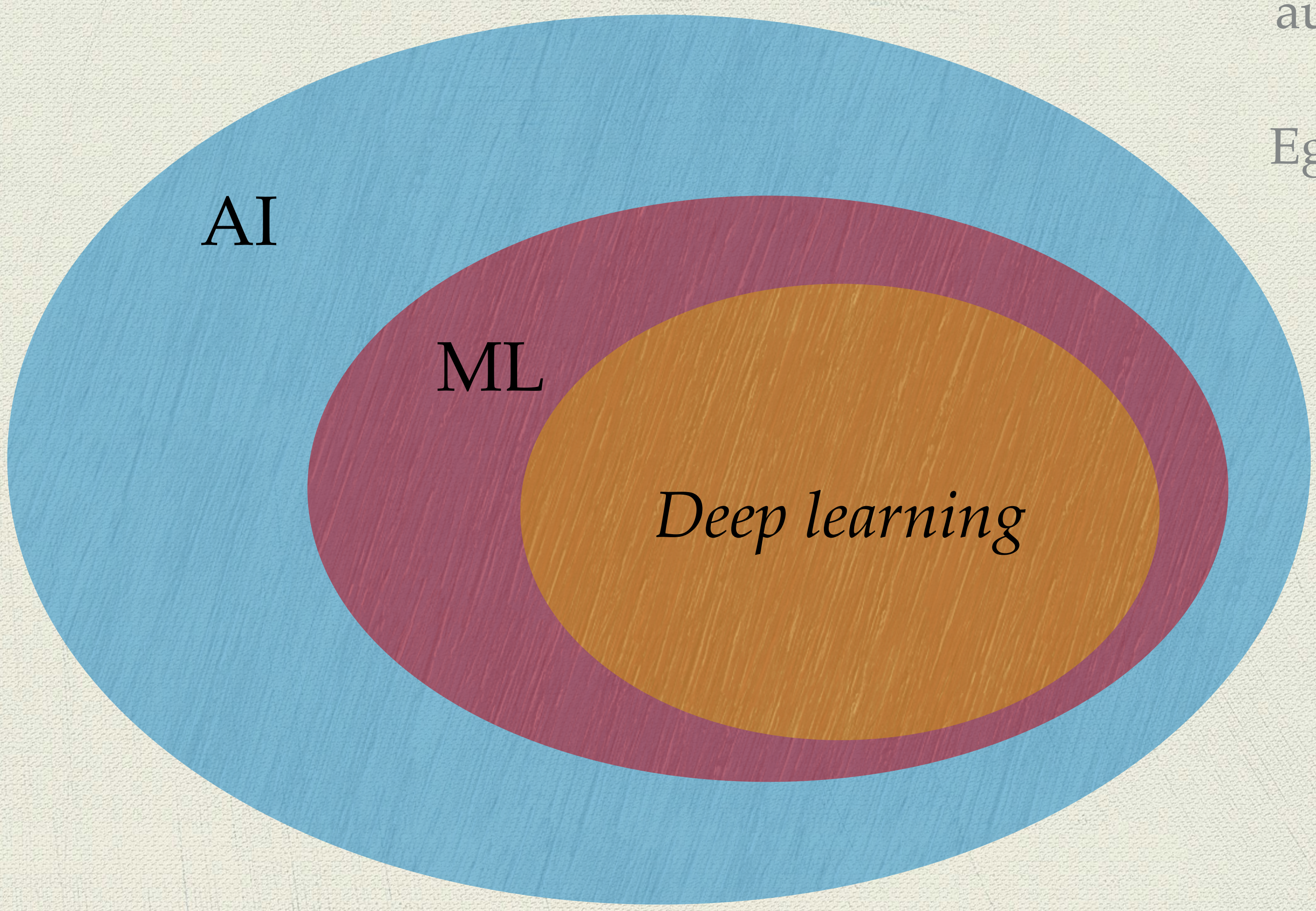
Borrow these methods from machine learning and reconstruct the large-scale structures of the Universe in 3D and also understand what it is reconstructing?

Example of restoration work: Elias Garcia Martinez's work 'Ecce Homo'

taken from:

Modern techniques - neural networks.

Machine learning and Neural nets?



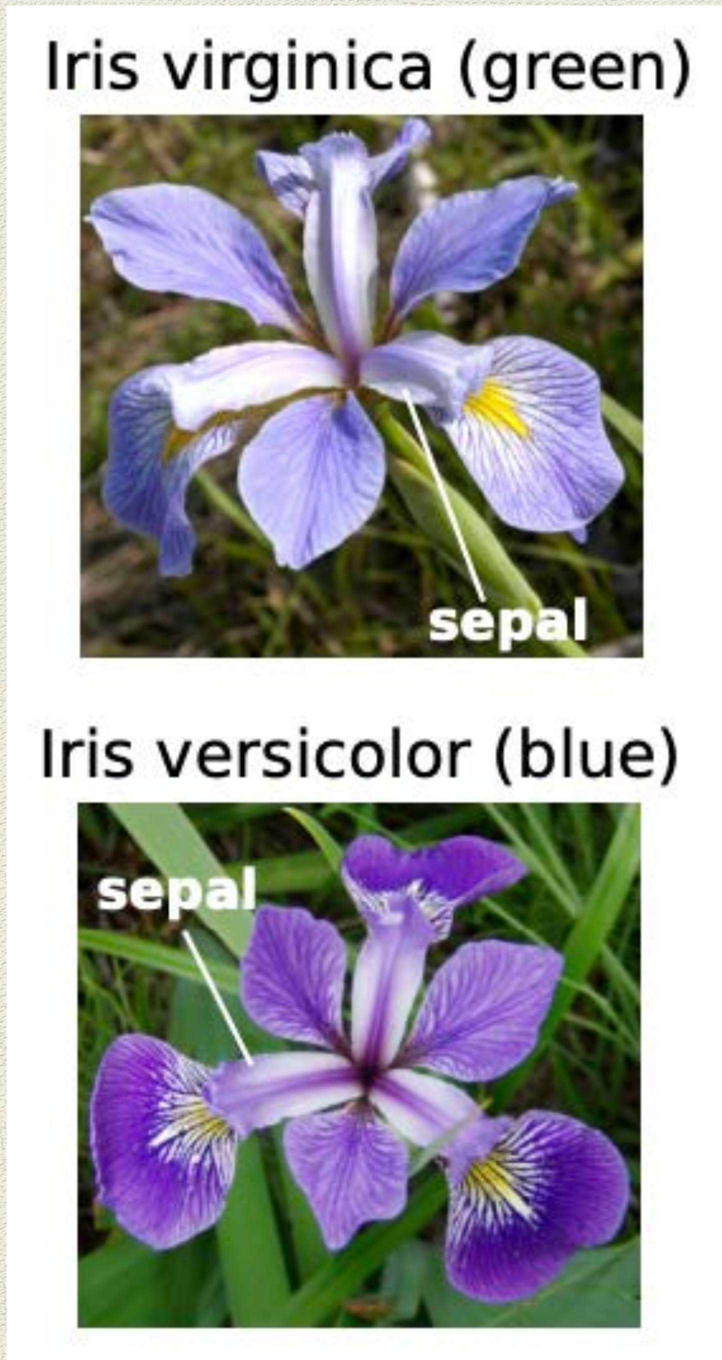
Artificial Intelligence: the effort to automate intellectual tasks normally performed by humans.
Eg: early chess programs, robot vacs, automated cars, etc

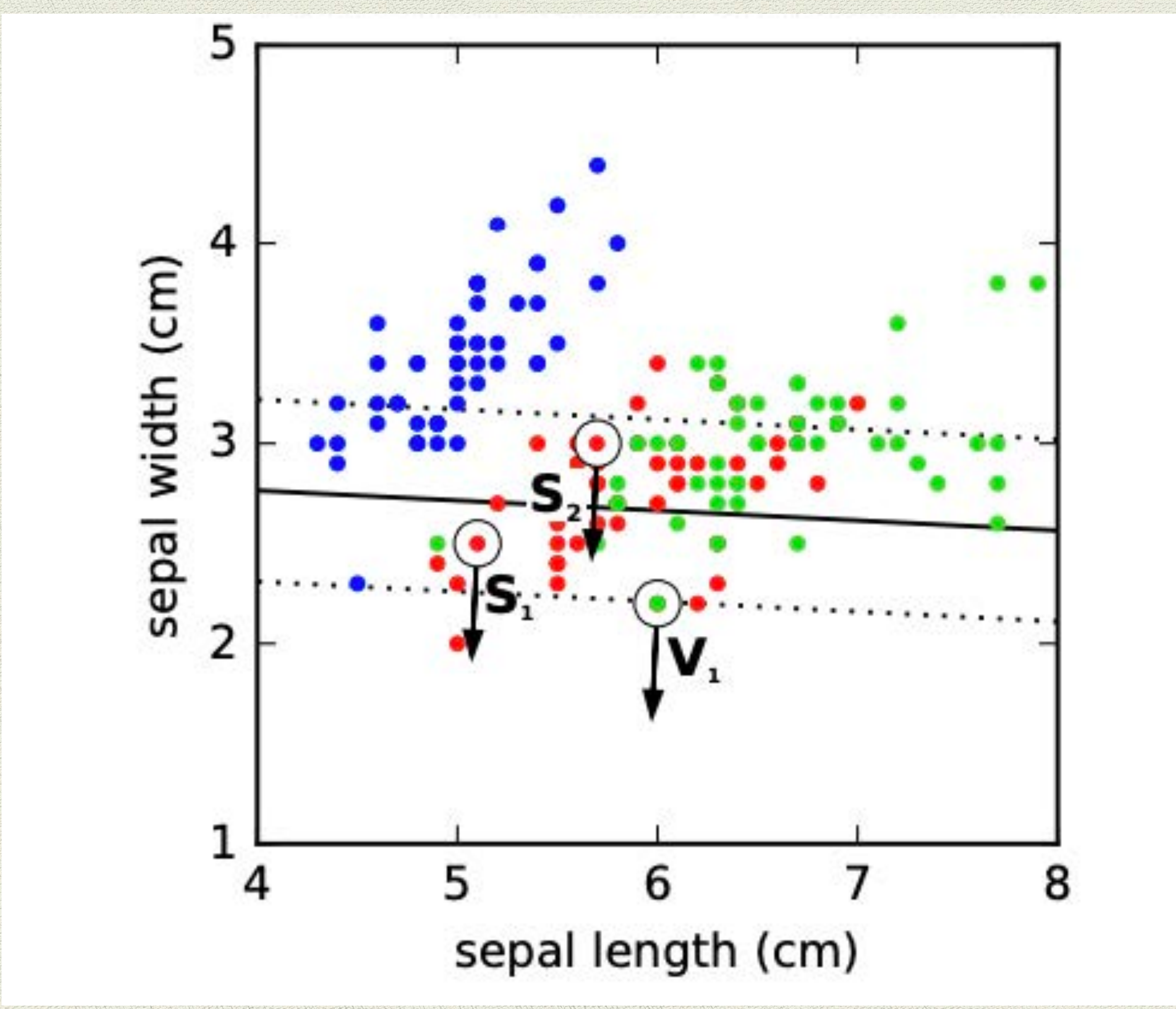
Machine Learning:
could a computer learn on its own how to perform a specified task?

Deep Learning: specific subfield of machine learning: learning representations from data that puts an emphasis on learning successive layers of increasingly meaningful representations

Adapted from the book: deep learning with python by François Chollet

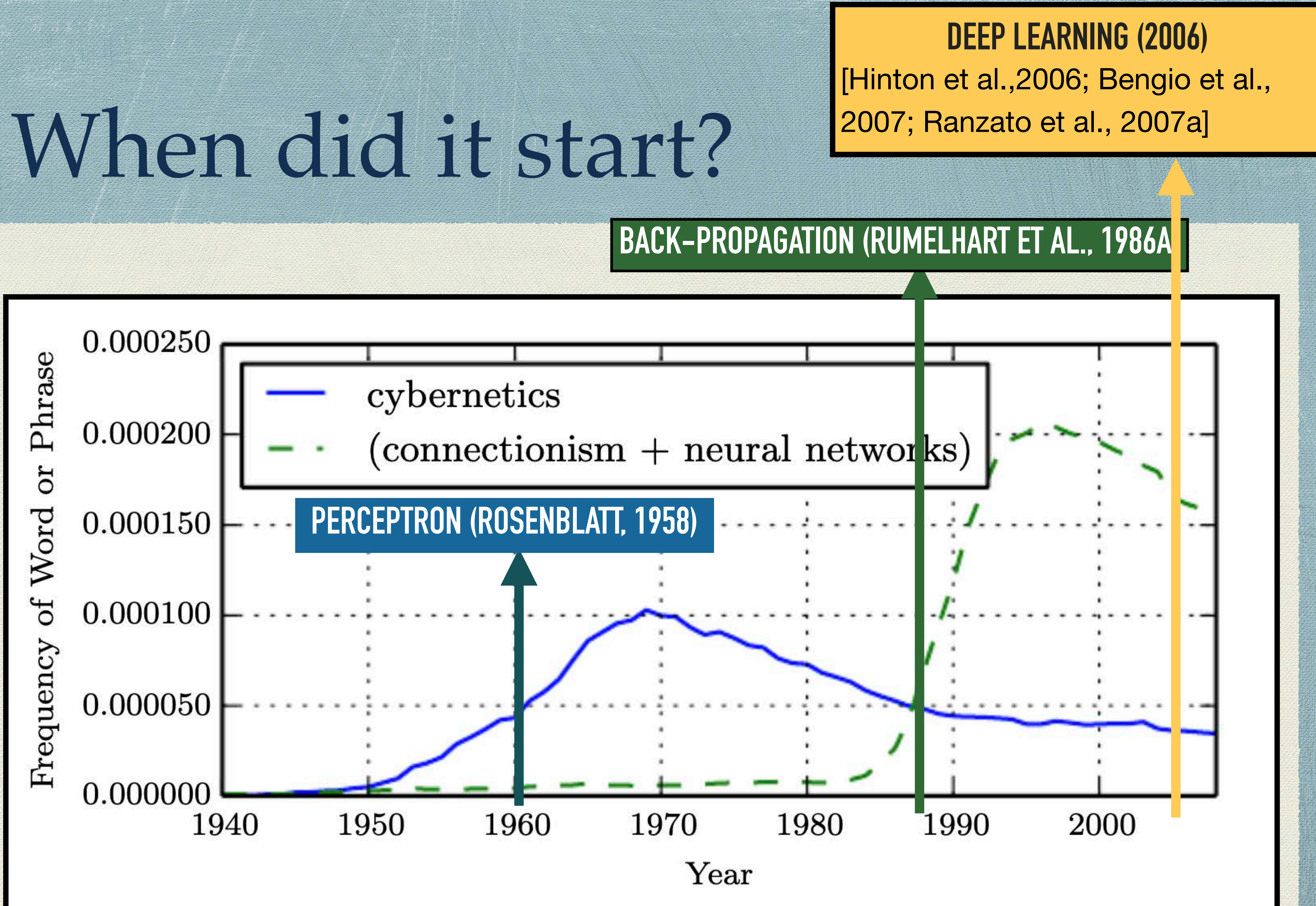
Machine learning v/s Deep learning?





Lapuschkin. S, et al 2019

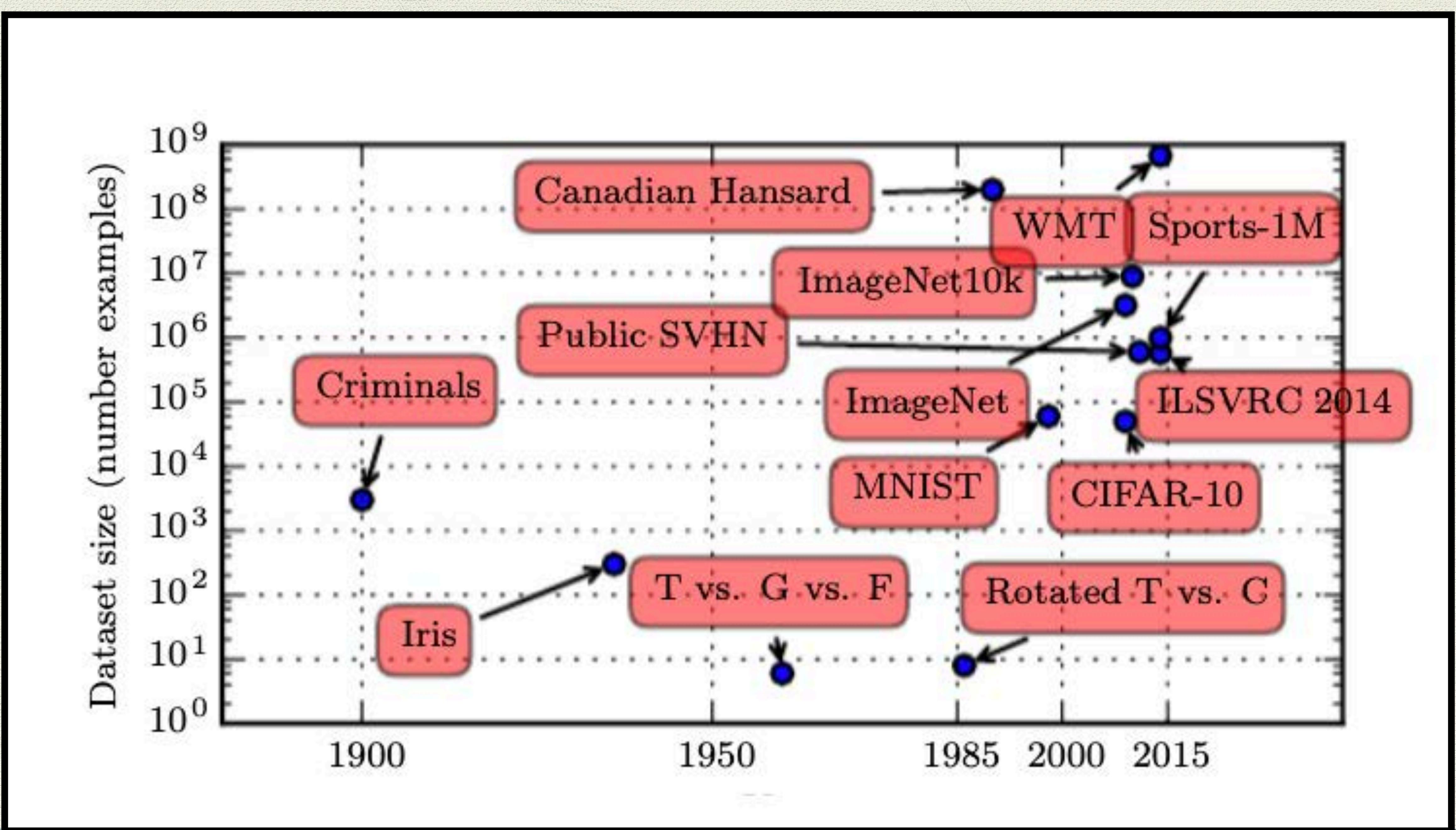
When did it start?



Adapted from the book: deep learning by Ian Goodfellow, Yoshura Bengio and Aaron Courville

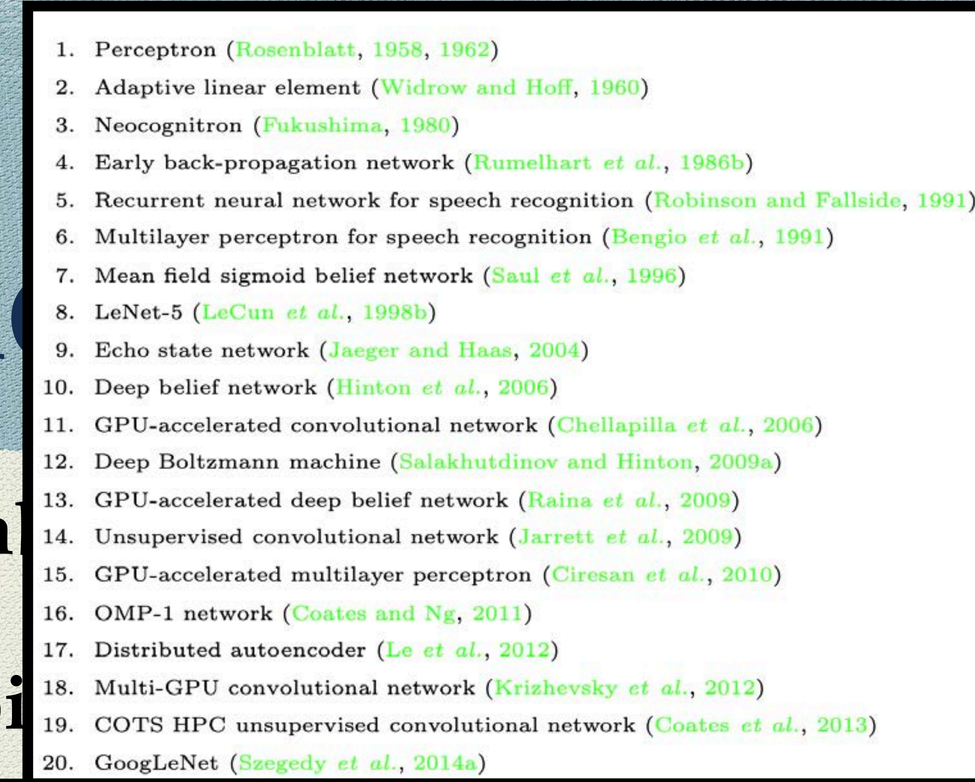
Why is deep learning popular now?

