Prospects for CMB lensing—galaxy clustering cross-correlations and modeling biased tracers

Marcel Schmittfull

IAS INSTITUTE FOR ADVANCED STUDY

IPMU, October 2019

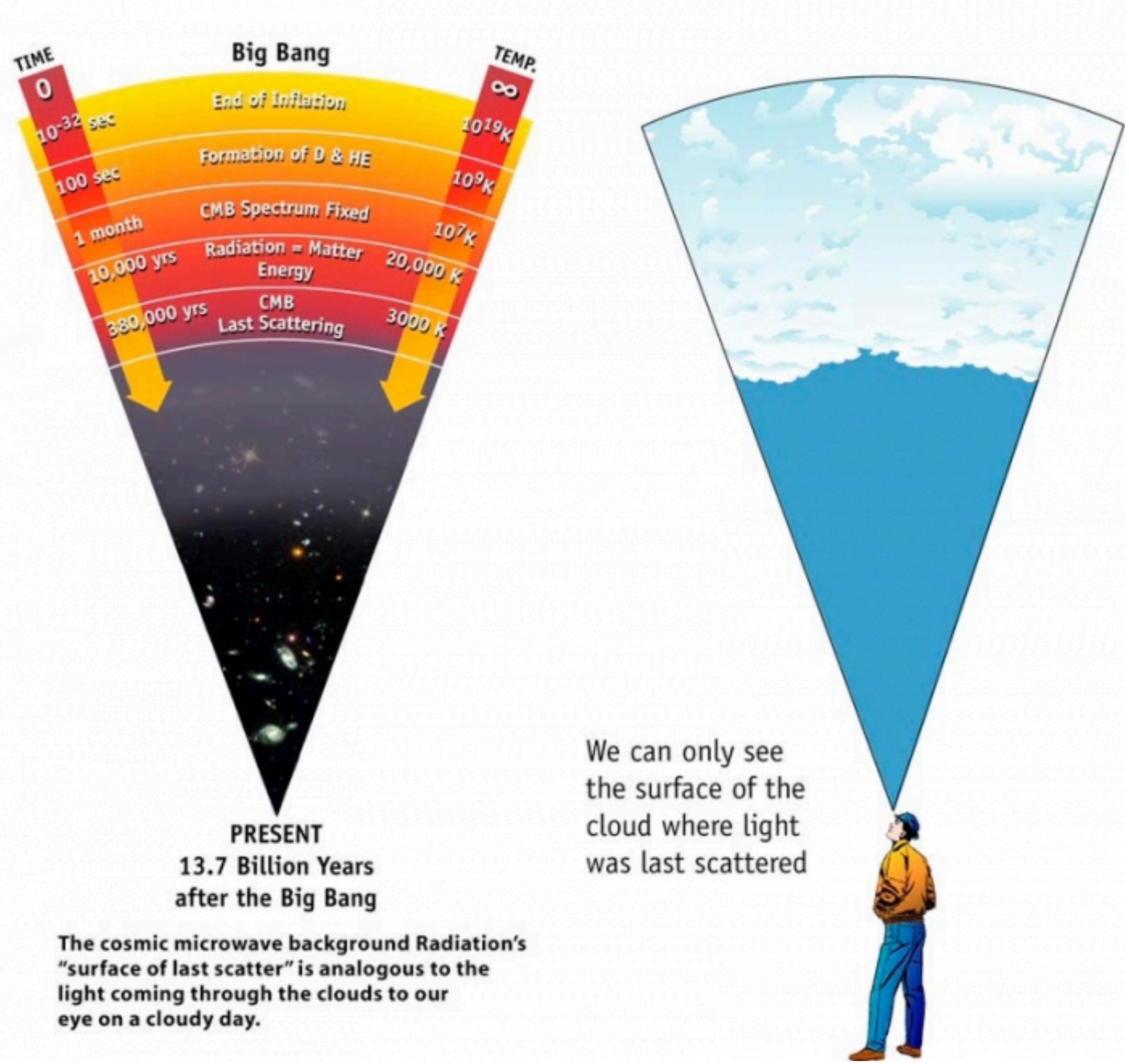
Cosmic Microwave Background (CMB)

Right after Big Bang light scatters frequently —> opaque

As Universe expands, turns transparent

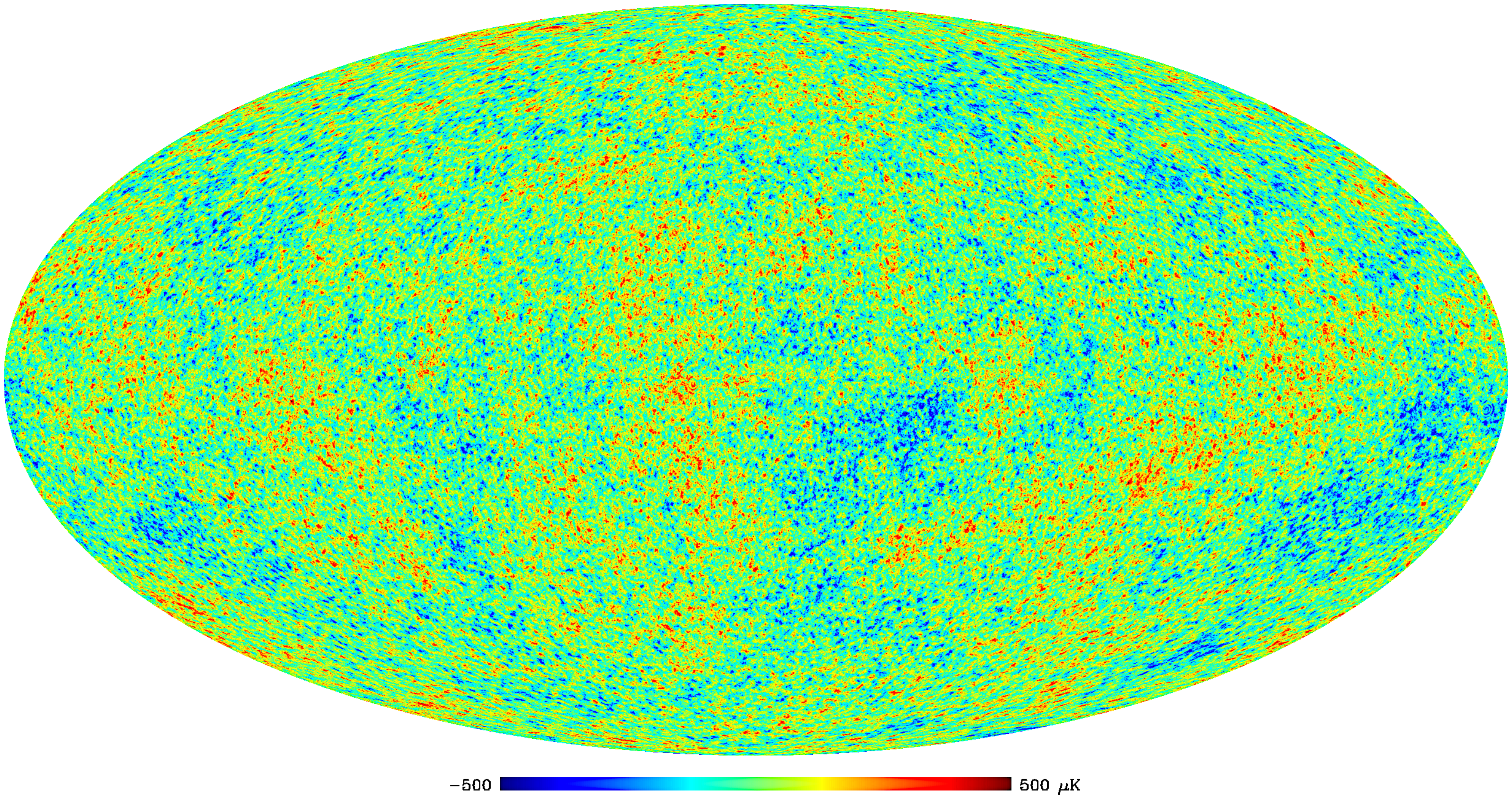
See surface where light last scattered — 13.6996 bn yrs ago

