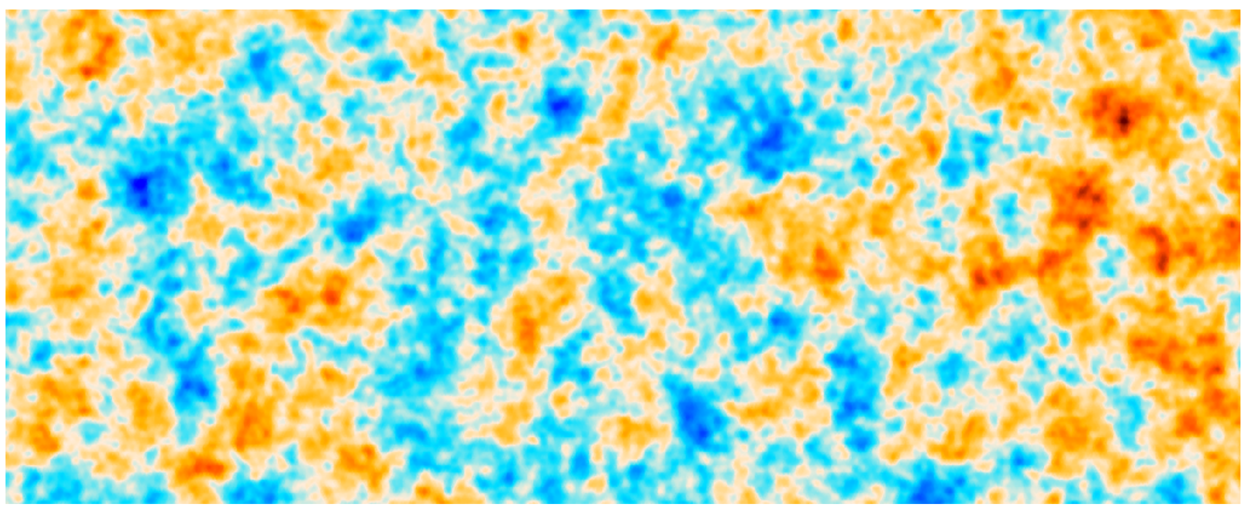
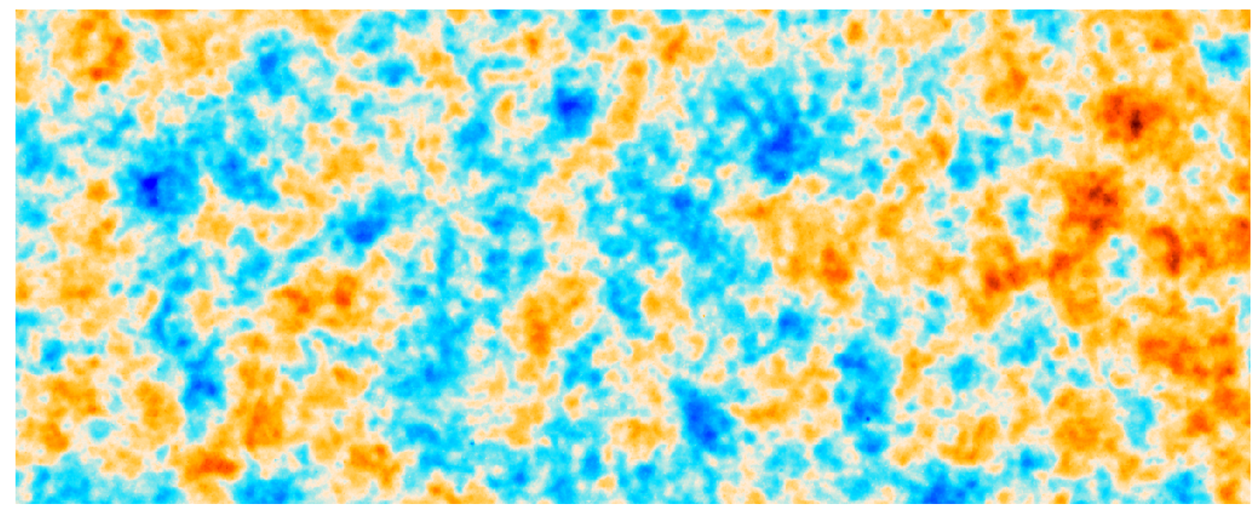
Mocking the Universe
& Machine Learning Approaches to Large-Scale Structure

George Stein



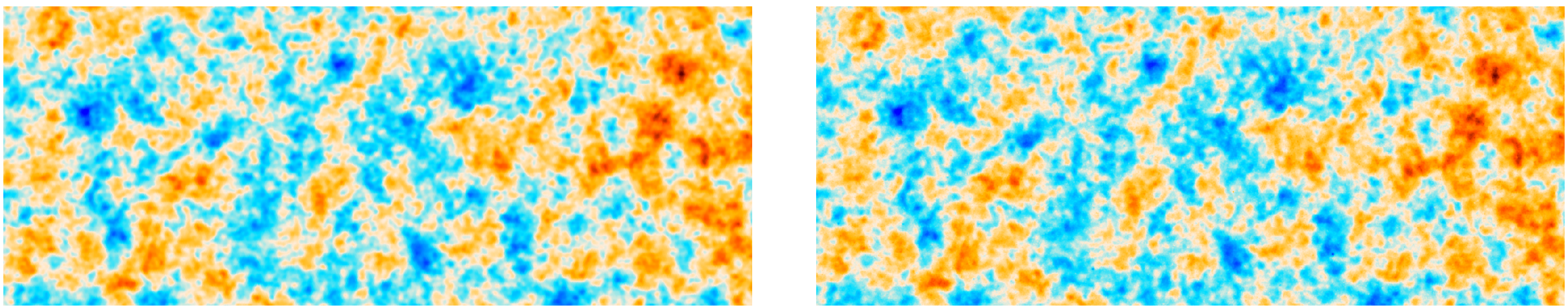


BERKELEY CENTER *for*
COSMOLOGICAL PHYSICS

Mocking the Universe ~~&~~ Machine Learning Approaches to Large-Scale Structure

With

George Stein



BERKELEY CENTER *for* COSMOLOGICAL PHYSICS

(a small sampling of)

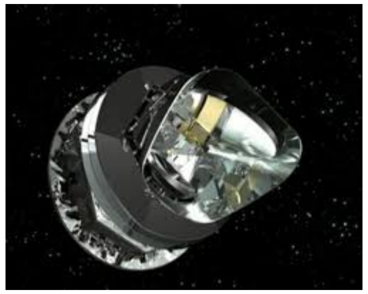
the Era of Large-Survey Precision Cosmology

Cosmic Microwave Background (CMB)

ACT

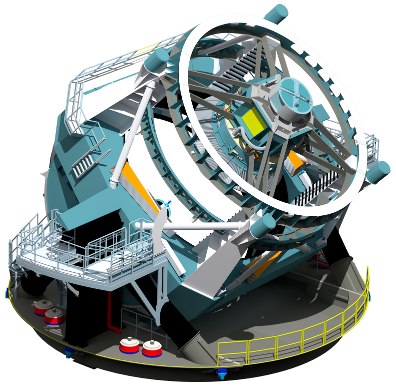
SPT

Planck



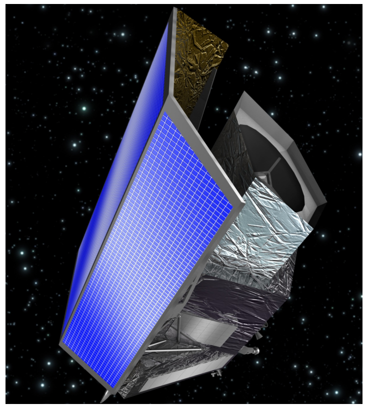
Large Scale Structure (LSS)

LSST



DESI

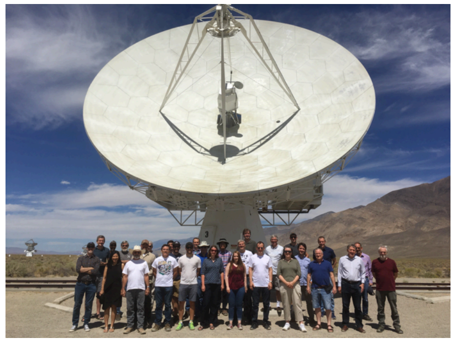
Euclid



Intensity Mapping (IM)

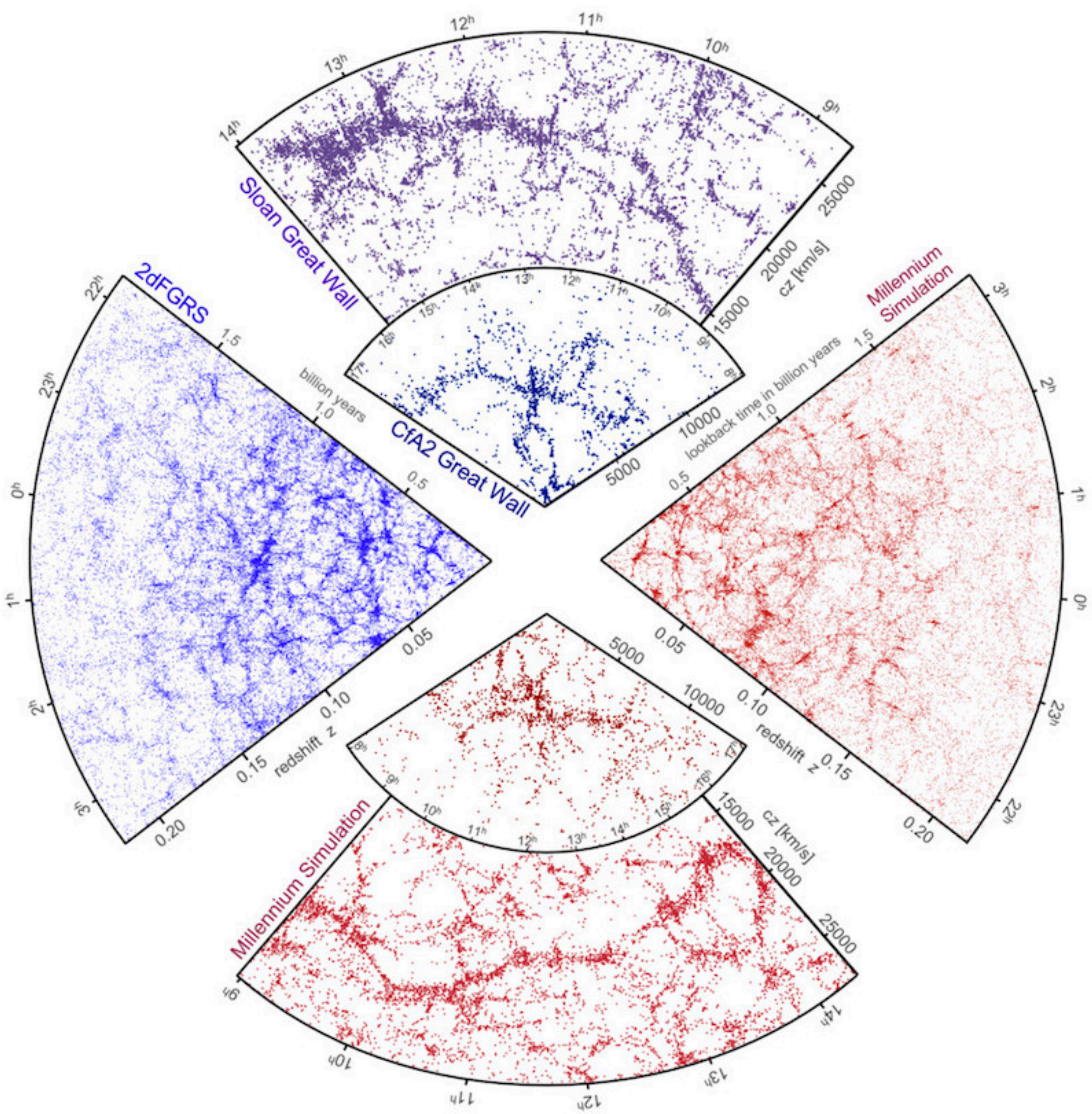
CHIME

COMAP



How do we interpret the data?

Theory/
Initial Conditions ⟷ Model ⟷ **Observables**



https://wwwmpa.mpa-garching.mpg.de/millennium/

Cosmological simulations are required to:

- forecast the expected effects of cosmological scenarios
- analyze current and near-future cosmological datasets
- develop and test operational pipelines

Modern Cosmological Simulations

Physical Scale

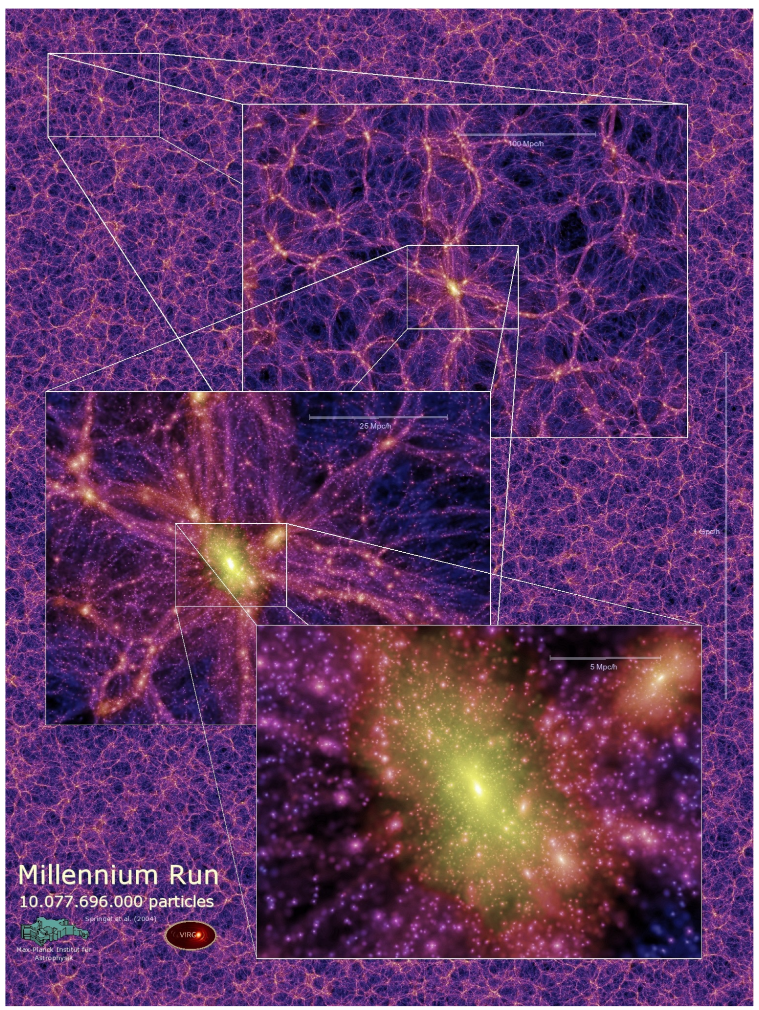
Realism

Modern Cosmological Simulations

Physical Scale

Realism

ultra large box, gravity only
Cosmic Web



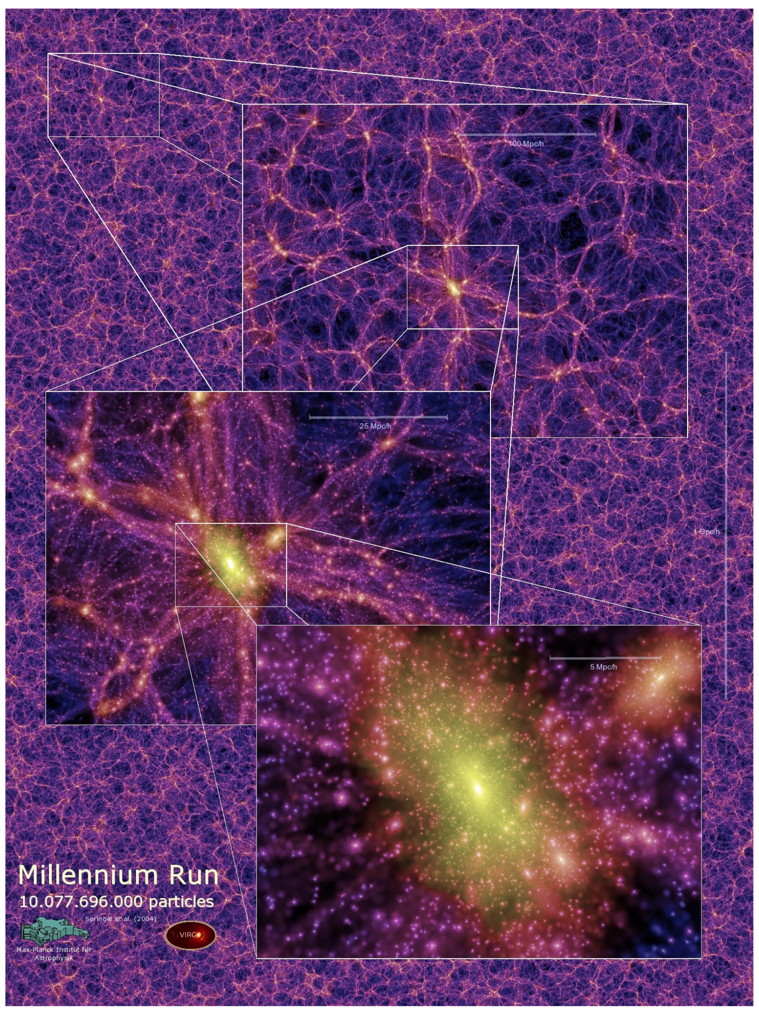
N-body or 'approximate methods'

Modern Cosmological Simulations

Physical Scale

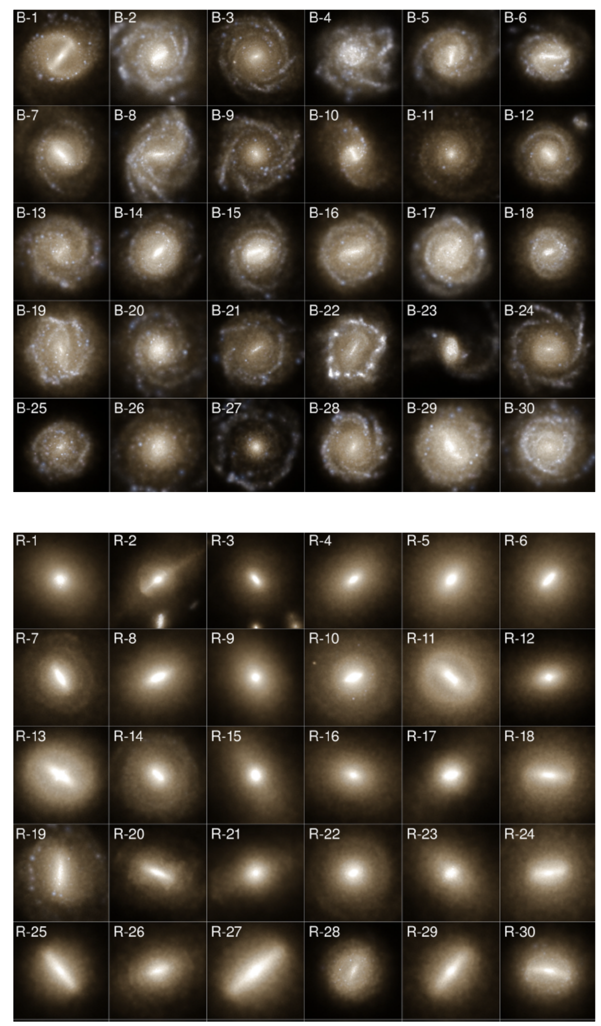
Realism

ultra large box, gravity only
Cosmic Web



N-body or 'approximate methods'

full physics, medium sized box
The properties of dark matter halos



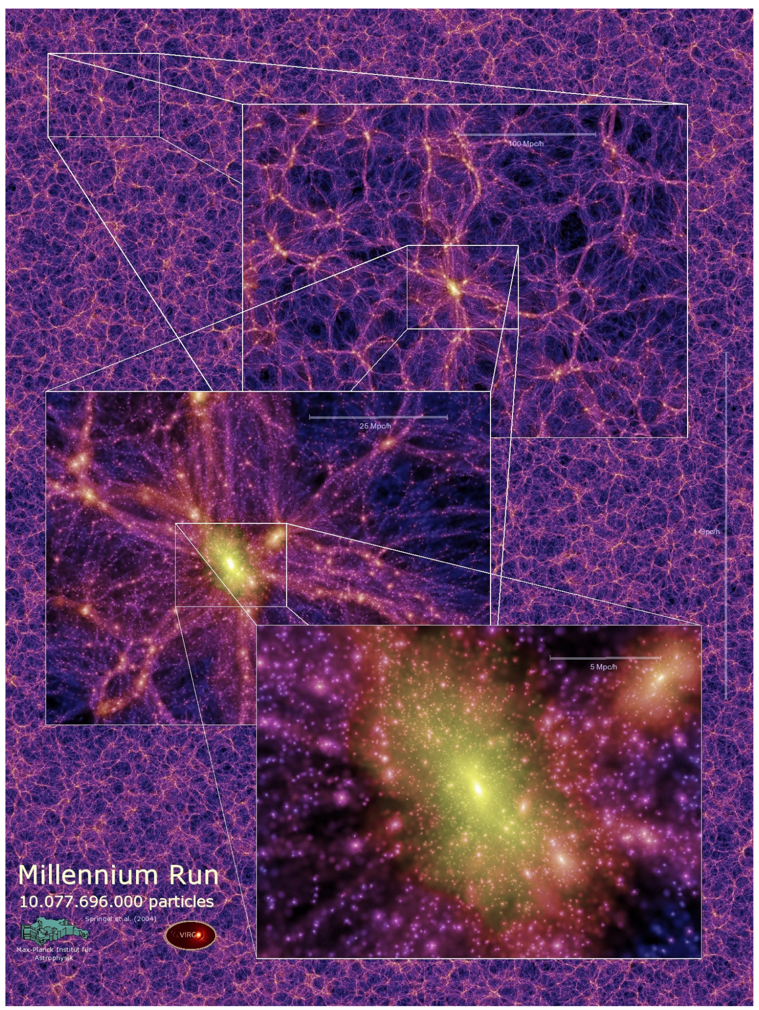
e.g. *Illustris*

Modern Cosmological Simulations

Physical Scale

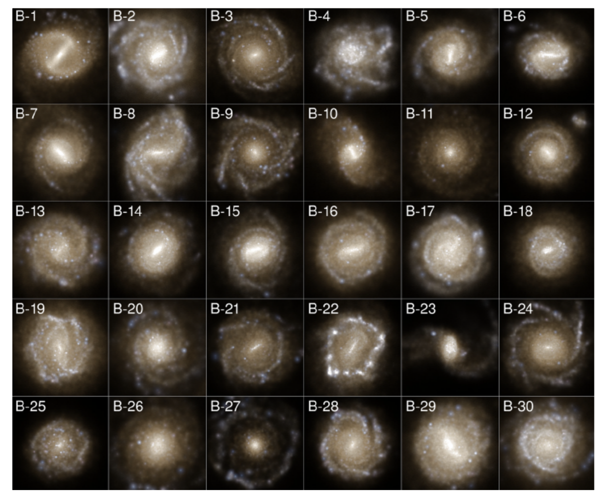
Realism

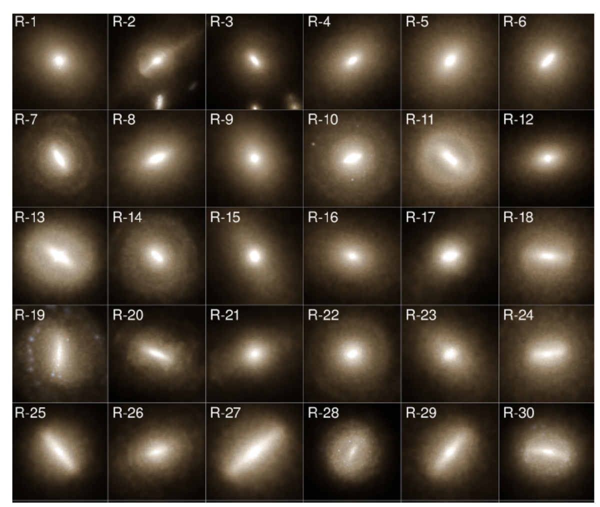
ultra large box, gravity only
Cosmic Web



N-body or 'approximate methods'

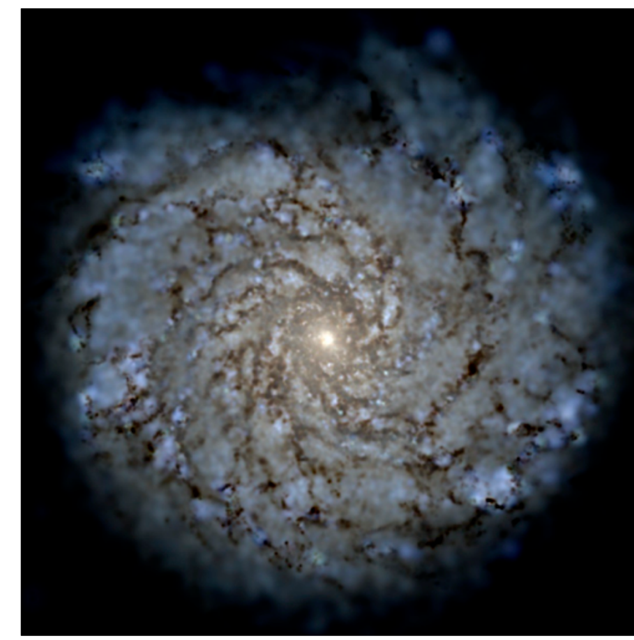
full physics, medium sized box
The properties of dark matter halos





e.g. *Illustris*

full physics, small sized box
Detailed substructures of halos & galaxies

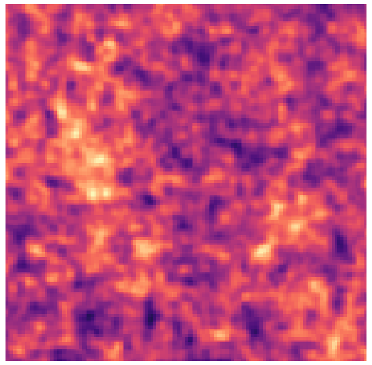


e.g. *FIRE*

Approximate Cosmological simulations

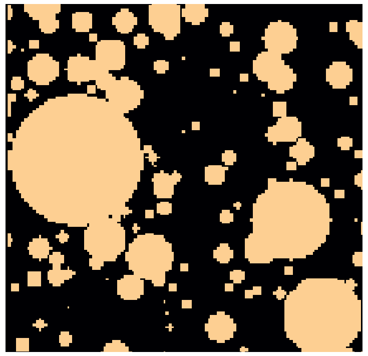
*map from initial conditions to final observable as fast as possible,
with an appropriate accuracy

Initial Conditions



Physical Approximation

Final Matter Density/Halo Catalogue



Approximate Cosmological simulations

*map from initial conditions to final observable as fast as possible, with an appropriate accuracy

1.) stochastic methods

1. **EZmocks (2014)** - Chuang et al.
   - Zeldovich + PDF mapping
2. **HALOGEN (2014)** - Avila et al.
   - LPT + stochastic halo biasing
3. **Log-normal (1991)** - Coles, Jones
   - +many others since
4. **PATCHY (2013)** - Kitaura, Yepes, Prada
   - APT + nonlinear stochastic biasing
5. **PThalos (2001)** - Scoccimarro, Sheth
   - 2LPT + $n(m,\delta \mid z)$ assignment using merger tree + NFW particle placement
   - ★ Manera et al. (2012)
     - 2LPT + fof (linking length = 0.38) + HMF
6. **QPM (2013)** - White, Tinker, McBride
   - low-res PM, halos sampled from particles based on local densities

Approximate Cosmological simulations

*map from initial conditions to final observable as fast as possible, with an appropriate accuracy

1.) stochastic methods

1. **EZmocks (2014)** - Chuang et al.
   - Zeldovich + PDF mapping
2. **HALOGEN (2014)** - Avila et al.
   - LPT + stochastic halo biasing
3. **Log-normal (1991)** - Coles, Jones
   - +many others since
4. **PATCHY (2013)** - Kitaura, Yepes, Prada
   - APT + nonlinear stochastic biasing
5. **PThalos (2001)** - Scoccimarro, Sheth
   - 2LPT + $n(m,\delta \mid z)$ assignment using merger tree + NFW particle placement
   - ★ Manera et al. (2012)
     - 2LPT + fof (linking length = 0.38) + HMF
6. **QPM (2013)** - White, Tinker, McBride
   - low-res PM, halos sampled from particles based on local densities

2.) abridged particle mesh methods

1. **COLA (2013)** - Tassev, Zaldarriaga, Eisenstein
   - Nobody in frame comoving with 2LPT observers
2. **FastPM (2016)** - Feng, Chu, Seljak
   - PM + broadband correction at each timestep to enforce correct linear evolution

Approximate Cosmological simulations

*map from initial conditions to final observable as fast as possible, with an appropriate accuracy

1.) stochastic methods

1. **EZmocks (2014)** - Chuang et al.
   - Zeldovich + PDF mapping
2. **HALOGEN (2014)** - Avila et al.
   - LPT + stochastic halo biasing
3. **Log-normal (1991)** - Coles, Jones
   - +many others since
4. **PATCHY (2013)** - Kitaura, Yepes, Prada
   - APT + nonlinear stochastic biasing
5. **PThalos (2001)** - Scoccimarro, Sheth
   - 2LPT + n(m,δ | z) assignment using merger tree + NFW particle placement
   - ★ Manera et al. (2012)
     - 2LPT + fof (linking length = 0.38) + HMF
6. **QPM (2013)** - White, Tinker, McBride
   - low-res PM, halos sampled from particles based on local densities

2.) abridged particle mesh methods

1. **COLA (2013)** - Tassev, Zaldarriaga, Eisenstein
   - Nobody in frame comoving with 2LPT observers
2. **FastPM (2016)** - Feng, Chu, Seljak
   - PM + broadband correction at each timestep to enforce correct linear evolution

3.) predictive methods

Physics Based

1. **Peak Patch (1993)** - Bond, Myers
   - Homogeneous ellipsoid collapse + LPT
   - ★ Stein, Alvarez, Bond (2018)
     - parallel HPC version
2. **PINOCCHIO (2002)** - Monaco et al.
   - LPT + Inverse collapse time for particles + fragmentation criteria to create distinct halos

Machine Learned

1. **D3M (2018)** - He et al.
   - CNN to find correction to LPT displacements
2. **Halonet (2018)** - Berger, Stein
   - CNN to find Lagrangian halos +2LPT

Which method to use?