- Scale of **digitised data**



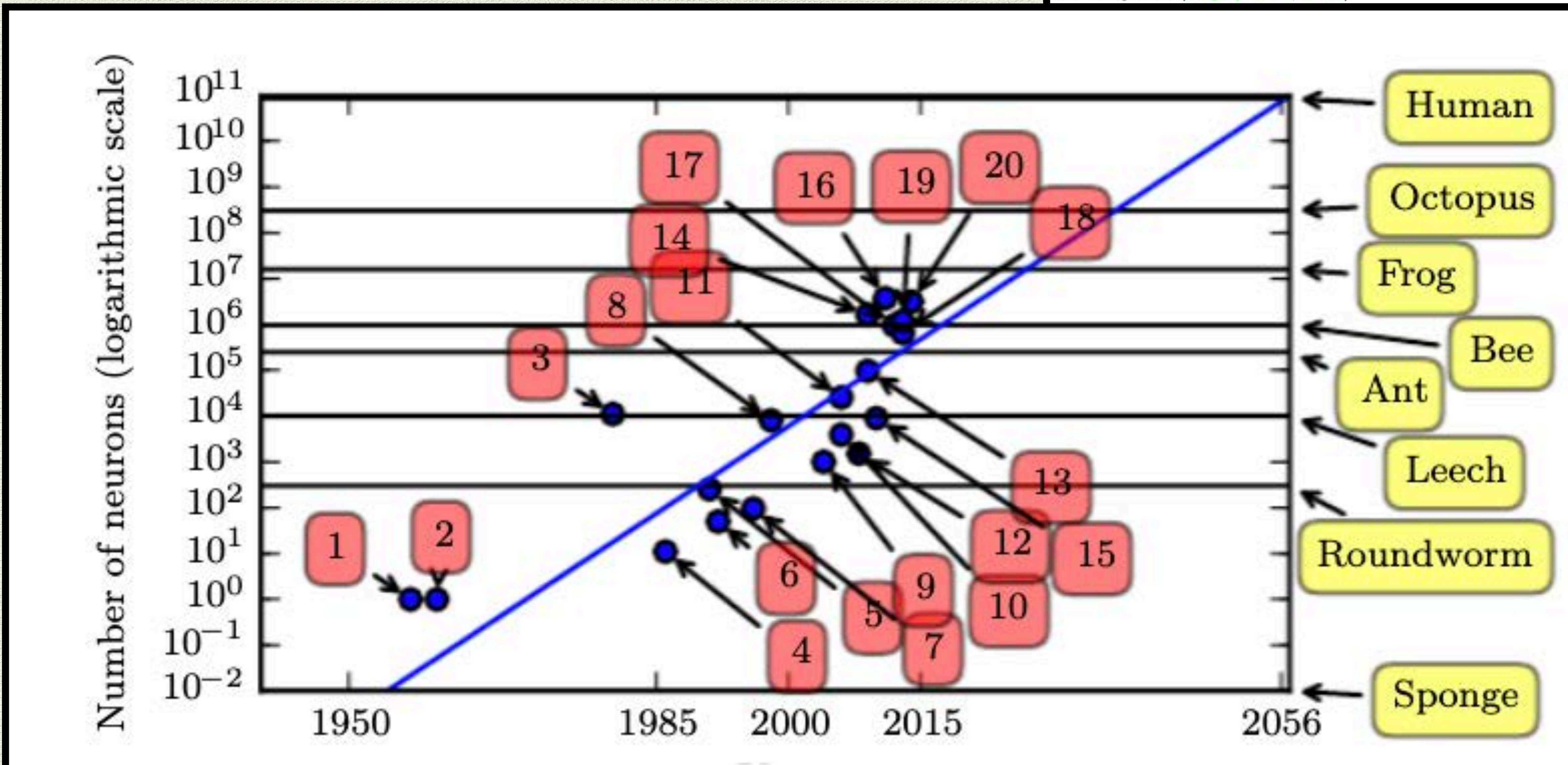
Adapted from the book: deep learning by Ian Goodfellow, Yoshura Bengio and Aaron Courville

Why is it popular n



1. Perceptron (Rosenblatt, 1958, 1962)
2. Adaptive linear element (Widrow and Hoff, 1960)
3. Neocognitron (Fukushima, 1980)
4. Early back-propagation network (Rumelhart et al., 1986b)
5. Recurrent neural network for speech recognition (Robinson and Fallside, 1991)
6. Multilayer perceptron for speech recognition (Bengio et al., 1991)
7. Mean field sigmoid belief network (Saul et al., 1996)
8. LeNet-5 (LeCun et al., 1998b)
9. Echo state network (Jaeger and Haas, 2004)
10. Deep belief network (Hinton et al., 2006)
11. GPU-accelerated convolutional network (Chellapilla et al., 2006)
12. Deep Boltzmann machine (Salakhutdinov and Hinton, 2009a)
13. GPU-accelerated deep belief network (Raina et al., 2009)
14. Unsupervised convolutional network (Jarrett et al., 2009)
15. GPU-accelerated multilayer perceptron (Ciresan et al., 2010)
16. OMP-1 network (Coates and Ng, 2011)
17. Distributed autoencoder (Le et al., 2012)
18. Multi-GPU convolutional network (Krizhevsky et al., 2012)
19. COTS HPC unsupervised convolutional network (Coates et al., 2013)
20. GoogLeNet (Szegedy et al., 2014a)

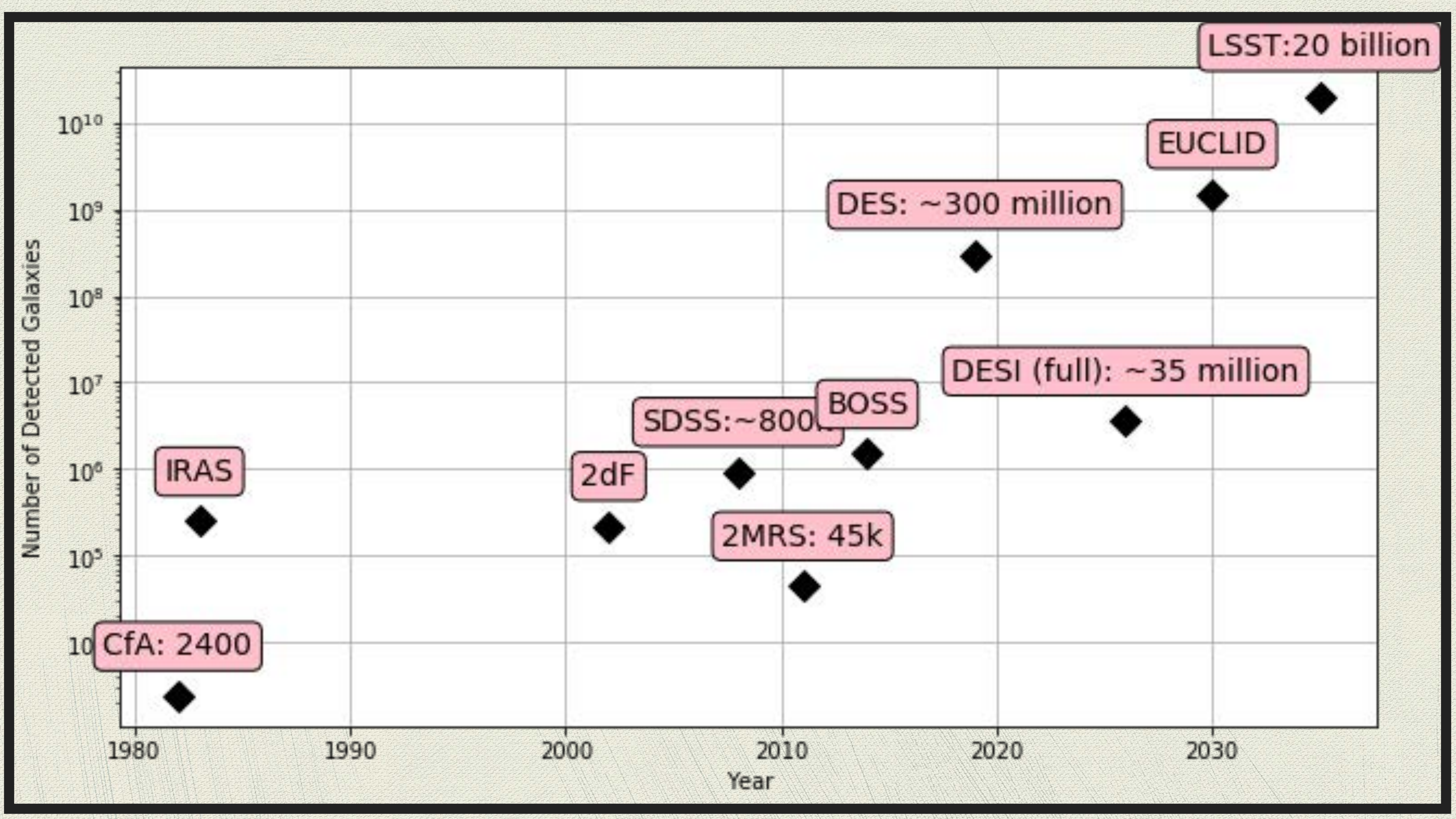
- **Scale of computation - hardware GPUs, num**
- **Algorithmic innovations - eg ReLU to sigmoi**



Adapted from the book: deep learning by Ian Goodfellow, Yoshura Bengio and Aaron Courville

What about in cosmology?

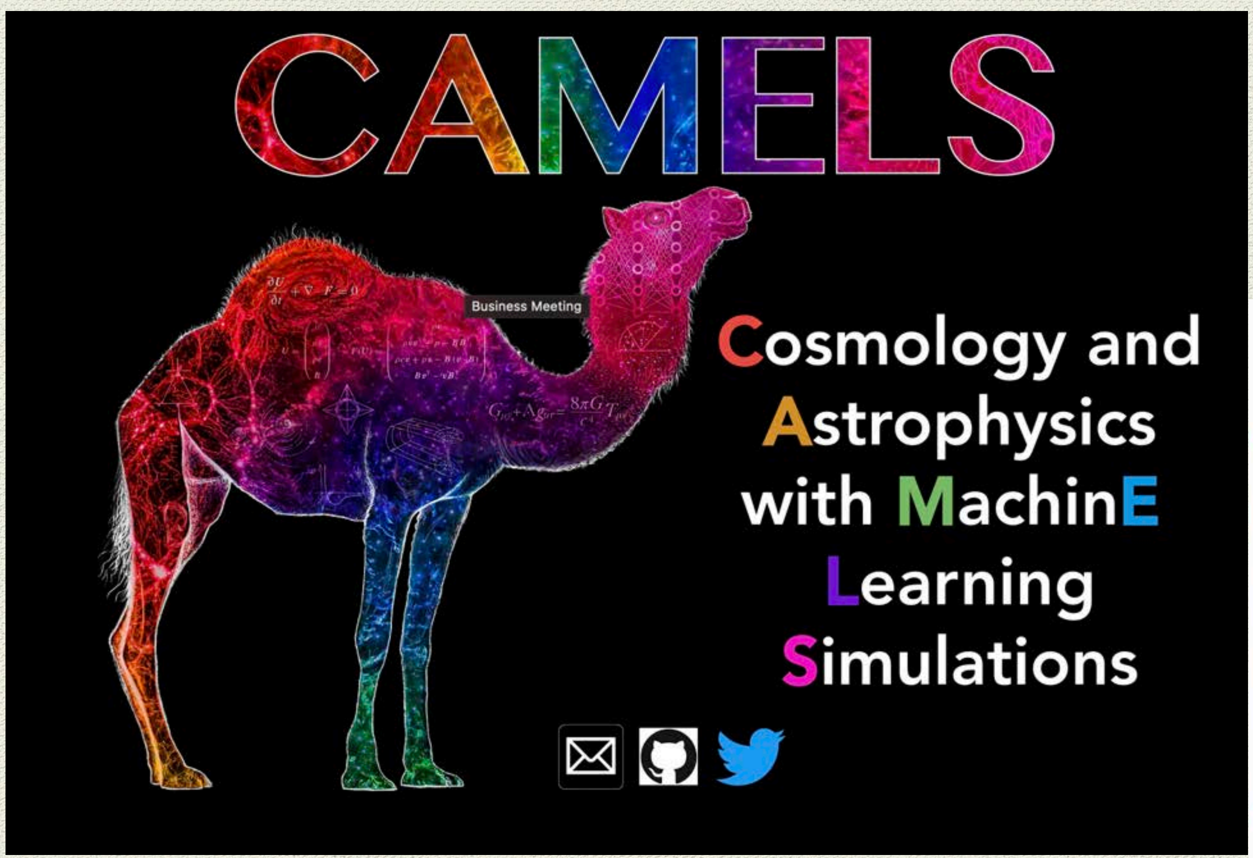
- Scale of observational data

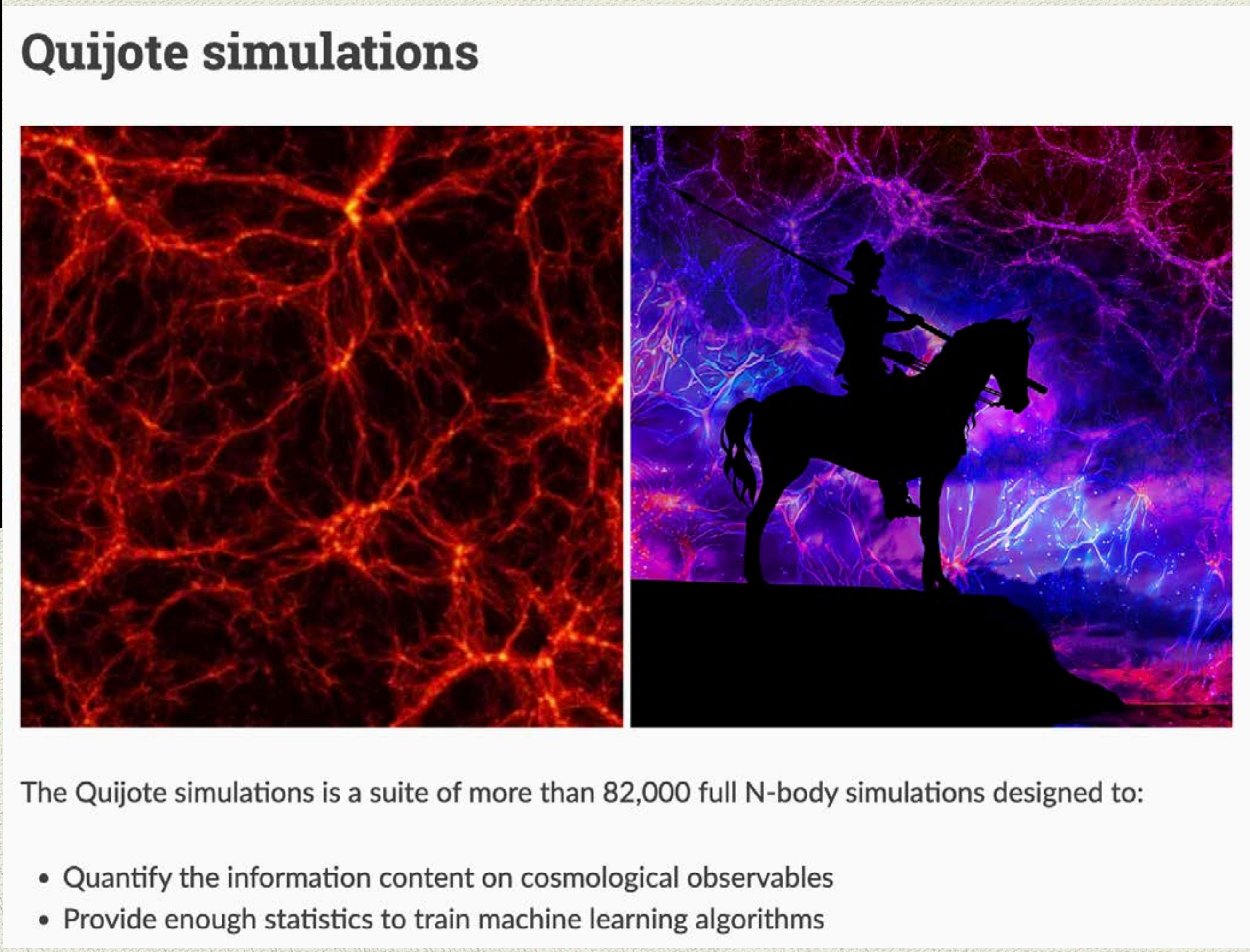


I made this plot from data available on survey websites. Please verify.

What about in cosmology?

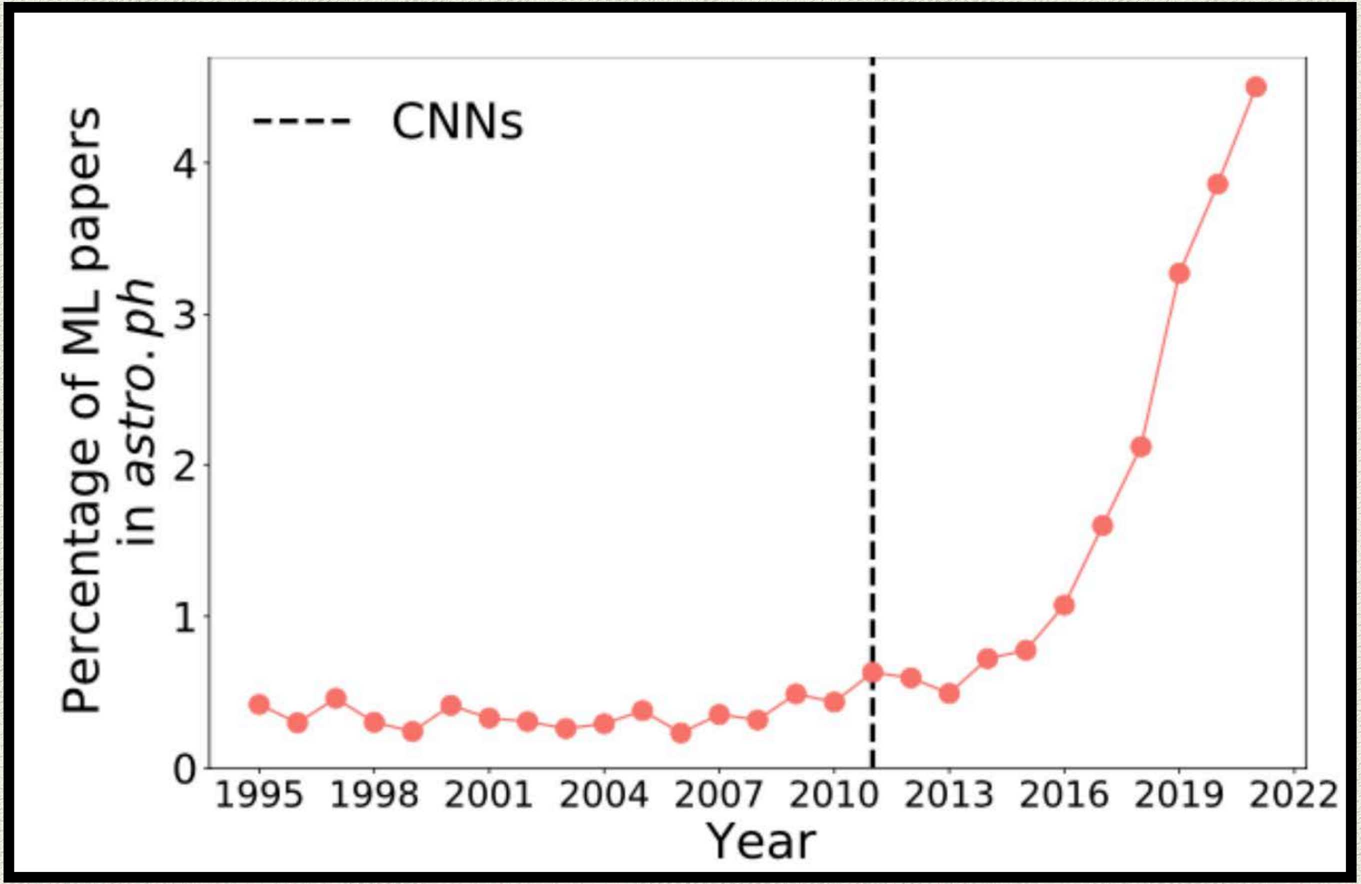
- Scale of simulated data - needs catching up!

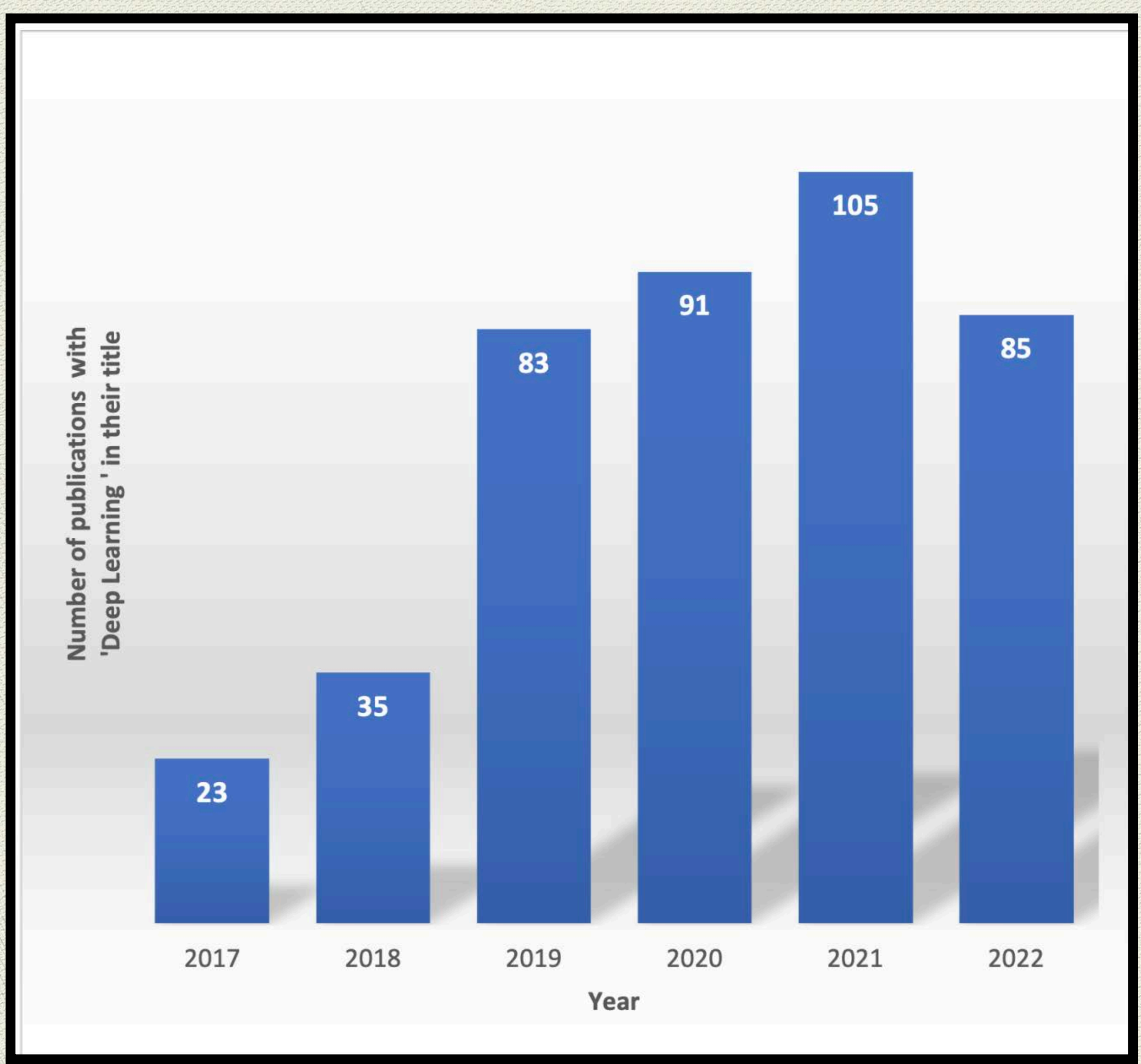




I made this plot from data available on survey websites. Please verify.