This is the CMB



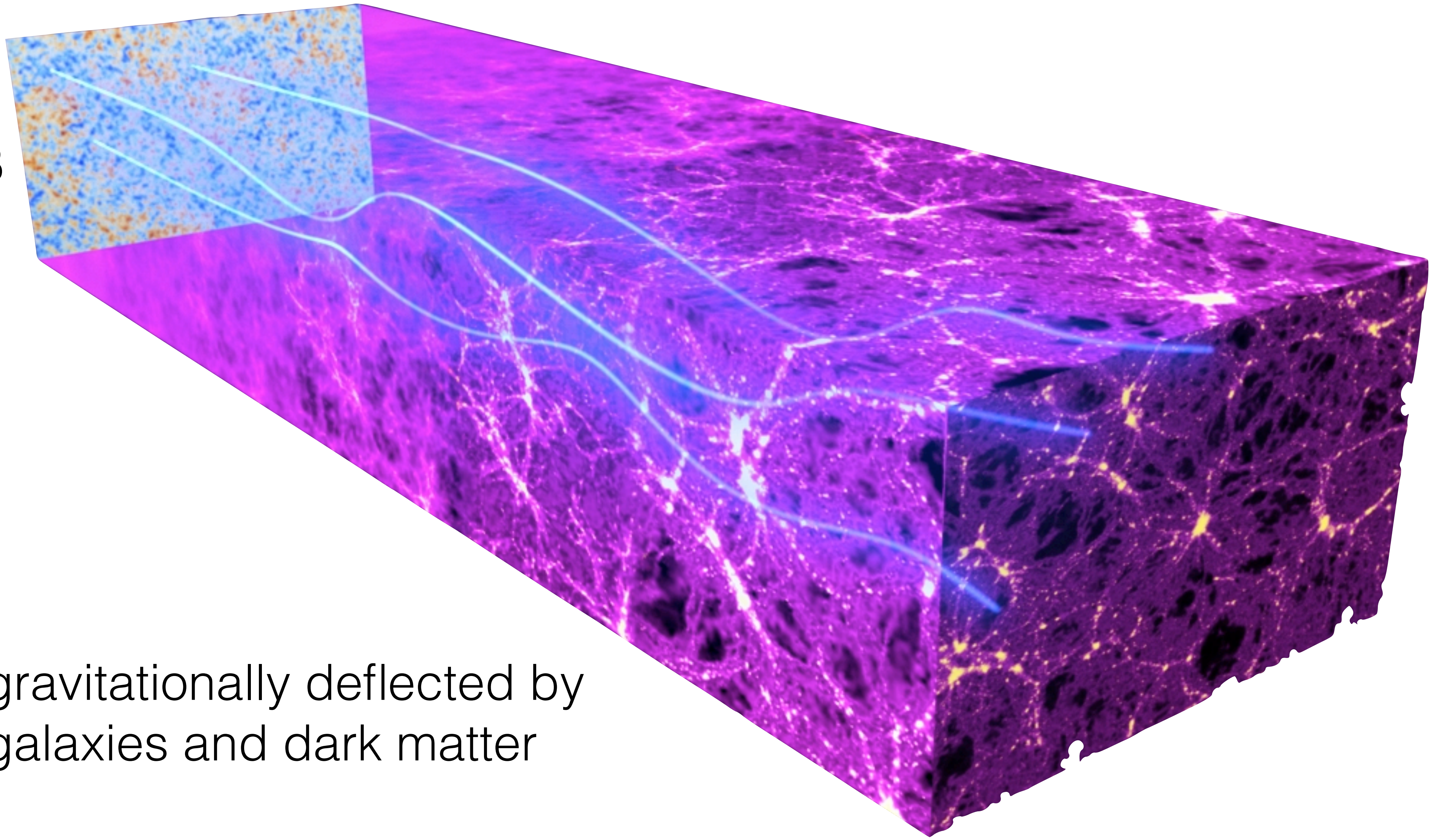
The cosmic microwave background Radiation's "surface of last scatter" is analogous to the light coming through the clouds to our eye on a cloudy day.

Cosmic Microwave Background (CMB)