Comparing approximate methods for mock catalogues and covariance matrices

300 1.5h^{-1}Gpc, 1000^3 particles z=1 simulations

– *I: correlation function* [1806.09477]
– *II: power spectrum multipoles* [1806.09497]
– *III: Bispectrum* [1806.09499]

Which method to use?

Comparing approximate methods for mock catalogues and covariance matrices

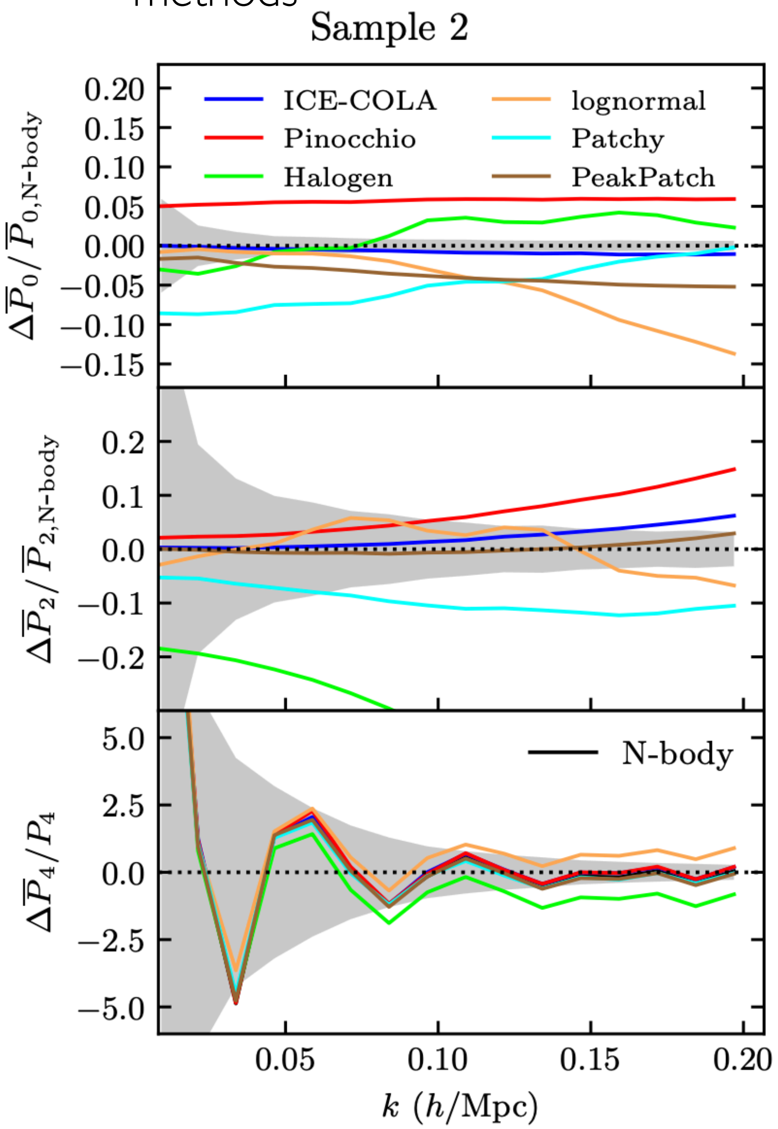
300 1.5h[-1]Gpc, 1000[3] particles z=1 simulations

– *I: correlation function* [1806.09477]

– *II: power spectrum multipoles* [1806.09497]

– *III: Bispectrum* [1806.09499]

- Power spectrum multipoles see ~20/40/50% difference between methods



George Stein

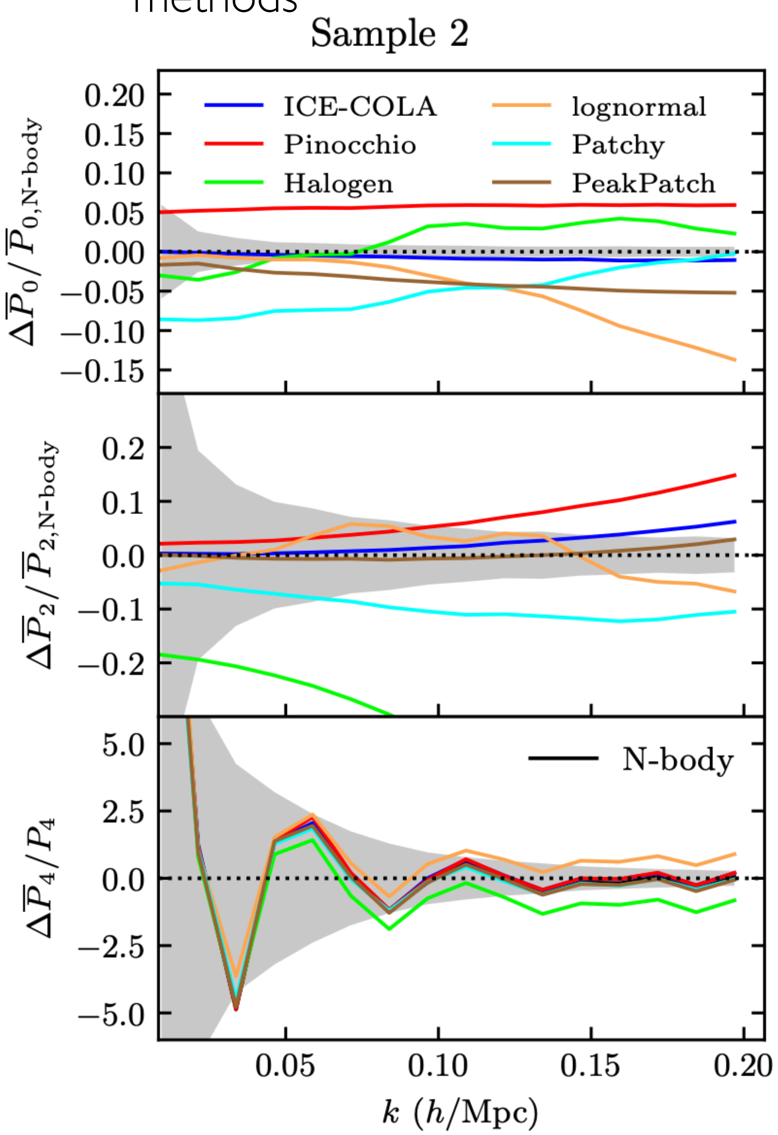
Which method to use?

Comparing approximate methods for mock catalogues and covariance matrices

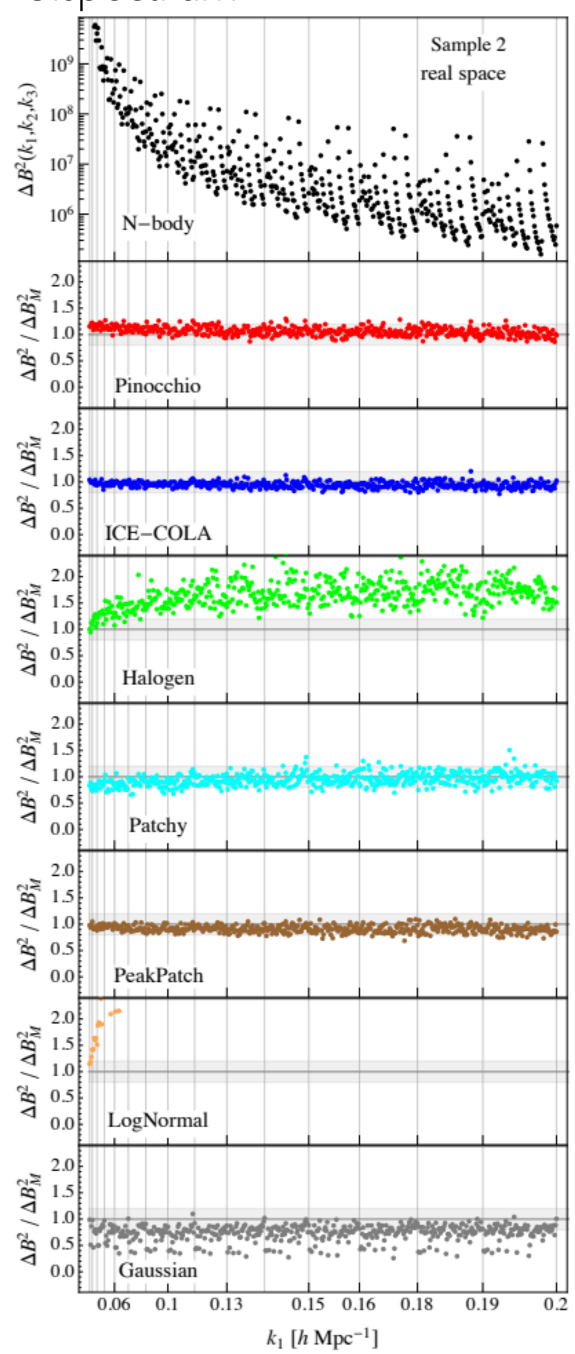
300 $1.5h^{-1}$Gpc, 1000^3 particles z=1 simulations

- *I: correlation function* [1806.09477](1806.09477)
- *II: power spectrum multipoles* [1806.09497](1806.09497)
- *III: Bispectrum* [1806.09499](1806.09499)

- Power spectrum multipoles see ~20/40/50% difference between methods



- Methods fit to reproduce power spectrum do not reproduce bispectrum



Which method to use?

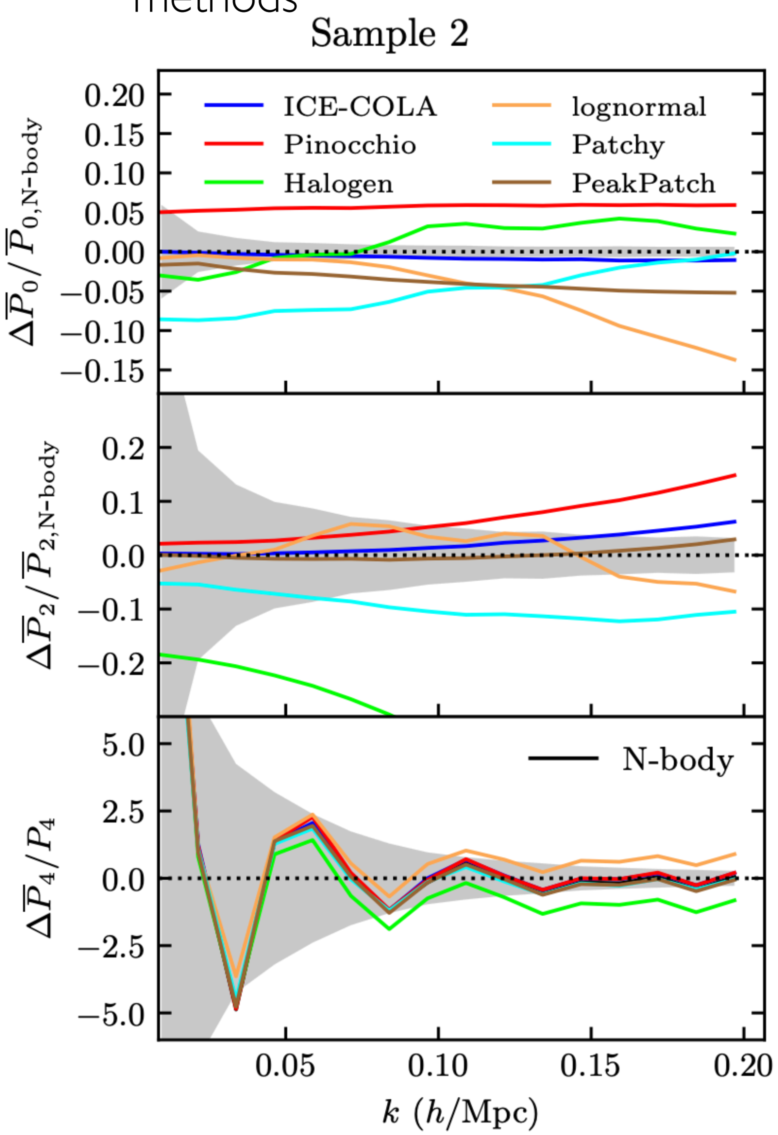
Comparing approximate methods for mock catalogues and covariance matrices

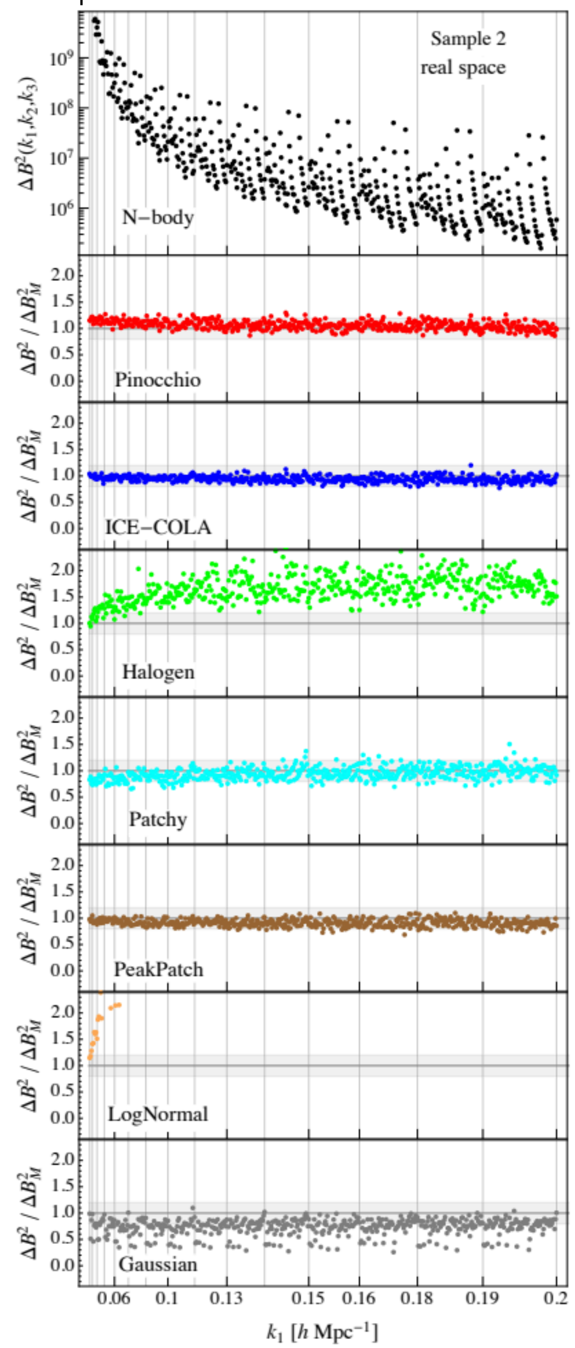
300 $1.5h^{-1}$Gpc, 1000^3 particles z=1 simulations

- *I: correlation function* [1806.09477](https://arxiv.org/abs/1806.09477)
- *II: power spectrum multipoles* [1806.09497](https://arxiv.org/abs/1806.09497)
- *III: Bispectrum* [1806.09499](https://arxiv.org/abs/1806.09499)

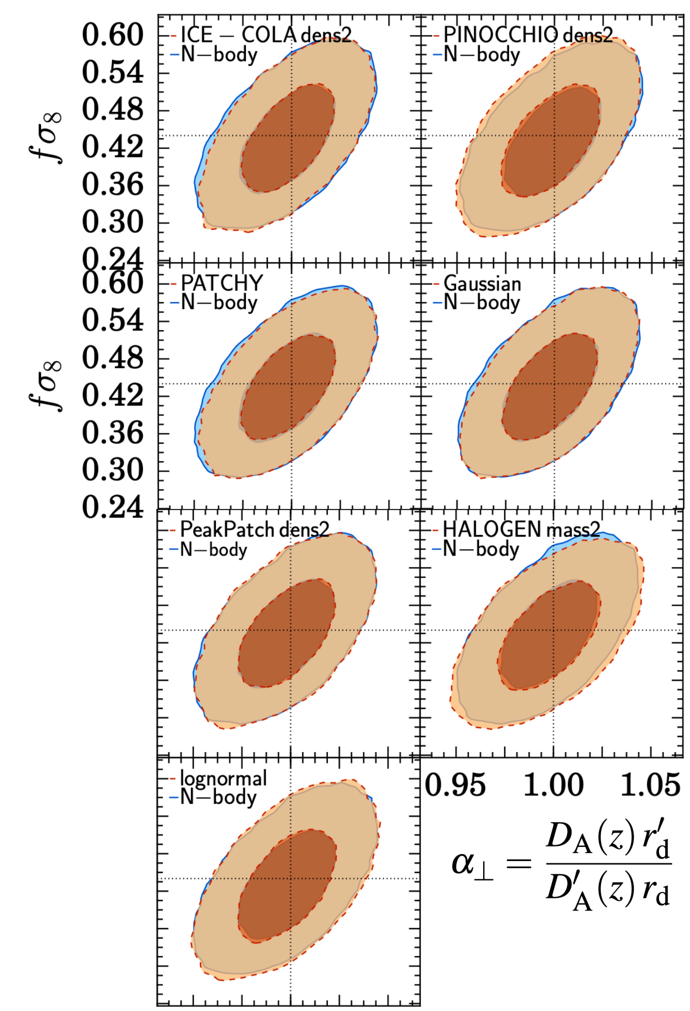
- Methods fit to reproduce power spectrum do not reproduce bispectrum

- Power spectrum multipoles see ~20/40/50% difference between methods





- covariance matrices produce similar cosmological parameter constraints



George Stein

Approximate Cosmological simulations

*map from initial conditions to final observable as fast as possible, with an appropriate accuracy

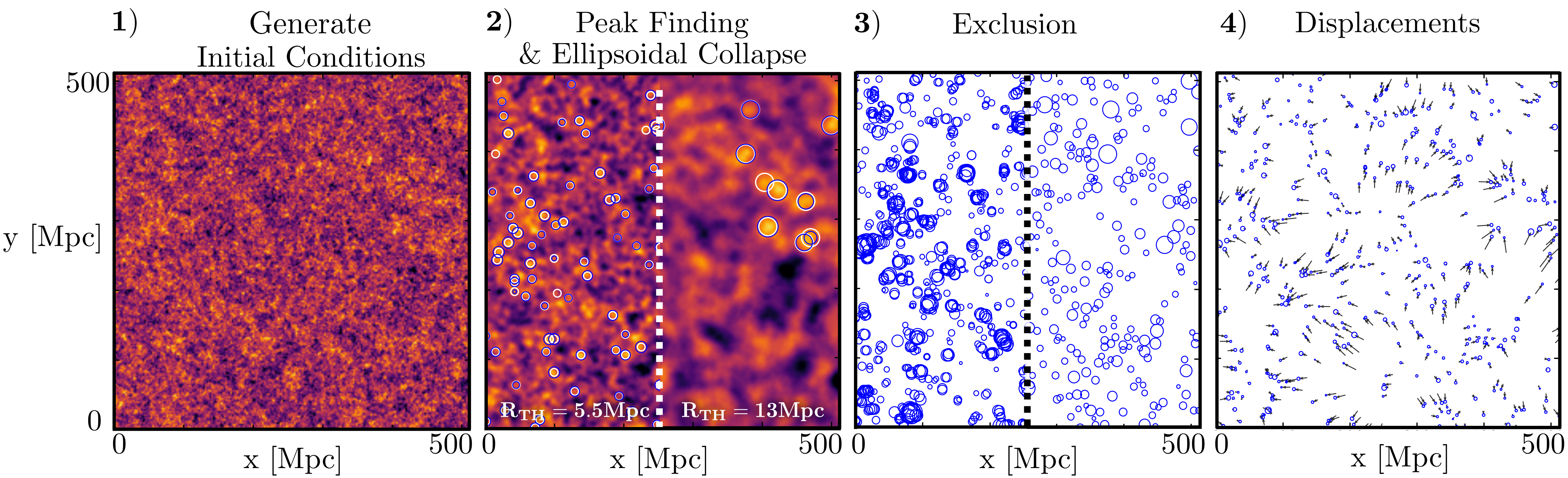
mass-Peak Patch

Stein, Alvarez, Bond, 1810.07727

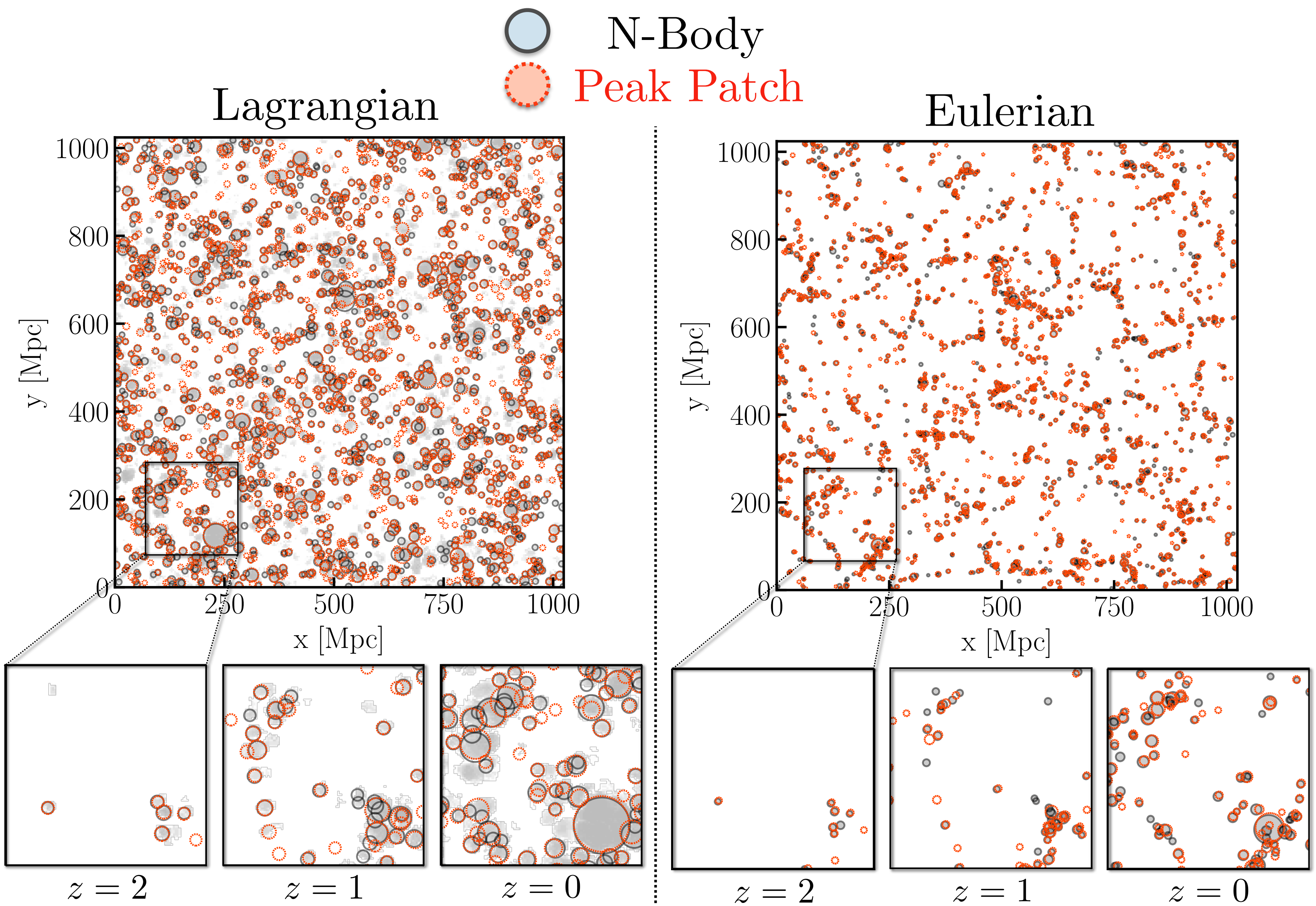
Bond & Myers 1996ApJS..103….1B, 1996ApJS..103…41B, 1996ApJS..103...63B

Based on Homogeneous Ellipsoid Collapse

$$\frac{\ddot{x}_i}{x_i} = \frac{\ddot{a}}{a} - \frac{1}{2}\Omega_m H^2 \left[b_i\bar{\delta} + c_i\bar{\delta}_{\rm lin}\right]$$



Comparisons to N-Body simulations

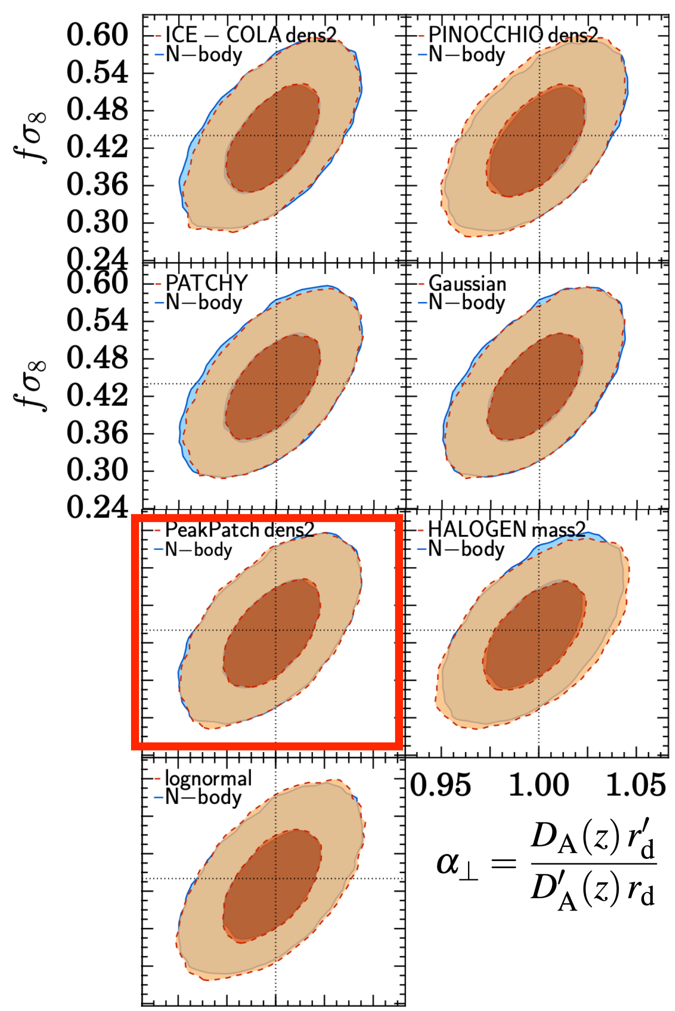


Which method to use?