ML papers in astro-ph

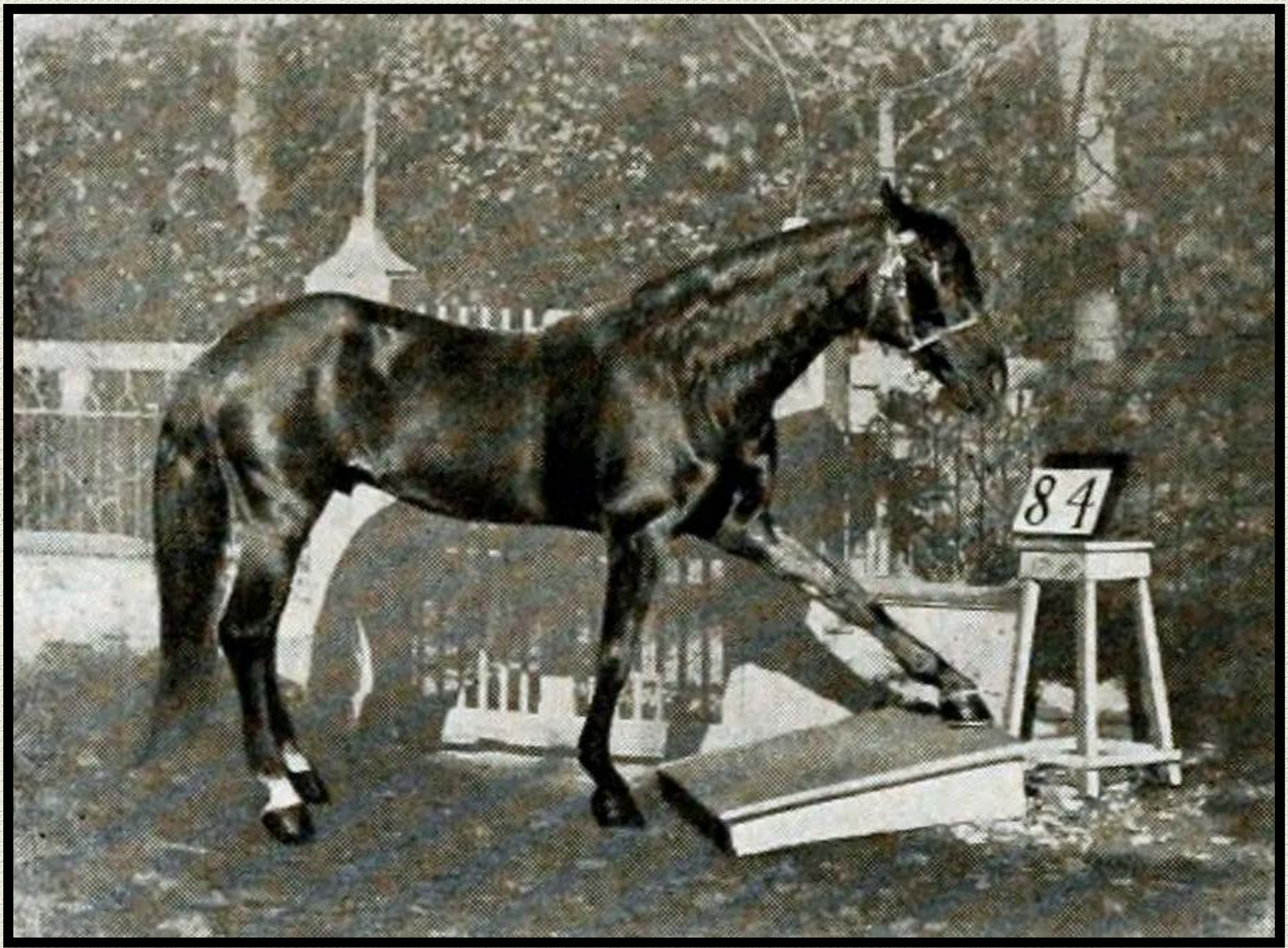




Plots taken from: Ofer Lahav 2023 (https://arxiv.org/pdf/2302.04324)

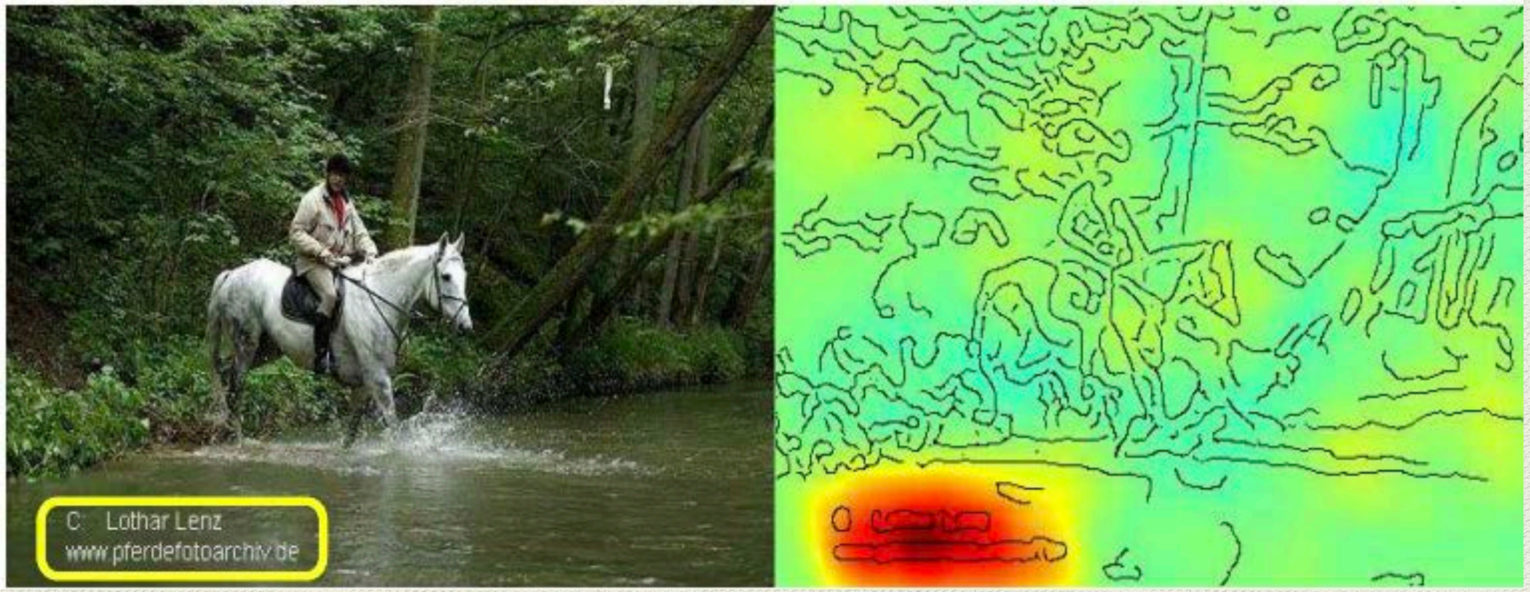
Caveats of machine learning

- **Clever Hans effect**: model might appear to perform well, but could be picking up on **spurious correlations or artefacts** in the data — not learning what we think it's learning



The Horse called Hans. [Image source: Karl Krall, Public domain, via Wikimedia Commons]

Caveats of machine learning

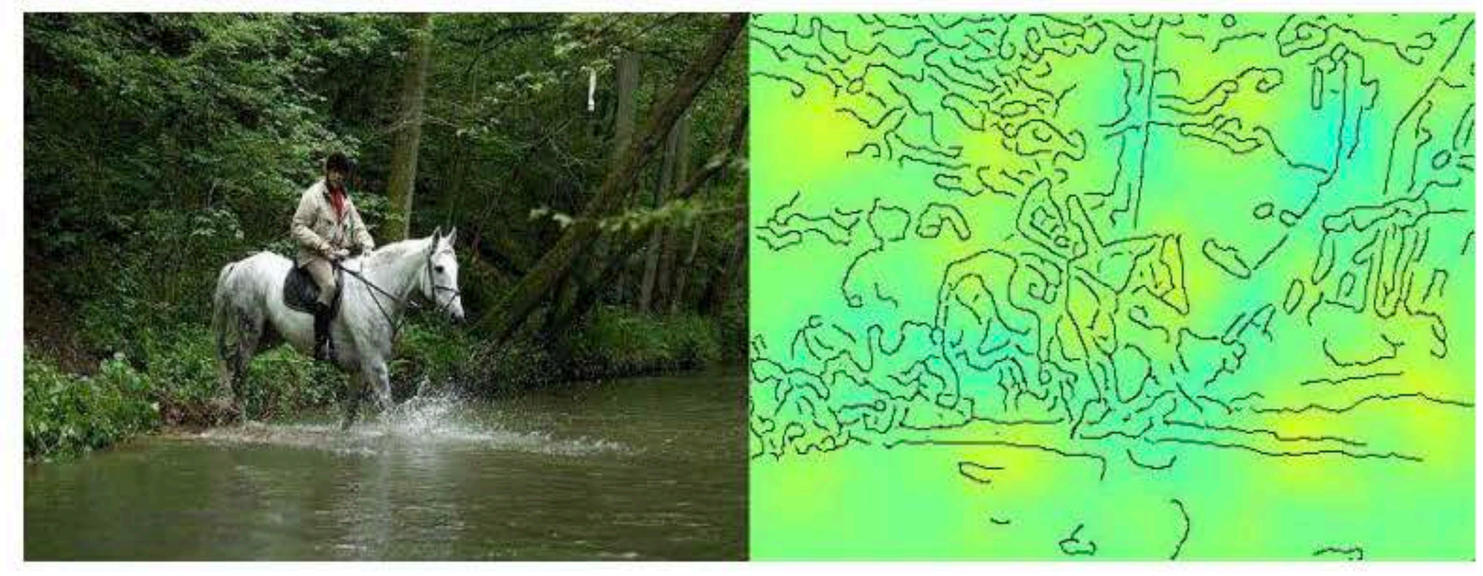


Classified as a Horse

Classified as a Horse

Lapuschkin. S, et al 2019

Caveats of machine learning



Not a Horse

Not a Horse

Lapuschkin. S, et al 2019

Caveats of machine learning

- Large representative dataset to train - method is only **as good as the data is.** (eg: Clever Hans phenomenon)

- **Start simple** with a small network and an understanding of what we are optimising for - **loss function**.

- Compare it with **already existing techniques**, to see if the performance is better or worse.

Back to reconstruction!

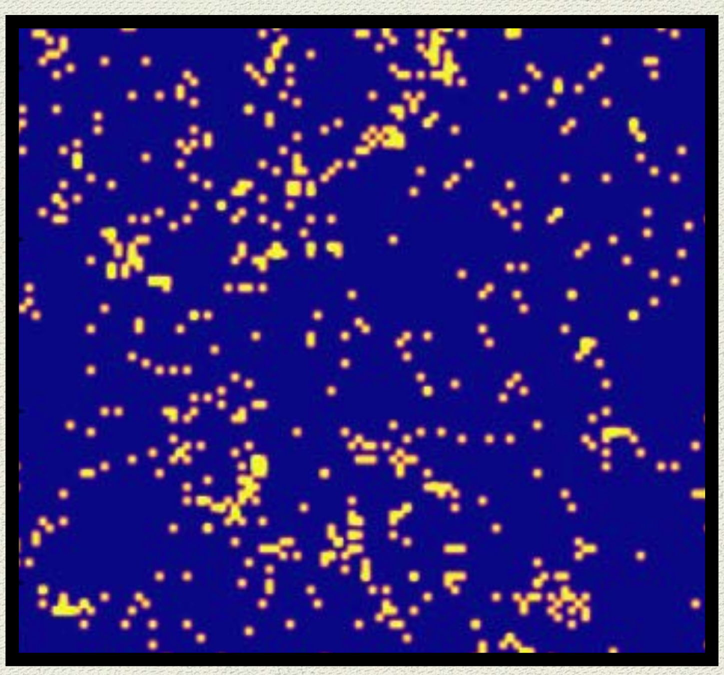
GOALS:

- Given a galaxy distribution in redshift space, reconstruct the true **underlying matter density** and **velocity fields** using **neural networks.**
- In the process, demystify machine learning - can we interpret what the machine does using **known statistical techniques?**
- Can we recover **Wiener filter** from **neural network** methods?
- Use a hybrid technique - physics+neural network
- Apply this technique to **2MRS data.**

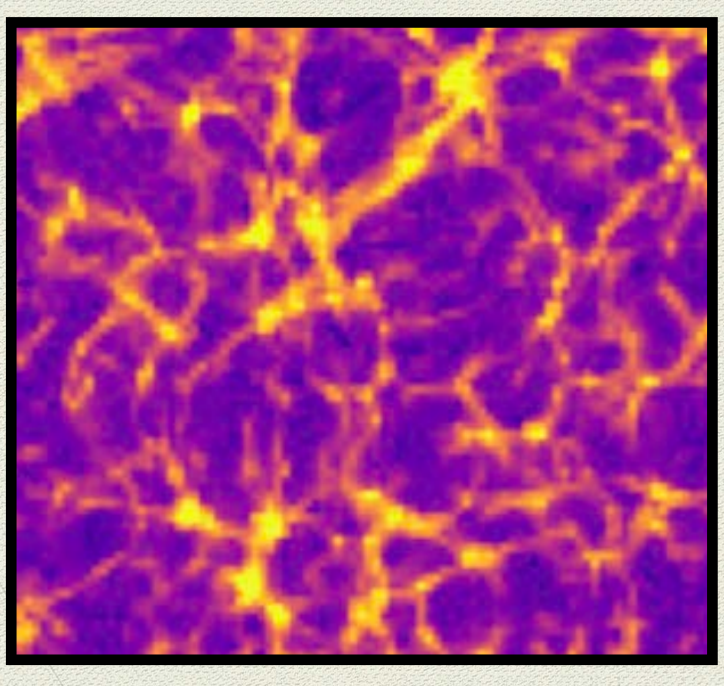
Main aim: map the density or velocity fields

Observed galaxy density field

True underlying matter density



Train a network to learn this mapping



Input field : I

Target field : T

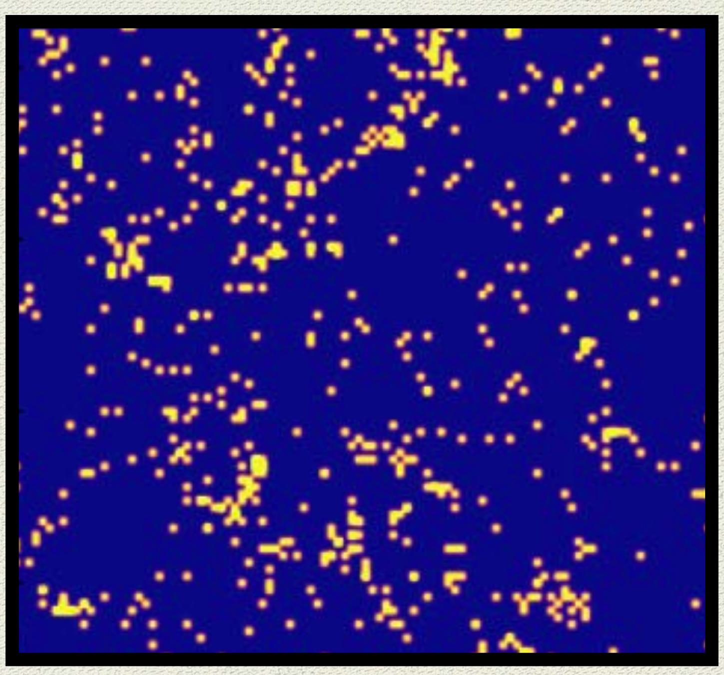
Other methods used so far for reconstructing LSS? (Eulerian reconstructions)

- **Wiener filter** - linear reconstruction - e.g *Zaurobi et al 1994, Lilow et al 2021*
- **Other reconstruction methods** - e. g. *Bertschinger & Dekel 1989; Yahil et al. 1991; Nusser & Davis 1994; Fisher et al. 1995; Bistolas & Hoffman 1998; Zaroubi et al. 1999; Kitaura et al. 2010; Jasche et al. 2010; Courtois et al. 2011; Kitaura 2013; Jasche & Wandelt 2013; Wang et al. 2013; Carrick et al. 2015; Lavaux 2016; Bos et al 2016,2018; Jasche & Lavaux 2019; Graziani et al. 2019; Kitaura et al. 2020; Zhu et al. 2020*

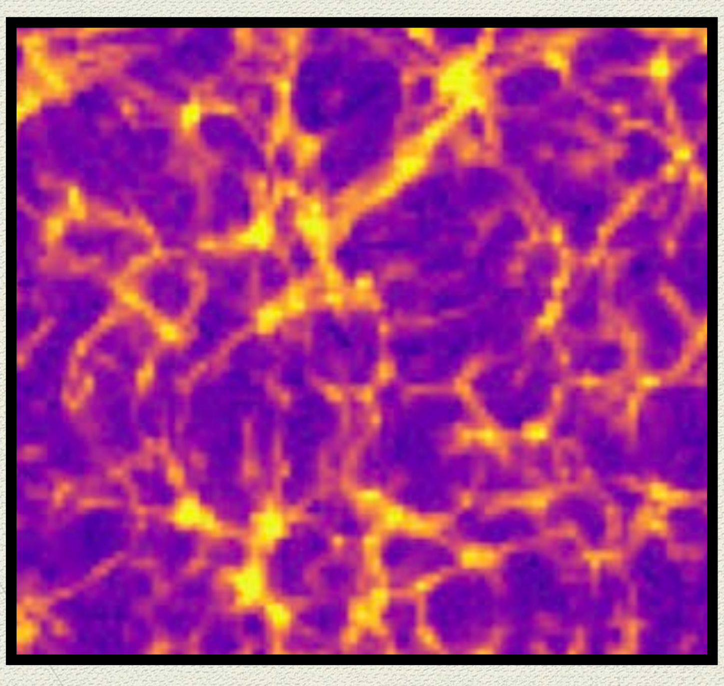
Main aim: map the density or velocity fields

Observed galaxy density field

True underlying matter density



Train a network to learn this mapping

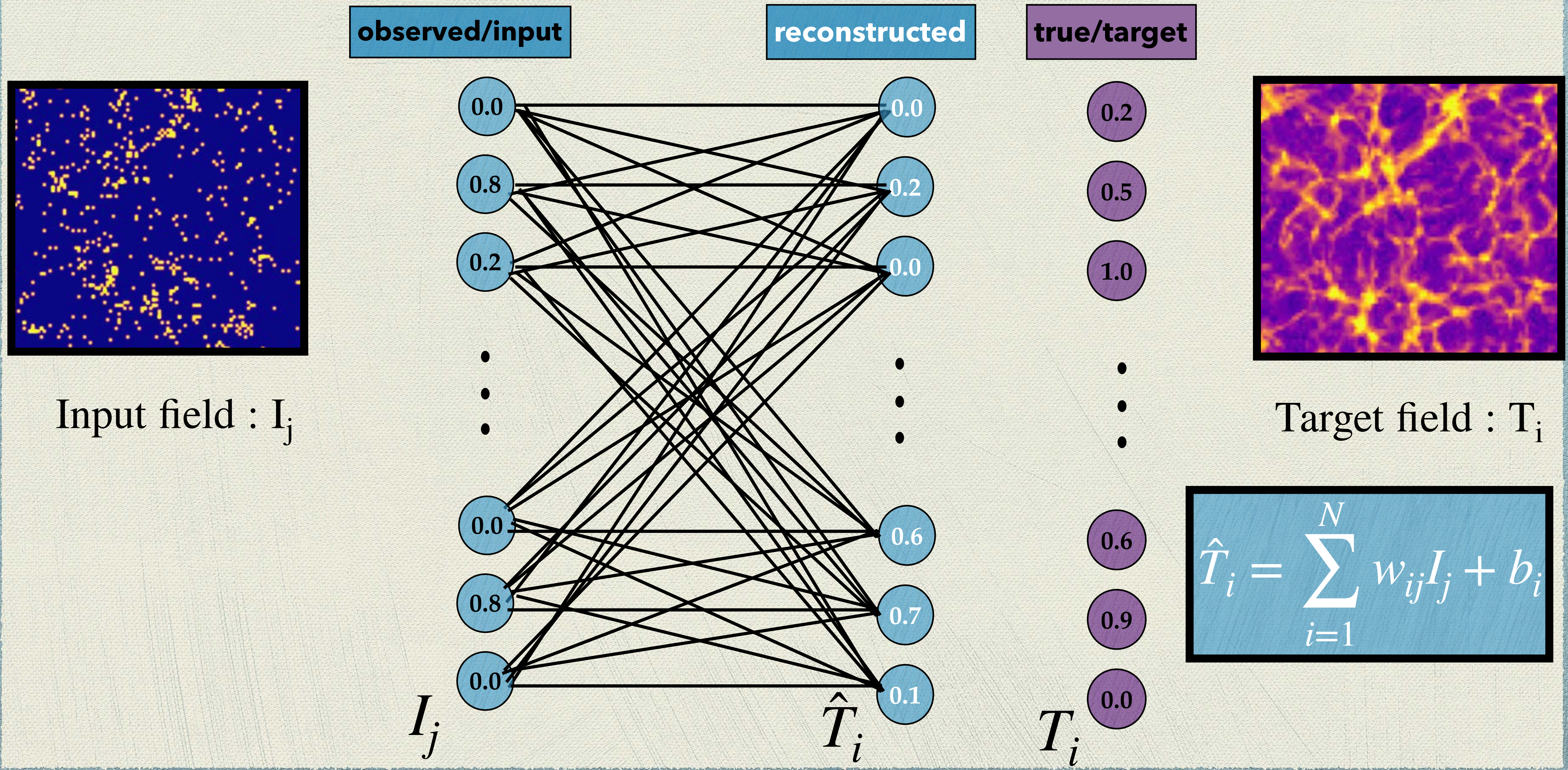


Input field : I

Target field : T

A simple network

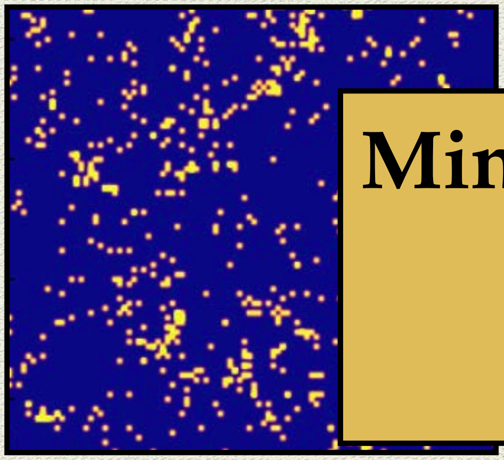
$$\hat{T}_1 = w_{11}I_1 + w_{12}I_2 + w_{13}I_3 + \ldots + w_{1n}I_n + b_1$$



$$\hat{T}_i = \sum_{i=1}^{N} w_{ij}I_j + b_i$$

Nonlinear network + MSE loss = Mean posterior estimate

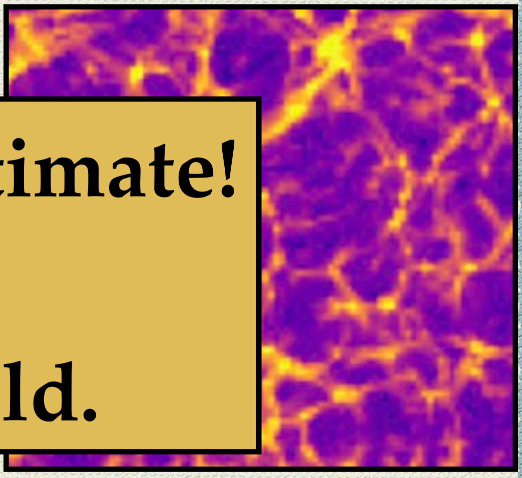
$$L^{\mathrm{MSE}}(\hat{\mathbf{T}}) = \frac{1}{MN}\sum_{\alpha=1}^{M}\sum_{j=1}^{N}\left(T_j^{\alpha} - \hat{T}_j(\mathbf{I}^{\alpha})\right)^2$$



Input field : I_j

Minimising MSE gives the mean posterior estimate!

Mean of true fields given the observed field.