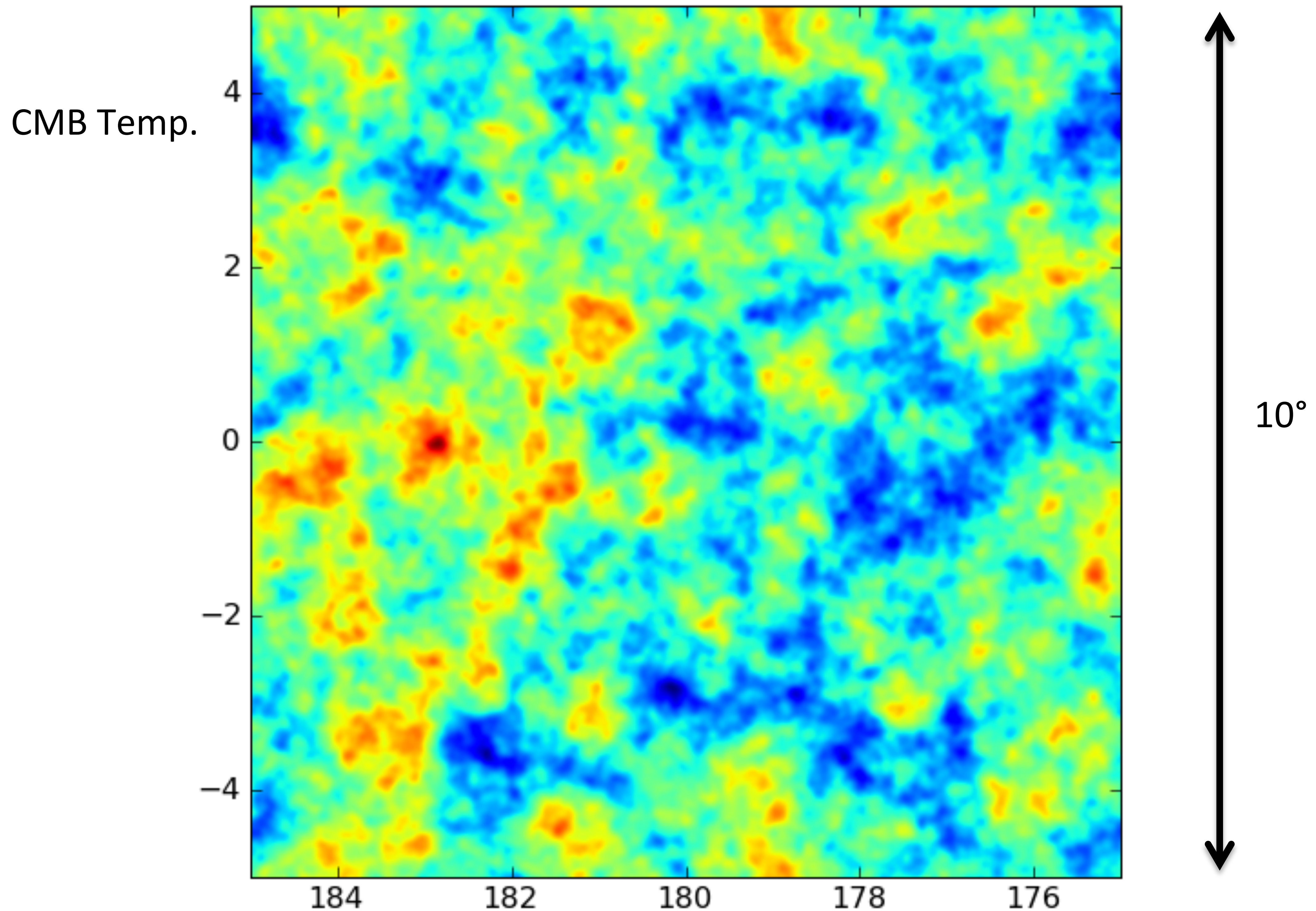


Hot and cold blobs: Picture of the Universe 13.6996 bn years ago

Gravitational lensing of the CMB



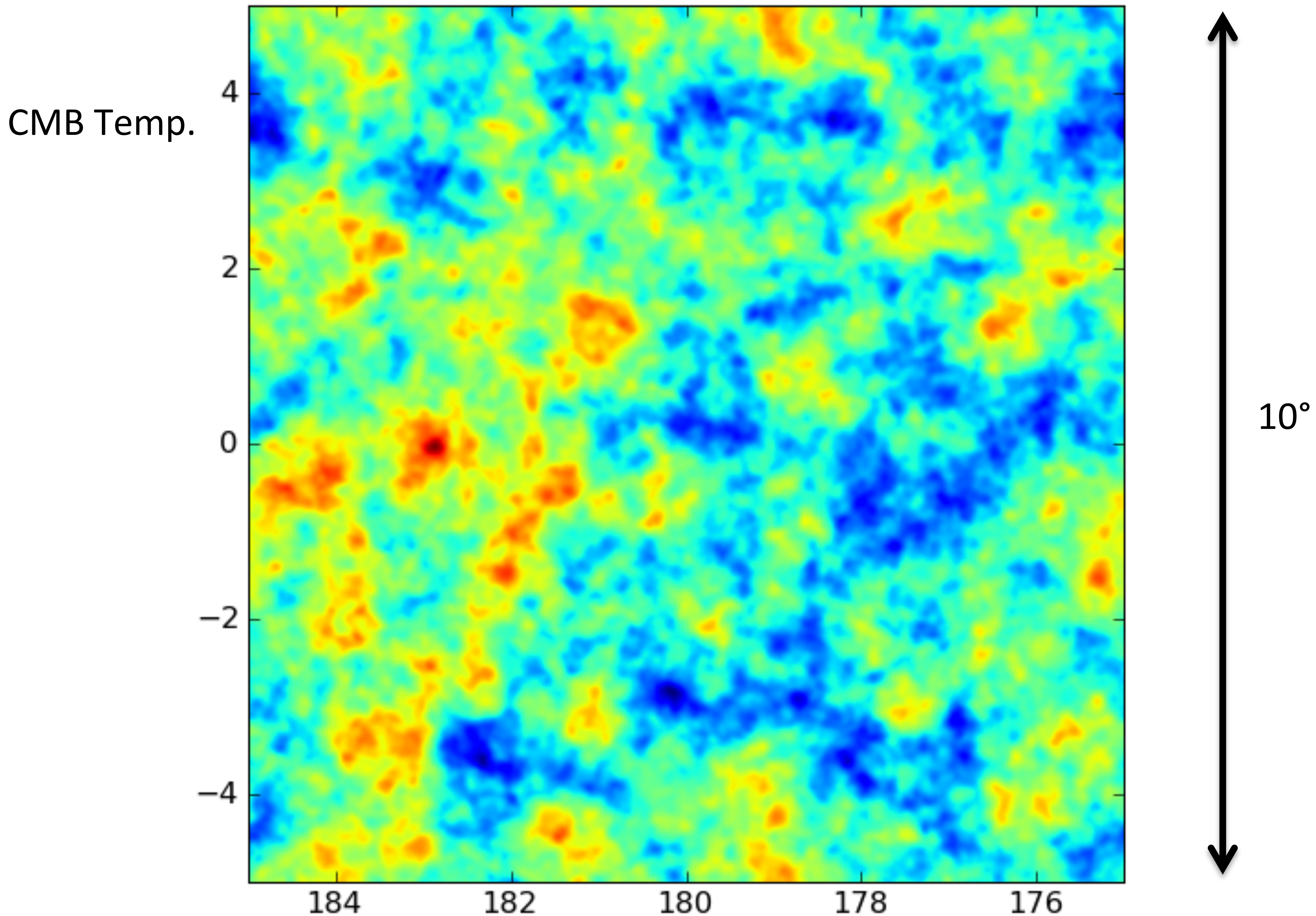
CMB before lensing



$T(\hat{\mathbf{n}})_{\mathrm{unlensed}}$

Slide credit: Blake Sherwin

CMB after lensing