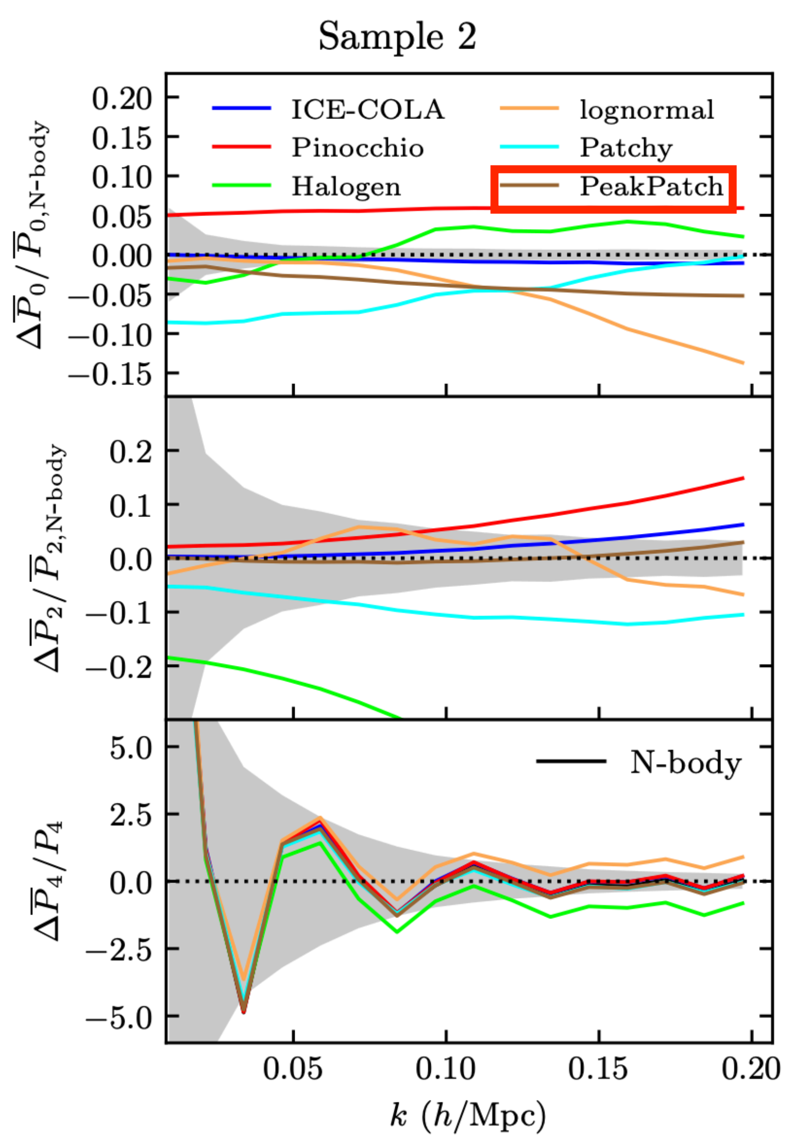
Comparing approximate methods for mock catalogues and covariance matrices

– *I: correlation function* [1806.09477]
– *II: power spectrum multipoles* [1806.09497]
– *III: Bispectrum* [1806.09499]

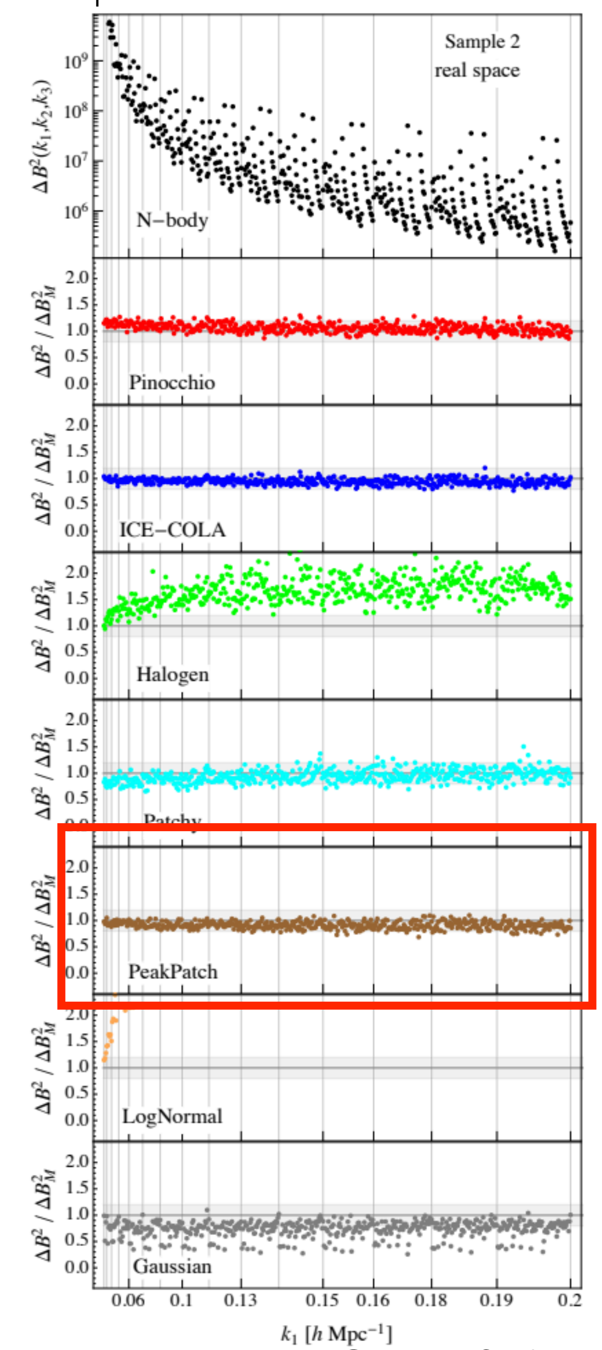
- covariance matrices produce similar cosmological parameter constraints



- Power spectrum multipoles see ~15% difference between methods



- Methods fit to reproduce power spectrum do not reproduce bispectrum



George Stein

Constructing the Microwave Sky

Desired CMB Mock:

Fullsky out to z~5

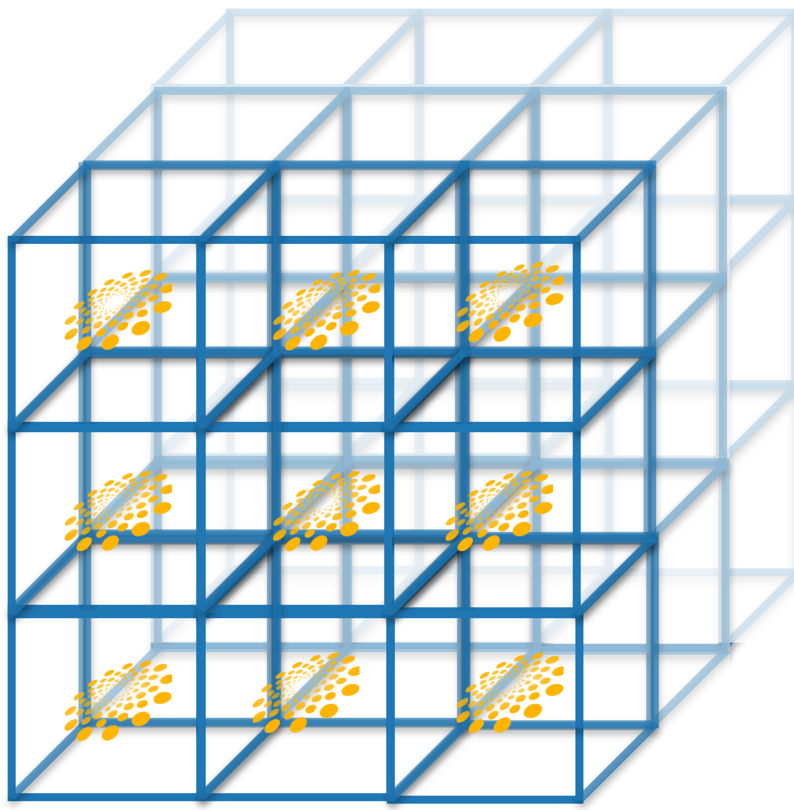
capable of resolving halos above $\sim 1 \times 10^{12}\ M_\odot$

Requires:

$\sim 13{,}000^3$ particle simulation in a box size of ~16 Gpc.

~1 billion halos

Computational Approaches



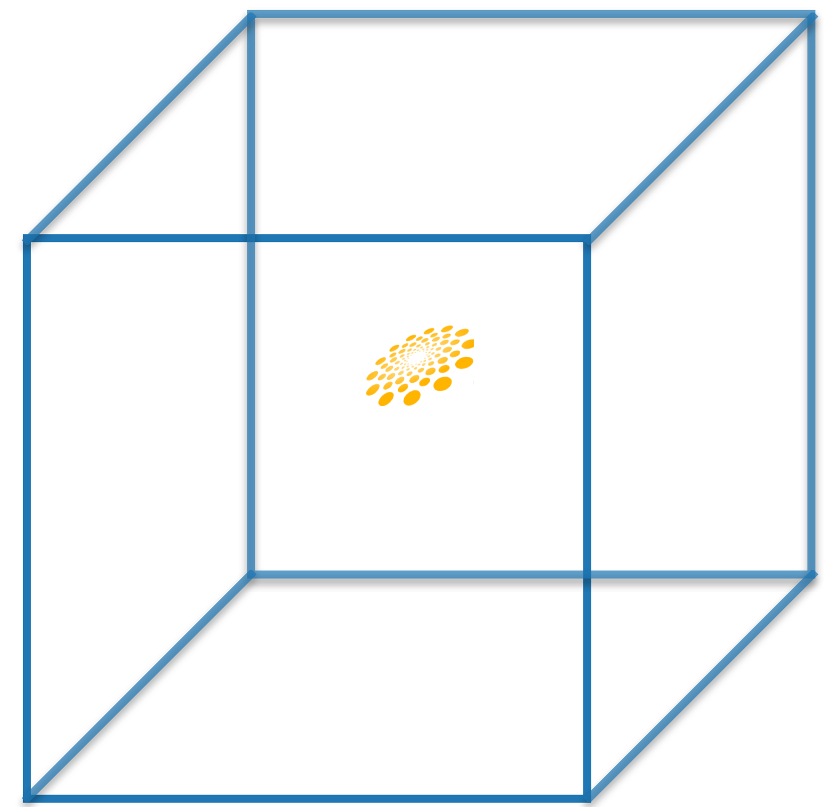
Small Box & Repeat

Repeats structures

Miss large scale density/velocity fluctuations

& rare objects in the universe

Computationally less demanding



Large Box

Includes large scale density/velocity fluctuations

& rare objects in the universe

Computationally more demanding

Dark Matter Lightcone

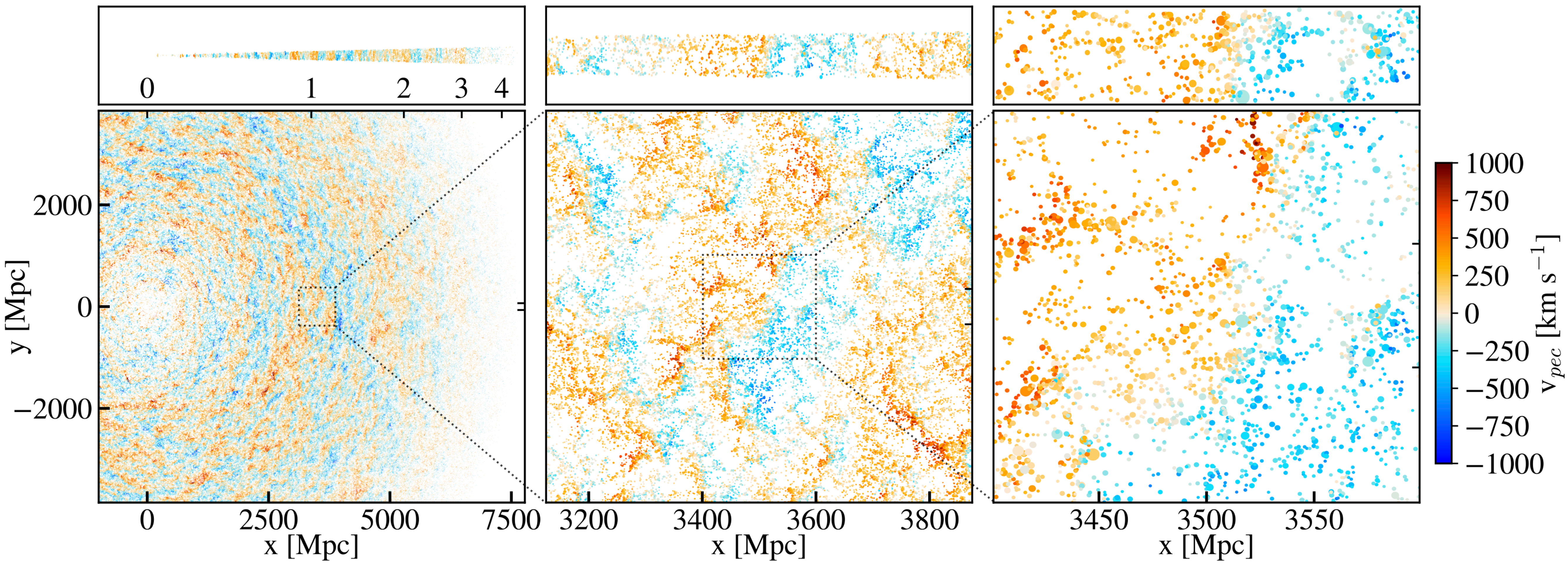
Simulation details:

Created using the mass-Peak Patch method *Stein, Alvarez, Bond - 1810.07727*

fullsky $z<4.6$, 9×10^8 halos,

effective 12,288 particles, 1900 Gpc^3 volume, from octants

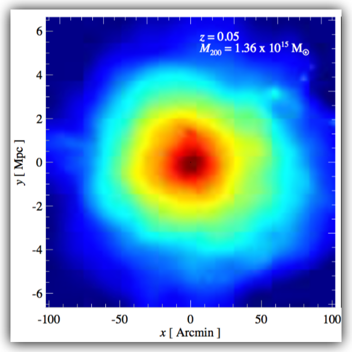
4336 CPU hours, 7.67 TB



Dark Matter to Baryonic Observables

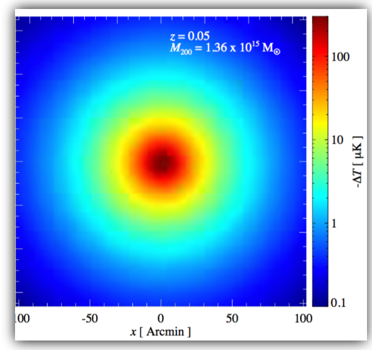
- Use spherical halo profiles fit to small box hydrodynamical simulations

Full Simulation

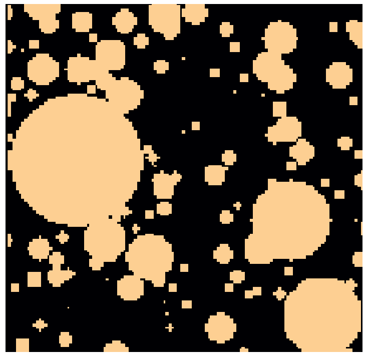


Hydro

Measure Fitting Functions

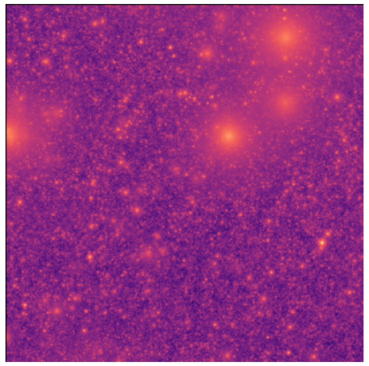


Fast Simulation



mass-Peak Patch

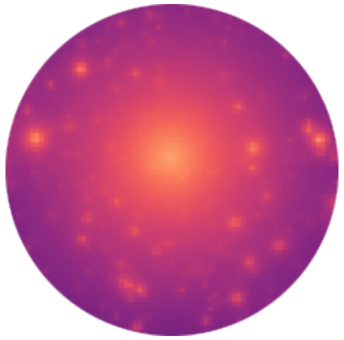
Mock Observation



Painting the Extragalactic Sky

thermal Sunyaev-Zel'dovich (tSZ)

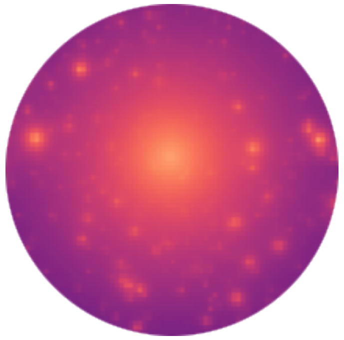
$$\propto \frac{k_B \sigma_T}{m_e c^2} \int d\chi (1+z)^{-1} P_{th}(\chi \hat{n})$$



Painting the Extragalactic Sky

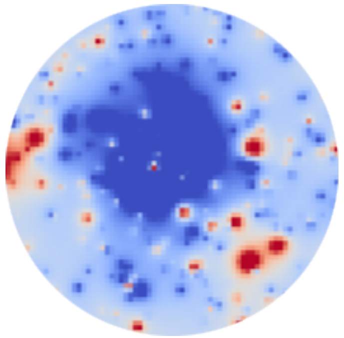
thermal Sunyaev-Zel'dovich (tSZ)

$$\propto \frac{k_B \sigma_T}{m_e c^2} \int d\chi (1+z)^{-1} P_{th}(\chi \hat{n})$$



kinetic Sunyaev-Zel'dovich (kSZ)

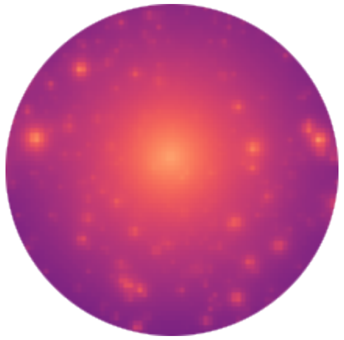
$$\propto \frac{\rho_{b,0} \sigma_T}{\mu_e m_p} \int d\chi (1+z)^2 \Delta_g(\chi \hat{n}) \mathbf{v}(\chi \hat{n}) \cdot \hat{n}$$



Painting the Extragalactic Sky

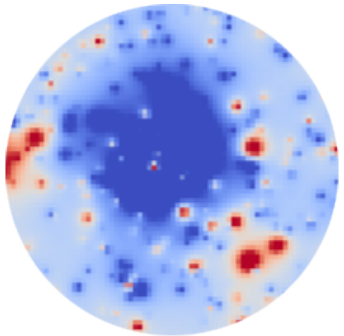
thermal Sunyaev-Zel'dovich (tSZ)

$$\propto \frac{k_B \sigma_T}{m_e c^2} \int d\chi (1+z)^{-1} P_{th}(\chi \hat{n})$$



kinetic Sunyaev-Zel'dovich (kSZ)

$$\propto \frac{\rho_{b,0} \sigma_T}{\mu_e m_p} \int d\chi (1+z)^2 \Delta_g(\chi \hat{n}) \mathbf{v}(\chi \hat{n}) \cdot \hat{n}$$

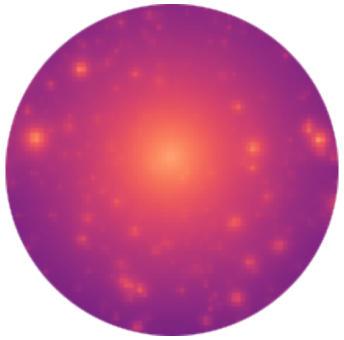


using gNFW profiles from *Battaglia, Bond, Pfrommer, Sievers (2012)*

Painting the Extragalactic Sky

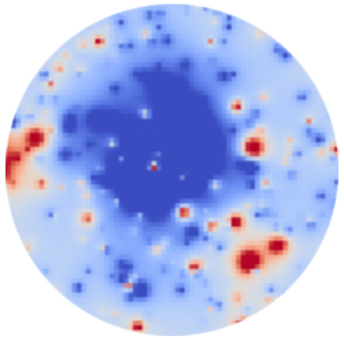
thermal Sunyaev-Zel'dovich (tSZ)

$$\propto \frac{k_B \sigma_T}{m_e c^2} \int d\chi (1+z)^{-1} P_{th}(\chi \hat{n})$$



kinetic Sunyaev-Zel'dovich (kSZ)

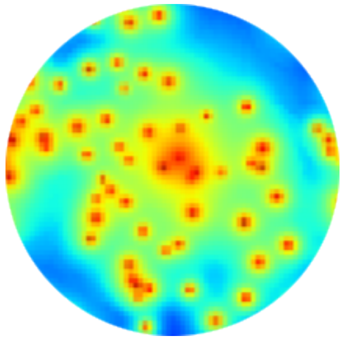
$$\propto \frac{\rho_{b,0} \sigma_T}{\mu_e m_p} \int d\chi (1+z)^2 \Delta_g(\chi \hat{n}) \mathbf{v}(\chi \hat{n}) \cdot \hat{n}$$



using gNFW profiles from *Battaglia, Bond, Pfrommer, Sievers (2012)*

Weak Lensing

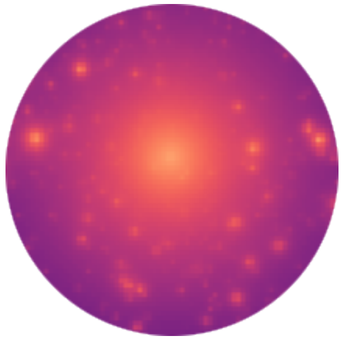
$$= \int d\chi W_\kappa(\chi) \delta(\chi \hat{n})$$



Painting the Extragalactic Sky

thermal Sunyaev-Zel'dovich (tSZ)

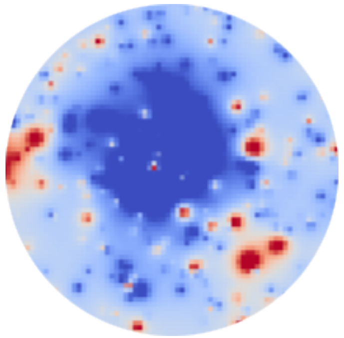
$$\propto \frac{k_B \sigma_T}{m_e c^2} \int d\chi (1+z)^{-1} P_{th}(\chi \hat{n})$$



using gNFW profiles from *Battaglia, Bond, Pfrommer, Sievers (2012)*

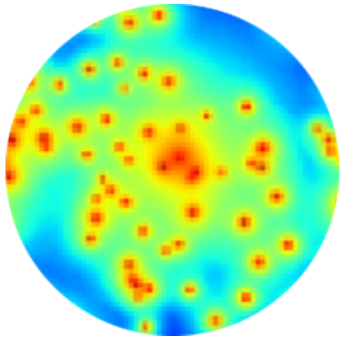
kinetic Sunyaev-Zel'dovich (kSZ)

$$\propto \frac{\rho_{b,0} \sigma_T}{\mu_e m_p} \int d\chi (1+z)^2 \Delta_g(\chi \hat{n}) \mathbf{v}(\chi \hat{n}) \cdot \hat{n}$$