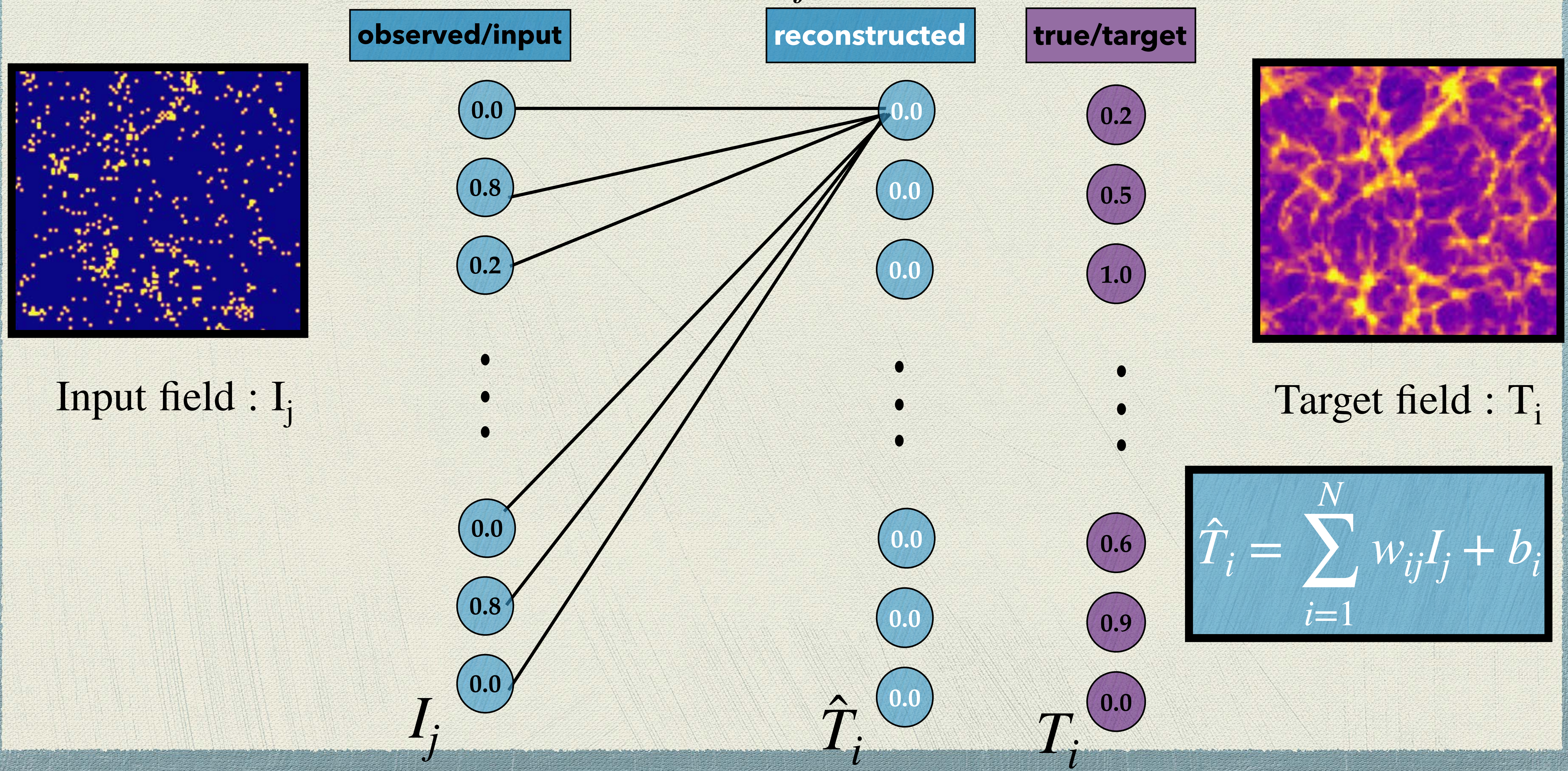


Target field : T_i

$$\hat{\mathbf{T}}_i^{\mathrm{MSE}}(\mathbf{I}) = \sum_{T} P(\mathbf{T}\,|\,\mathbf{I})\, T_i = \langle T_i\,|\,\mathbf{I}\rangle,$$

A simple linear network+MSE = Wiener filter
Complex network with linear activation+MSE=WF

$$L^{\mathrm{MSE}} = \frac{1}{N} \sum_{j=1}^{N} \left(T_j - \hat{T}_j(\mathbf{I})\right)^2$$



Wiener filtering for galaxy distributions

[Zaroubi et al 1994]

- Observed density field —-> True density field
- Reconstructed field is a linear combination of the observed field $\hat{T}_i^{WF(I)} = \sum_j w_{ij}^{WF} I_j + b_i^{WF}$
- Minimum variance estimator: minimise MSE.

$$T^{WF} = \langle TI \rangle \langle II \rangle^{-1} I$$

Wiener filtering for galaxy distributions

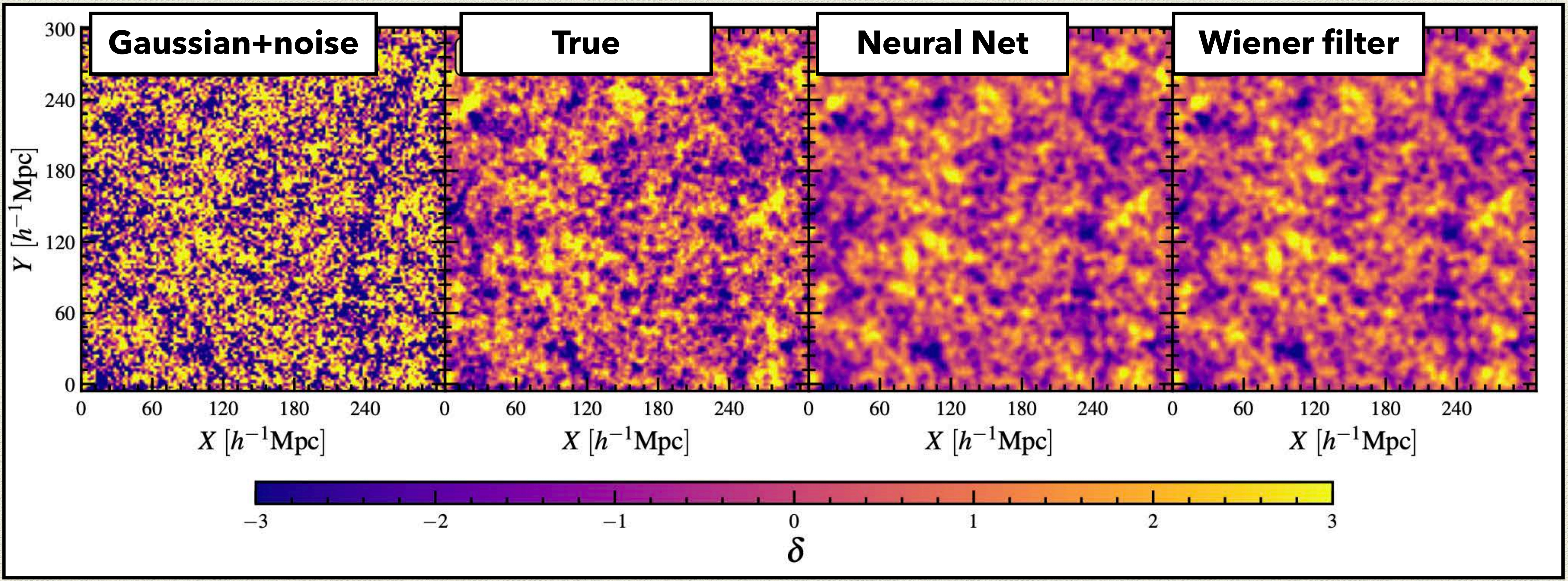
[Zaroubi et al 1994]

1. A **neural network** with an input and output layer and **linear activation** is equivalent to a **WF**.

2. When the field to be reconstructed is **Gaussian**, WF (min. var) and non-linear NN (mean posterior) estimates should both be the same!

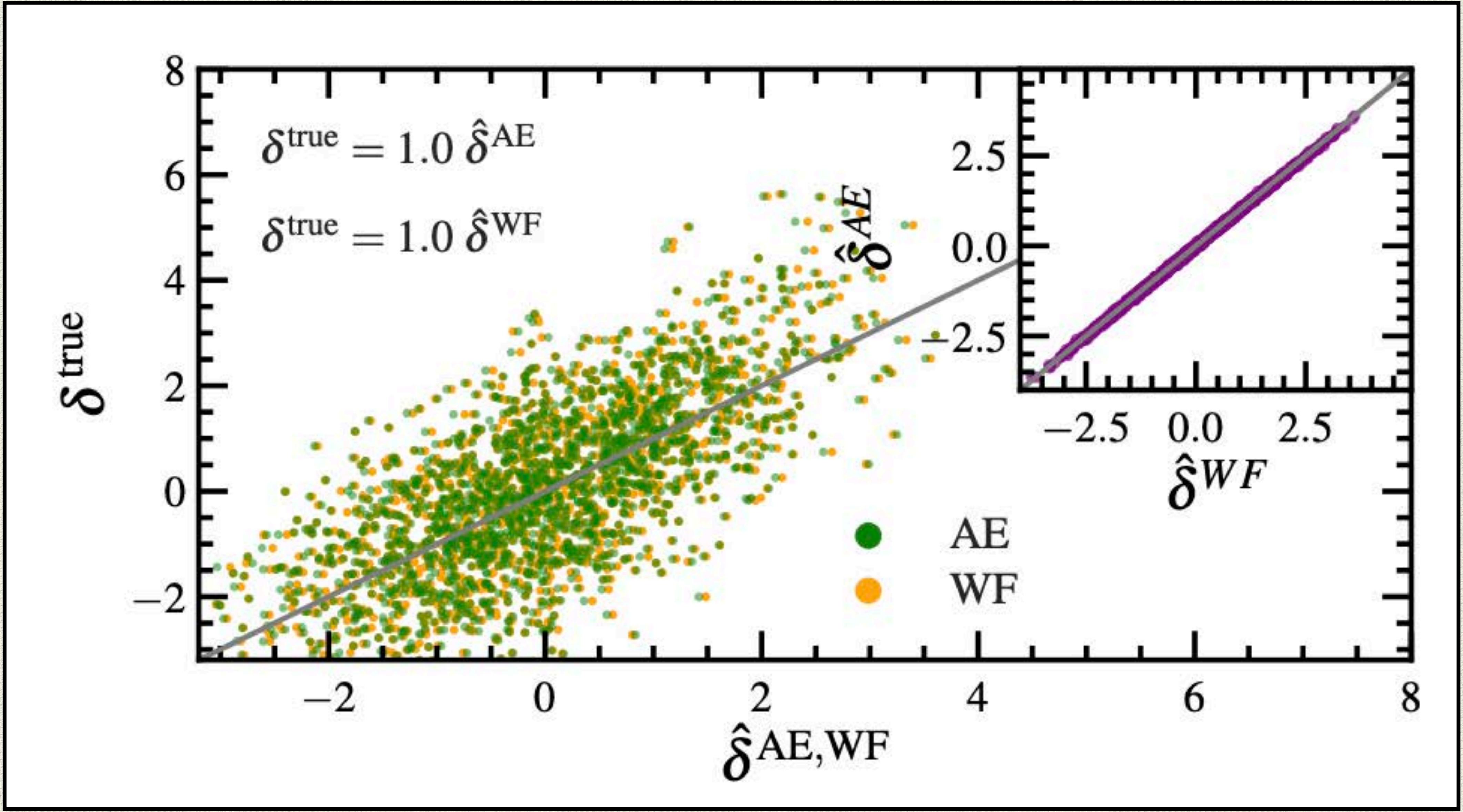
$$T^{WF} = \langle TI \rangle \langle II \rangle^{-1} I$$

Gaussian fields



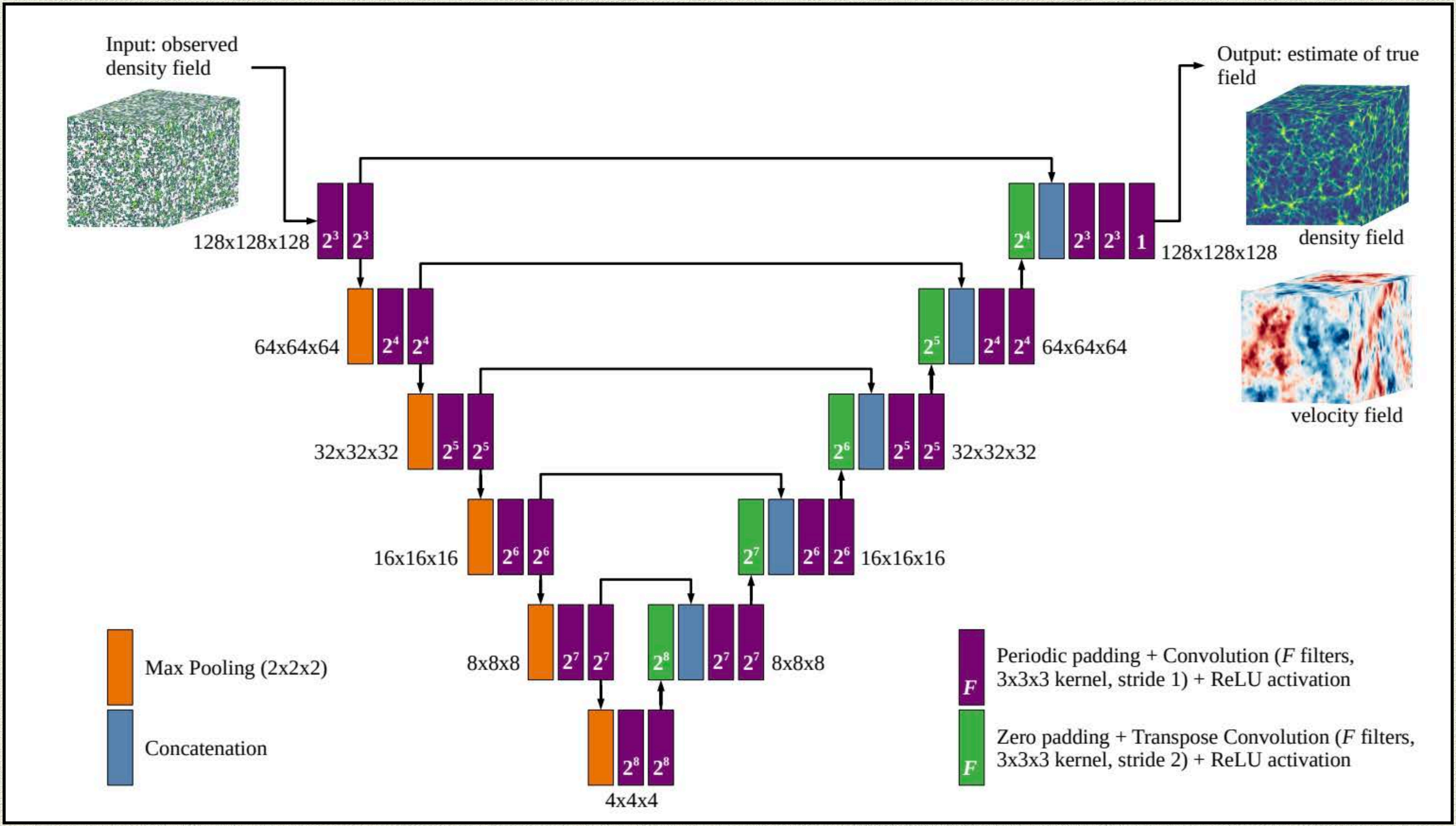
P.Ganeshaiah Veena, R.Lilow & A. Nusser 2024

Gaussian fields

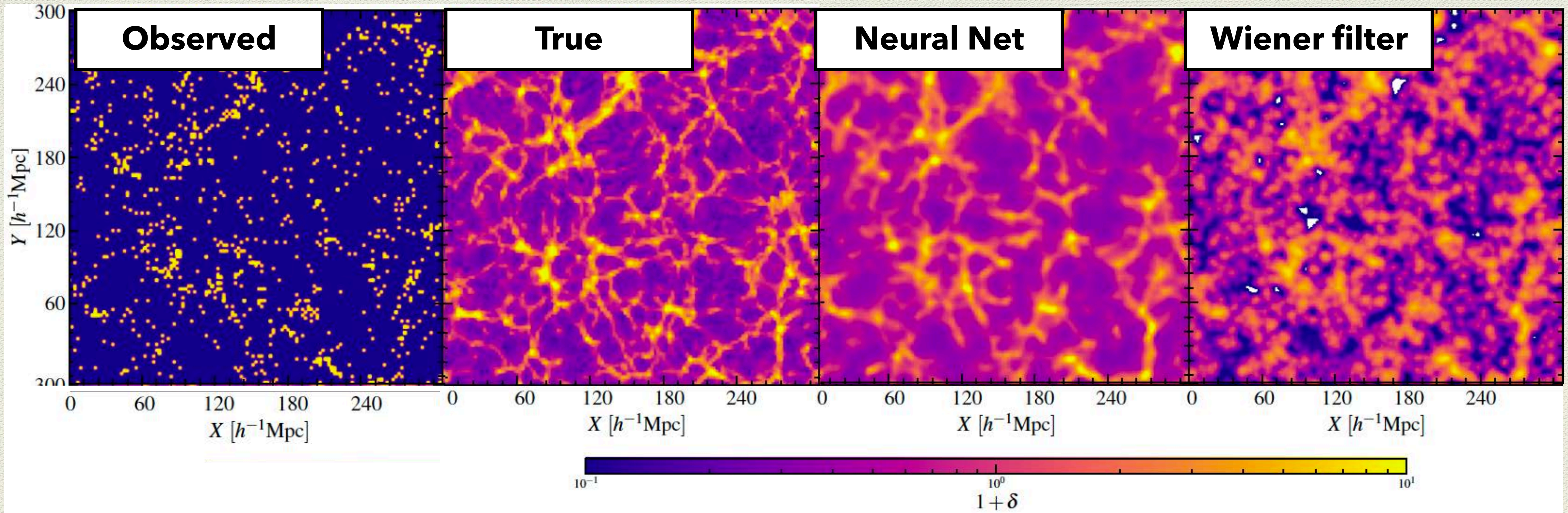


Wiener filter and Neural Network give the same result for Gaussian fields.

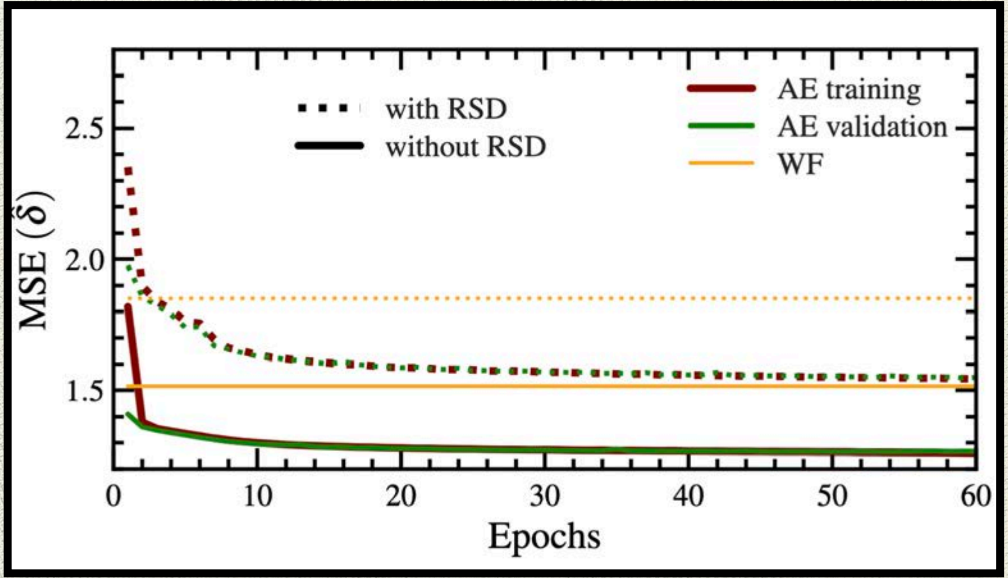
For 3D data, use convolutions: Autoencoder