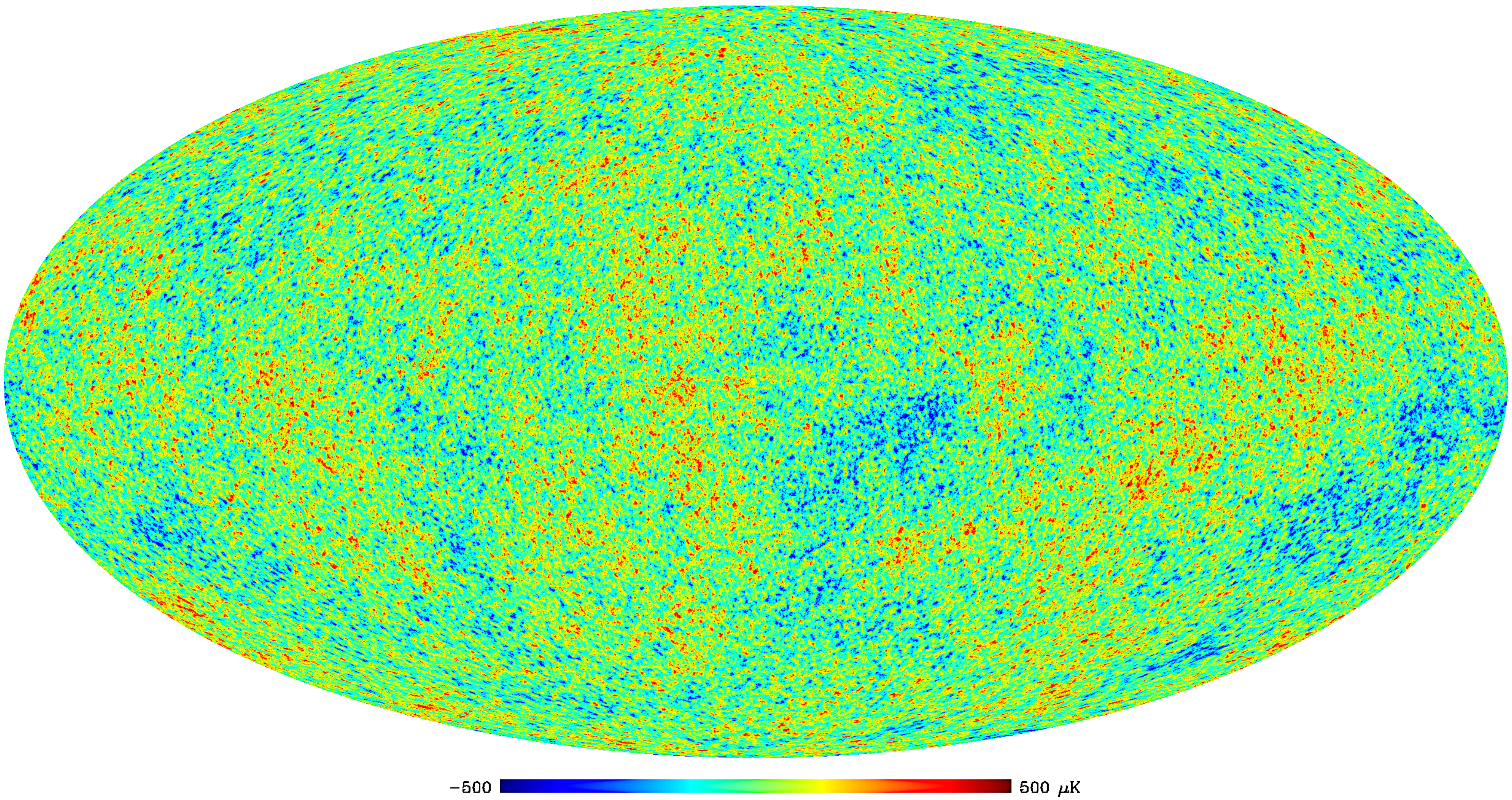
$$T(\hat{\mathbf{n}})_{\rm lensed} = T(\hat{\mathbf{n}} + \mathbf{d}(\hat{\mathbf{n}}))_{\rm unlensed}$$

Slide credit: Blake Sherwin

Statistics of the CMB

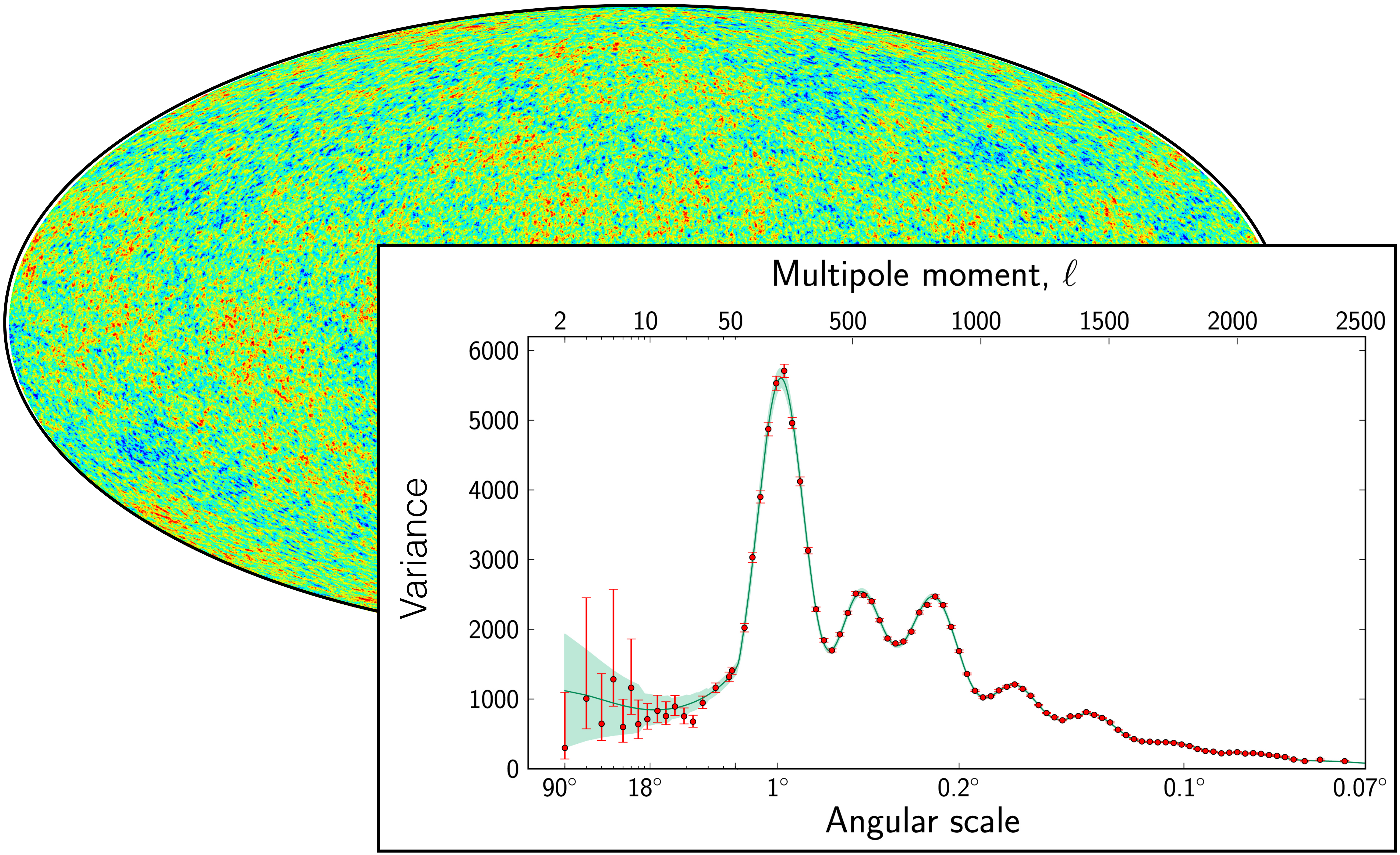
What are the statistics of the CMB before and after lensing?