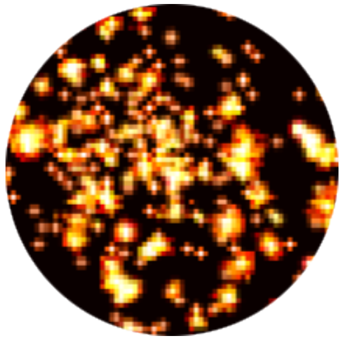
Weak Lensing

$$= \int d\chi W_\kappa(\chi) \delta(\chi \hat{n})$$



Cosmic Infrared Background (CIB)

$$L_{(1+z)\nu}(M,z) = L_0 \Phi(z) \Sigma(M,z) \Theta[(1+z)\nu, T_d(z)]$$

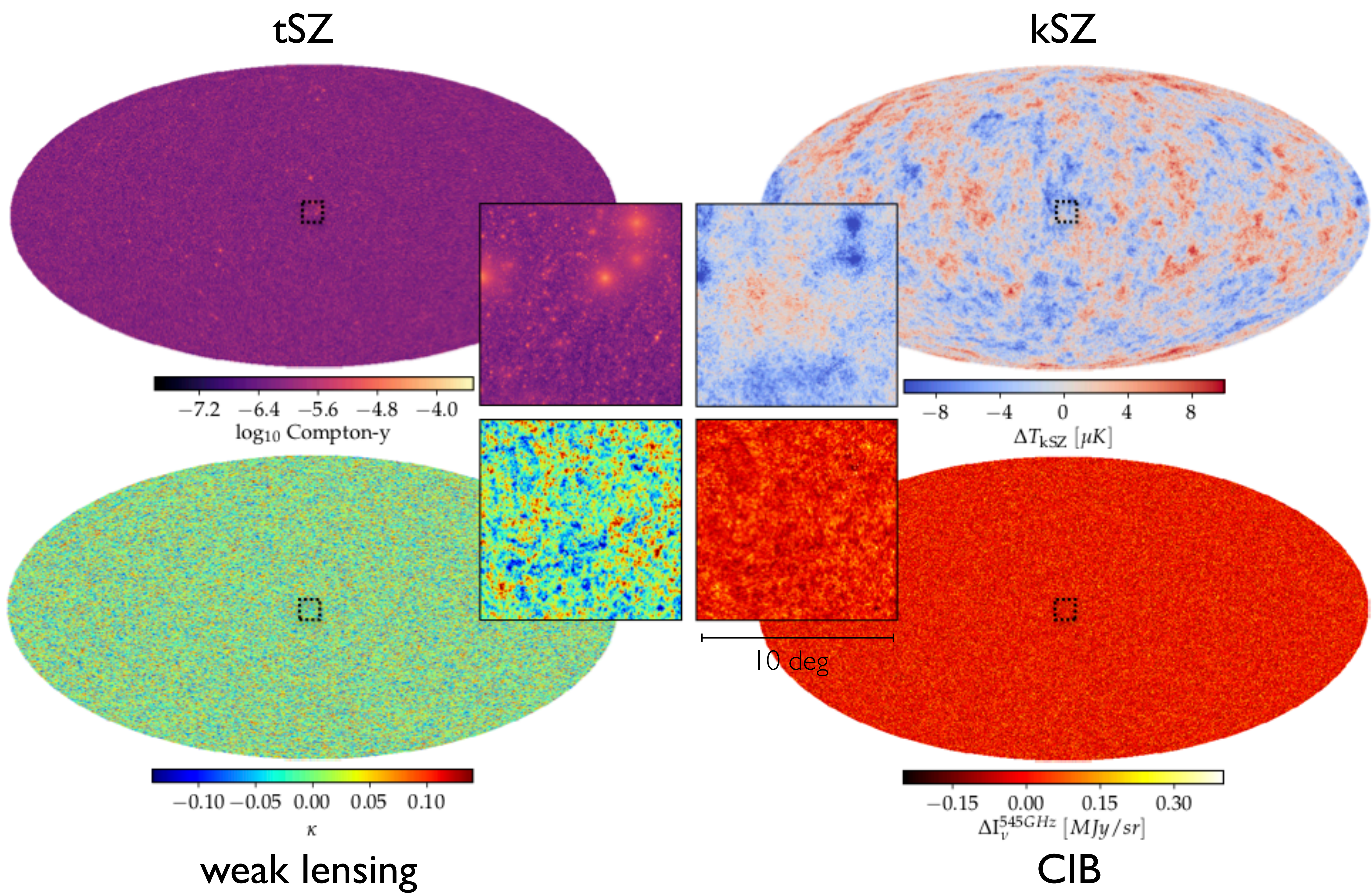


$\Phi(z)$: normalization of the L-M relation

$\Sigma(M,z)$: dependence of the luminosity on halo mass

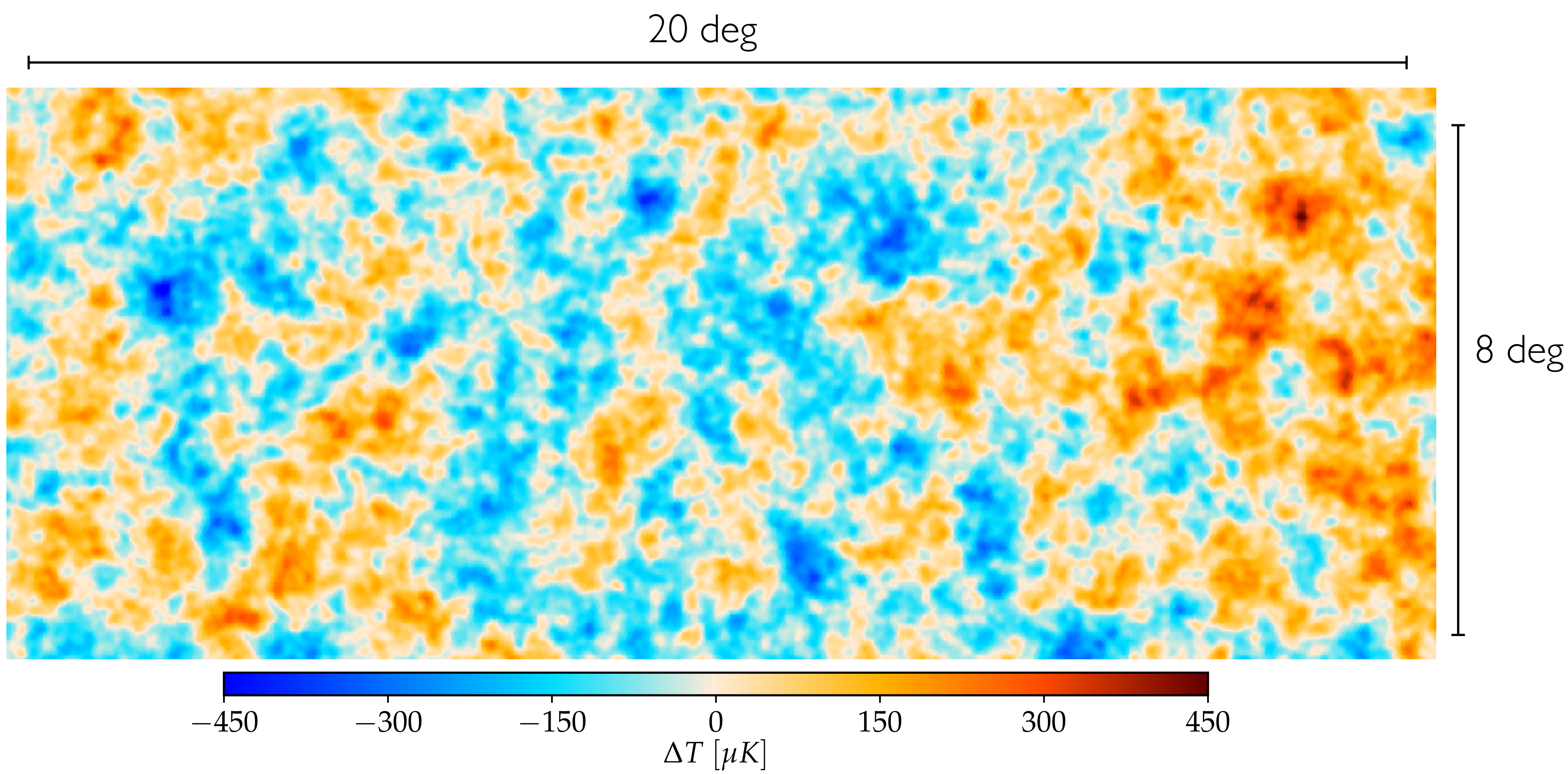
$\Theta[(1+z)\nu, T_d(z)]$: spectral energy distribution