Density field reconstructions in 3D

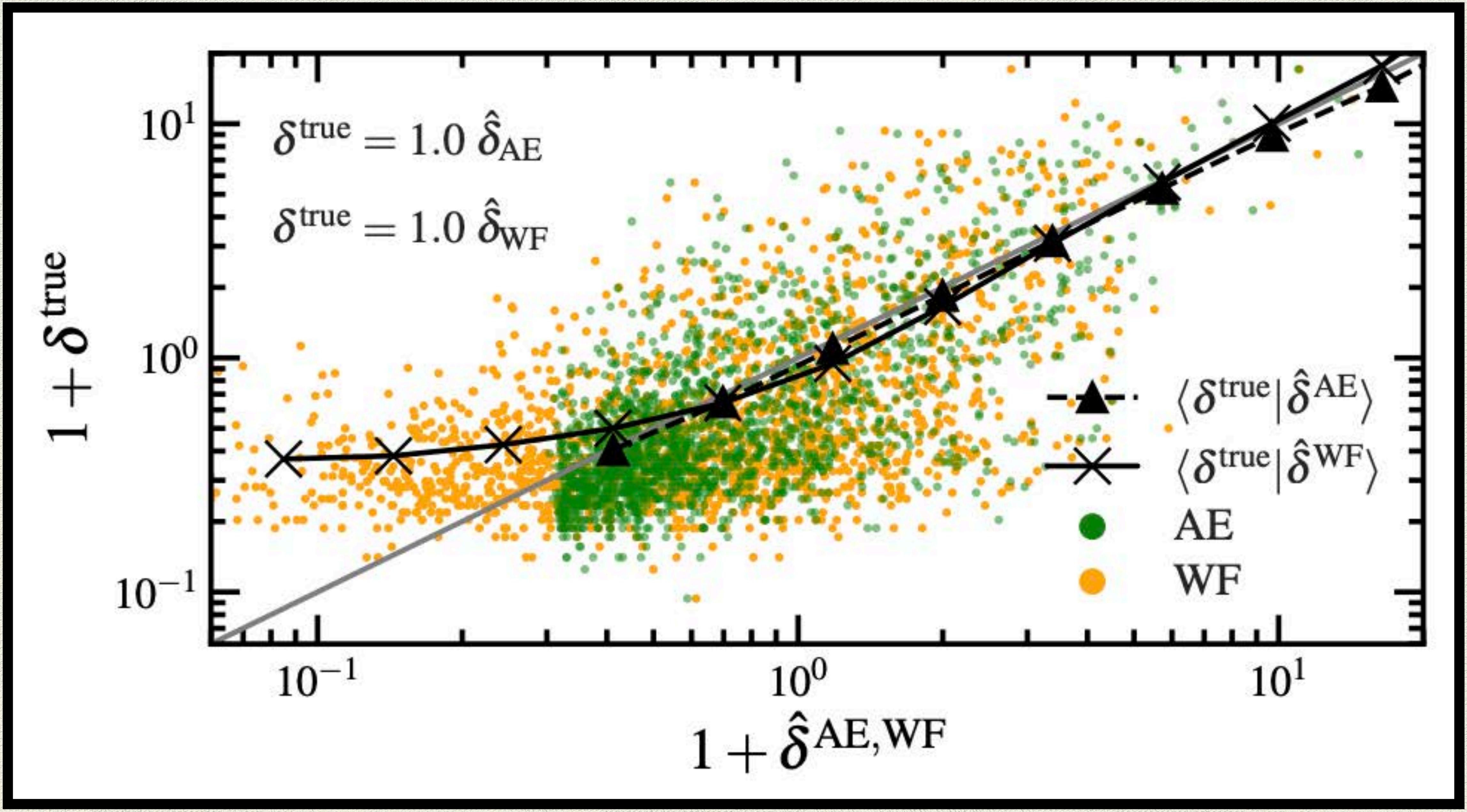


$$\delta(x) = \frac{\rho(x) - \bar{\rho}}{\bar{\rho}}$$



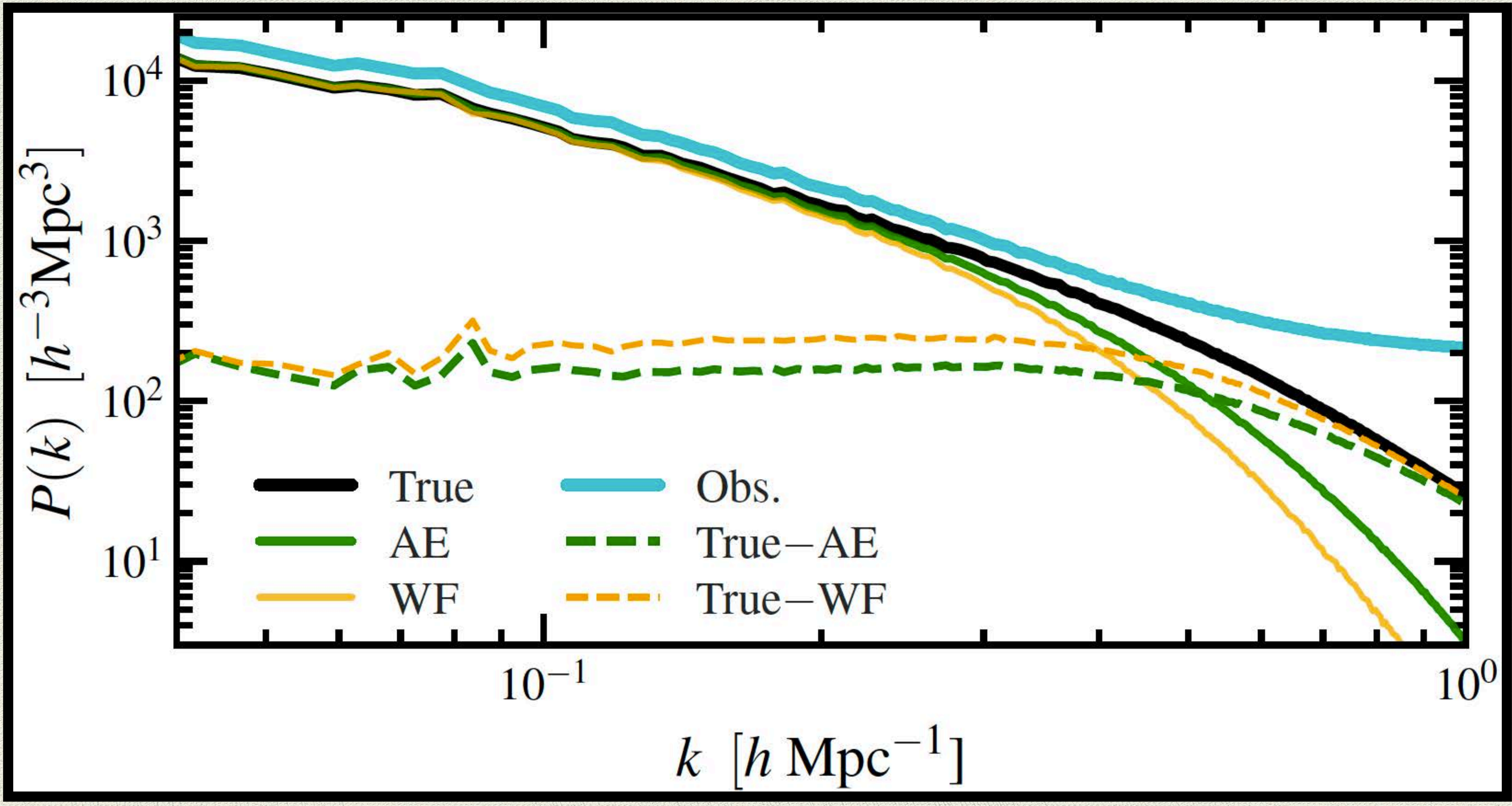
P.Ganeshaiah Veena, R.Lilow & A. Nusser 2023

Density field reconstructions



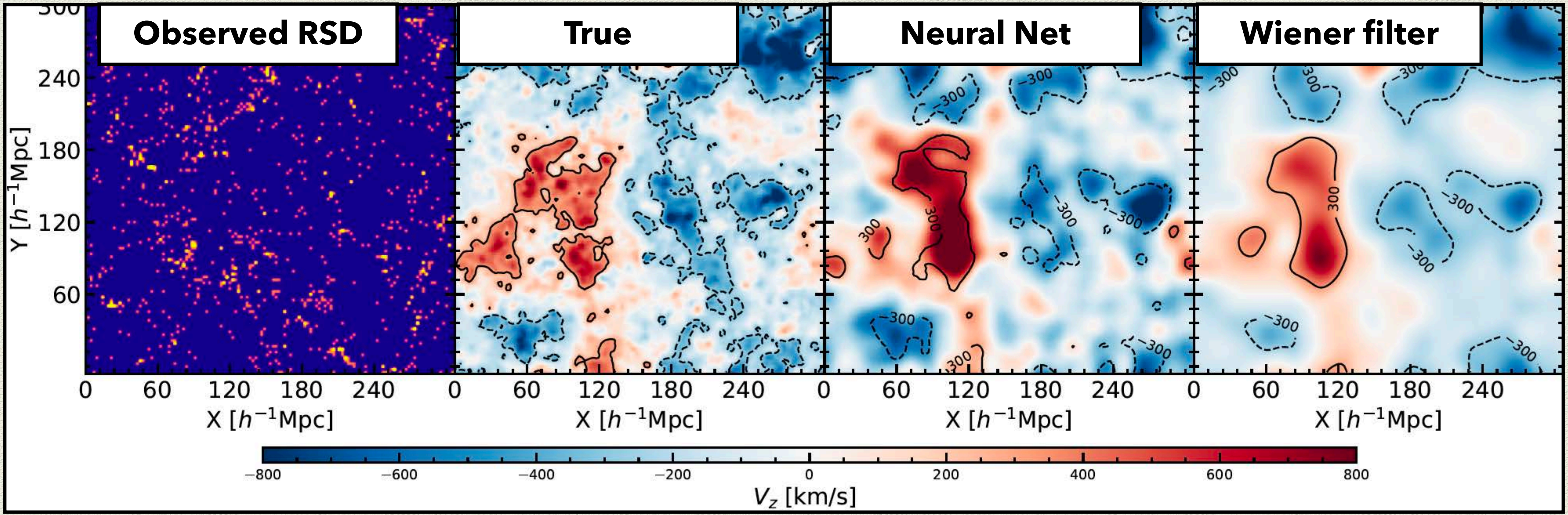
P.Ganeshaiah Veena, R.Lilow & A. Nusser 2023

Density field reconstructions - with RSD



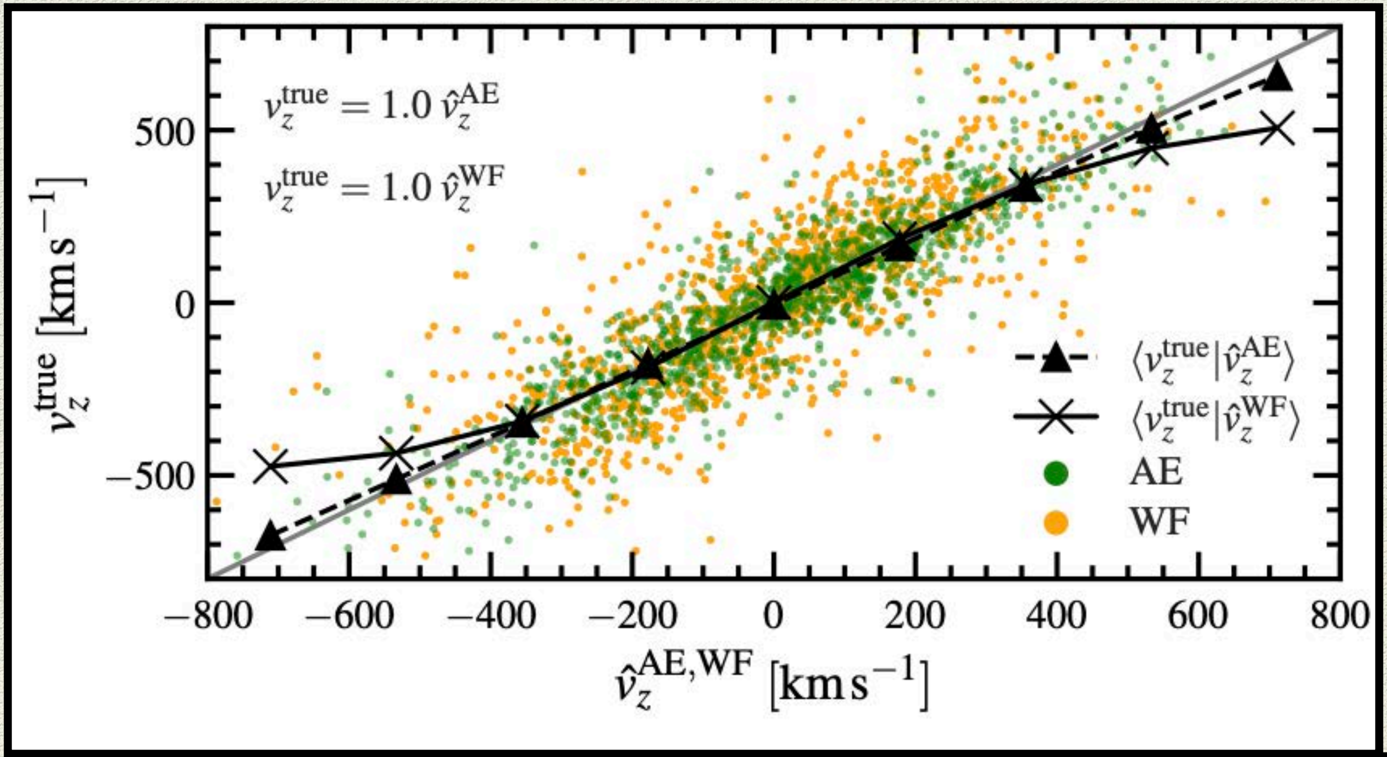
P.Ganeshaiah Veena, R.Lilow & A. Nusser 2023

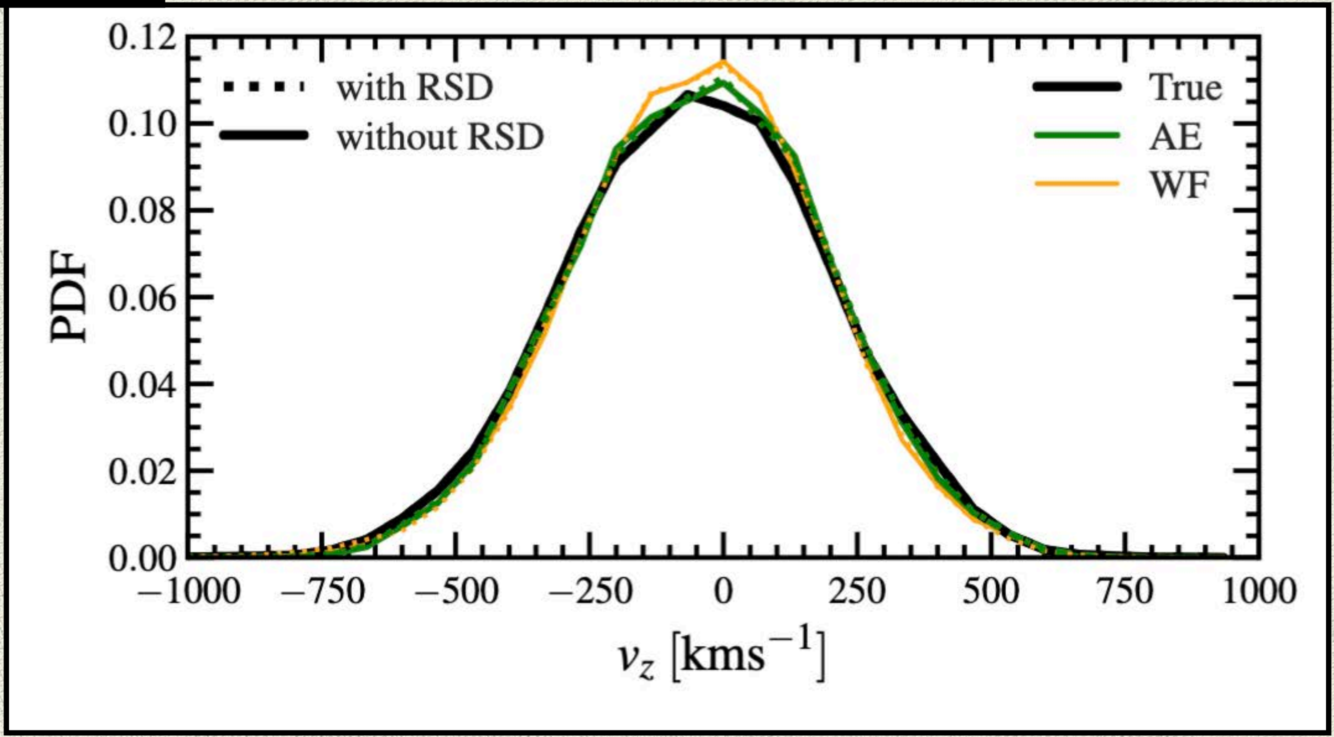
Velocity field reconstructions.



P.Ganeshaiah Veena, R.Lilow & A. Nusser 2023

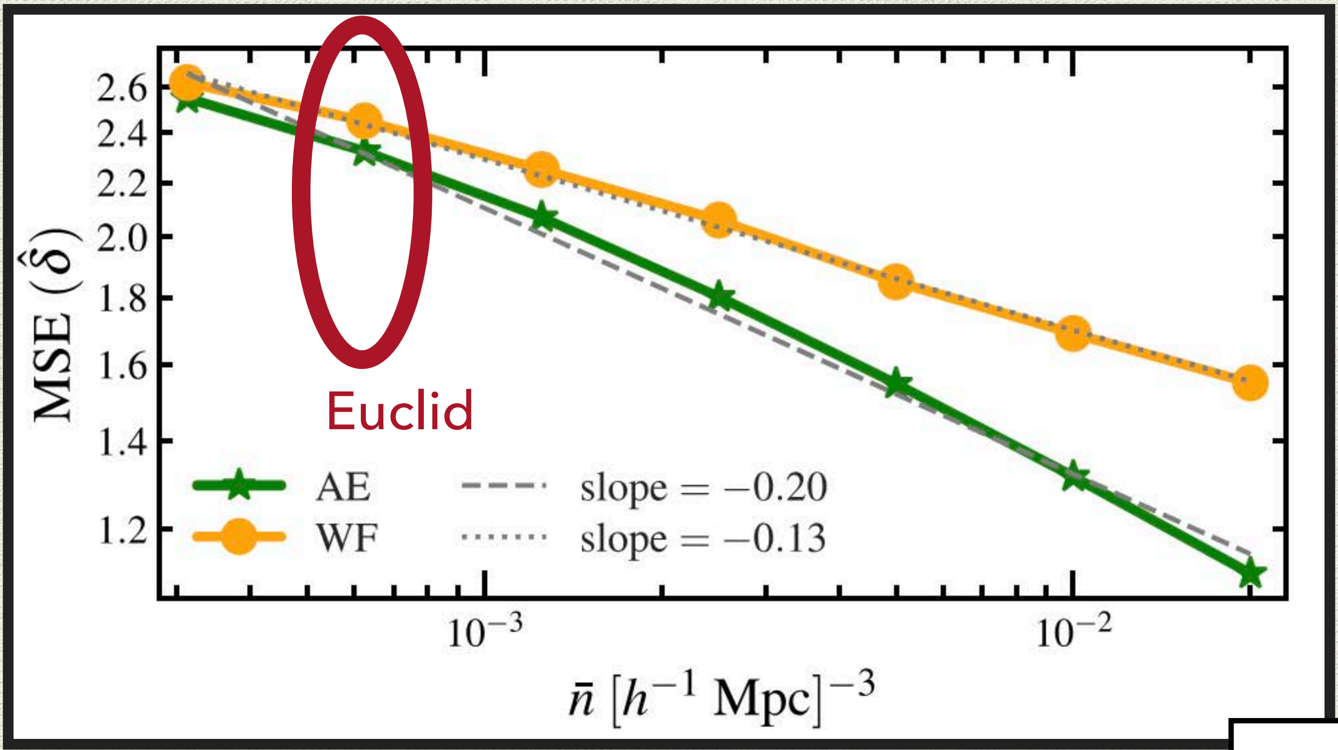
Velocity field reconstructions

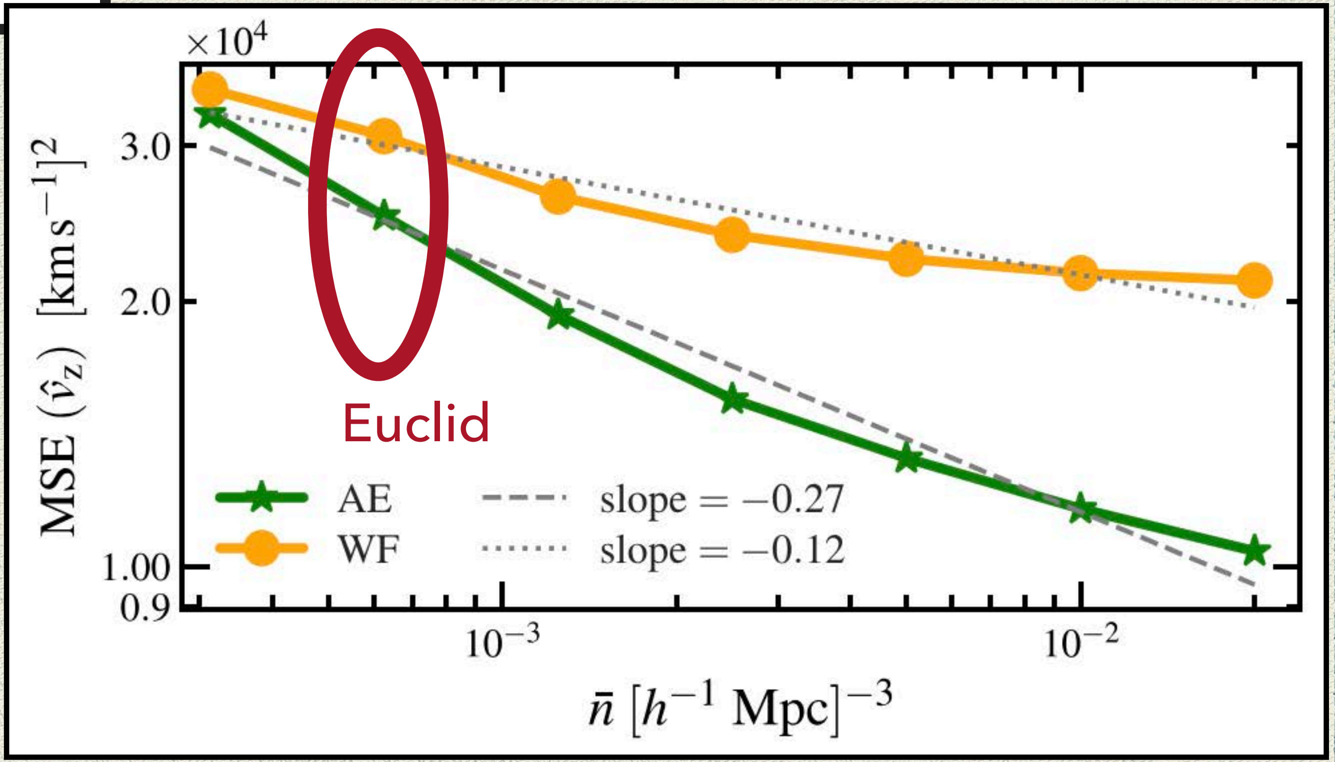




P.Ganeshaiah Veena, R.Lilow & A. Nusser 2023

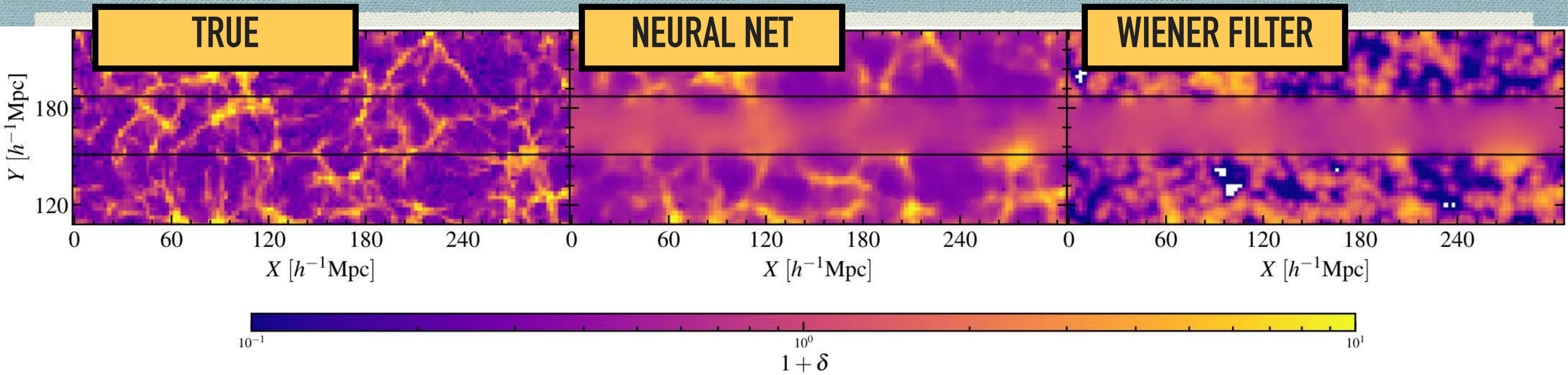
Reconstruction for different galaxy number densities