Statistics of the CMB before lensing

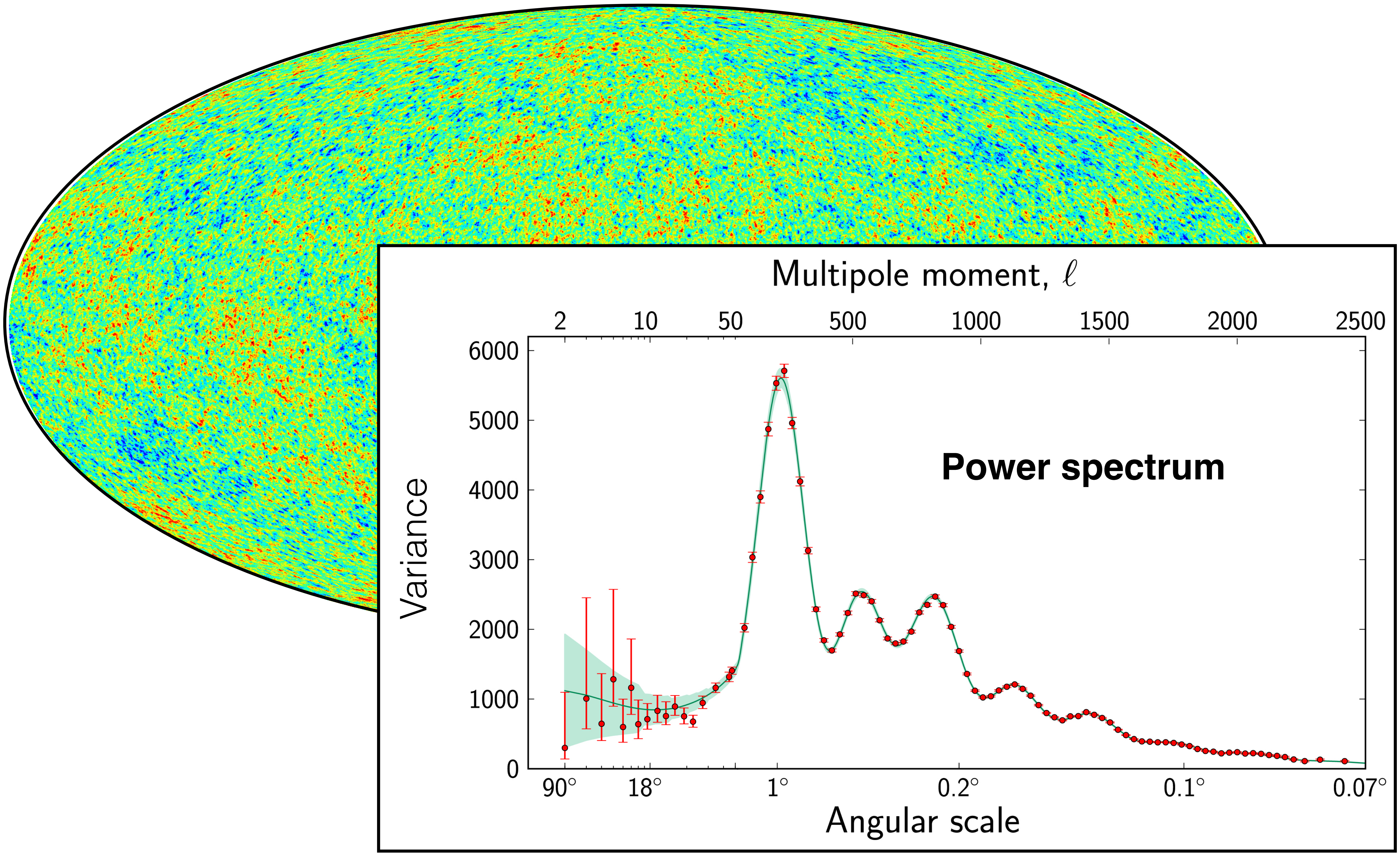


Normally distributed as far as we can tell

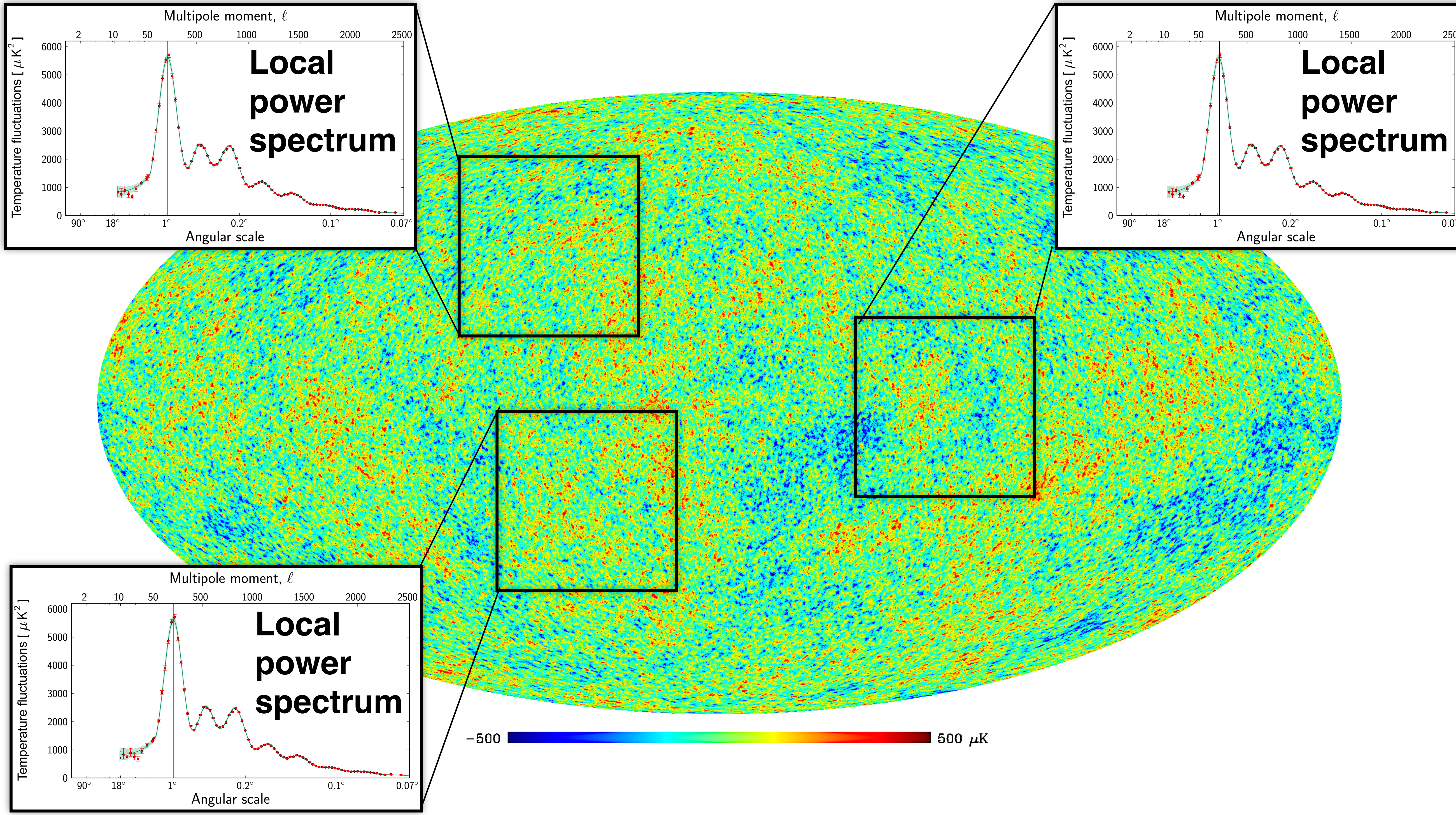
Statistics of the CMB before lensing



Statistics of the CMB before lensing

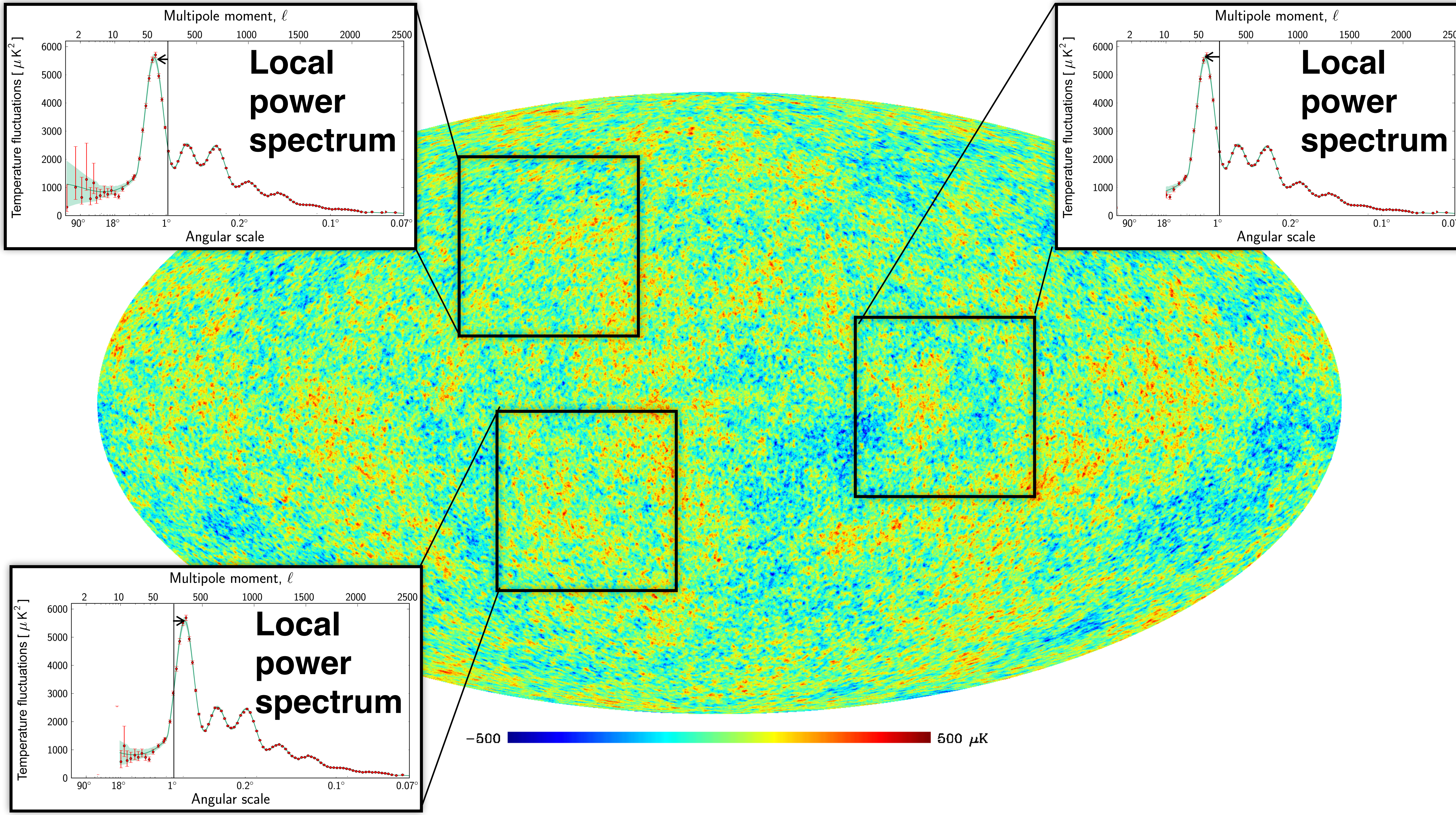


Statistics of the CMB before lensing



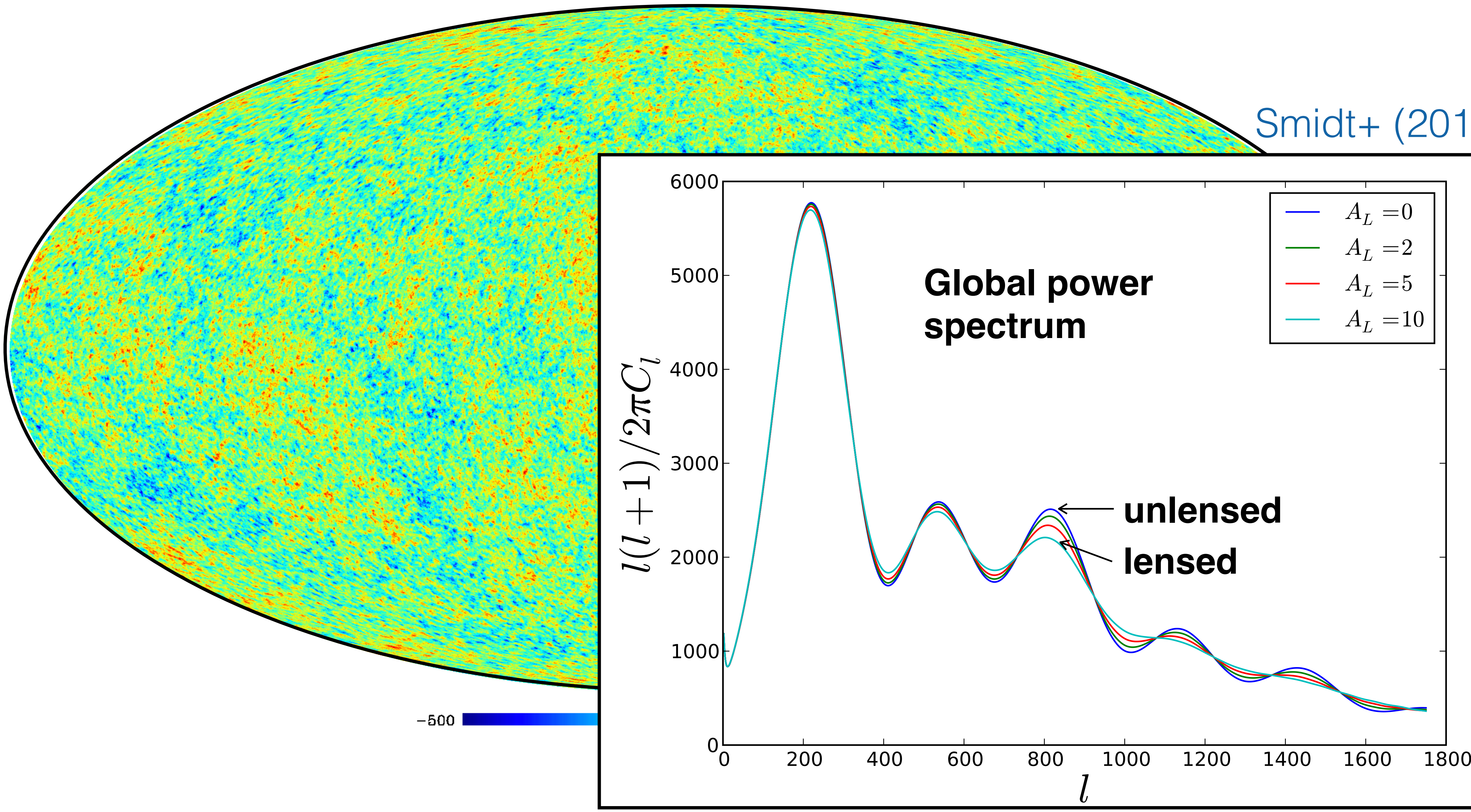
Local power spectrum is the same in each patch