The Websky Extragalactic Maps



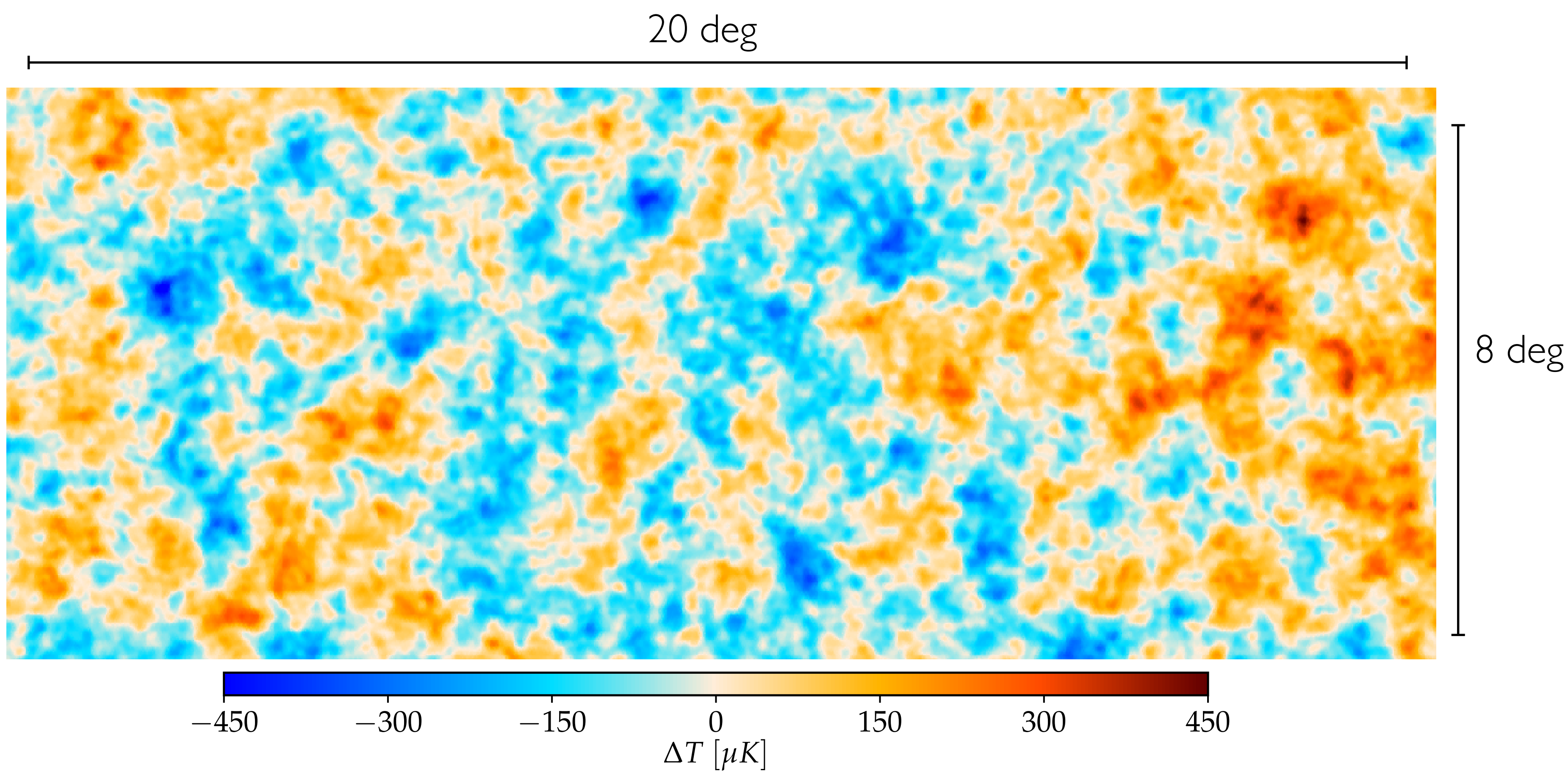
Extragalactic Foregrounds

CMB

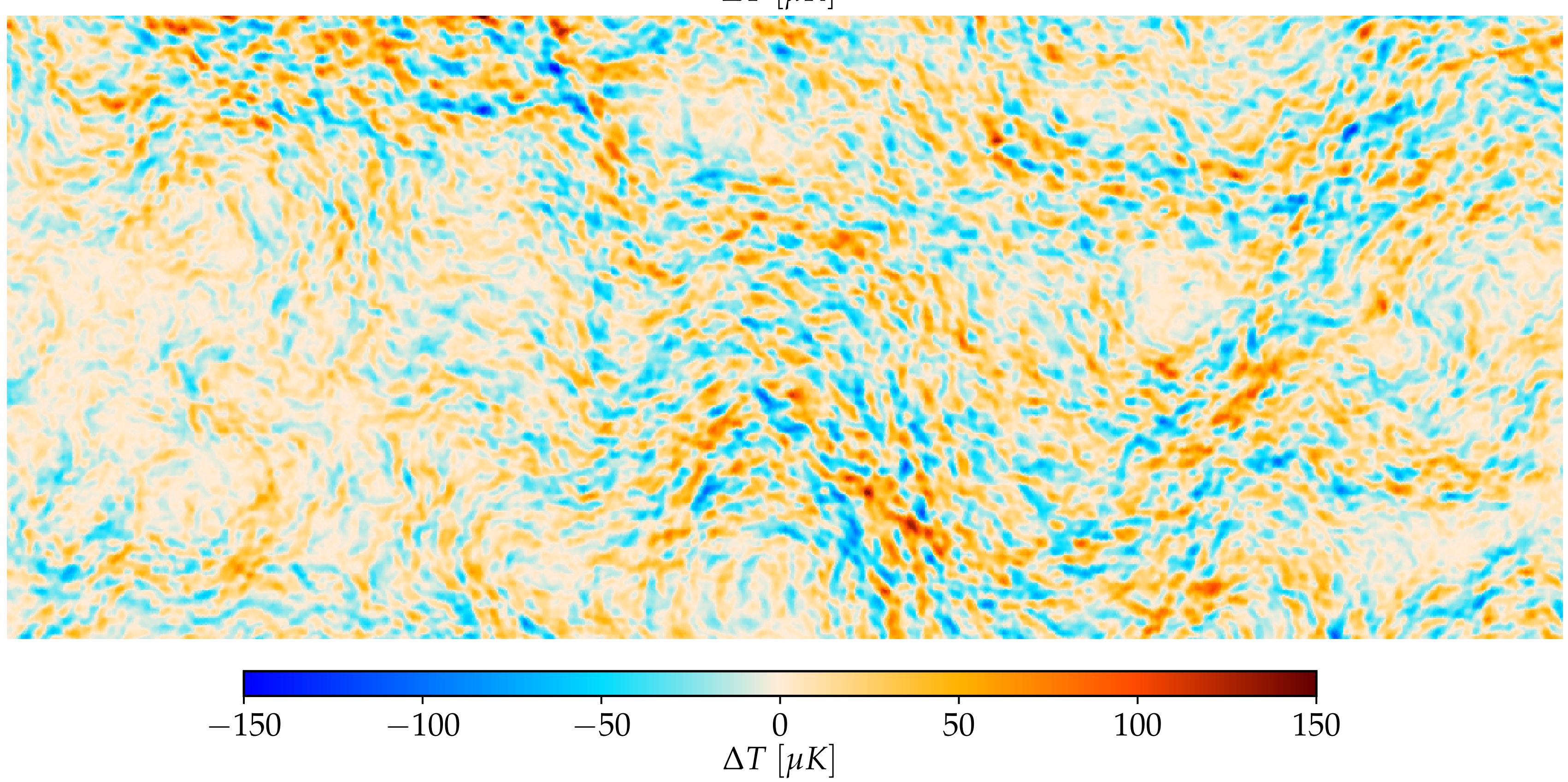


Extragalactic Foregrounds

Observed CMB

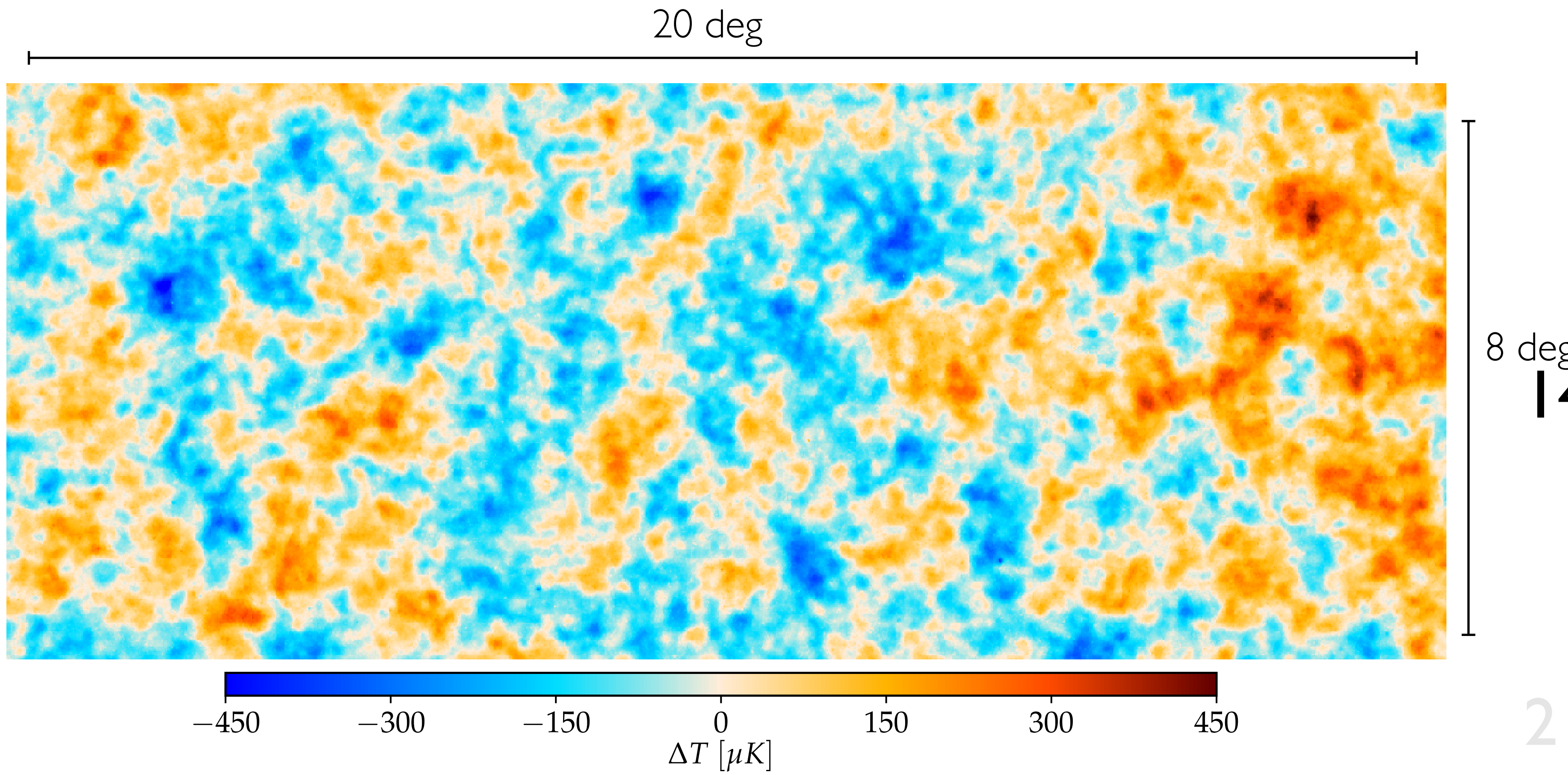


Lensing



Extragalactic Foregrounds

Observed CMB

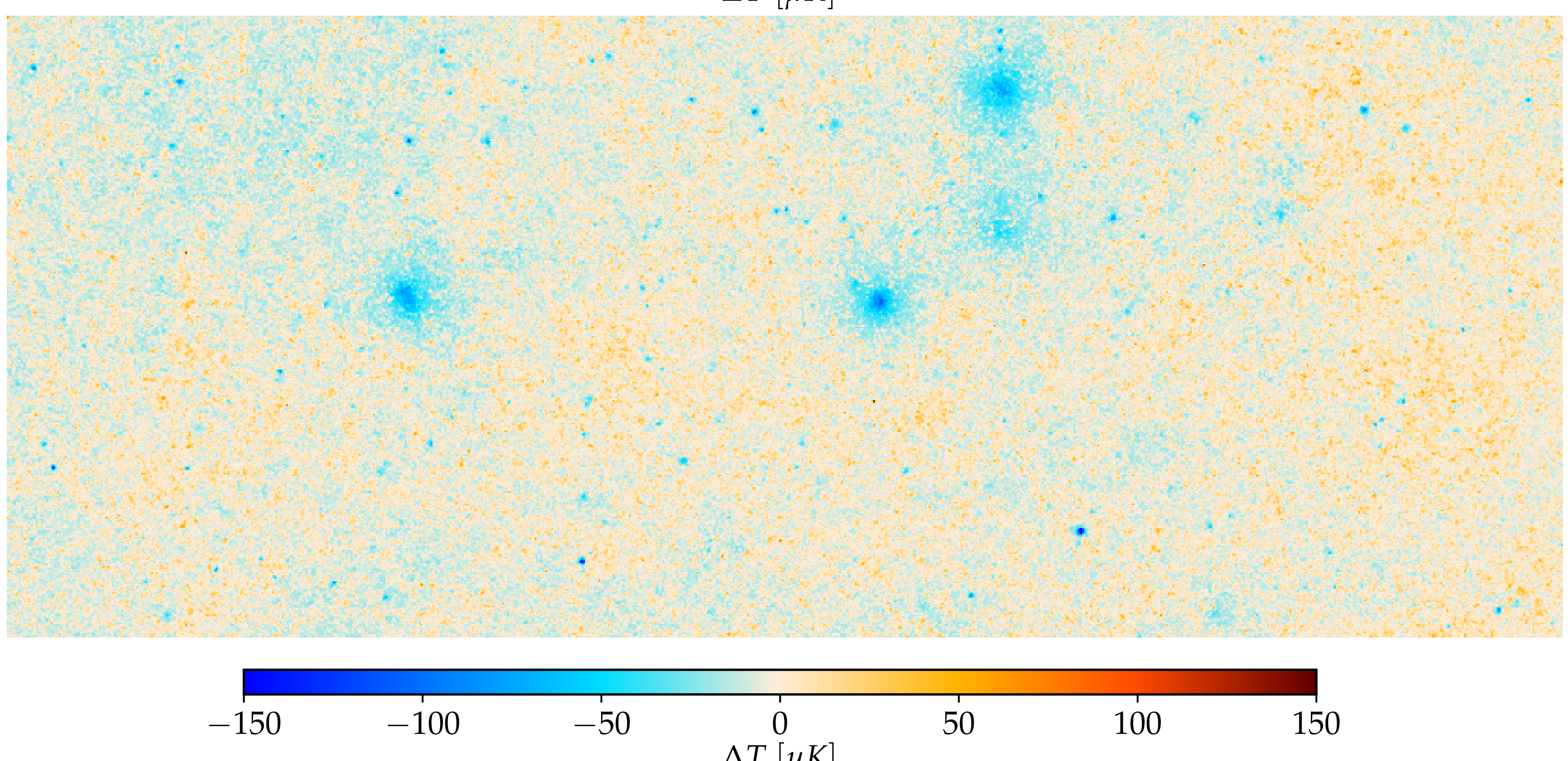


143 GHz

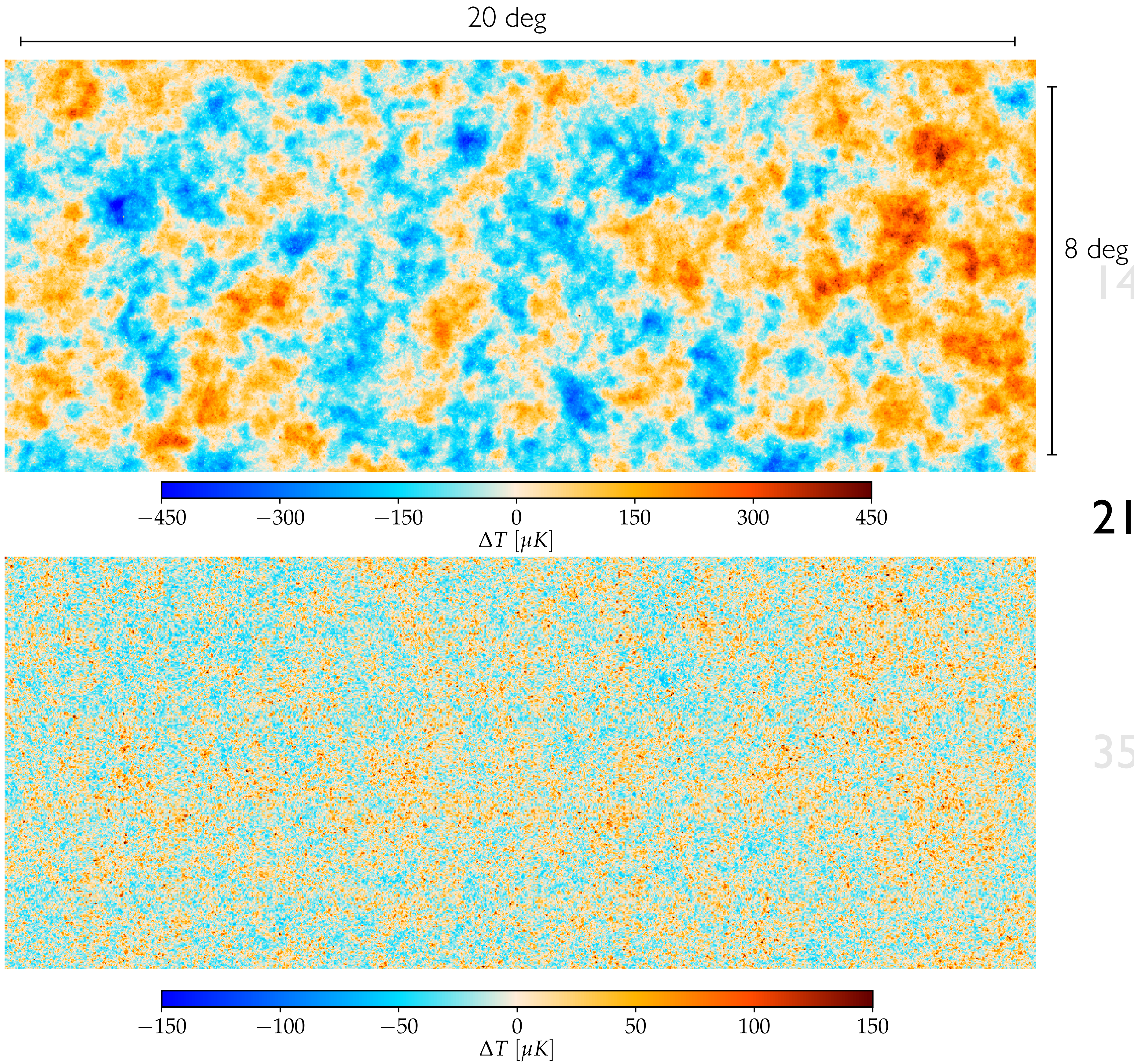
217 GHz

353 GHz

tSZ + kSZ + CIB



Extragalactic Foregrounds



Observed CMB

tSZ + kSZ + CIB

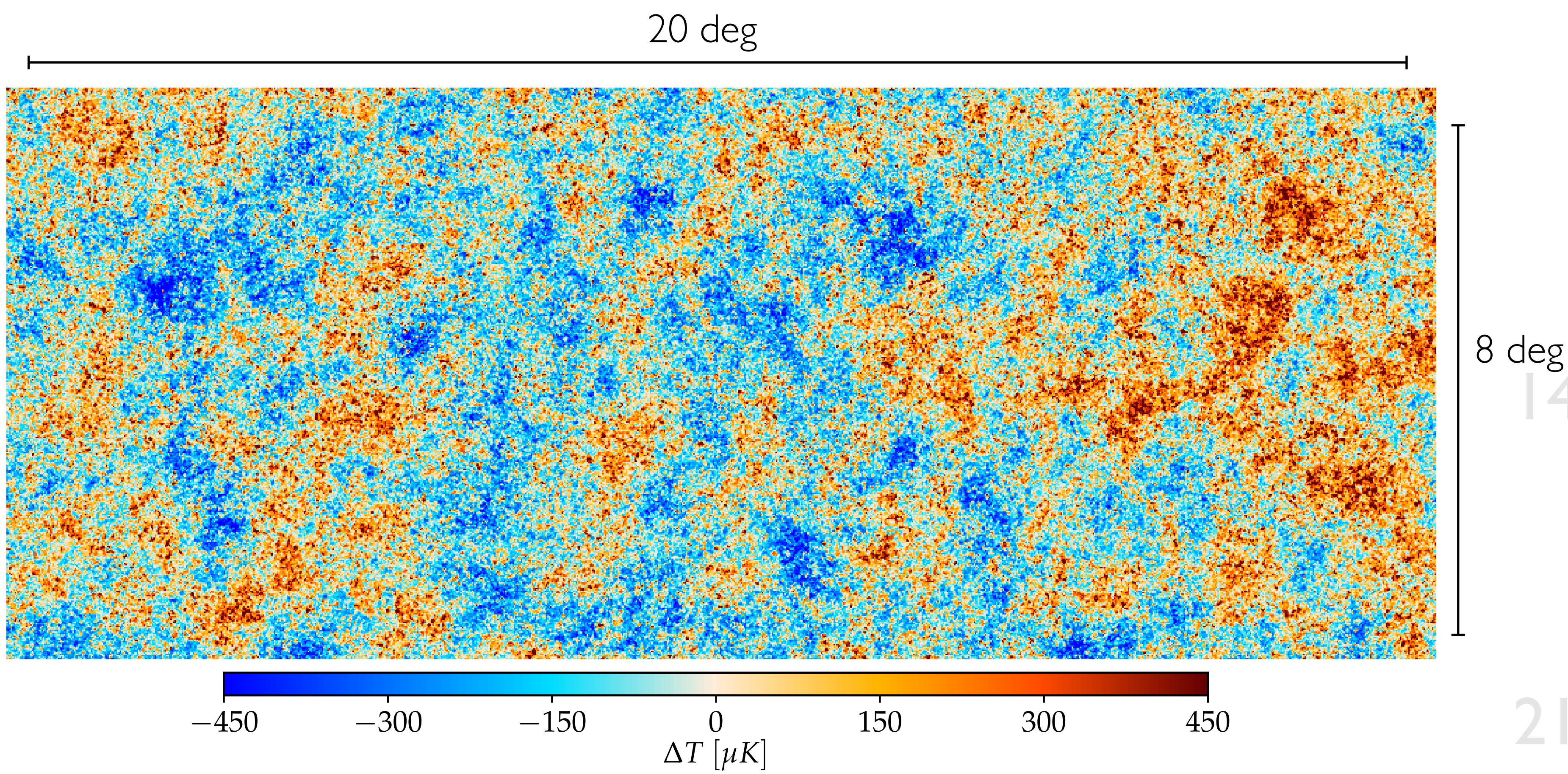
143 GHz

217 GHz

353 GHz

Extragalactic Foregrounds

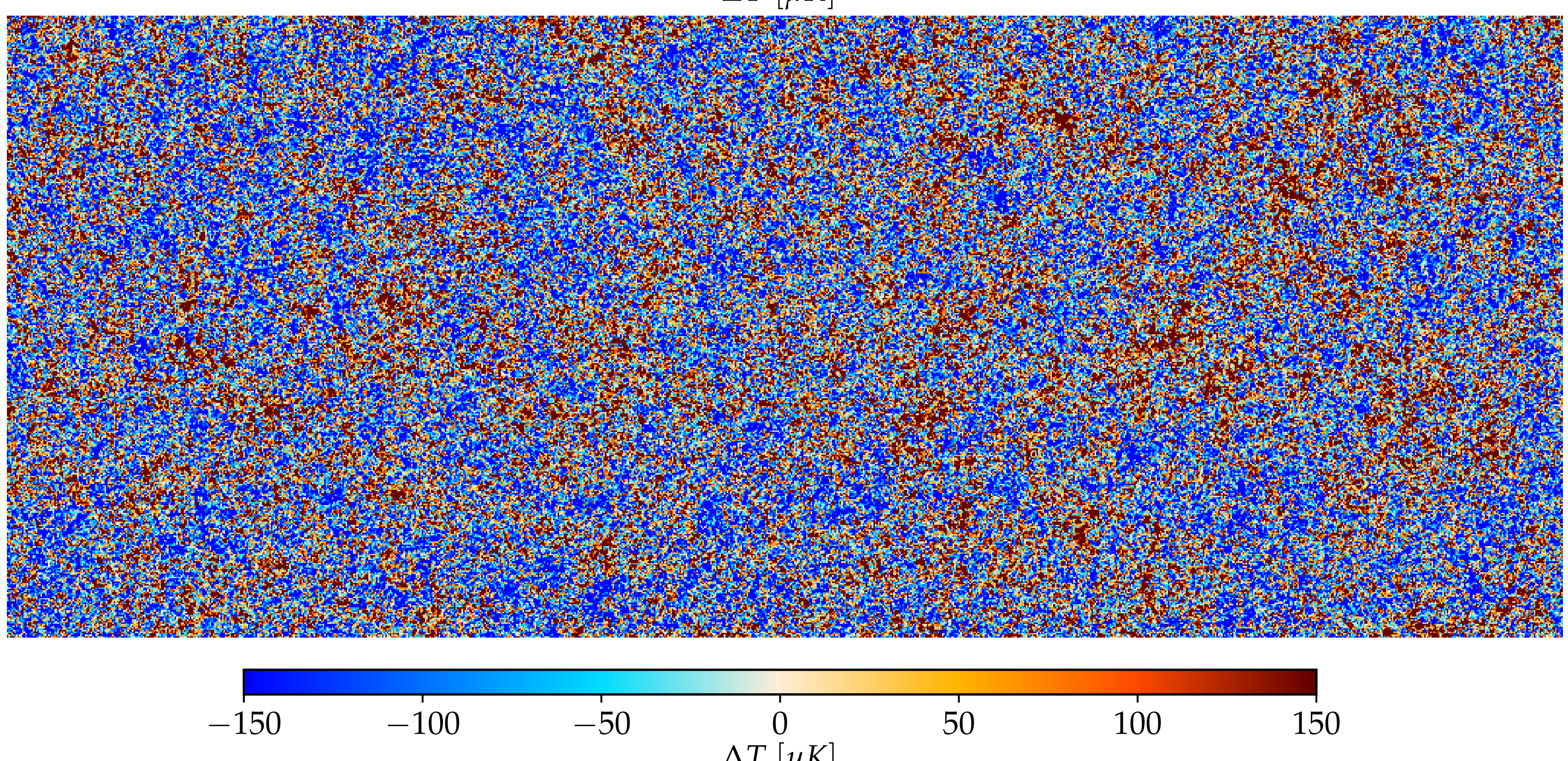
Observed CMB



143 GHz

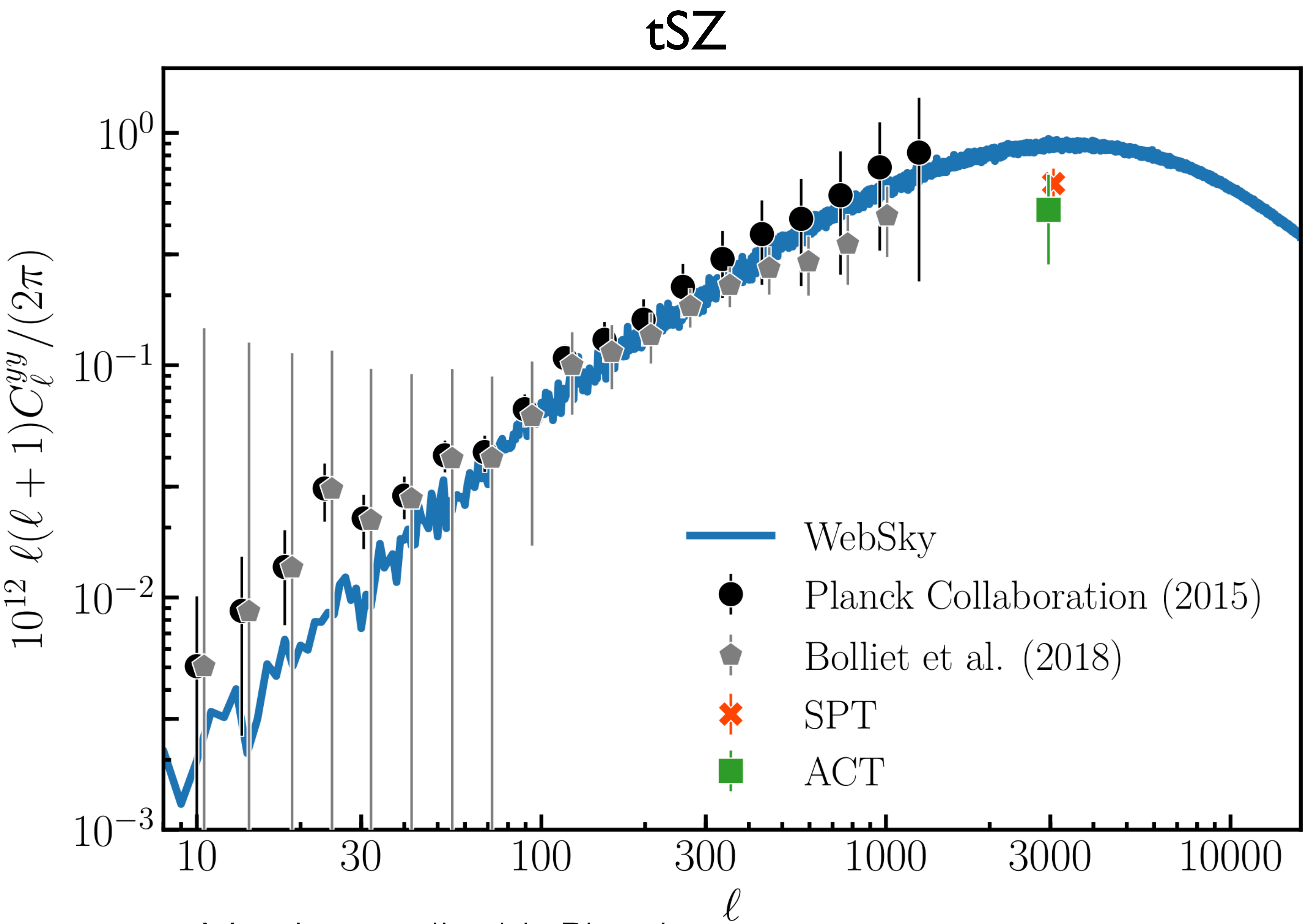
217 GHz

tSZ + kSZ + CIB



353 GHz

Validations: tSZ



- Matches well with Planck
- ~30% disagreement with ACT and SPT
  - small-scale suppression in the tSZ power?
  - systematic effects in the SPT and ACT measurements?

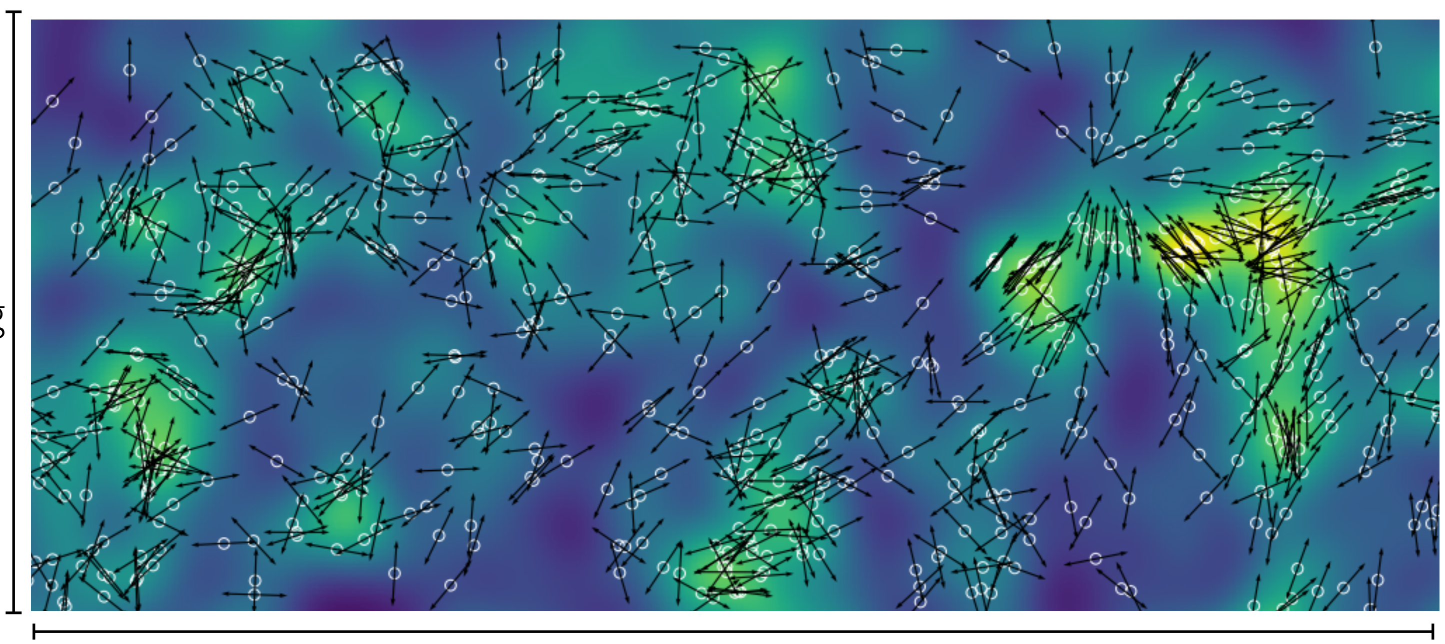
Cluster applications: tSZ stacking on DES Clusters

Preliminary

Martine Lokken, Mat Madhavacheril, Renee Hlozek, …

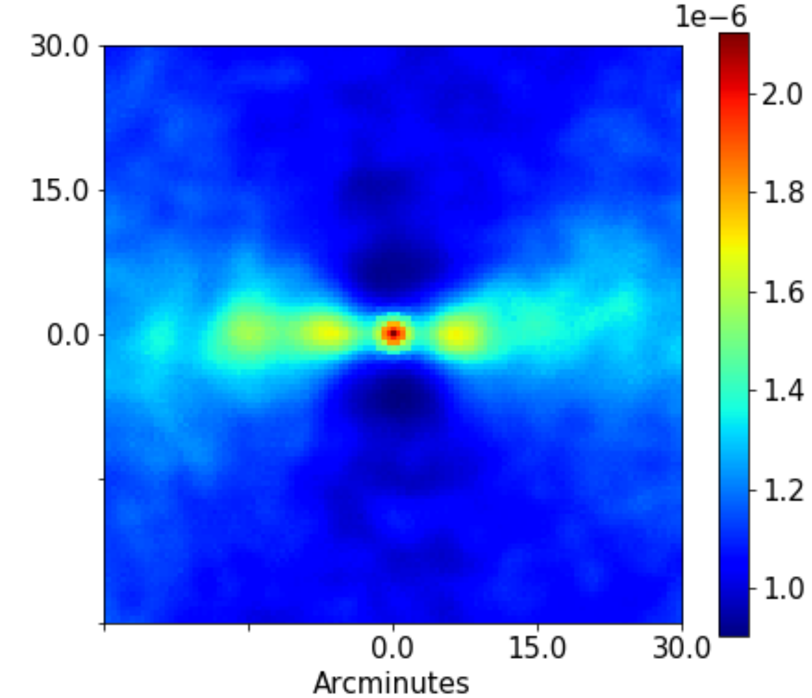
Halos

Orientation w.r.t LSS

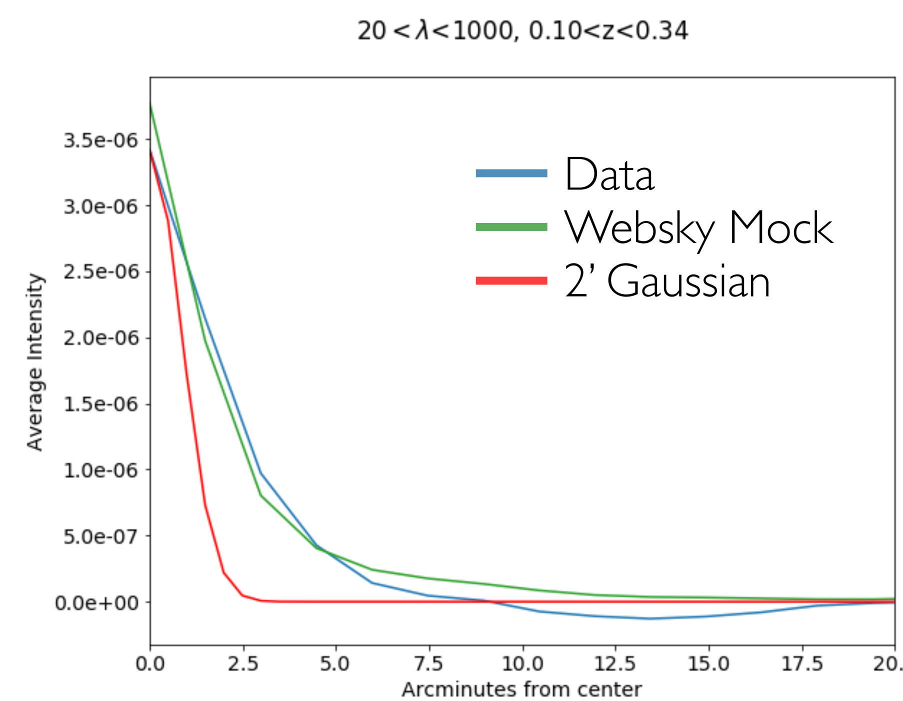


Websky halos

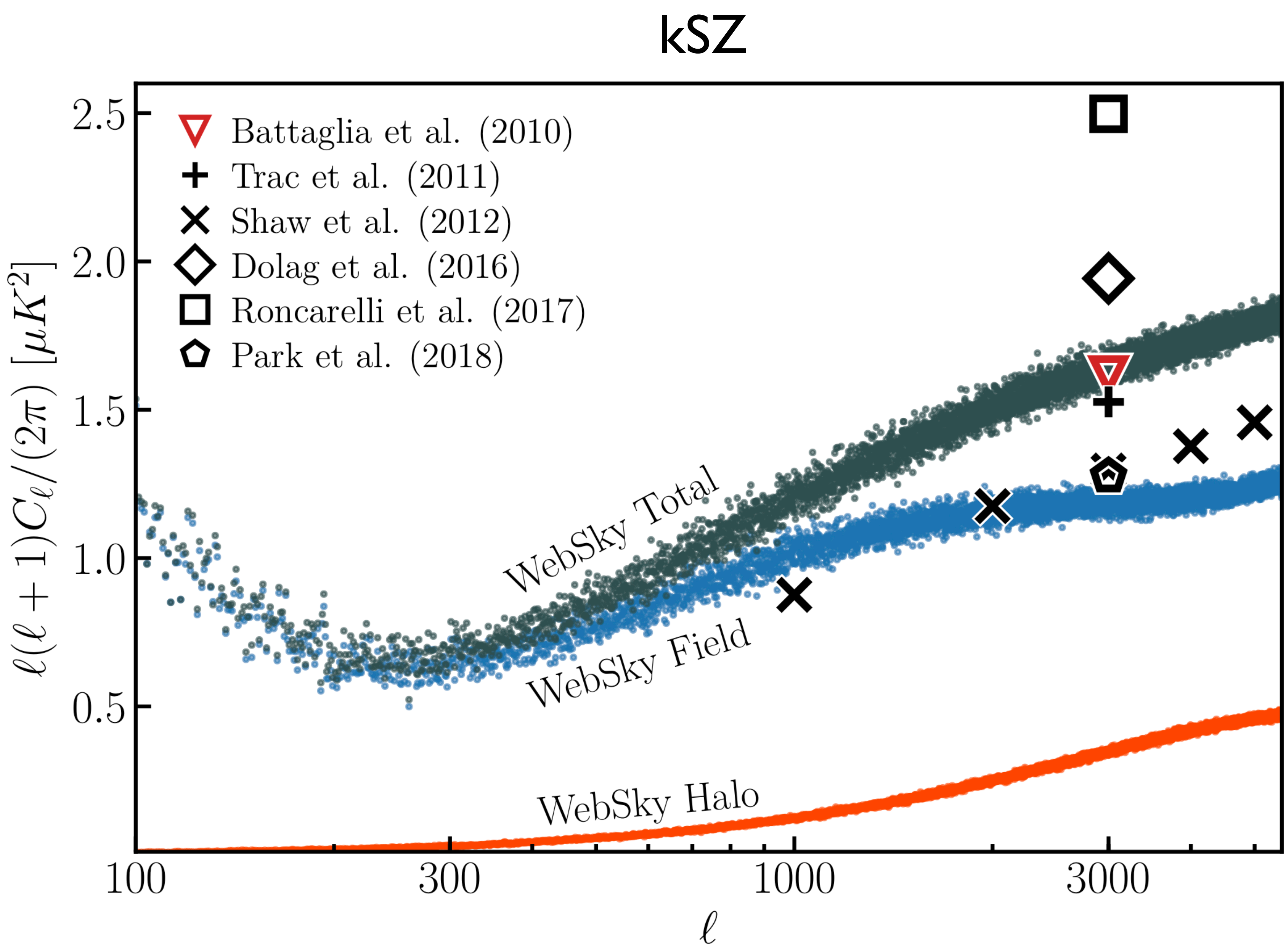
tSZ signal of Oriented Stack



Radial profile compared to data

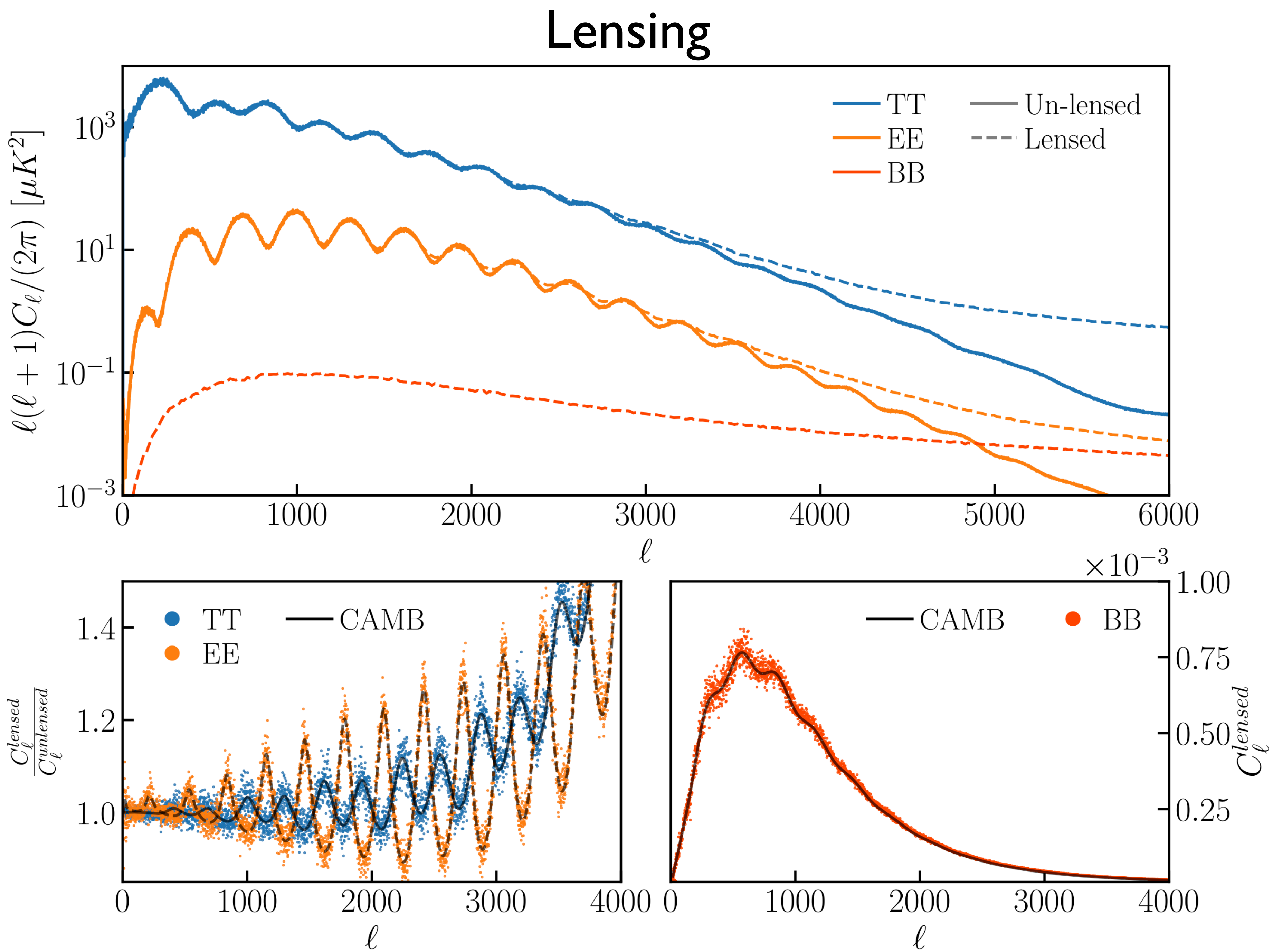


Validations: kSZ



- Agrees with various hydrodynamical studies
- Includes late-time contribution only. Early time kSZ can be added independently

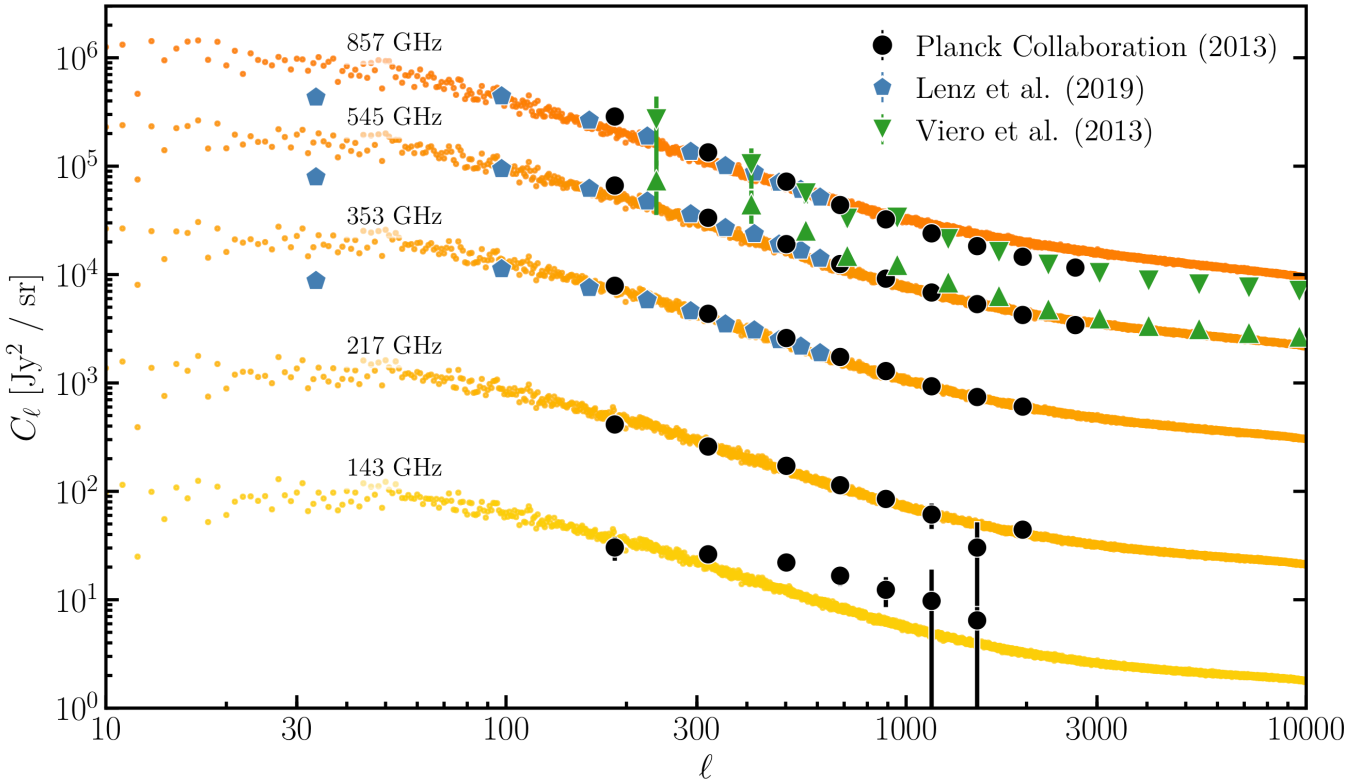
Validations: Weak Lensing



- Peak Smearing well reproduced
- r=0 in input CMB - feel free to lens your own with the convergence map