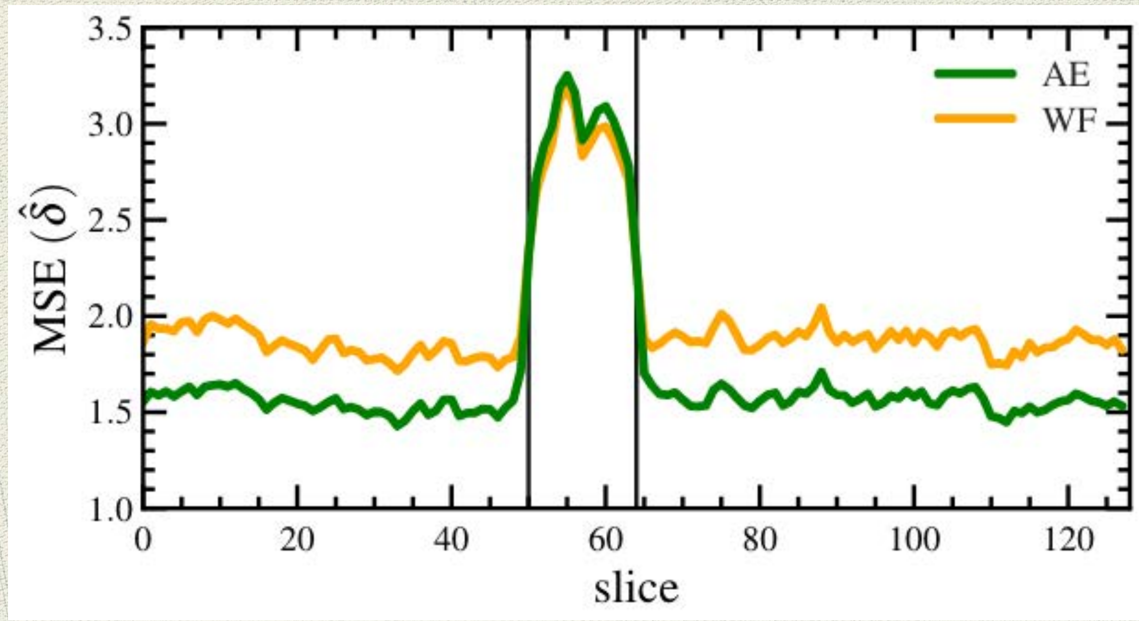




P.Ganeshaiah Veena, R.Lilow & A. Nusser 2023

Reconstruction in gaps

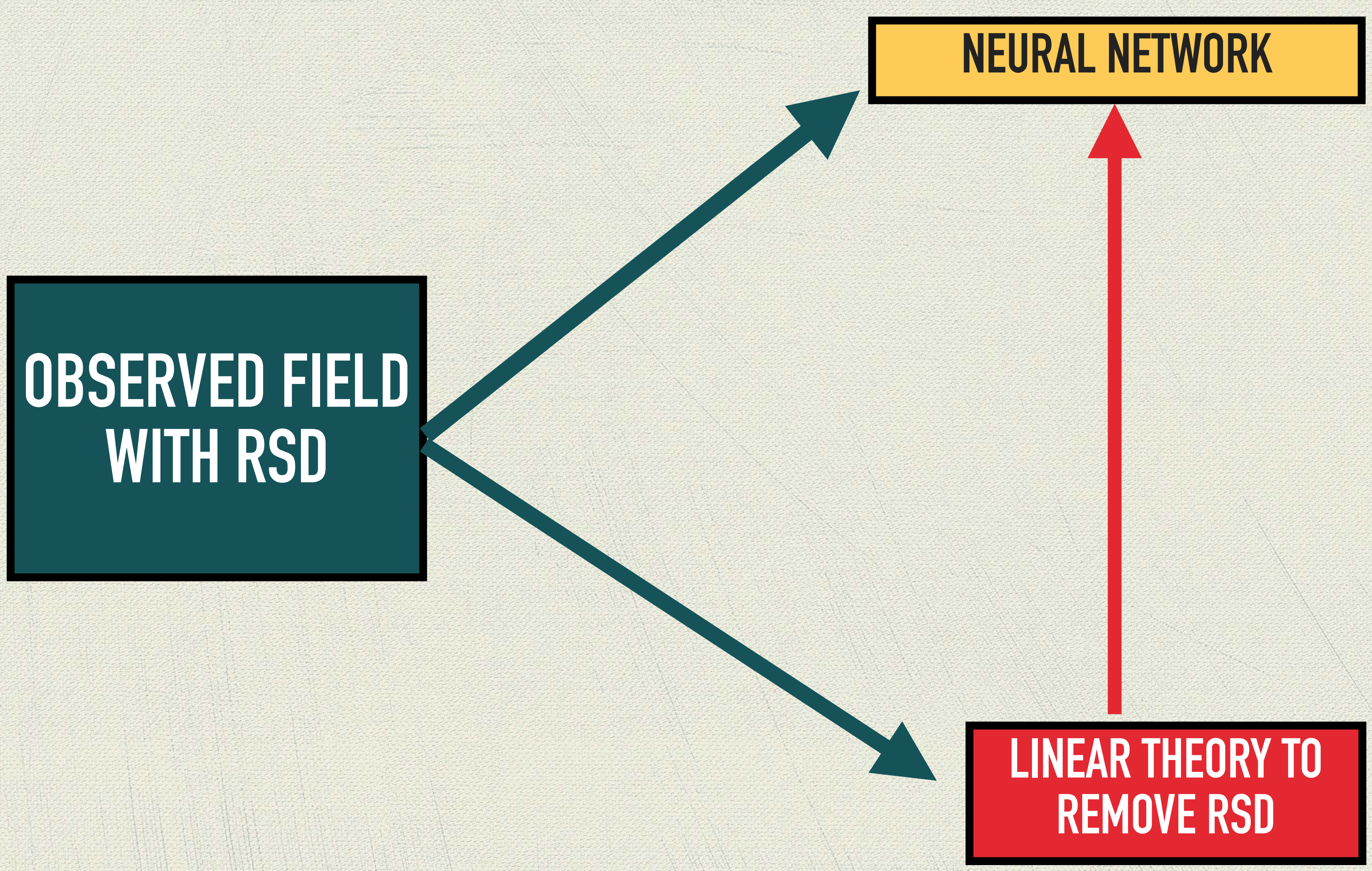




P.Ganeshaiah Veena, R.Lilow & A. Nusser 2023

Informed learning:
first Linear Theory and then Neural network

NN+Linear Theory for removing RSD



$$\nabla . \Psi + \beta \nabla \cdot (\Psi . \hat{r})\hat{r} = - \frac{\delta_s}{b}$$

E.Maragliano, P. Ganeshaiah Veena, G. Degni & E. Branchini in prep: expected to arrive soon!

Linear Theory

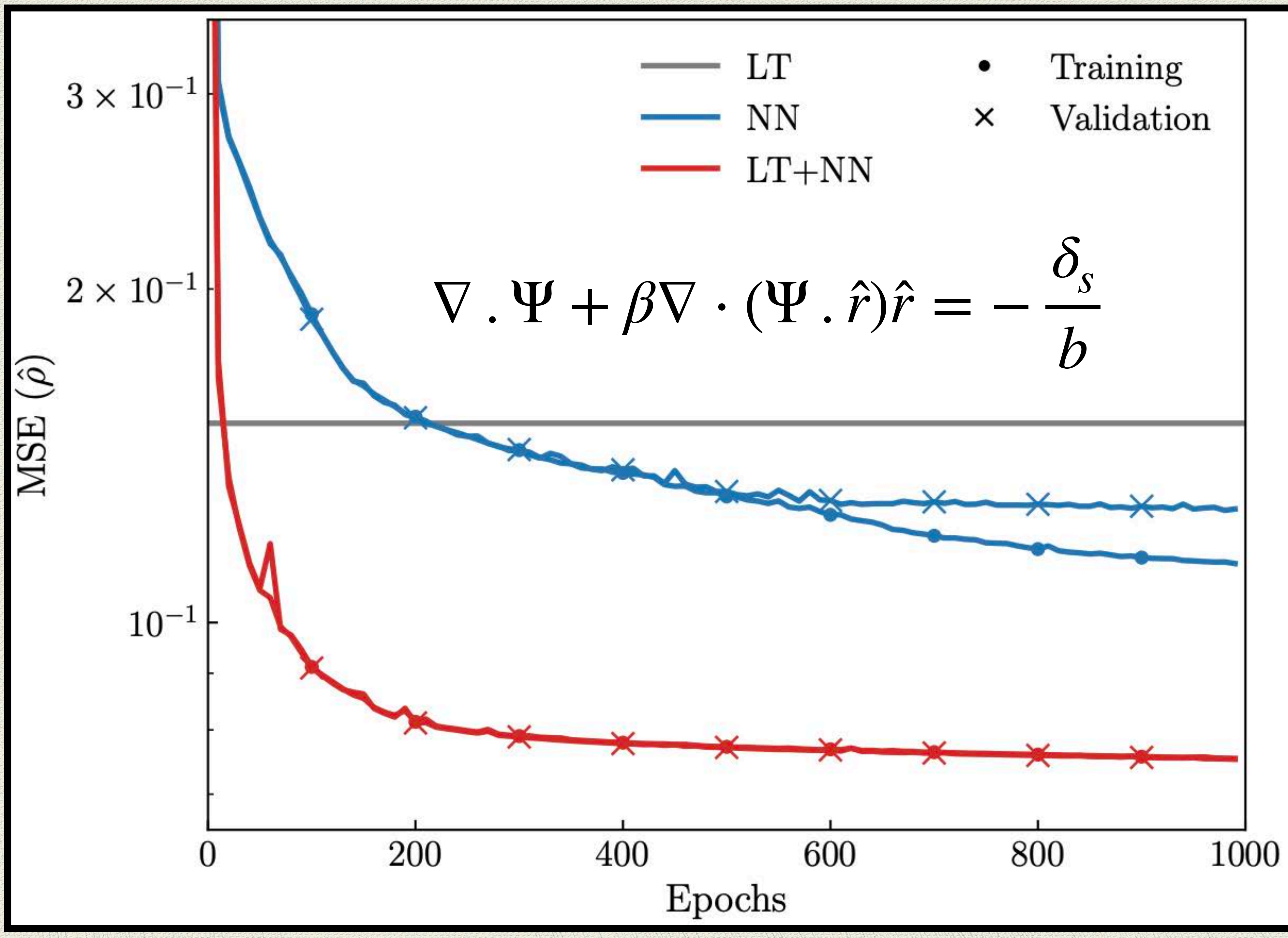
$$\mathbf{\Psi} = \mathbf{x} - \mathbf{q}$$

$$\nabla . \Psi + \beta \nabla \cdot (\Psi . \hat{r})\hat{r} = -\frac{\delta_s}{b}$$

$$\delta_{\rm s}(k) = \exp\left[-\frac{1}{2}\left(\frac{k}{R_s}\right)^2\right]\delta_{\rm obs}(k)$$

E.Maragliano, P. Ganeshaiah Veena, G. Degni & E. Branchini in prep: expected to arrive soon!

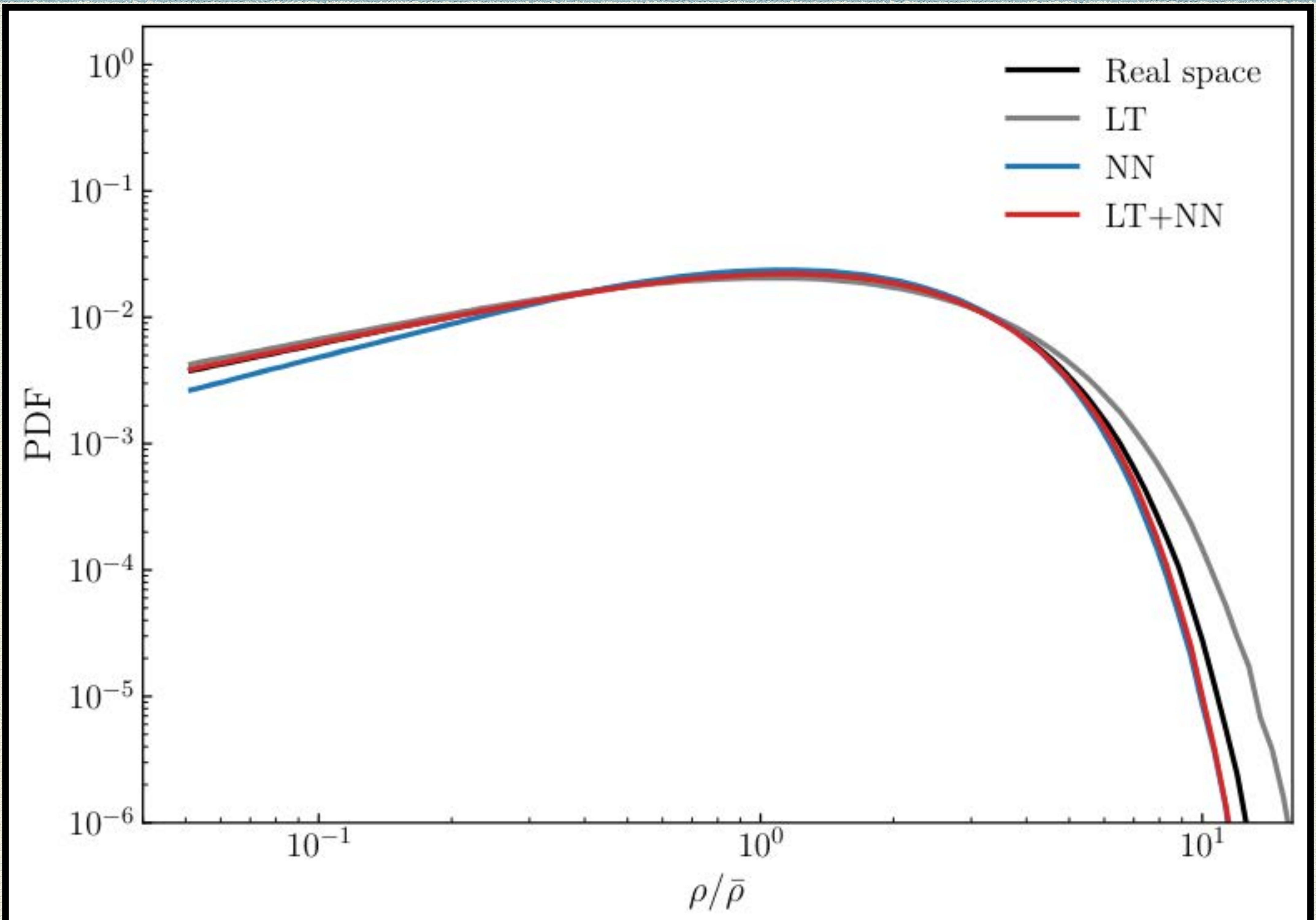
NN+Linear Theory for removing RSD



$$\nabla \, . \, \Psi + \beta \nabla \cdot (\Psi \, . \, \hat{r})\hat{r} = -\frac{\delta_s}{b}$$

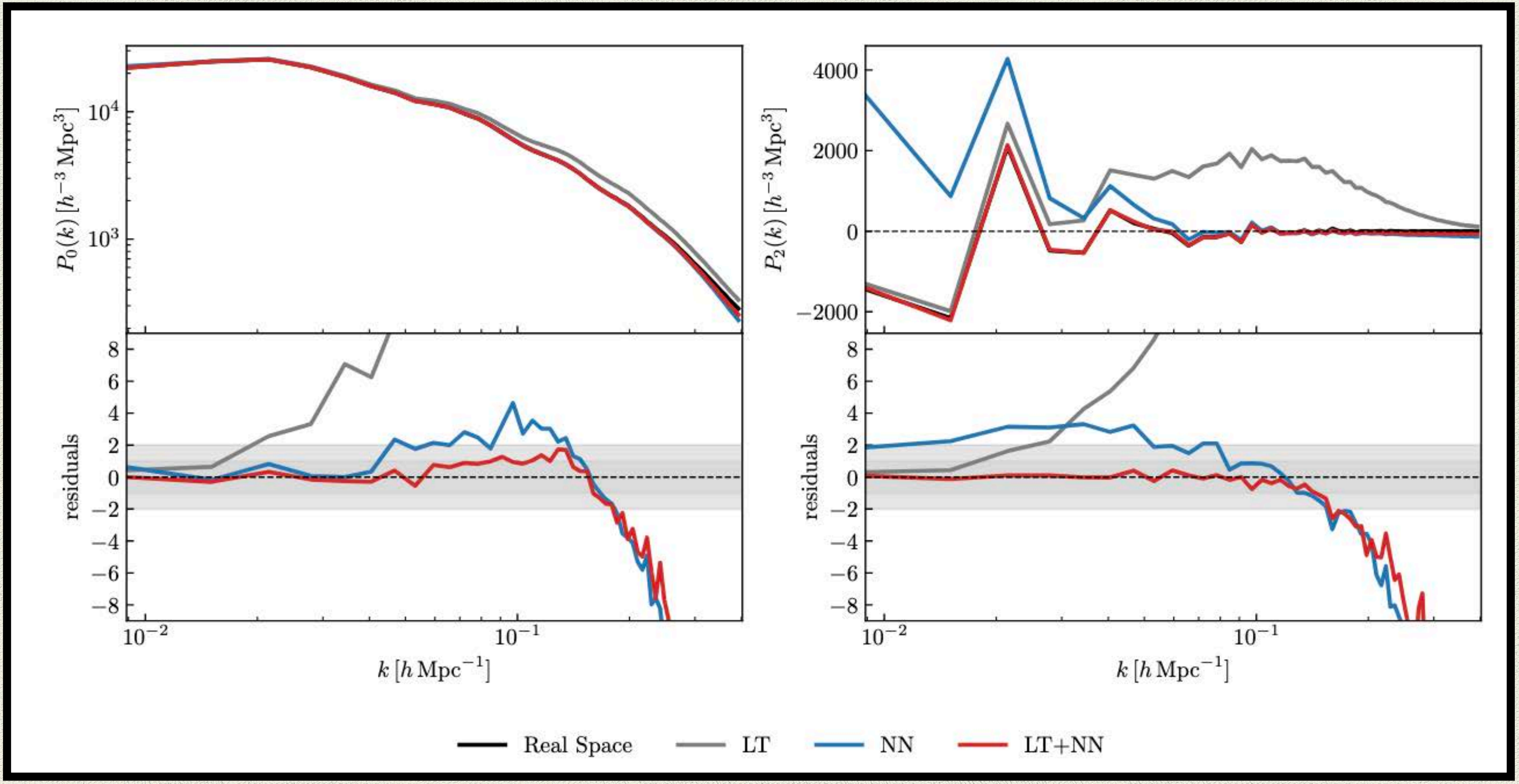
E.Maragliano, P. Ganeshaiah Veena, G. Degni & E. Branchini in prep: expected to arrive soon!

Density distribution



E.Maragliano, P. Ganeshaiah Veena, G. Degni & E. Branchini in prep: expected to arrive soon!

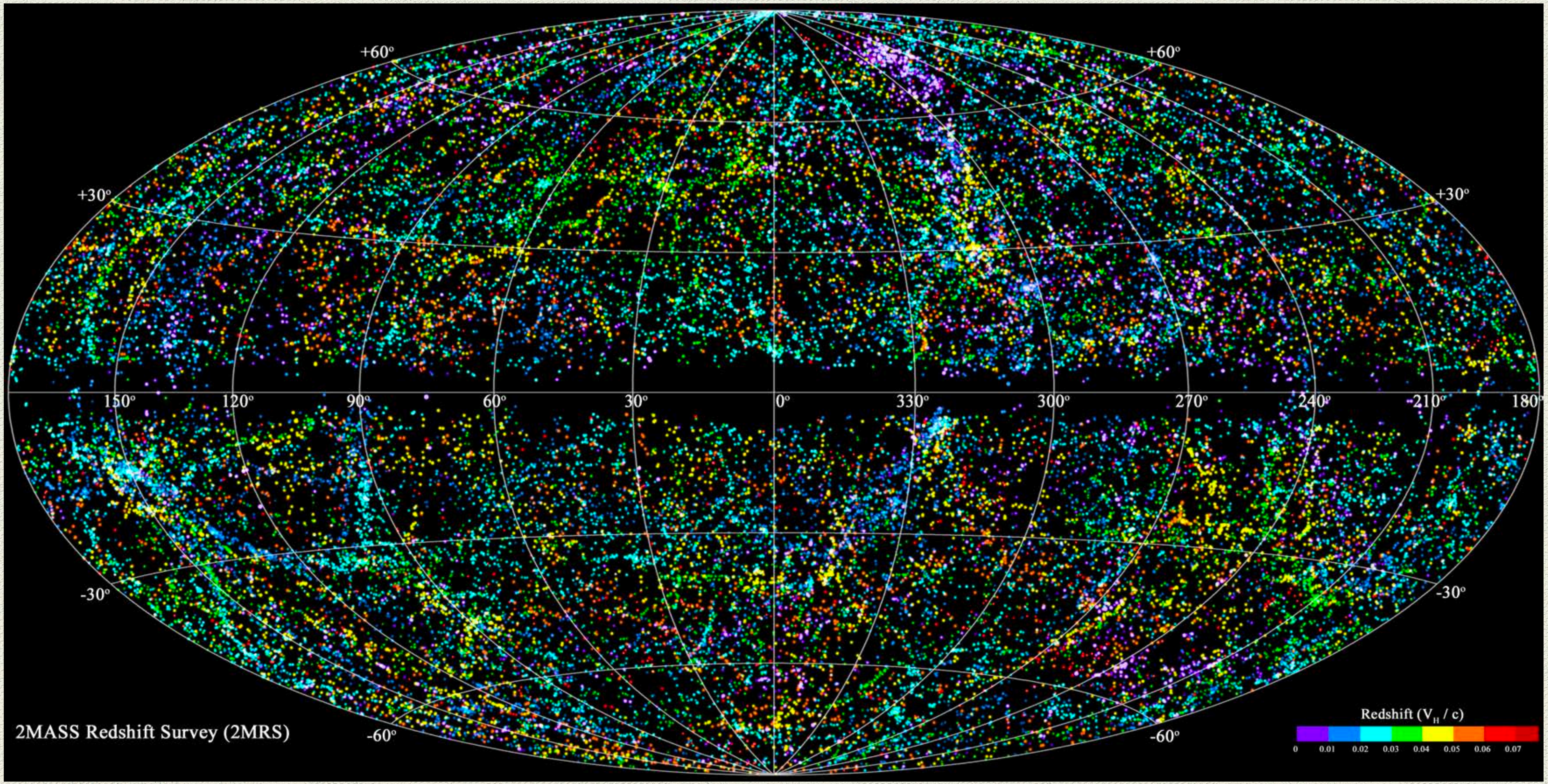
NN+Linear Theory for removing RSD



E.Maragliano, P. Ganeshaiah Veena, G. Degni & E. Branchini in prep: expected to arrive soon!