Statistics of the CMB after lensing



Local power is magnified or de-magnified

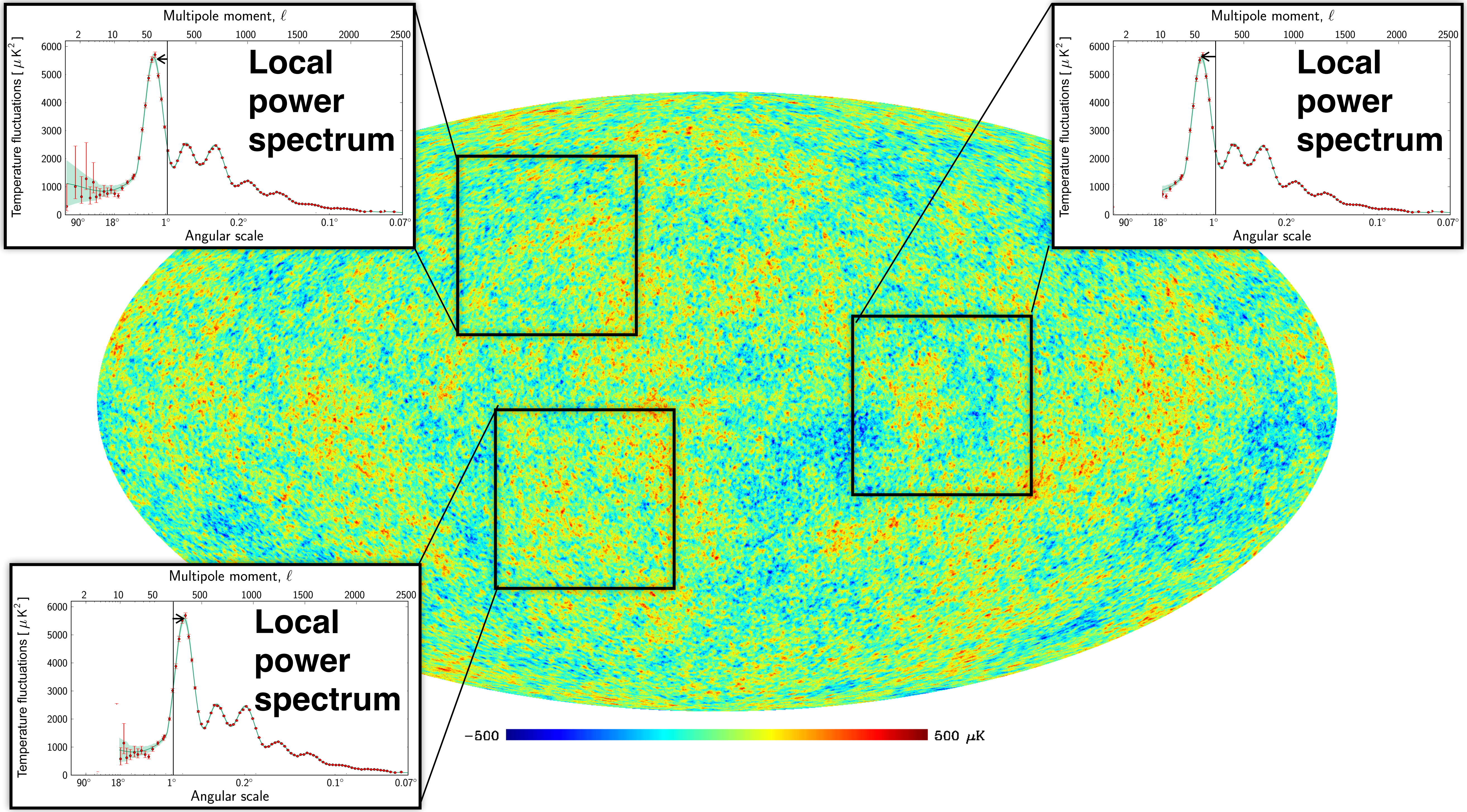
Statistics of the CMB after lensing



Smidt+ (2010)

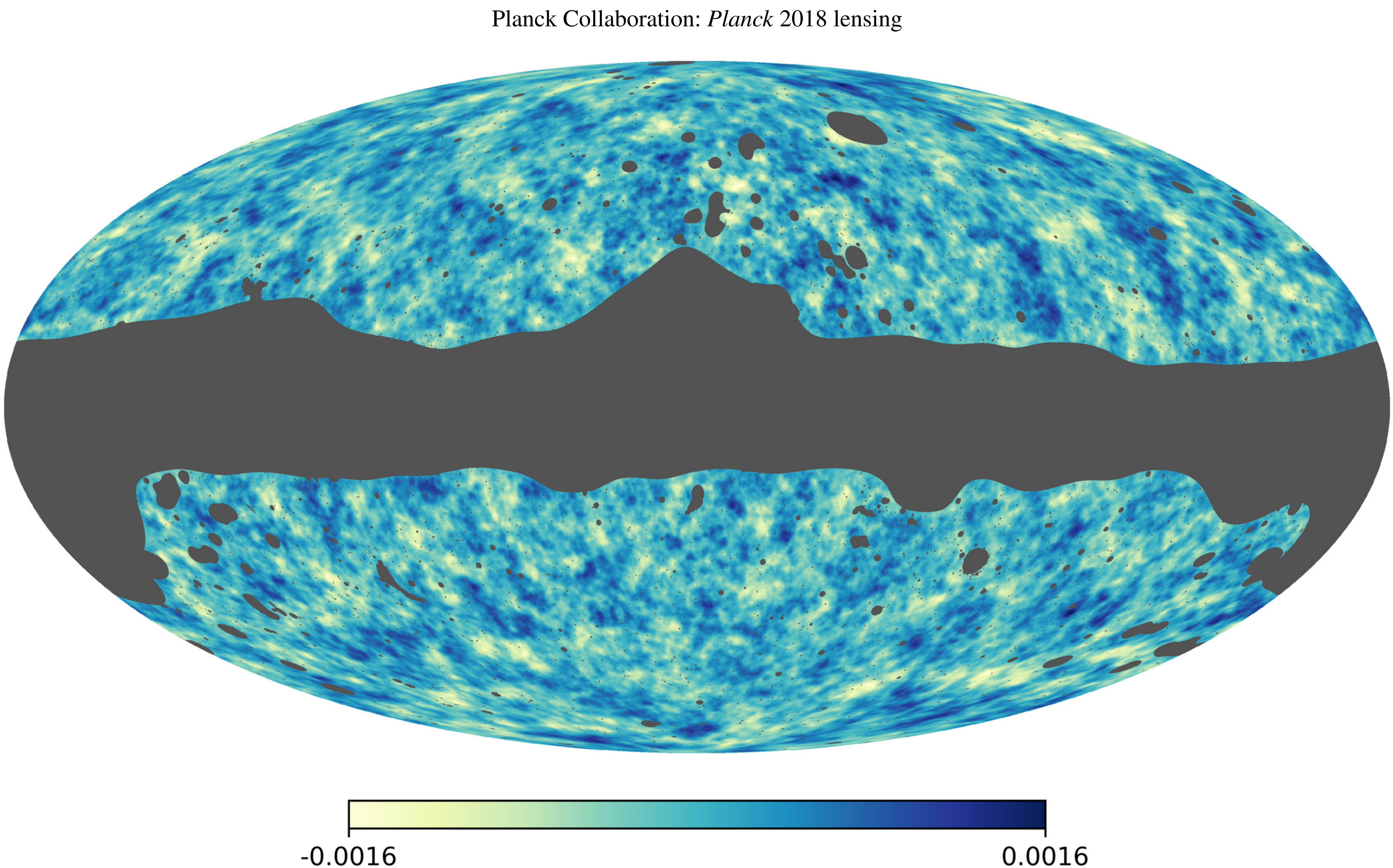
Peaks of global power are smeared out

Statistics of the CMB after lensing



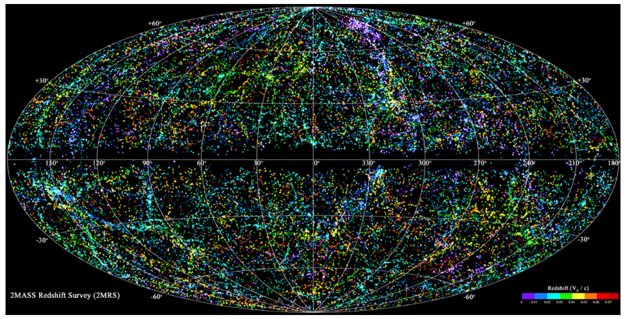
Rather than averaging the modulation, measure it as a signal
—> magnification map