Validations: CIB

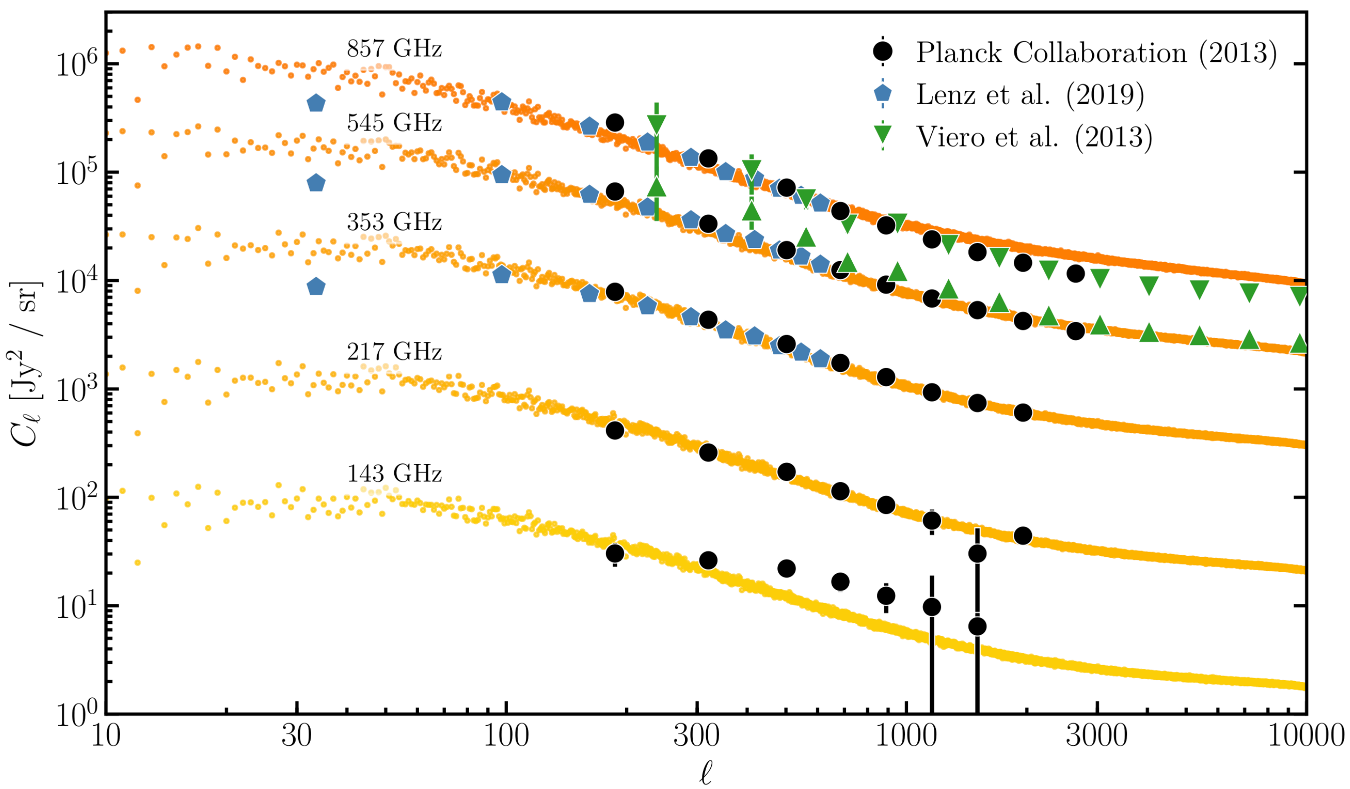
CIB Powerspectra



- Slight excess in small scale power at 857 GHz
  - $M_{h,min} \sim 1 \times 10^{12}\ M_\odot$

Validations: CIB

CIB Powerspectra



CIB Decoherence

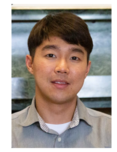
		857	545	353	217
	Websky	1	0.933 ±0.017	0.882 ±0.021	0.838 ±0.026
857	*Planck*	1	0.949 ±0.005	0.911 ±0.003	0.85 ±0.05
	Lenz et al.	1	0.96 ±0.01	0.91 ±0.01	0.85 ±0.05
	Websky	...	1	0.960 ±0.014	0.935 ±0.018
545	*Planck*	...	1	0.983 ±0.007	0.90 ±0.05
	Lenz et al.	...	1	0.98 ±0.01	
	Websky	...	...	1	0.968 ±0.014
353	*Planck*	...	...	1	0.91 ±0.05
	Lenz et al.	...	...	1	

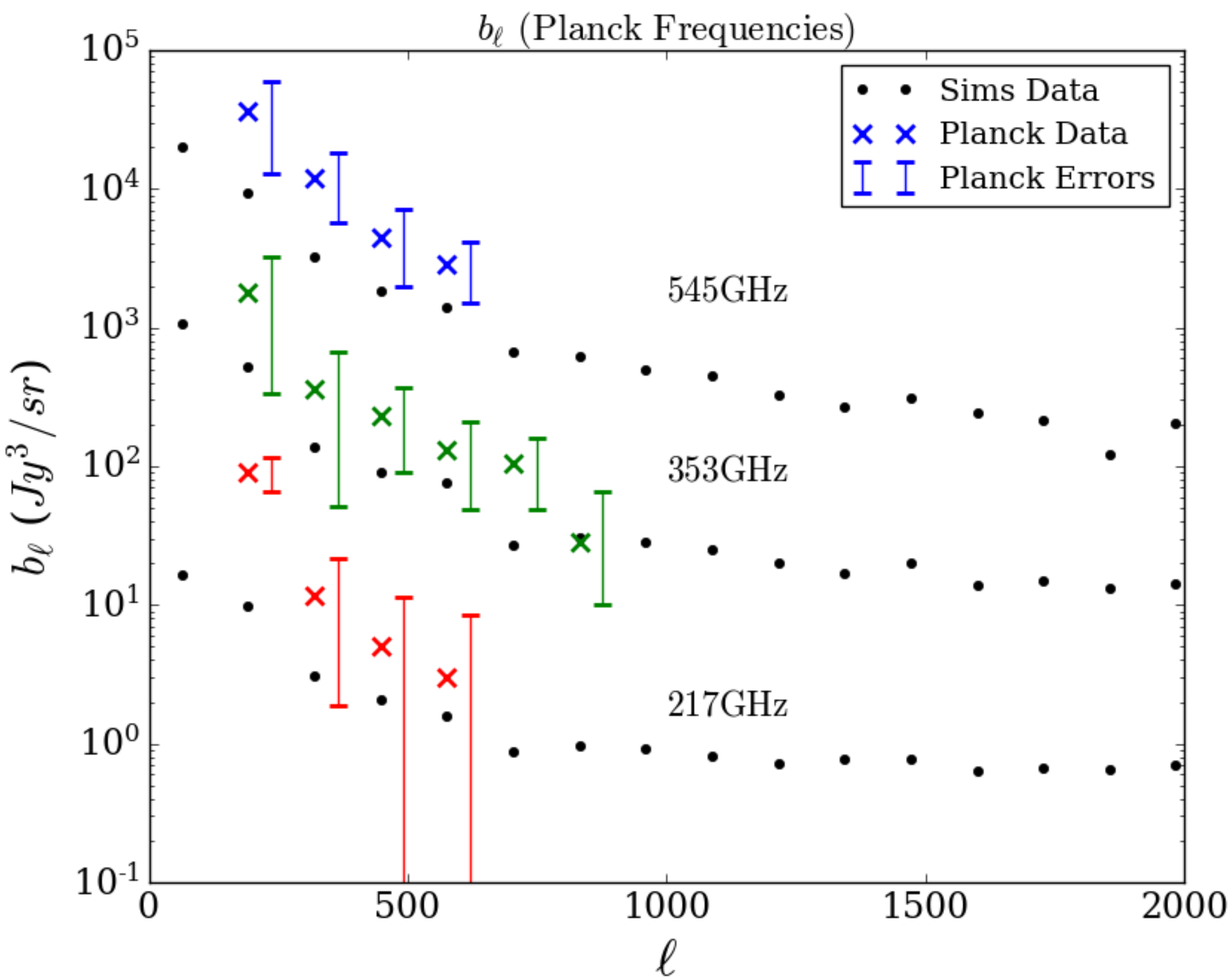
Table 1. Frequency decoherence of the CIB measured by averaging $C_\ell^{\nu\nu'}/(C_\ell^{\nu\nu}C_\ell^{\nu'\nu'})^{1/2}$ over the range $150 < \ell < 1000$. Error bars correspond to the standard deviation in this range. We include the Planck measurements of [16] and the Lenz et al. measurements of [100].

- Slight excess in small scale power at 857 GHz
  - $M_{h,min} \sim 1\times 10^{12}\ M_\odot$

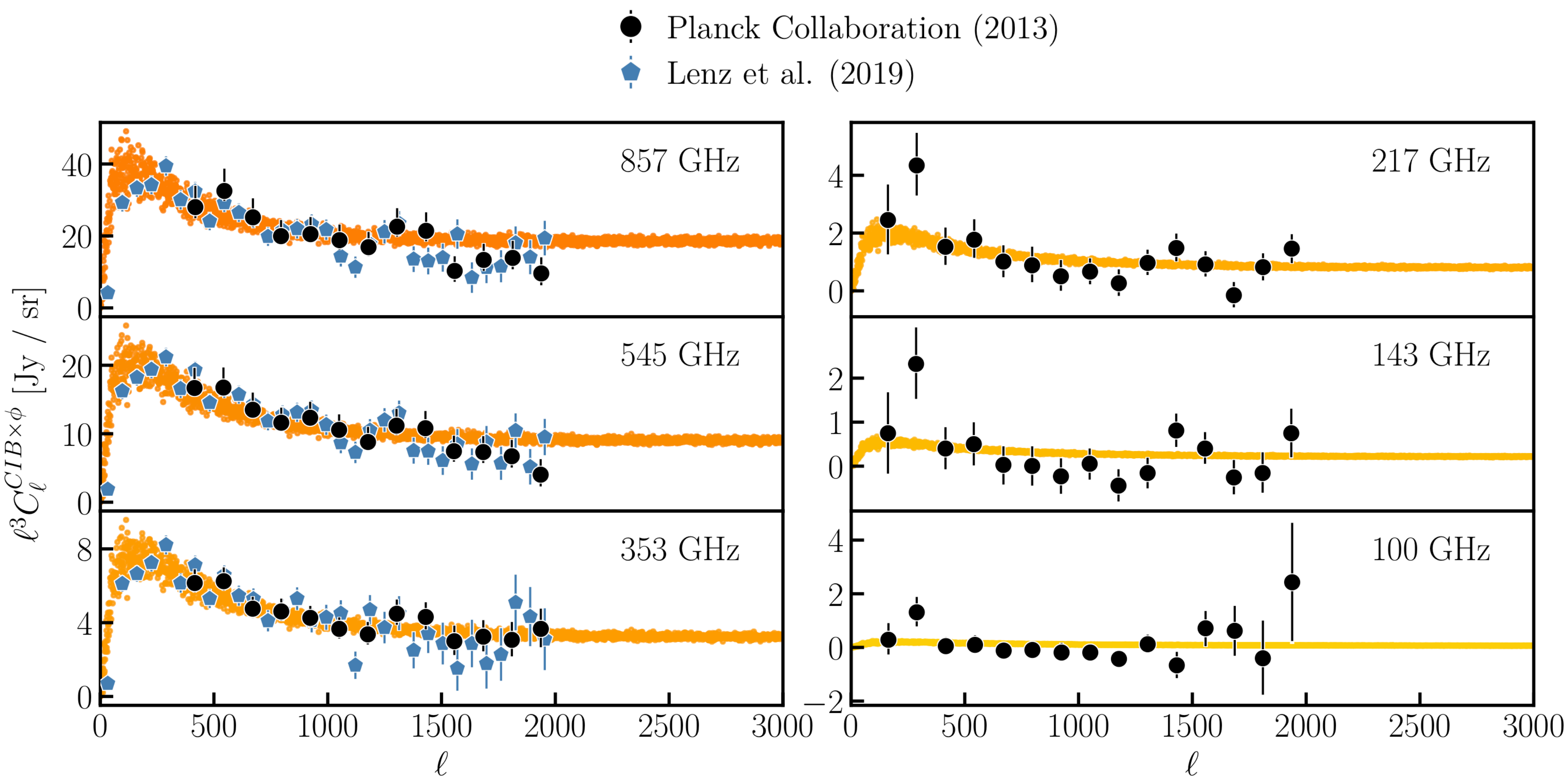
- Frequency dependence not just an amplitude scaling

Validations: CIB Bispectrum

Preliminary  **Jason Lee**, Pavel Motloch, …



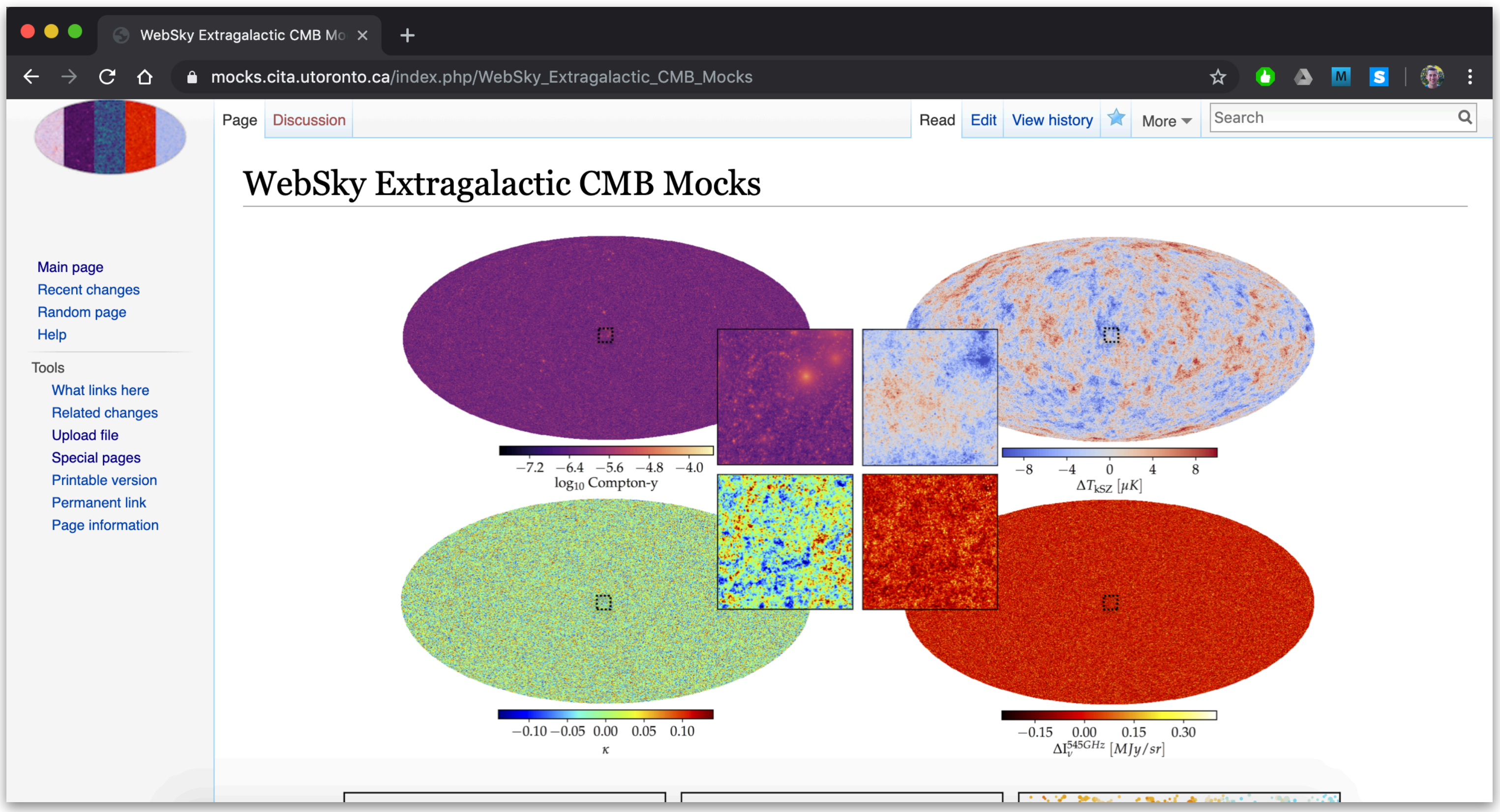
Validations: CIB x lensing



- Cross correlations to date look good

The Current Generation of CMB Extragalactic Simulations

All data publicly available @
mocks.cita.utoronto.ca



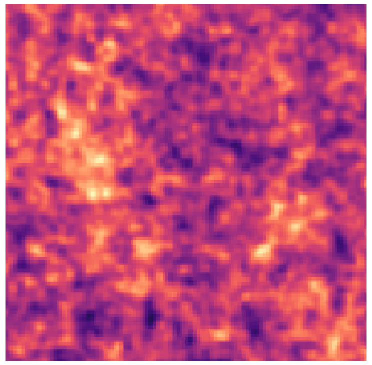
Ask away @
websky-cmb

slack

The Current Generation of CMB Extragalactic Simulations

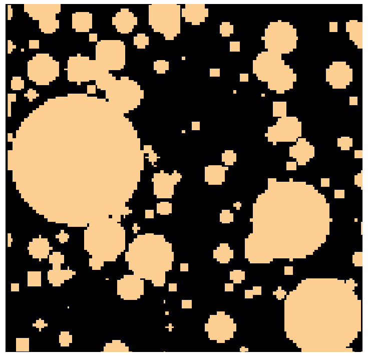
- Minimum halo mass of $\sim 1 \times 10^{12}\ M_\odot$ for full-sky simulation

Initial Conditions



Physical Approximation

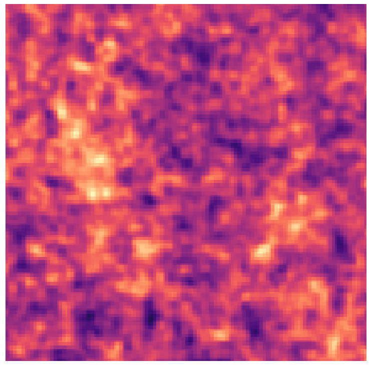
Final Halos/Matter Field



The Current Generation of CMB Extragalactic Simulations

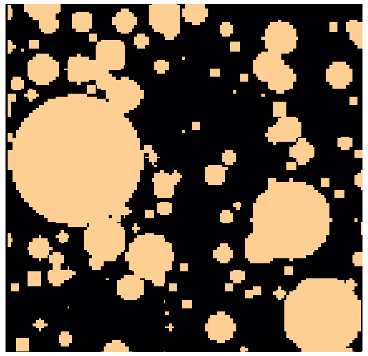
- Minimum halo mass of ~$1 \times 10^{12}\ M_{\odot}$ for full-sky simulation

Initial Conditions



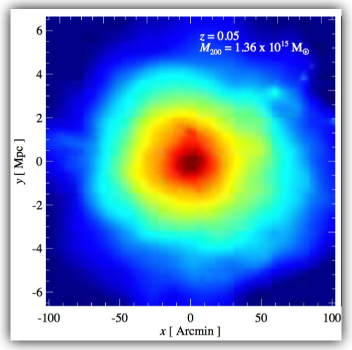
Physical Approximation

Final Halos/Matter Field

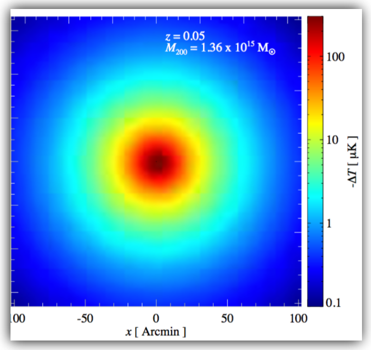


- Spherical halo profiles fit to small box hydrodynamical simulations used

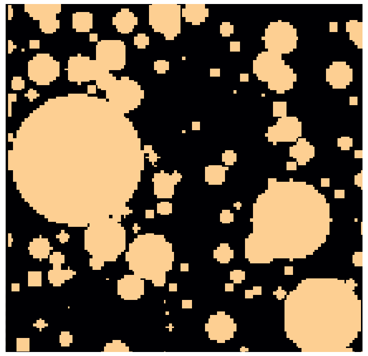
Full Simulation



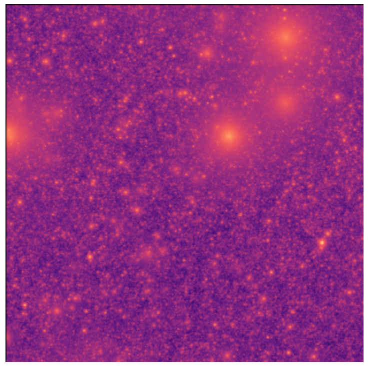
Measure Fitting Functions



Fast Simulation



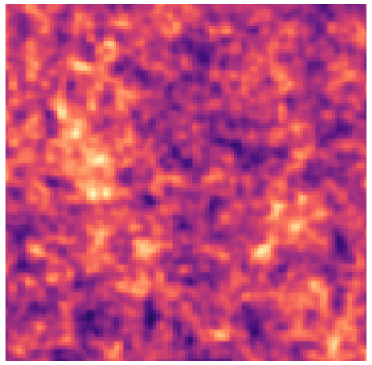
Mock Observation

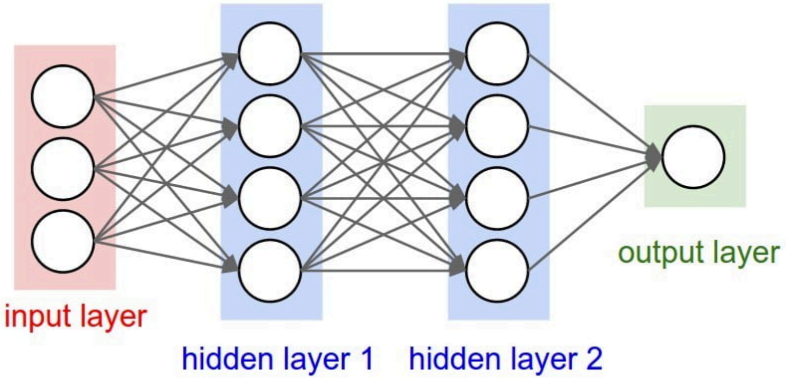


The Next Generation of Simulations?

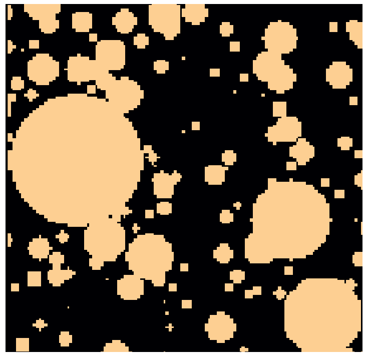
- Minimum halo mass of << $1\times 10^{12}\,M_\odot$ for full-sky simulation

Initial Conditions





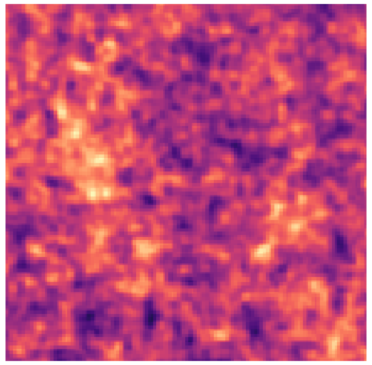
Final Conditions

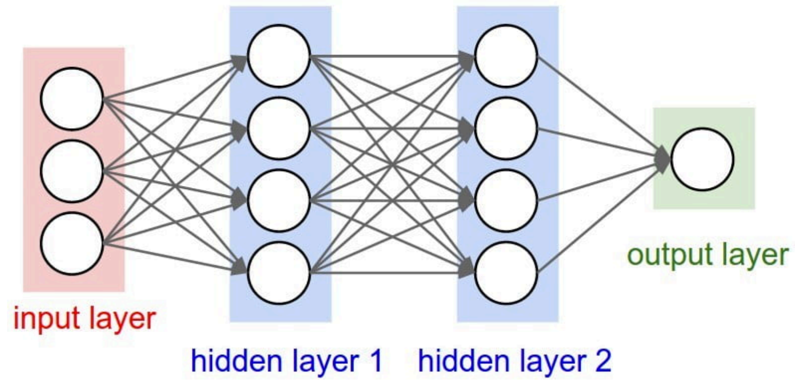


The Next Generation of Simulations?