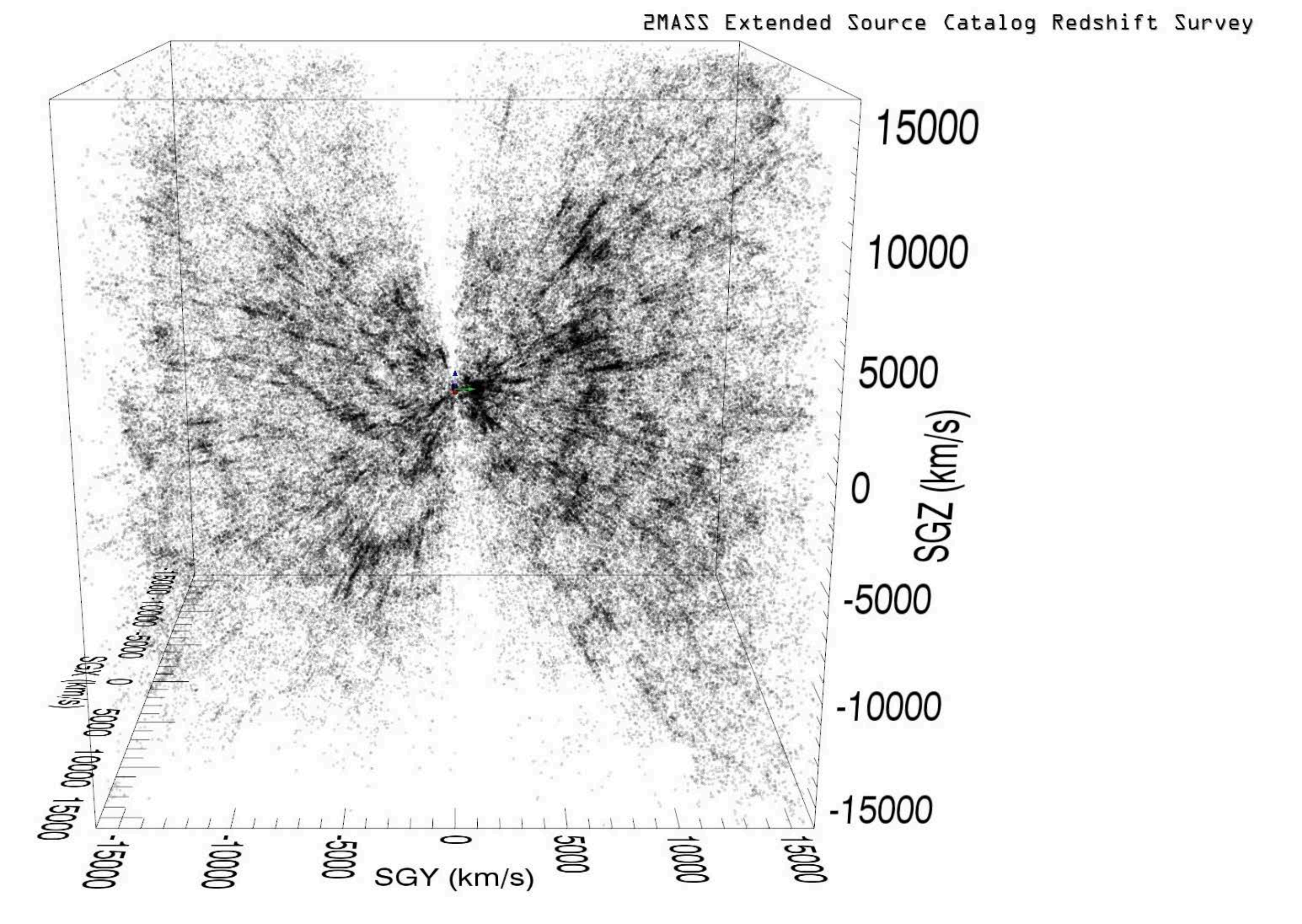
Apply this to real data: 2MRS

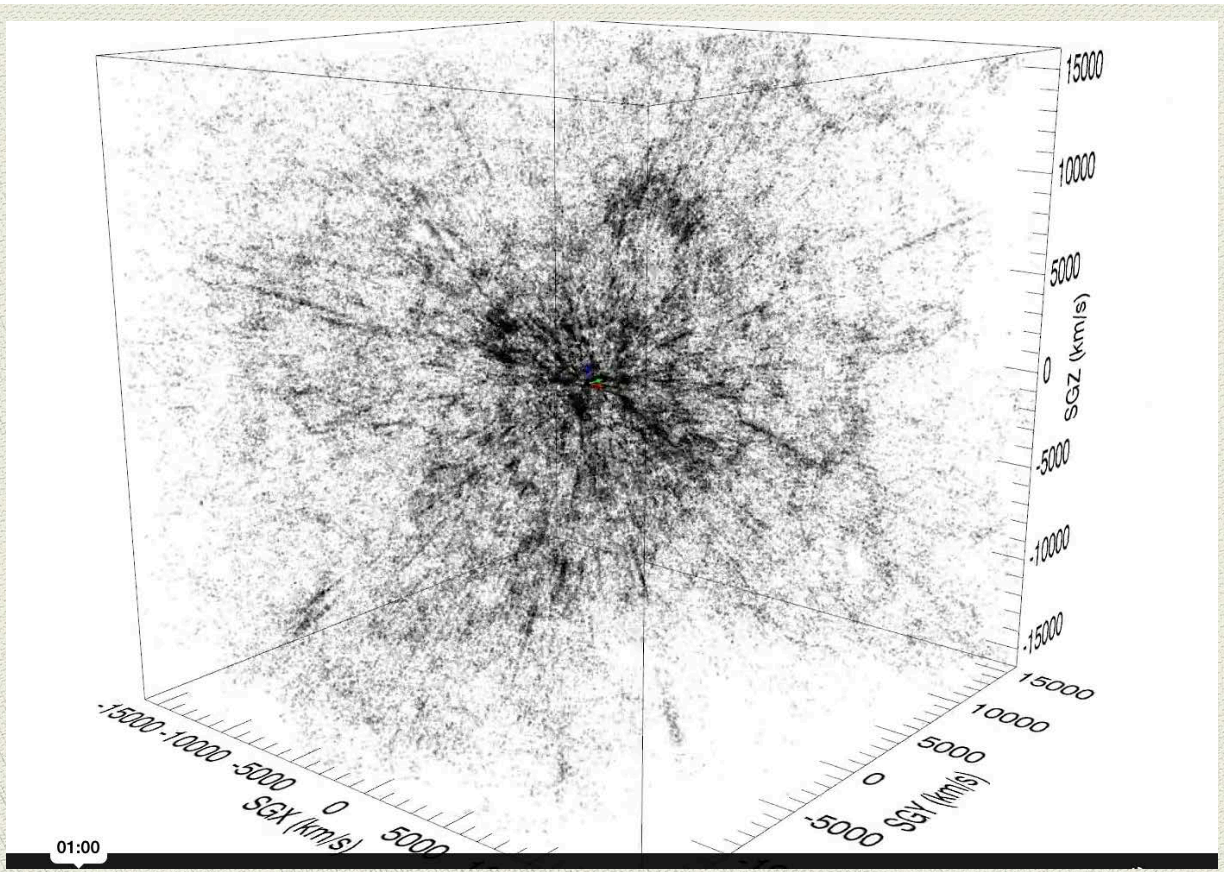
Create a 3D map of the local Universe using real galaxies.



2MRS data

- Flux limited survey, Ks-band magnitude of Ks ≤ 11.75
- Sky positions and spectroscopic redshifts for 44, 572 galaxies
- Survey footprint covers 91% of the sky, only missing the Zone of Avoidance (ZoA)
- For our work: spherical volume with a radius of 200 Mpc/h, encompassing 98% of all the galaxies





R.Lilow, P.Ganeshaiah Veena, & A. Nusser 2024

Mocks from Quijote simulations

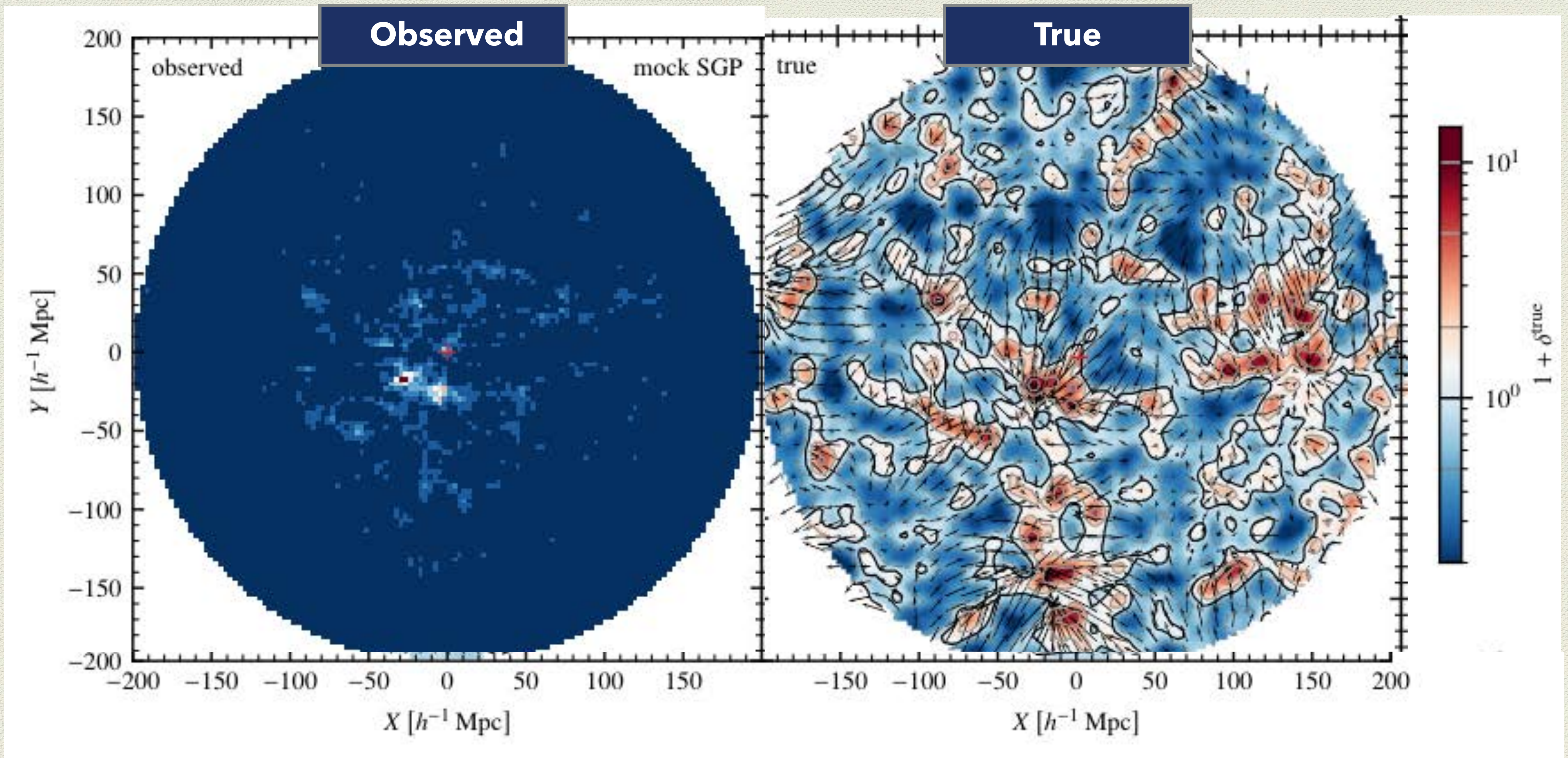
- Mocks include:
  - survey selection function
  - bias
  - redshift space distortions
  - zone of avoidance.
- 6400 mocks: 5760 for training and 640 for test+validation.
- Loss is MSE scaled with the selection function:
  - $\text{Loss}\left(\hat{\delta}^{\text{NN}}\right) = \frac{1}{M_{\text{train}} M_{\text{grid}}} \sum_{\alpha=1}^{M_{\text{train}}} \sum_{j=1}^{M_{\text{grid}}} \phi(r_j) \left(\delta_j^{\text{true},\alpha} - \hat{\delta}_j^{\text{NN},\alpha}\right)^2$

Loss functions

- $$\mathrm{Loss}\left(\hat{\delta}^{\mathrm{NN}}\right) = \frac{1}{M_{\mathrm{train}} M_{\mathrm{grid}}} \sum_{\alpha=1}^{M_{\mathrm{train}}} \sum_{j=1}^{M_{\mathrm{grid}}} \phi(r_j) \left(\delta_j^{\mathrm{true},\alpha} - \hat{\delta}_j^{\mathrm{NN},\alpha}\right)^2$$

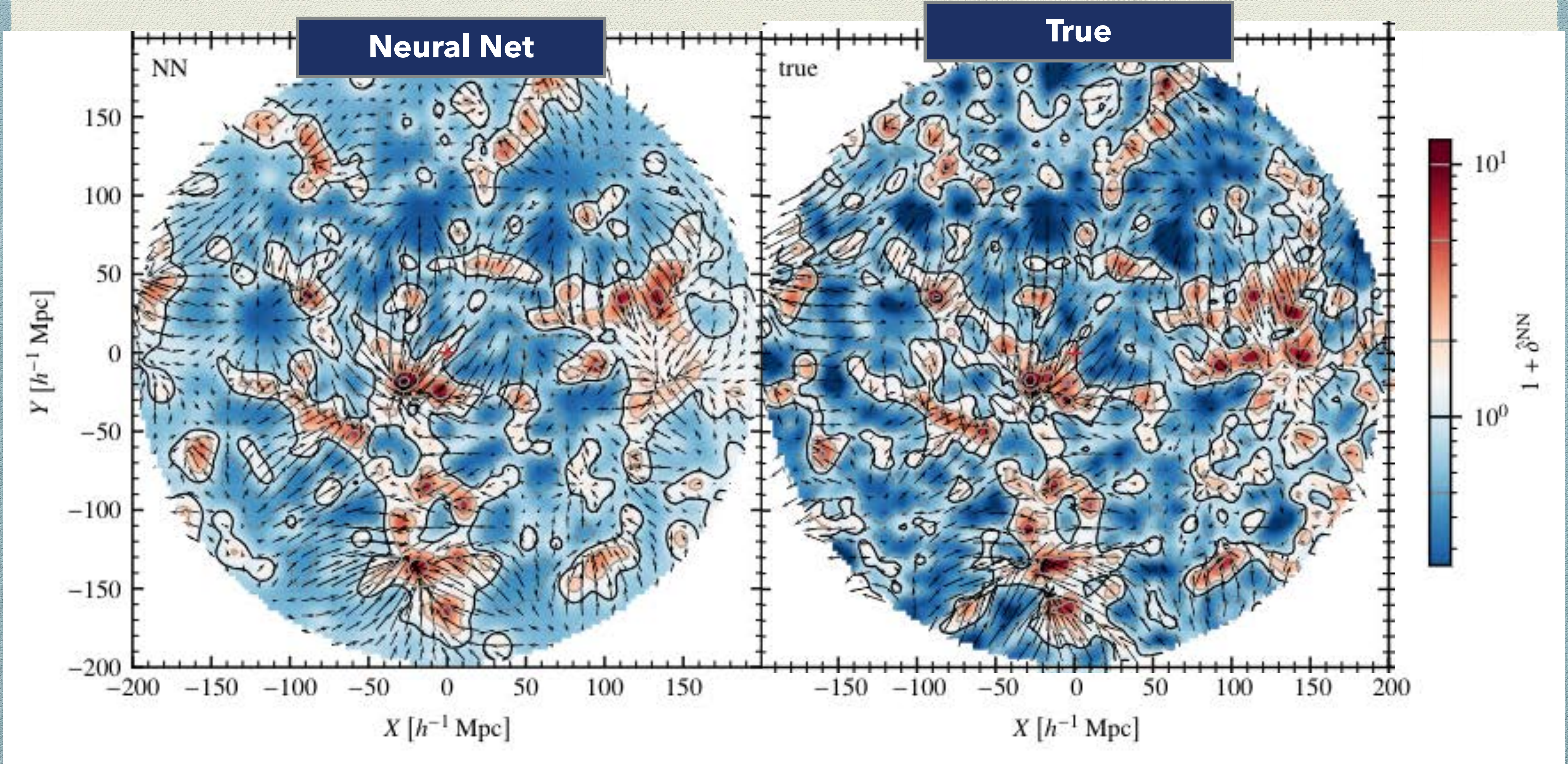
$$\mathrm{Loss}\left(\hat{\Psi}^{\mathrm{NN}}\right) = \frac{1}{M_{\mathrm{train}} M_{\mathrm{grid}}} \sum_{\alpha=1}^{M_{\mathrm{train}}} \sum_{j=1}^{M_{\mathrm{grid}}} \frac{\phi(r_j)}{r_j} \left(v_j^{\mathrm{true},\alpha} - \nabla\hat{\Psi}_j^{\mathrm{NN},\alpha}\right)^2$$

Observed fields- Quijote mocks



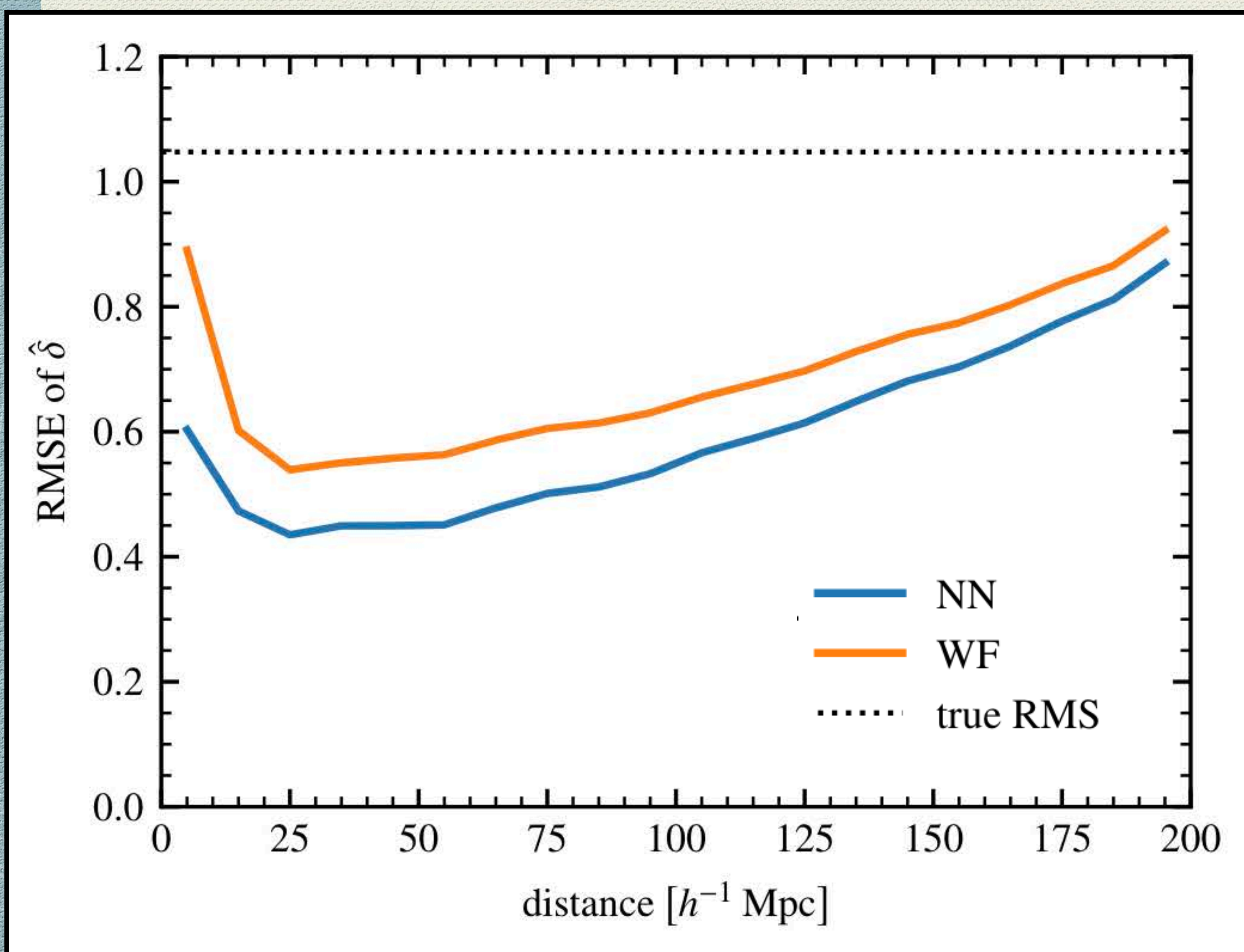
R.Lilow, P.Ganeshaiah Veena, & A. Nusser 2024