Measured lensing magnification

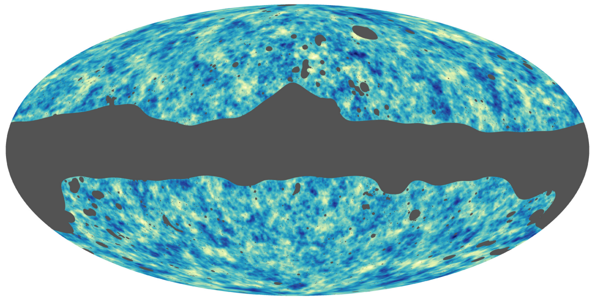


Cross-correlate with galaxy catalogs

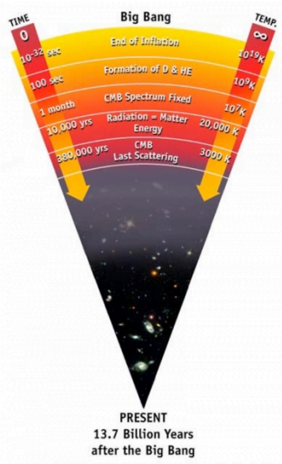
Galaxy catalog is sensitive to *tails* of the distribution of matter (galaxies form at peaks)



Lensing magnification probes *entire* distribution (light is deflected by any mass)



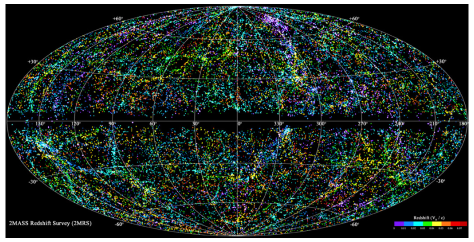
Some models of the Big Bang enhance peaks, so can test them



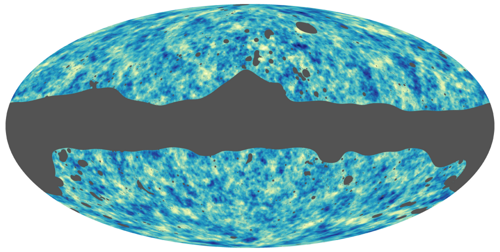
Cross-correlations particularly well suited: Can determine ratio of two perfectly correlated random variables with infinite precision

Sample variance cancellation

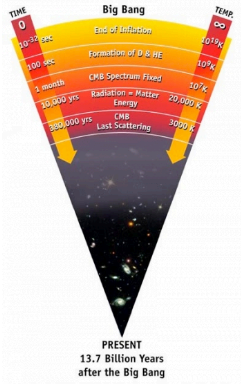
Galaxy number density: $g = (1 + f_{\mathrm{NL}})\kappa$



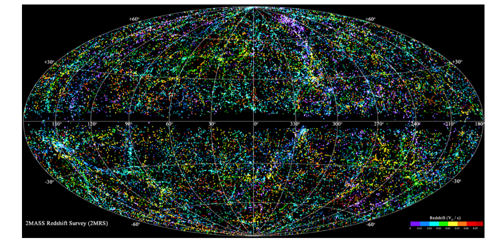
CMB magnification: $\kappa \sim \mathcal{N}(0, \sigma_\kappa^2)$



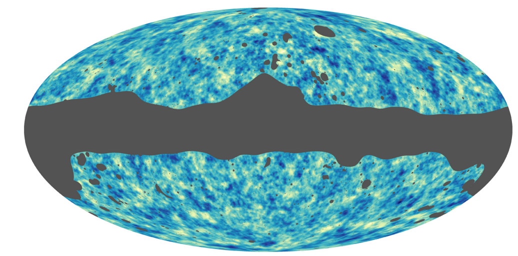
Big Bang models are characterized by parameter f_{NL}



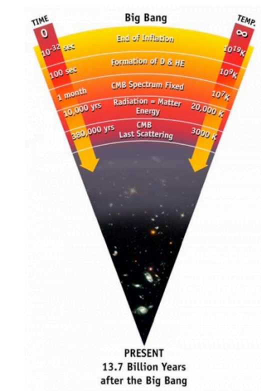
Sample variance cancellation



Galaxy number density: $g = (1 + f_{\mathrm{NL}})\kappa$



CMB magnification: $\kappa \sim \mathcal{N}(0, \sigma_\kappa^2)$



Big Bang models are characterized by parameter f_{NL}

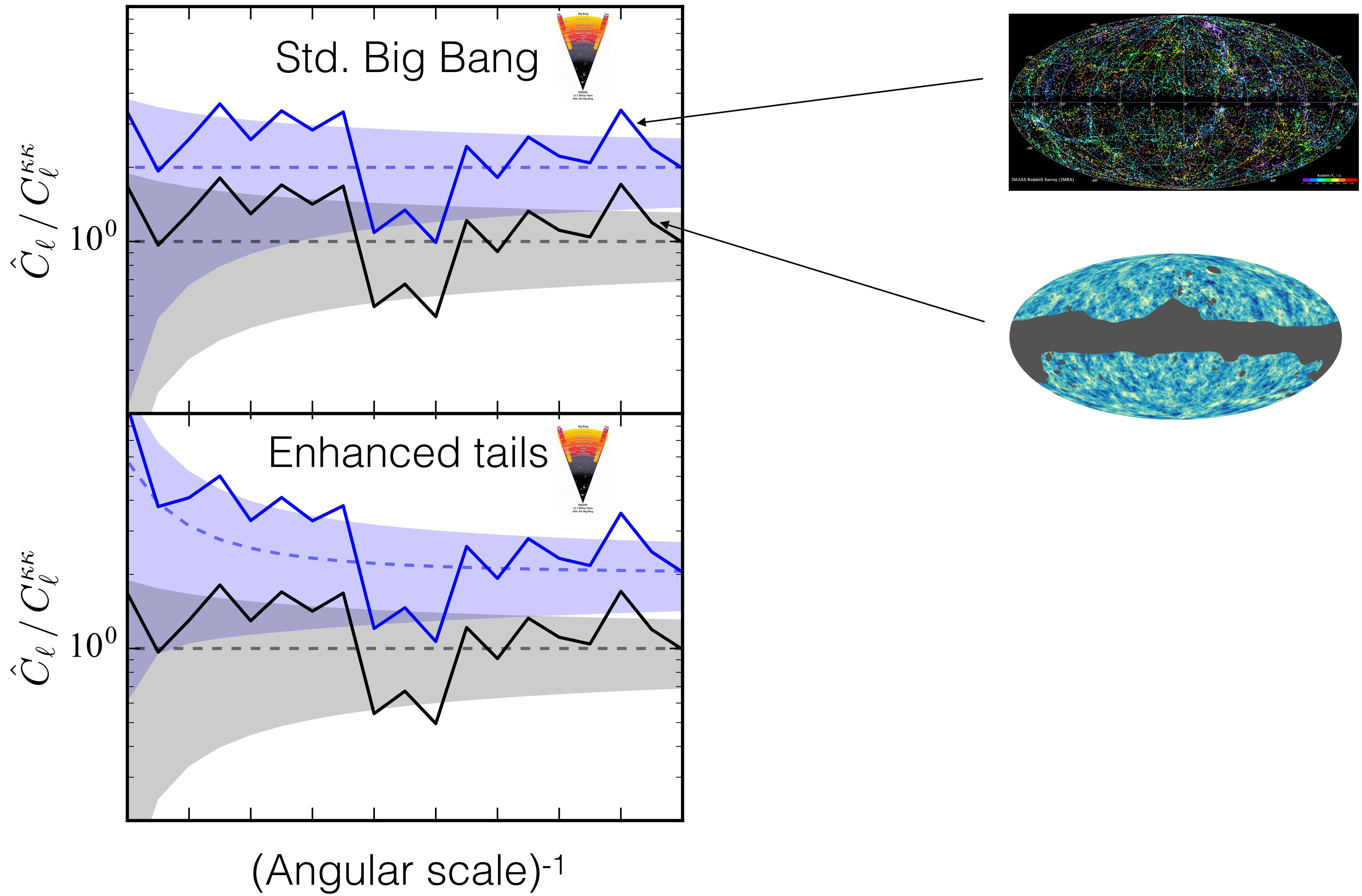
If we measure only g:

$$\mathrm{var}(\hat{f}_{\mathrm{NL}}) \sim \sigma_\kappa^2$$

If we measure g and κ:

$$\hat{f}_{\mathrm{NL}} = \frac{g}{\kappa} - 1 \quad \Rightarrow \quad \mathrm{var}(\hat{f}_{\mathrm{NL}}) = 0$$

Sample variance cancellation



Dalal+ (2008), Seljak (2009), McDonald & Seljak (2009), MS & Seljak (2018)

Forecast for future experiments

Forecast for future experiments

Compute Fisher information matrix assuming Gaussian likelihood (curvature of the likelihood near maximum = Hessian)

$$F_{ab} = -\left\langle \frac{\partial^2 \ln \mathcal{L}(\mathbf{d}|\theta)}{\partial\theta_a \partial\theta_b} \right\rangle$$