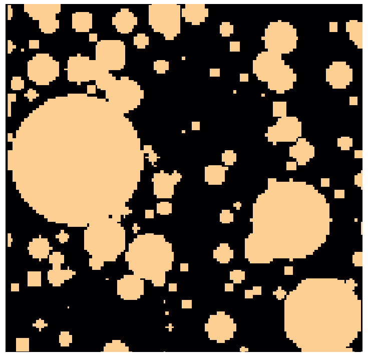
- Minimum halo mass of $<< 1\times 10^{12}\,M_\odot$ for full-sky simulation

Initial Conditions



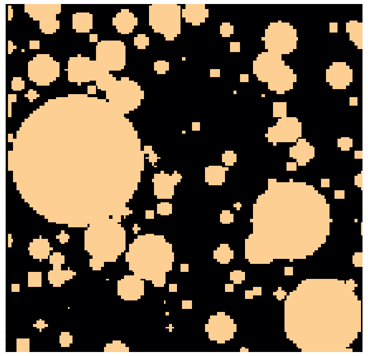


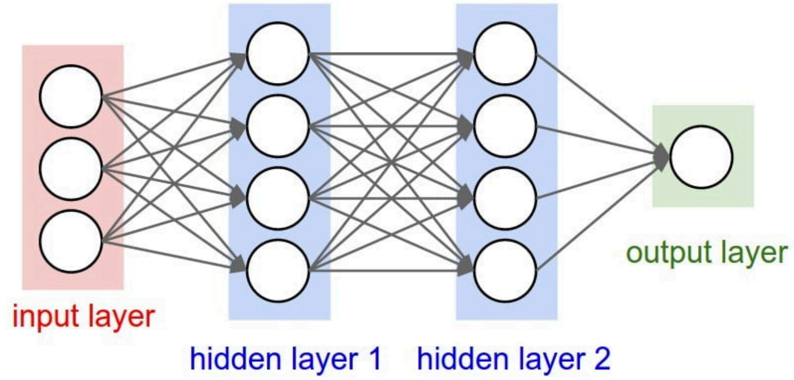
Final Conditions



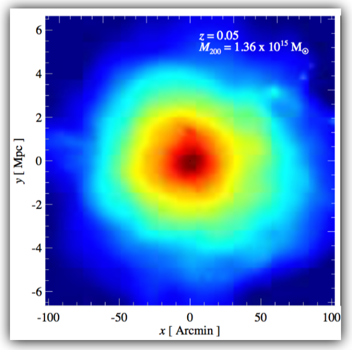
- Direct approximate method to mock observable mapping

Fast Simulation

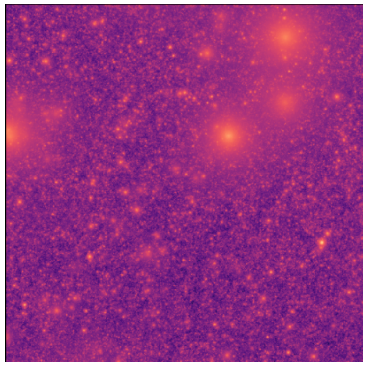




Full Simulation



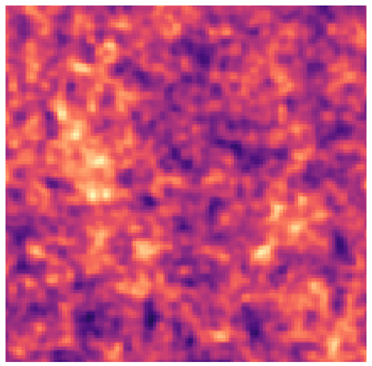
Mock Observation

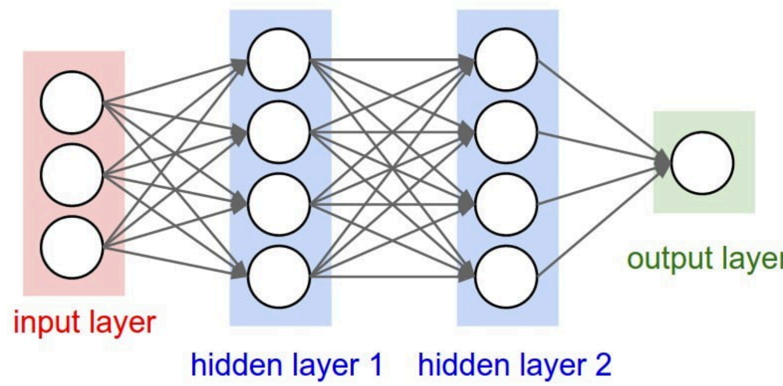


The Next Generation of Simulations?

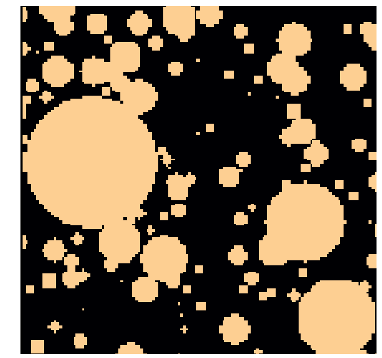
- Minimum halo mass of << $1 \times 10^{12}\ M_\odot$ for full-sky simulation

Initial Conditions

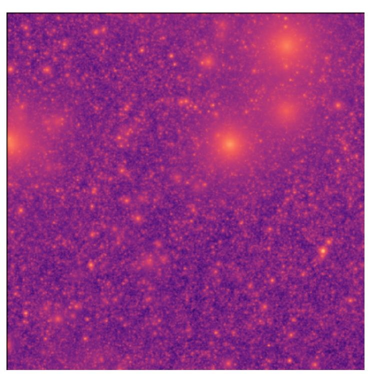




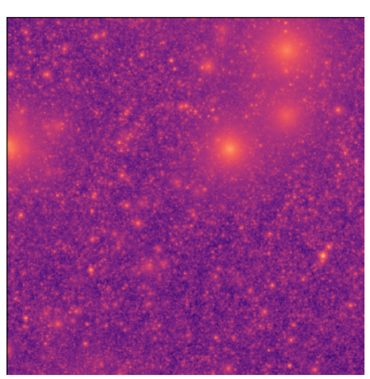
Final Conditions

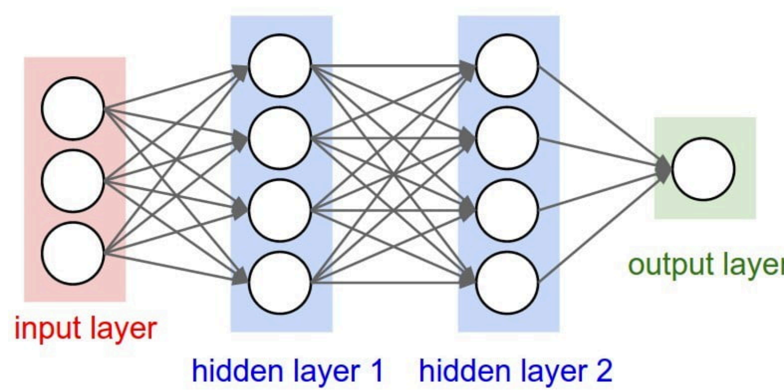


Mock Observation



True Observation

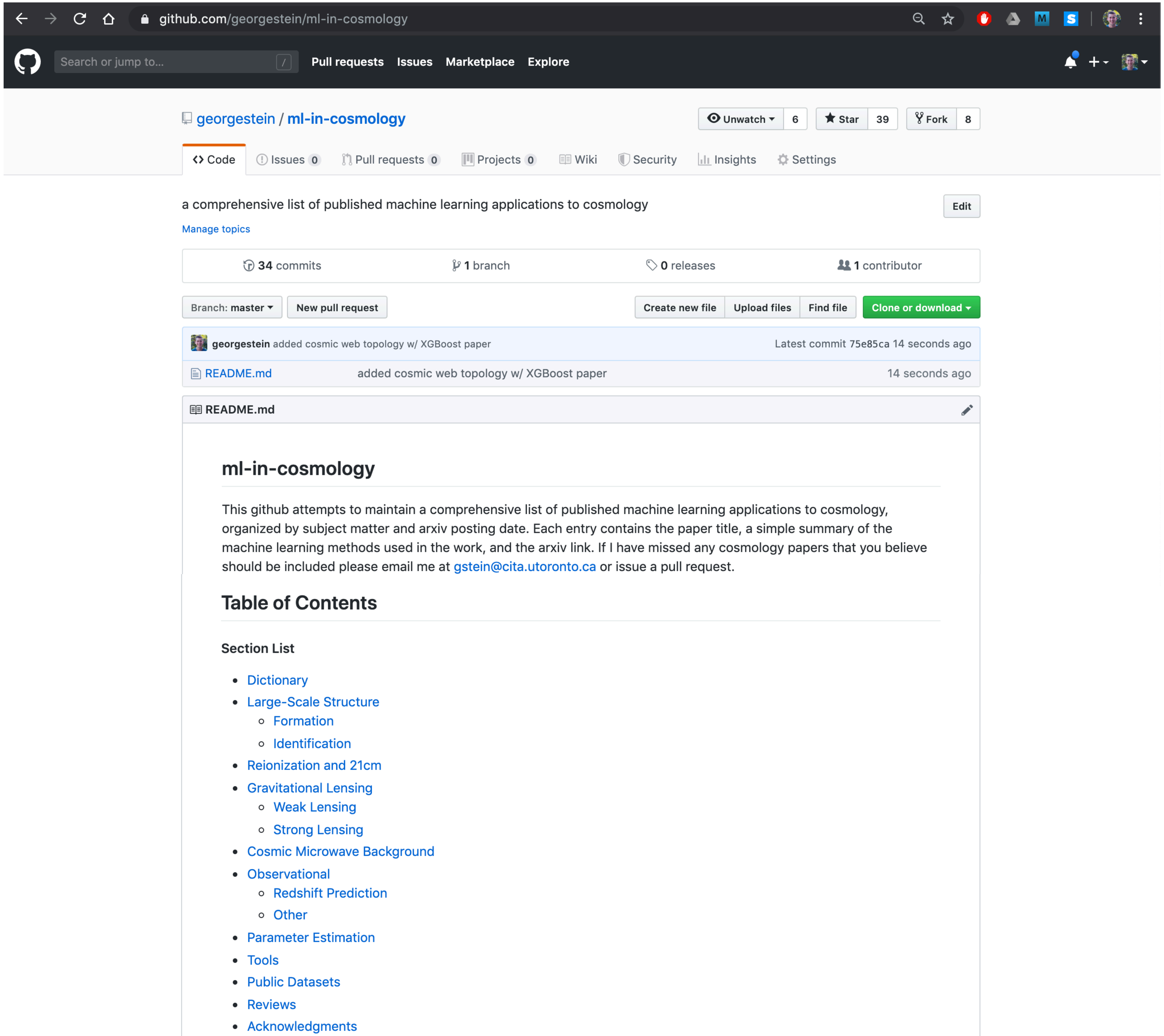




Cosmological Information

Machine Learning in Cosmology https://github.com/georgestein/ml-in-cosmology

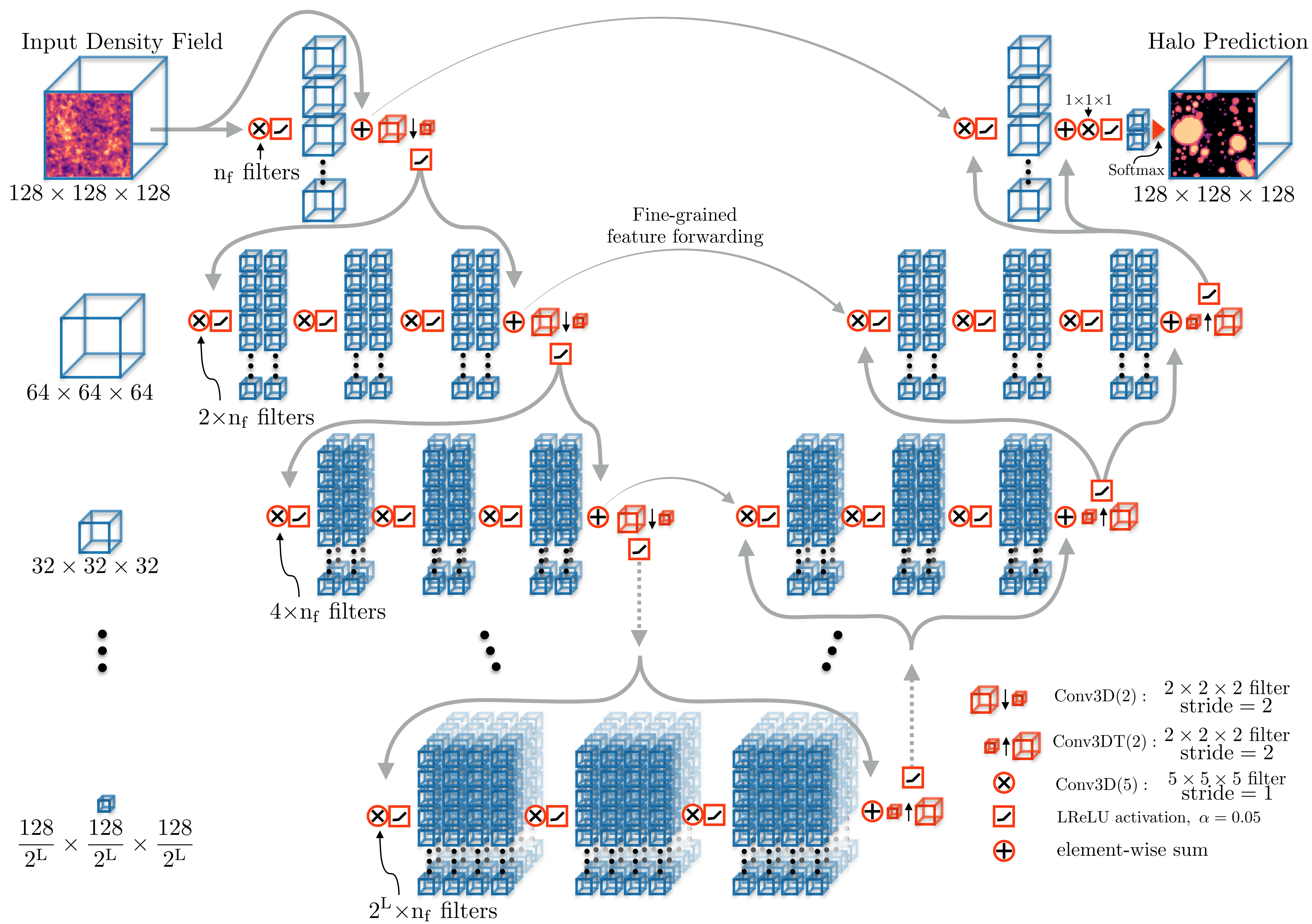
~140 papers



Convolutional Neural Networks for LSS

HaloNet - First application of a volumetric deep Convolutional Neural Network (CNN) for large scale structure simulations
Berger & Stein - arXiv 1805.04537

V-Net = 3D U-Net
Milletari, Navab, Ahmadi: 1606.04797
U-Net
Ronneberger, Fischer, Brox: 1505.04597

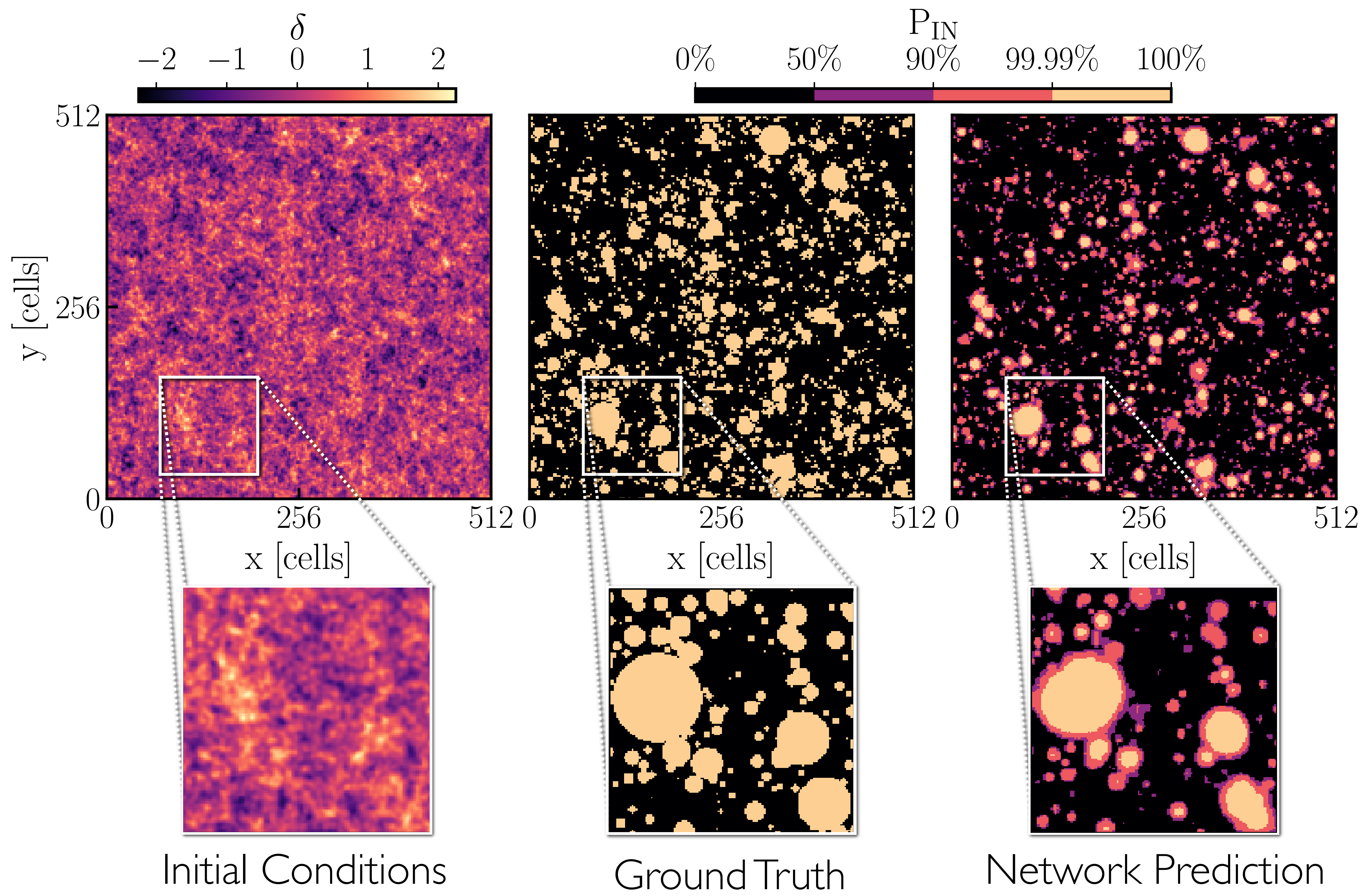


George Stein

Accurately find halos/collapsed regions in Lagrangian space

HaloNet - First application of a volumetric deep Convolutional Neural Network (CNN) for large scale structure simulations

Berger & Stein - arXiv 1805.04537



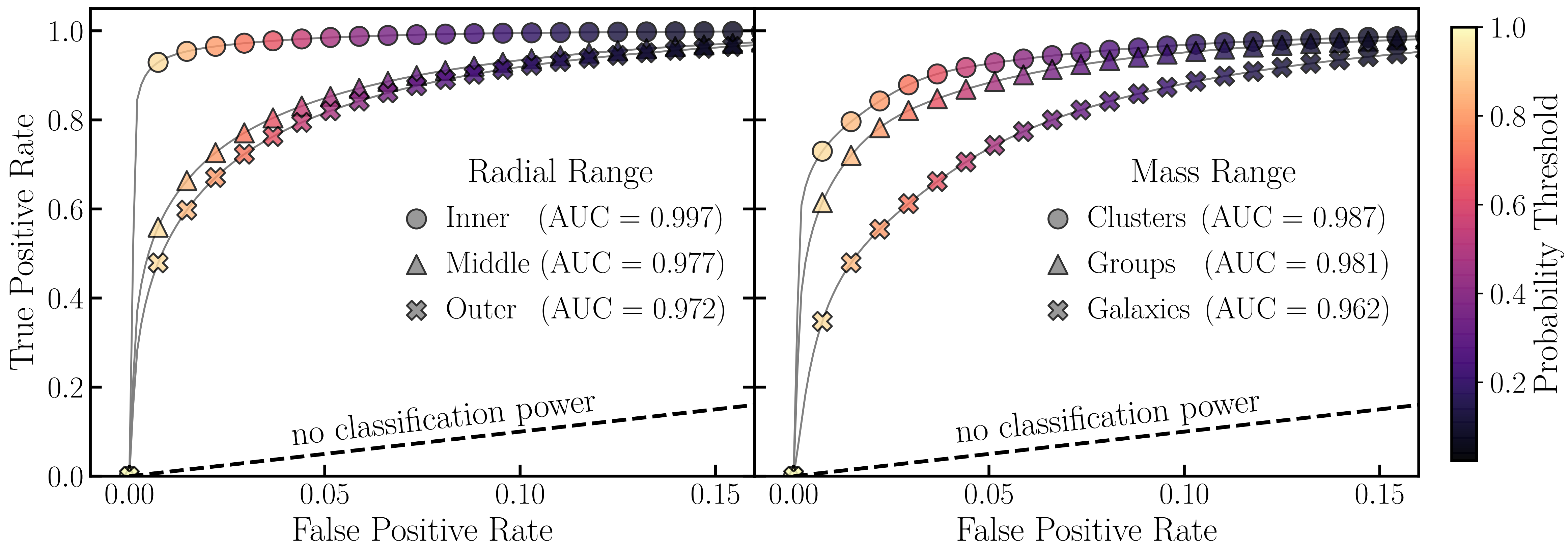
George Stein

Quantifying the network accuracy

receiver operating characteristic (ROC) curve

$$\text{True Positive Rate} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad \text{(Sensitivity)}$$

$$\text{False Positive Rate} = \frac{\text{False Positives}}{\text{False Positives} + \text{True Negatives}} \quad \text{(1- Specificity)}$$

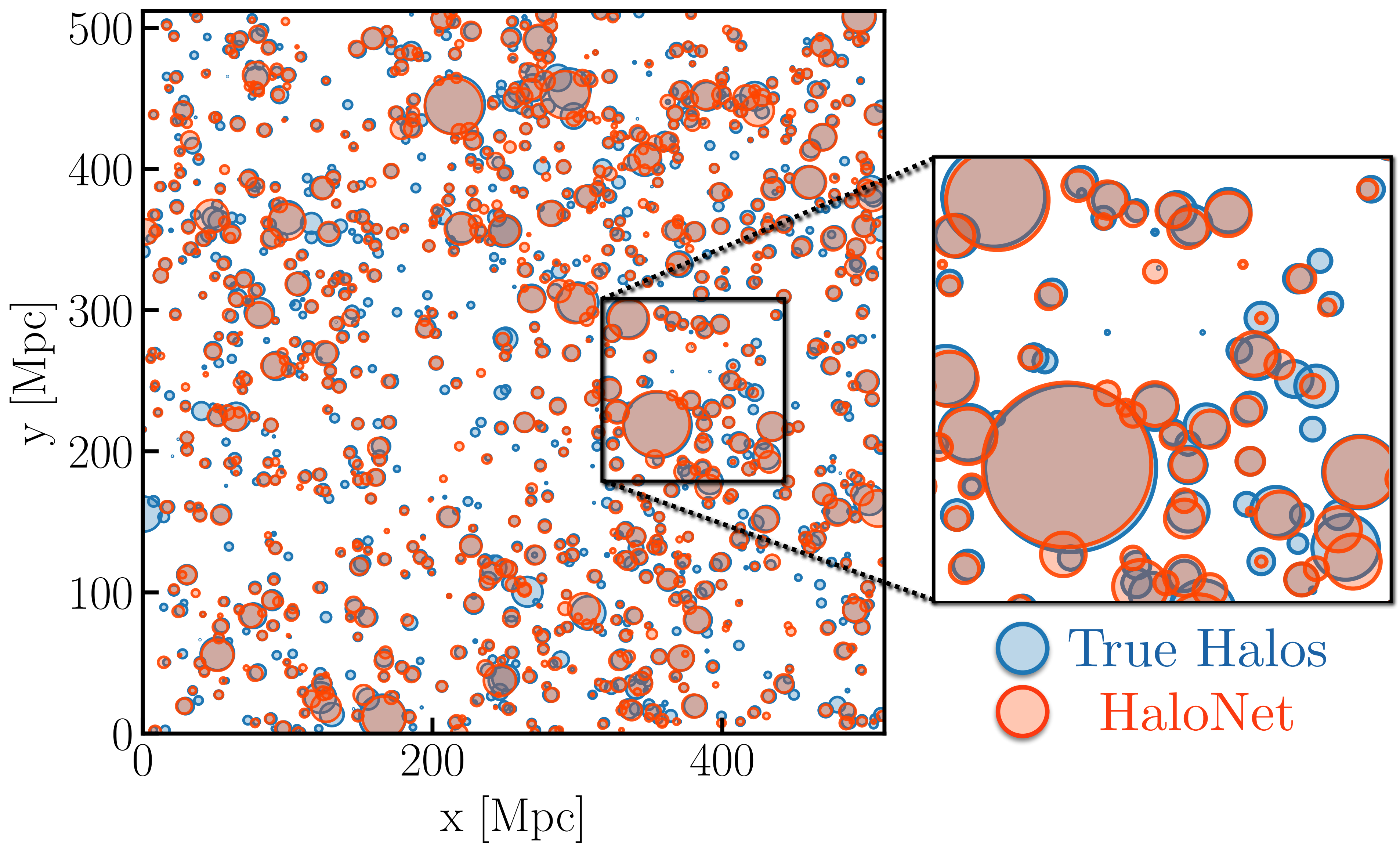


- Quick predictive halo finder, move halos with LPT for final halo catalogue, simple predictive method

Apply Lagrangian halo finder

- Developed simple Lagrangian halo finder
- Move halos with LPT for final halo catalogue,

= simple predictive method



Things to overcome for many current ML LSS methods

- Single (or low) simulation resolution
- Single redshift
- Single set of cosmological/astrophysical parameters

...			
ICs to halos	*Cosmological Reconstruction From Galaxy Light: Neural Network Based Light-Matter Connection*	NN	https://arxiv.org/abs/1805.02247
	A volumetric deep Convolutional Neural Network for simulation of mock dark matter halo catalogues	V-Net	https://arxiv.org/abs/1805.04537
corrections to LPT displacements	*Learning to Predict the Cosmological Structure Formation*	V-Net	https://arxiv.org/abs/1811.06533
	deepCool: Fast and Accurate Estimation of Cooling Rates in Irradiated Gas with Artificial Neural Networks	NN, RF, kNN	https://arxiv.org/abs/1901.01264
DM to galaxies	*From Dark Matter to Galaxies with Convolutional Networks*	V-Net	https://arxiv.org/abs/1902.05965
DM to halos	*Painting halos from 3D dark matter fields using Wasserstein mapping networks*	GAN	https://arxiv.org/abs/1903.10524
DM to pressure (2D)	*Painting with baryons: augmenting N-body simulations with gas using deep generative models*	GAN, VAE	https://arxiv.org/abs/1903.12173
HI cube generation	*HIGAN: Cosmic Neutral Hydrogen with Generative Adversarial Networks*	GAN	https://arxiv.org/abs/1904.12846
	A deep learning model to emulate simulations of cosmic reionization	CNN	https://arxiv.org/abs/1905.06958
	An interpretable machine learning framework for dark matter halo formation	BDT	https://arxiv.org/abs/1906.06339
DM cube generation + resolution upscaling	*Cosmological N-body simulations: a challenge for scalable generative models*	GAN	https://arxiv.org/abs/1908.05519

Things to overcome for many current ML LSS methods

- Single (or low) simulation resolution
- Single redshift
- Single set of cosmological/astrophysical parameters

...			
	Cosmological Reconstruction From Galaxy Light: Neural Network Based Light-Matter Connection	NN	https://arxiv.org/abs/1805.02247
ICs to halos	*A volumetric deep Convolutional Neural Network for simulation of mock dark matter halo catalogues*	V-Net	https://arxiv.org/abs/1805.04537
corrections to LPT displacements	*Learning to Predict the Cosmological Structure Formation*	V-Net	https://arxiv.org/abs/1811.06533
	deepCool: Fast and Accurate Estimation of Cooling Rates in Irradiated Gas with Artificial Neural Networks	NN, RF, kNN	https://arxiv.org/abs/1901.01264
DM to galaxies	*From Dark Matter to Galaxies with Convolutional Networks*	V-Net	https://arxiv.org/abs/1902.05965
DM to halos	*Painting halos from 3D dark matter fields using Wasserstein mapping networks*	GAN	https://arxiv.org/abs/1903.10524
DM to pressure (2D)	*Painting with baryons: augmenting N-body simulations with gas using deep generative models*	GAN, VAE	https://arxiv.org/abs/1903.12173
HI cube generation	*HIGAN: Cosmic Neutral Hydrogen with Generative Adversarial Networks*	GAN	https://arxiv.org/abs/1904.12846
	A deep learning model to emulate simulations of cosmic reionization	CNN	https://arxiv.org/abs/1905.06958
	An interpretable machine learning framework for dark matter halo formation	BDT	https://arxiv.org/abs/1906.06339
DM cube generation + resolution upscaling	*Cosmological N-body simulations: a challenge for scalable generative models*	GAN	https://arxiv.org/abs/1908.05519

Simulation to Simulation

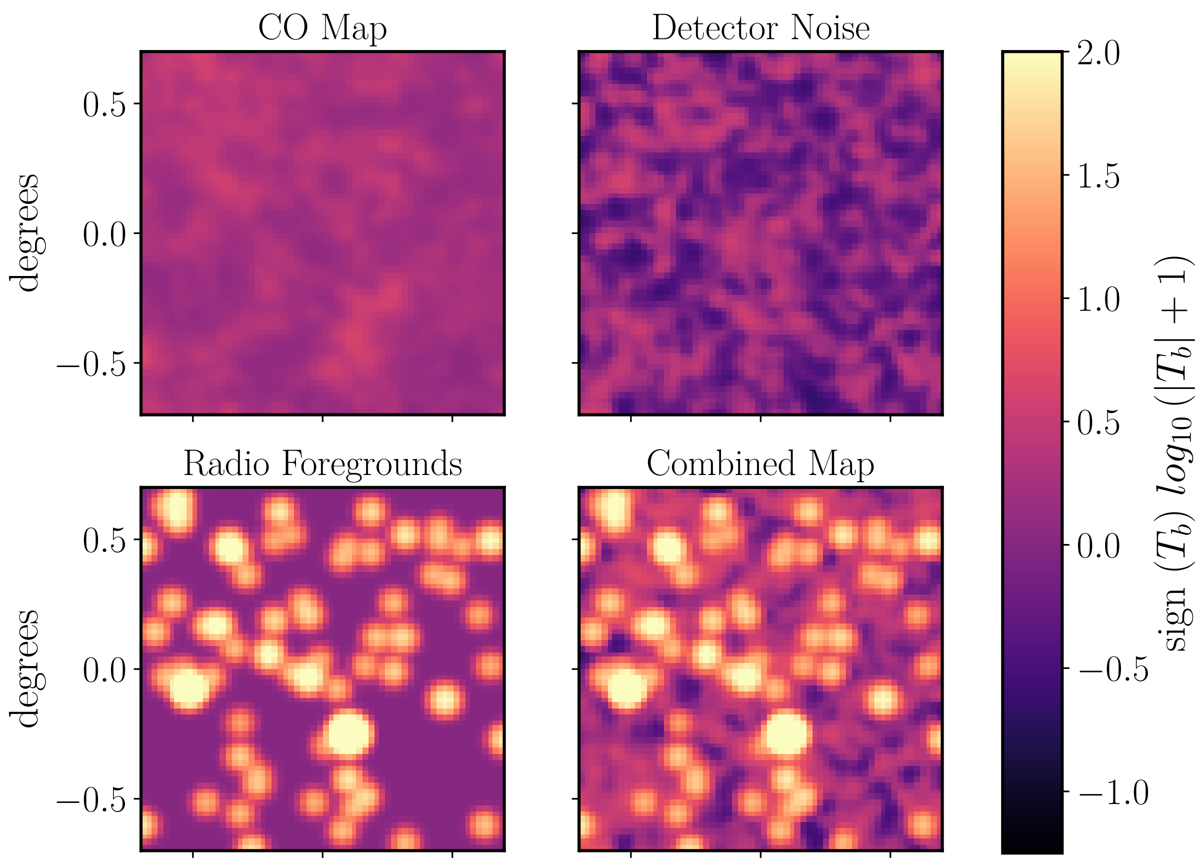
True answer known, full control, and ability to validate.