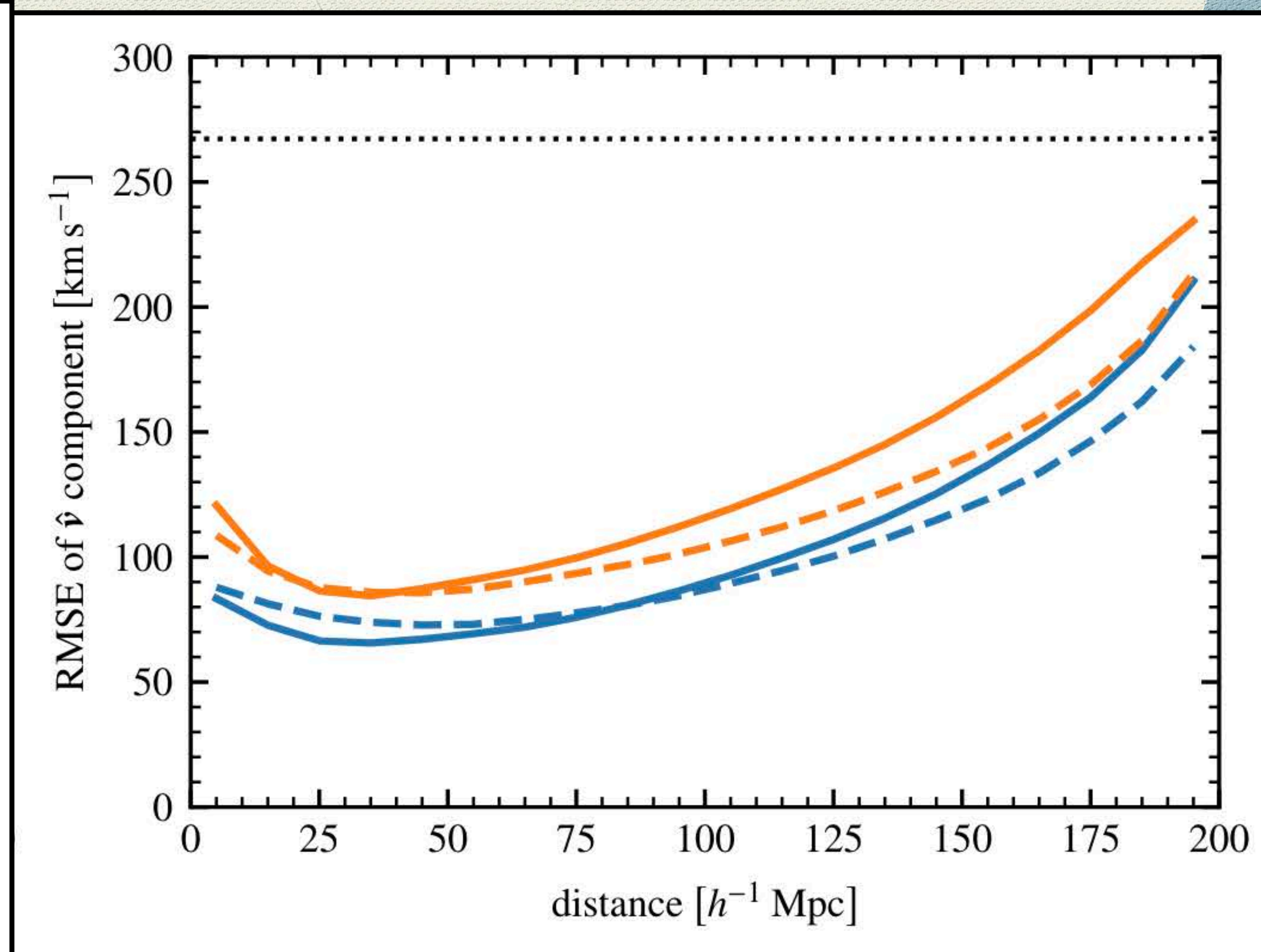
Reconstructed fields - Quijote mocks



R.Lilow, P.Ganeshaiah Veena, & A. Nusser 2024

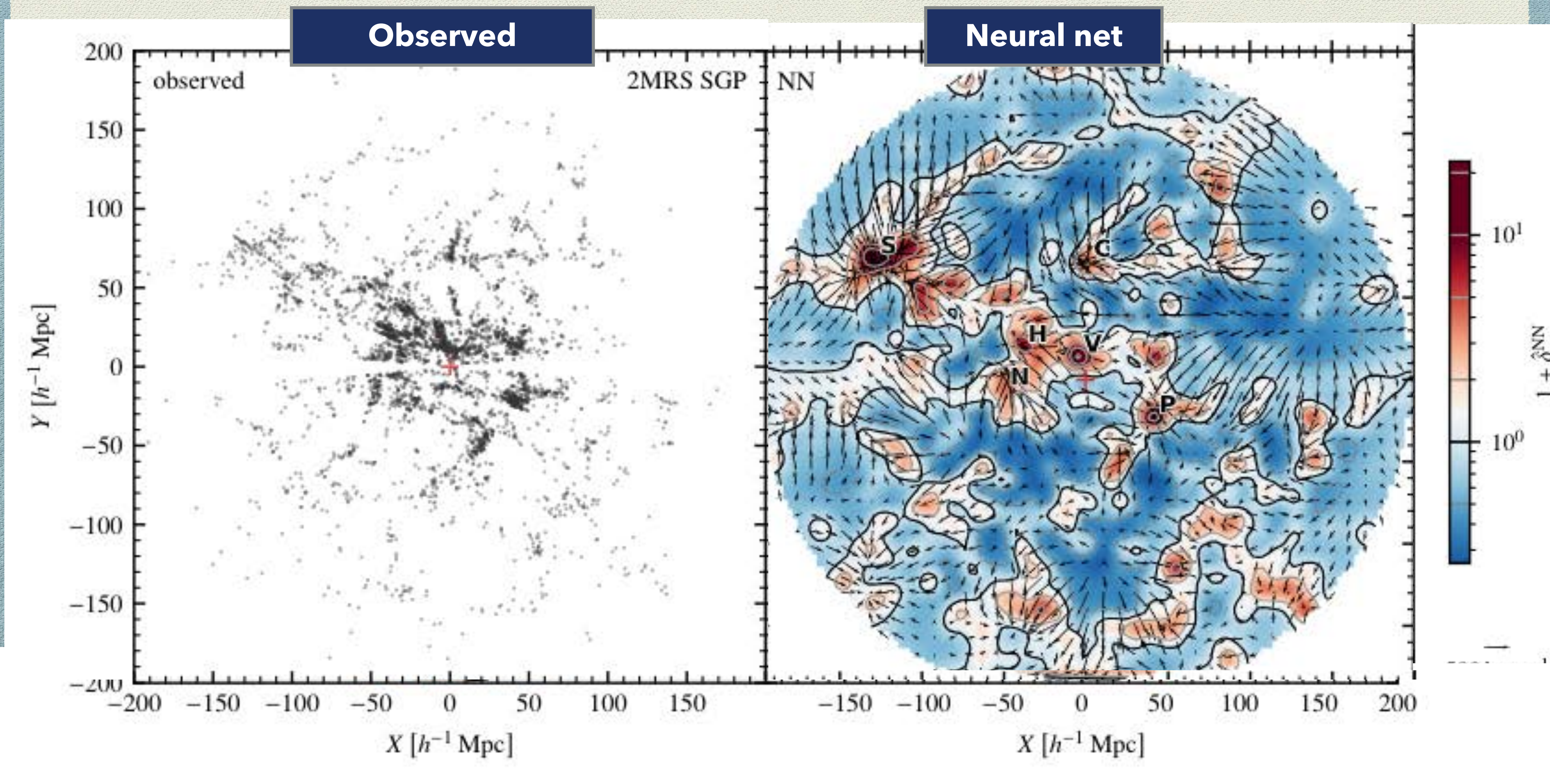
MSE - Quijote mocks





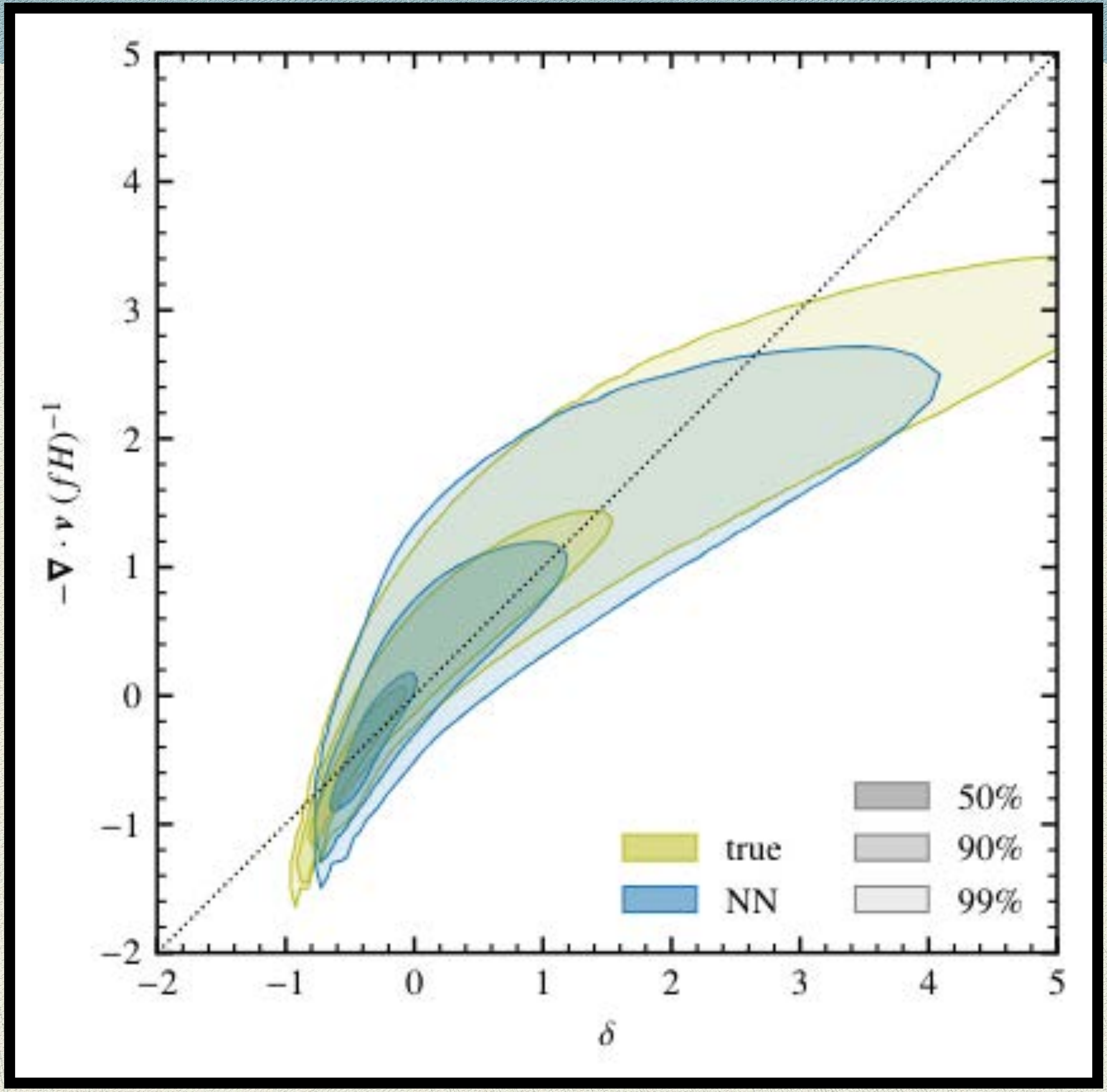
R.Lilow, P.Ganeshaiah Veena, & A. Nusser 2024

2MRS reconstruction



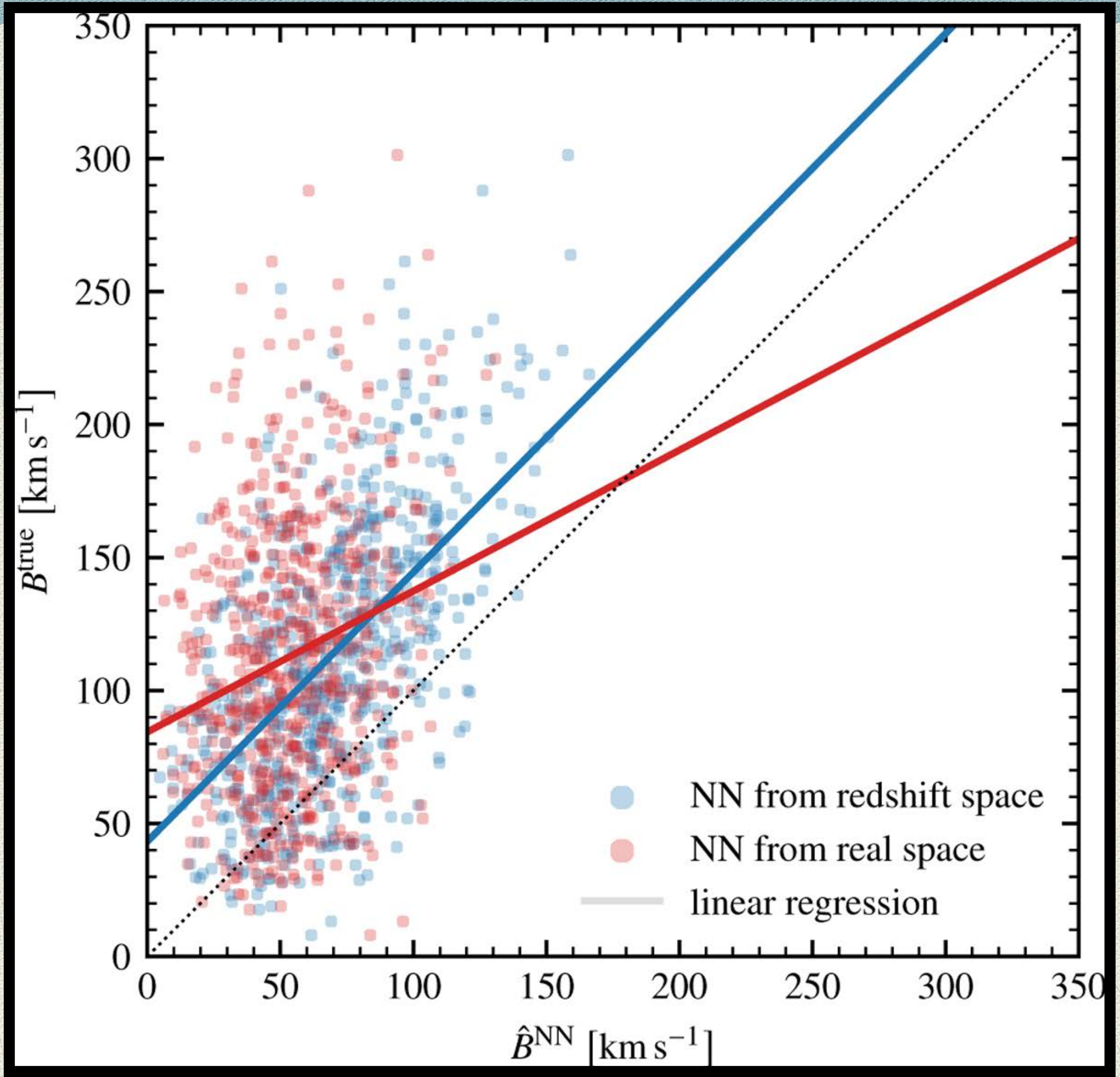
R.Lilow, P.Ganeshaiah Veena, & A. Nusser 2024

Density velocity relation- Quijote mocks



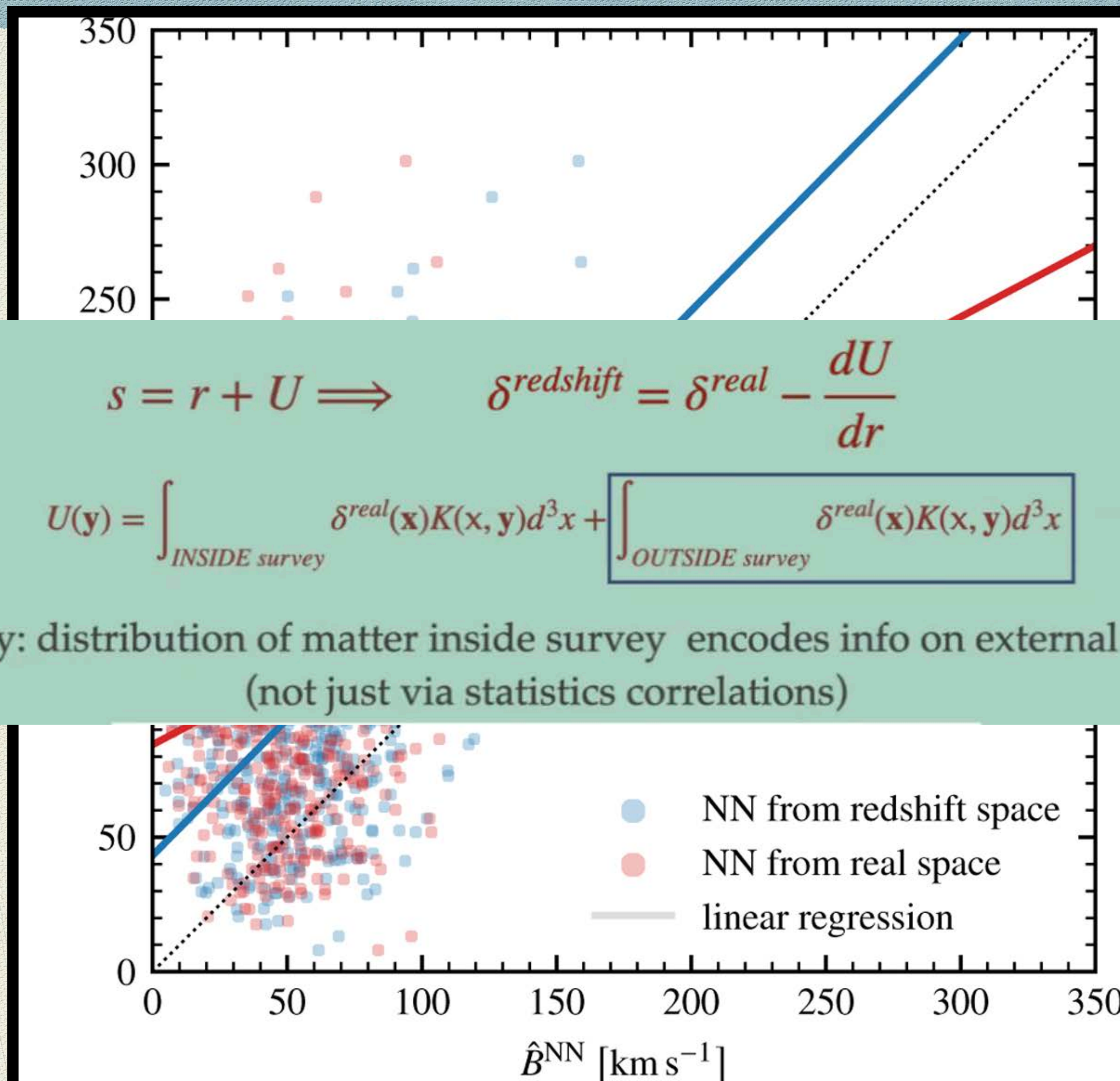
R.Lilow, P.Ganeshaiah Veena, & A. Nusser 2024

Probing "super survey" scales



R.Lilow, P.Ganeshaiah Veena, & A. Nusser 2024

Probing "super survey" scales



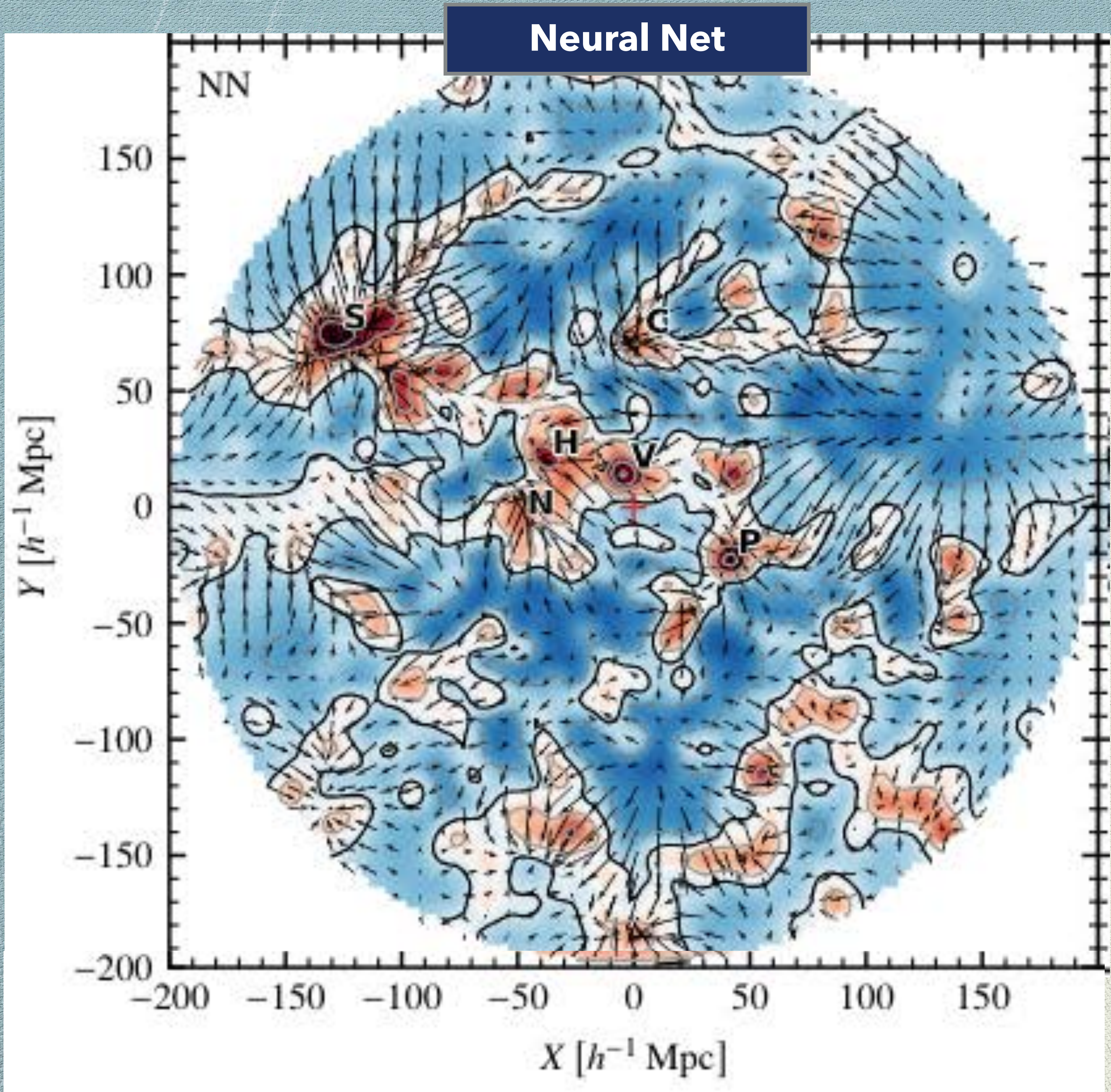
$$s = r + U \Longrightarrow \qquad \delta^{redshift} = \delta^{real} - \frac{dU}{dr}$$

$$U(\mathbf{y}) = \int_{INSIDE\ survey} \delta^{real}(\mathbf{x}) K(\mathrm{x}, \mathbf{y}) d^3x + \boxed{\int_{OUTSIDE\ survey} \delta^{real}(\mathbf{x}) K(\mathrm{x}, \mathbf{y}) d^3x}$$

Corollary: distribution of matter inside survey encodes info on external matter
(not just via statistics correlations)

R.Lilow, P.Ganeshaiah Veena, & A. Nusser 2024

2MRS reconstructions - cosmography



Clusters/ Superclusters

Shapley

Coma

Hydra-Centaurus

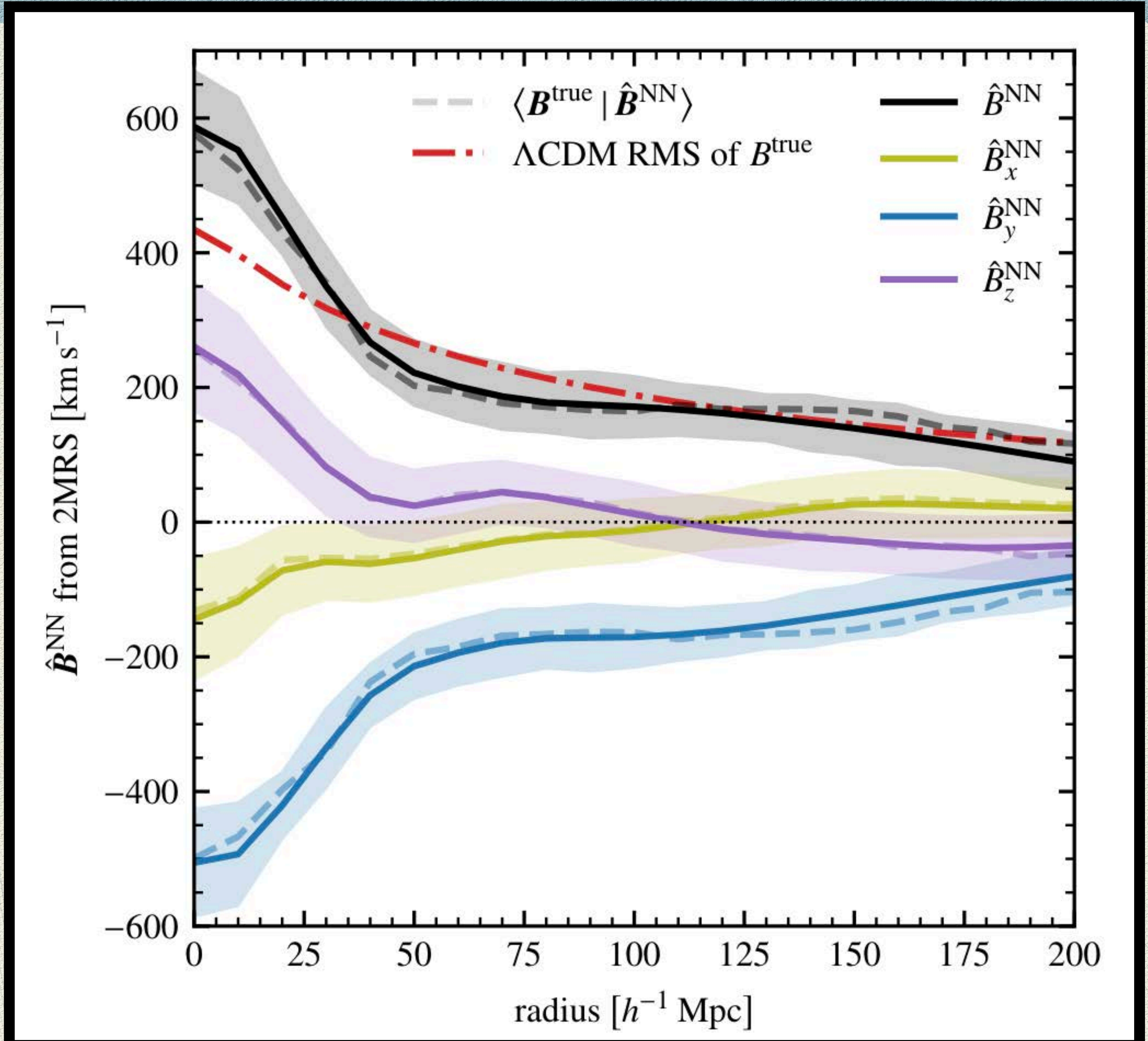
Virgo

Norma

Perseus-Pisces

R.Lilow, P.Ganeshaiah Veena, & A. Nusser 2024

Bulk velocity from NN



Thank you!