Simulation to Data to Paramaters

Making the Network robust to input

Deconfusing Intensity Maps with Neural Networks
Pfeffer, Breysse, Stein - arXiv 1905.10376

1.) Construct suite of mock CO intensity maps using one luminosity model and various levels of gaussian noise and foregrounds



Simulation to Data to Paramaters

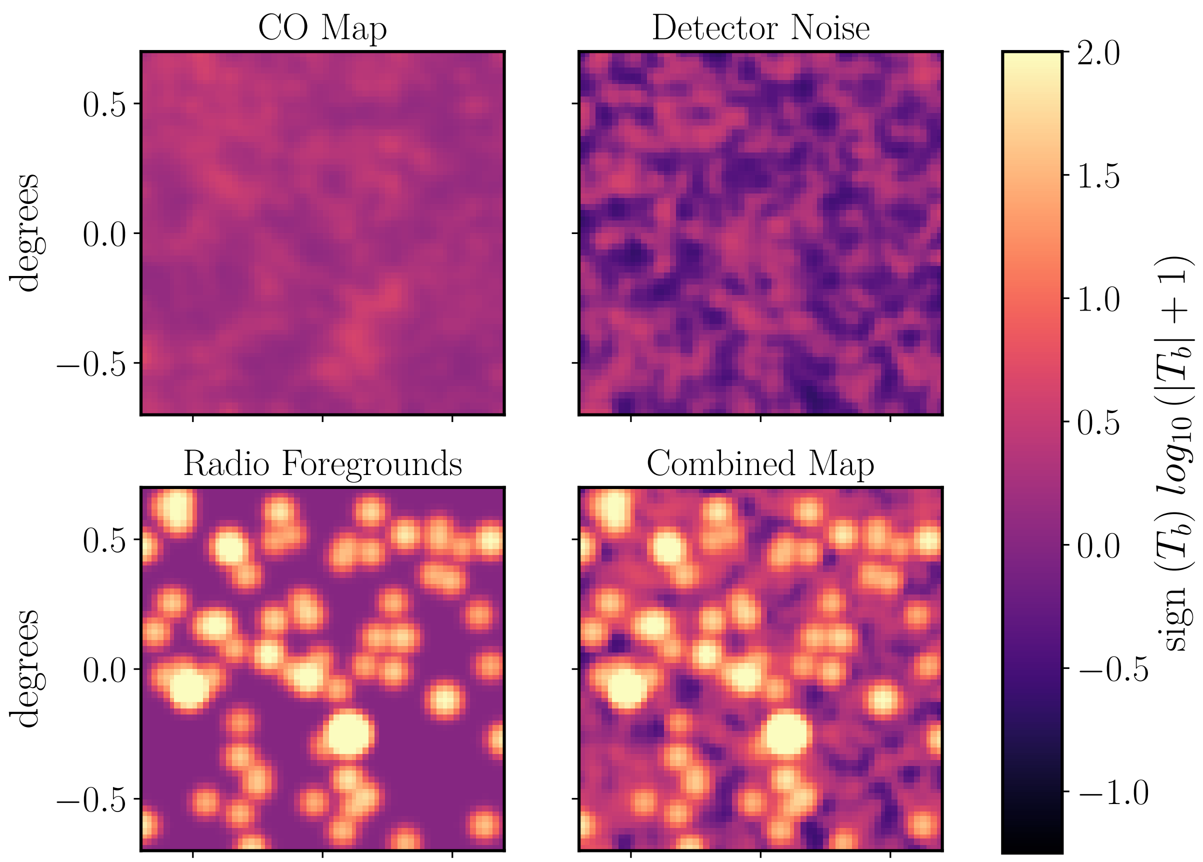
Making the Network robust to input

Deconfusing Intensity Maps with Neural Networks

Pfeffer, Breysse, Stein - arXiv 1905.10376

1.) Construct suite of mock CO intensity maps using one luminosity model and various levels of gaussian noise and foregrounds

2.) Train ResNet to pull out luminosity function from mocks



Simulation to Data to Paramaters

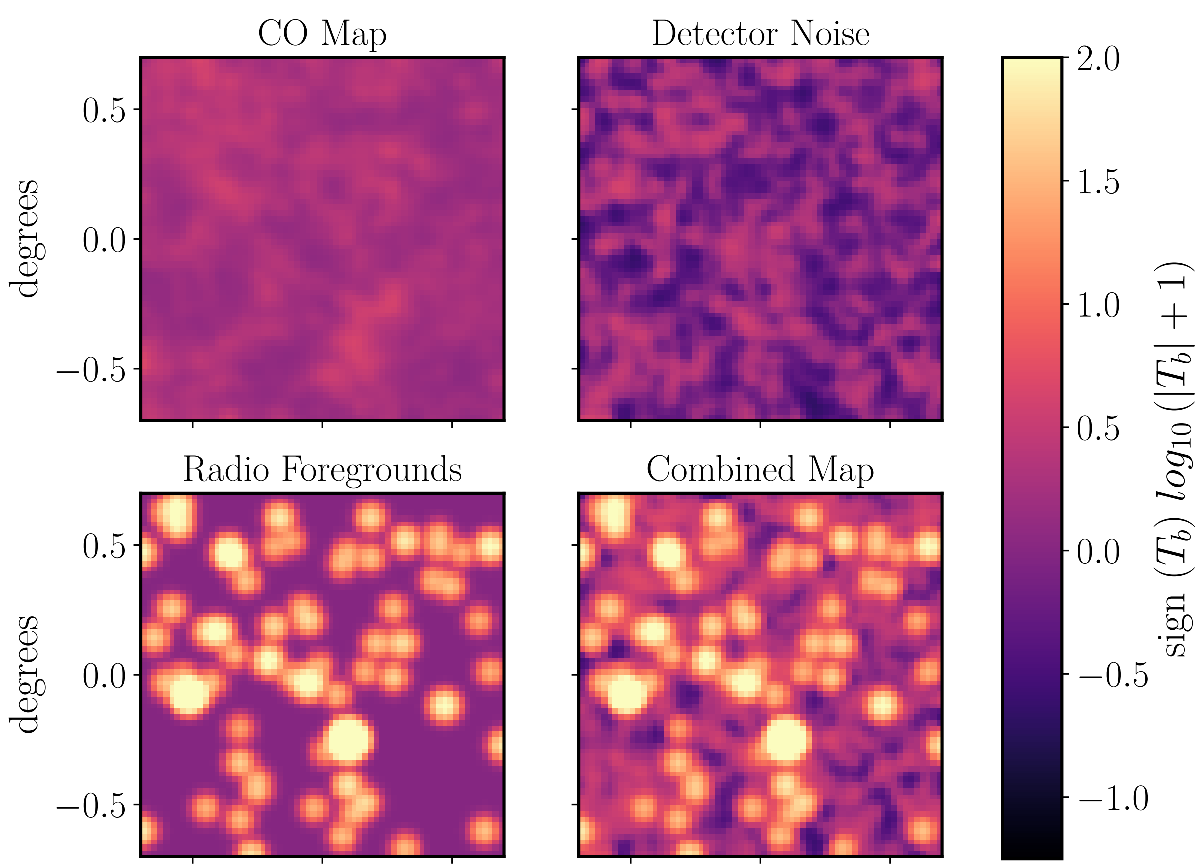
Making the Network robust to input

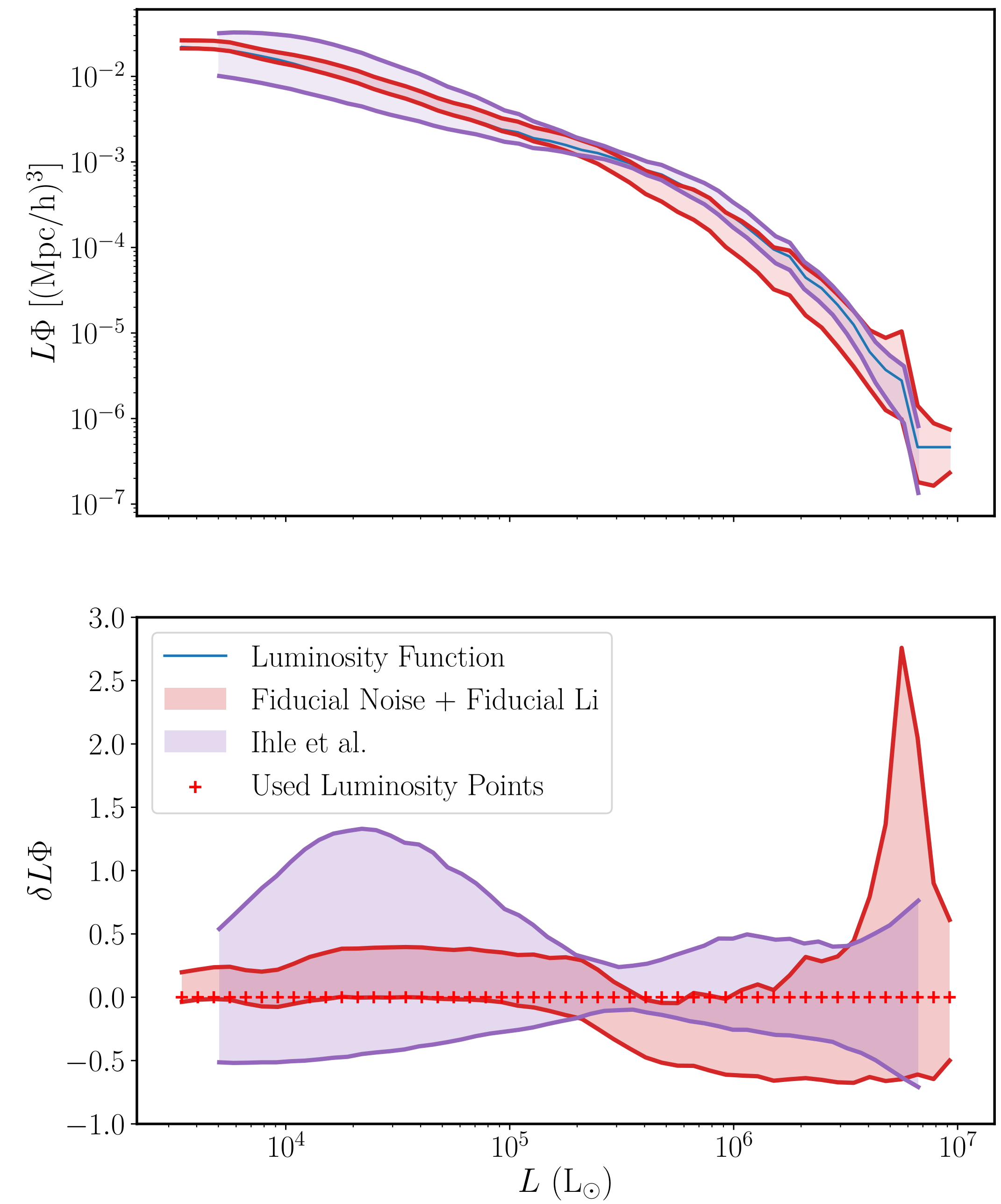
Deconfusing Intensity Maps with Neural Networks
Pfeffer, Breysse, Stein - arXiv 1905.10376

1.) Construct suite of mock CO intensity maps using one luminosity model and various levels of gaussian noise and foregrounds

2.) Train ResNet to pull out luminosity function from mocks

3.) Extract luminosity function from mocks





comparable with standard P(k) + PDF methods

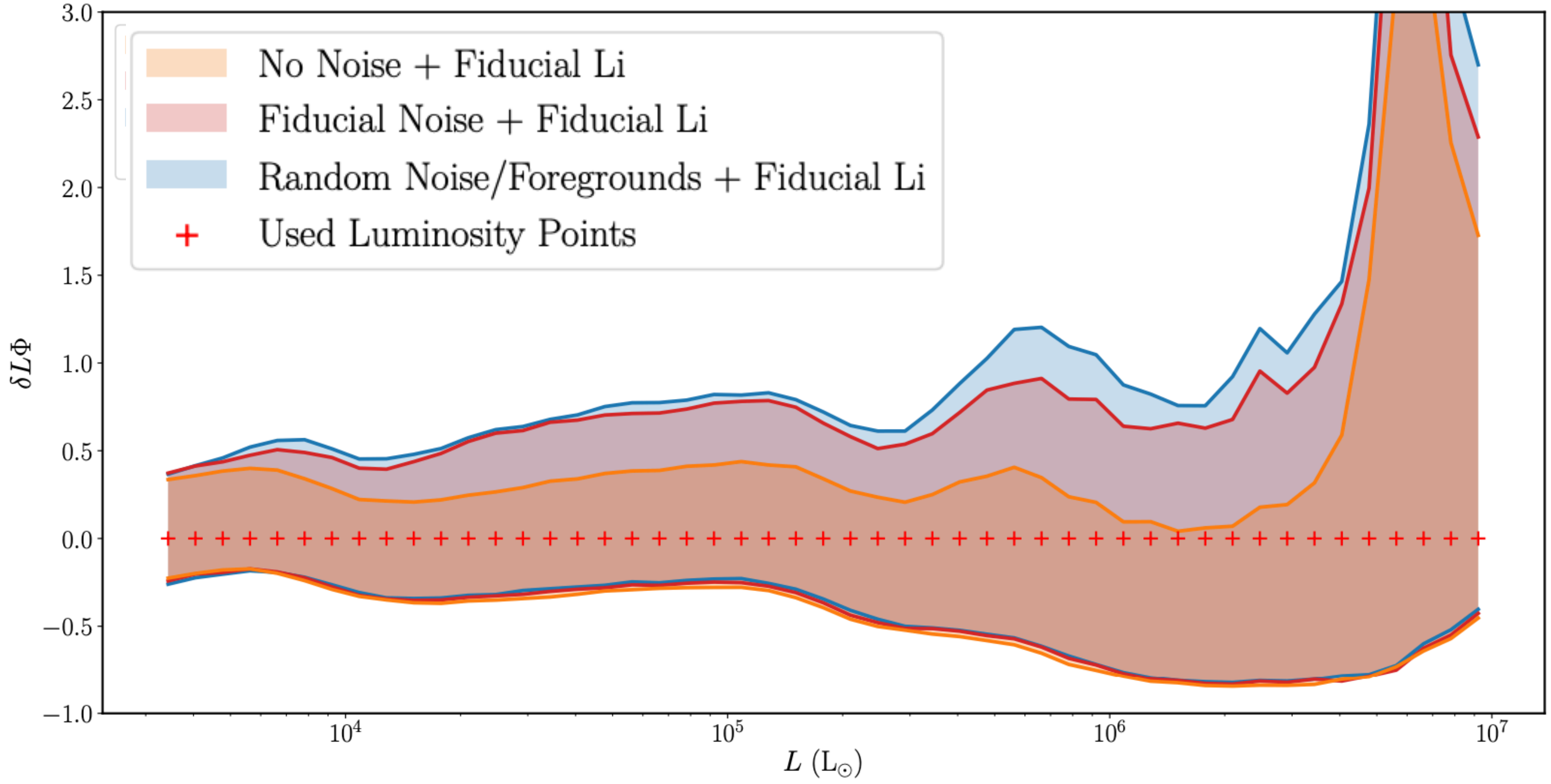
Simulation to Data to Paramaters

Making the Network robust to input

Deconfusing Intensity Maps with Neural Networks

Pfeffer, Breysse, Stein - arXiv 1905.10376

Network can still perform with various levels of noise and foregrounds that were included in training set



George Stein

Simulation to Data to Paramaters

Making the Network robust to input

Deconfusing Intensity Maps with Neural Networks
Pfeffer, Breysse, Stein - arXiv 1905.10376

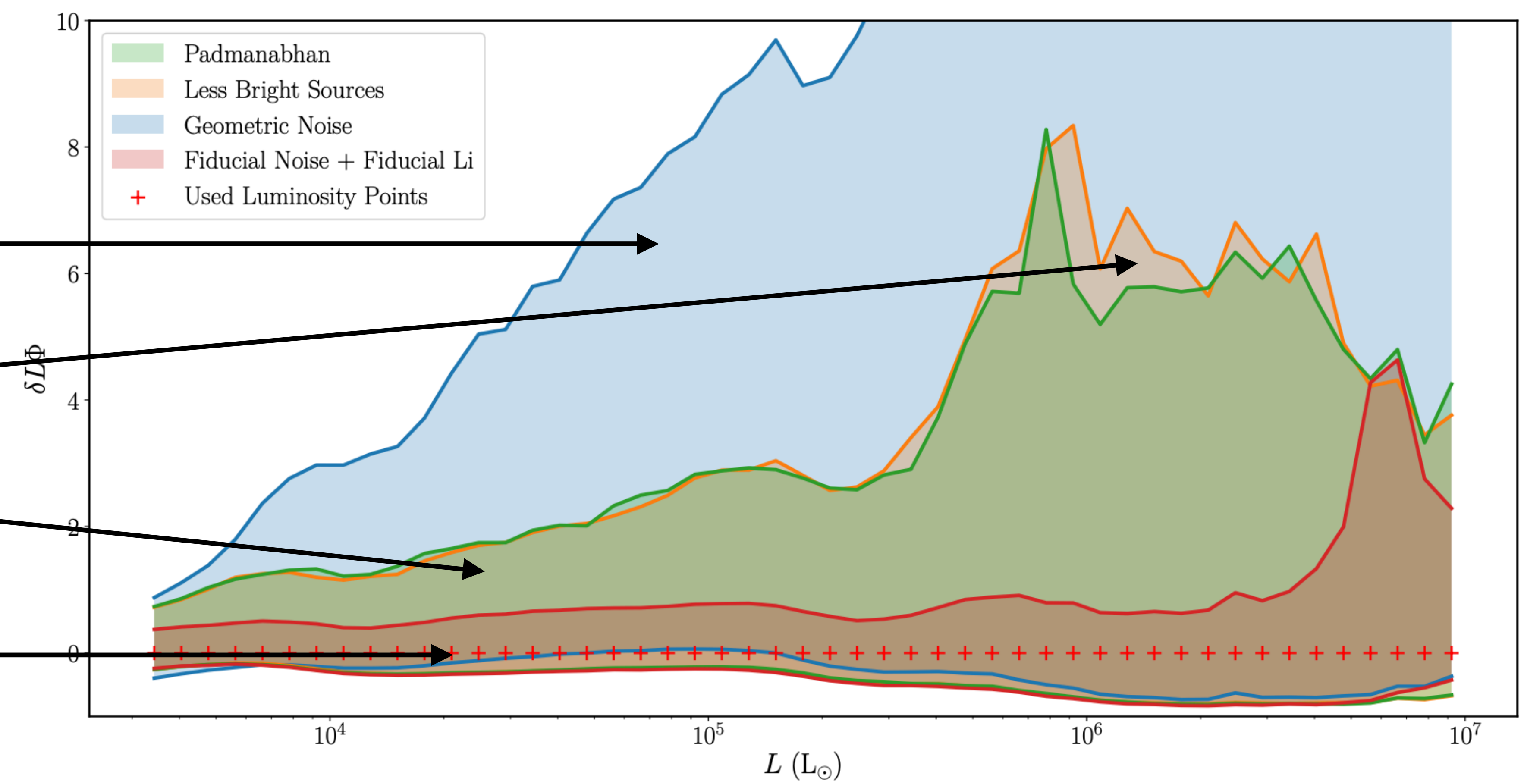
Network completely fails for scenarios not included in training set

Add more noise near edges of survey

Pick special cases of training set

Change input Luminosity Model

In training set



Simulation to Data to Paramaters

Making the Network robust to input

Deconfusing Intensity Maps with Neural Networks
Pfeffer, Breysse, Stein - arXiv 1905.10376

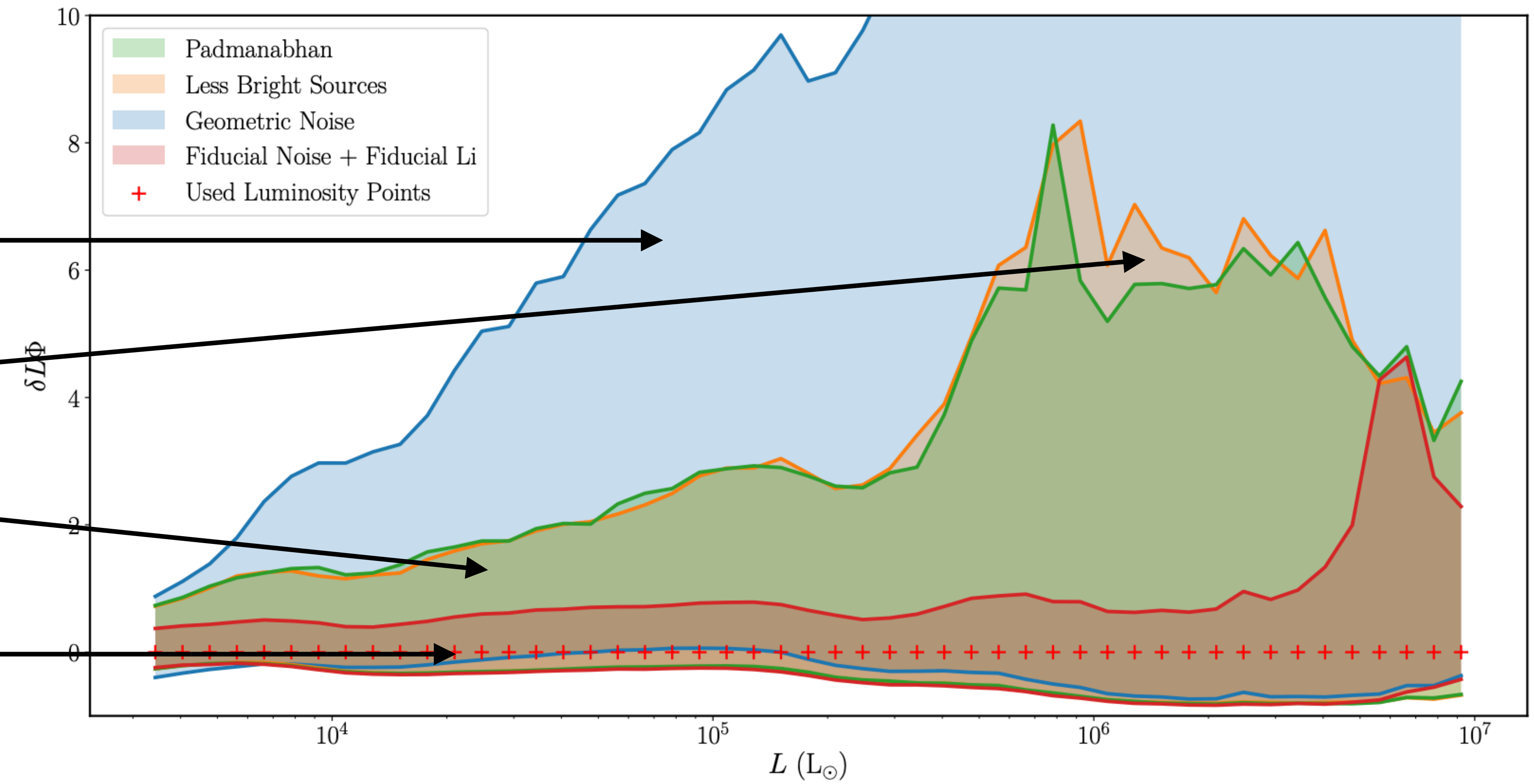
Network completely fails for scenarios not included in training set

Add more noise near edges of survey

Pick special cases of training set

Change input Luminosity Model

In training set



find that, in the ideal case where the mock data capture all of the features of the real data, the CNN performs comparably to or better than conventional analyses. However, the network's accuracy degrades considerably when tested on signals and systematics outside of those it was trained on. For both intensity mapping and cosmology as a whole, this motivates a broad-based study of whether simulated data can ever be generated with sufficient detail to realize the enormous potential of machine learning methods.

Come find me in office 341,

or email me at gstein@berkeley.edu