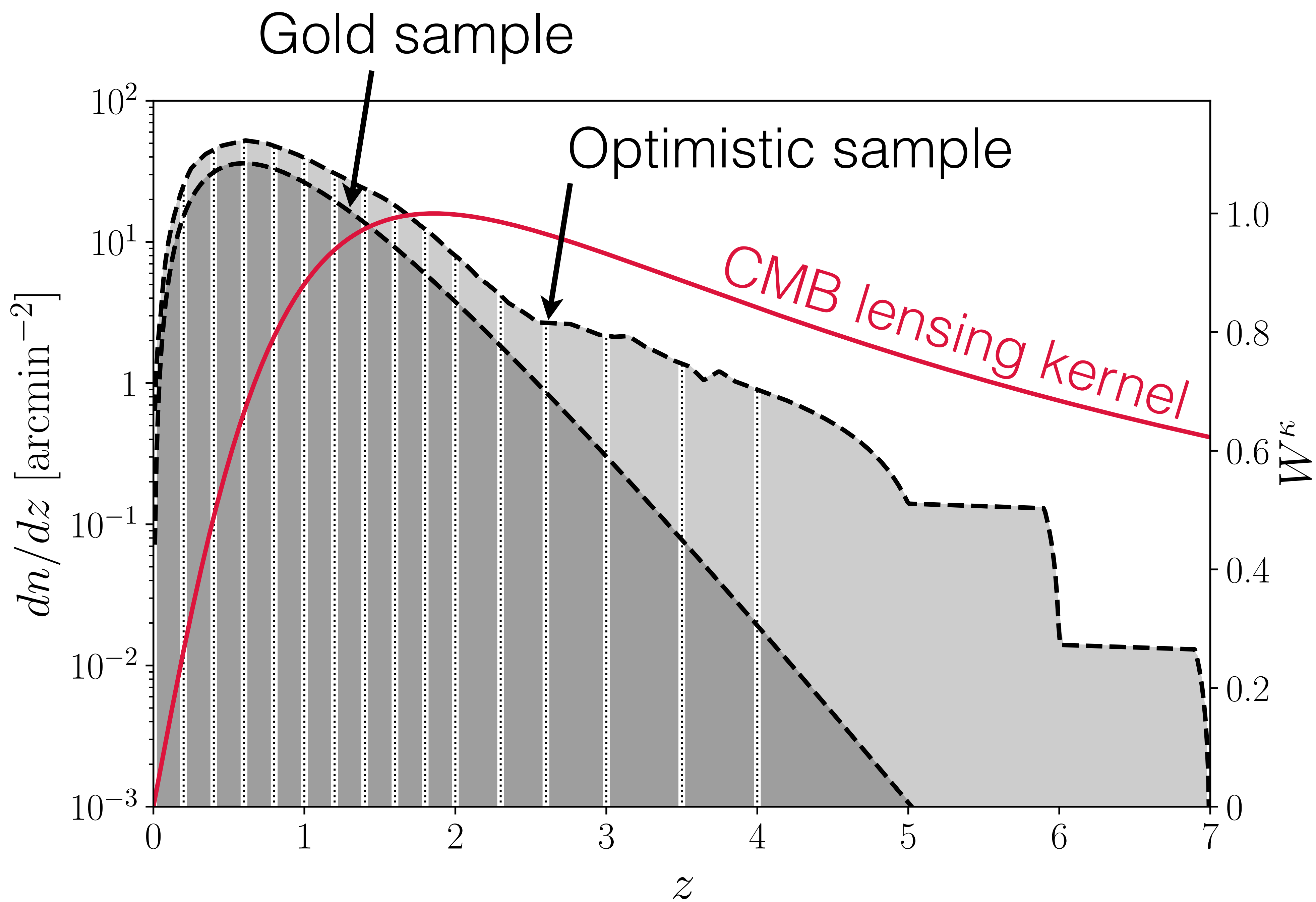
$$= \sum_{\ell=\ell_{\rm min}}^{\ell_{\rm max}} \frac{\partial \mathbf{d}_\ell}{\partial\theta_a} [\mathrm{cov}(\mathbf{d}_\ell, \mathbf{d}_\ell)]^{-1} \frac{\partial \mathbf{d}_\ell}{\partial\theta_b}.$$

Inverse gives lower bound on parameter error bars

Optimistic setting

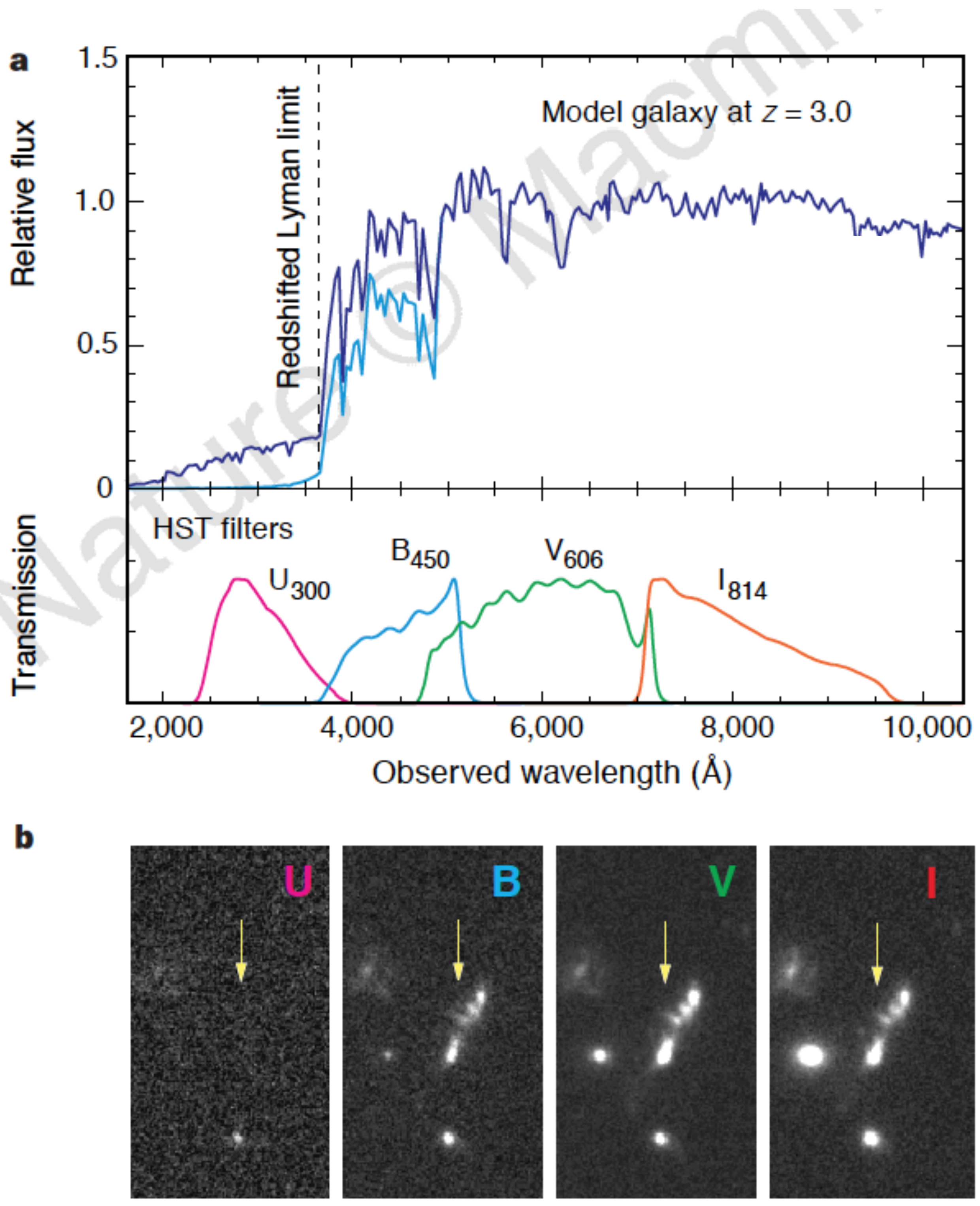
If models work and systematics under control, what can we hope for?

LSST number density



At low z, use clustering redshifts (Gorecki+ 2014)
At high z, add Lyman-break galaxies (dropouts; extrapolated from HSC observations)
Total of 66 galaxies per arcmin^2 (MS & Seljak 2018)

Lyman-break galaxies (LBGs) at z>3



Ellis 1998

Photons blueward of 912 Å ionize neutral hydrogen in young star-forming galaxies
⇒ Lyman break at (1+z) 912 Å

Blue band ‘dropout’

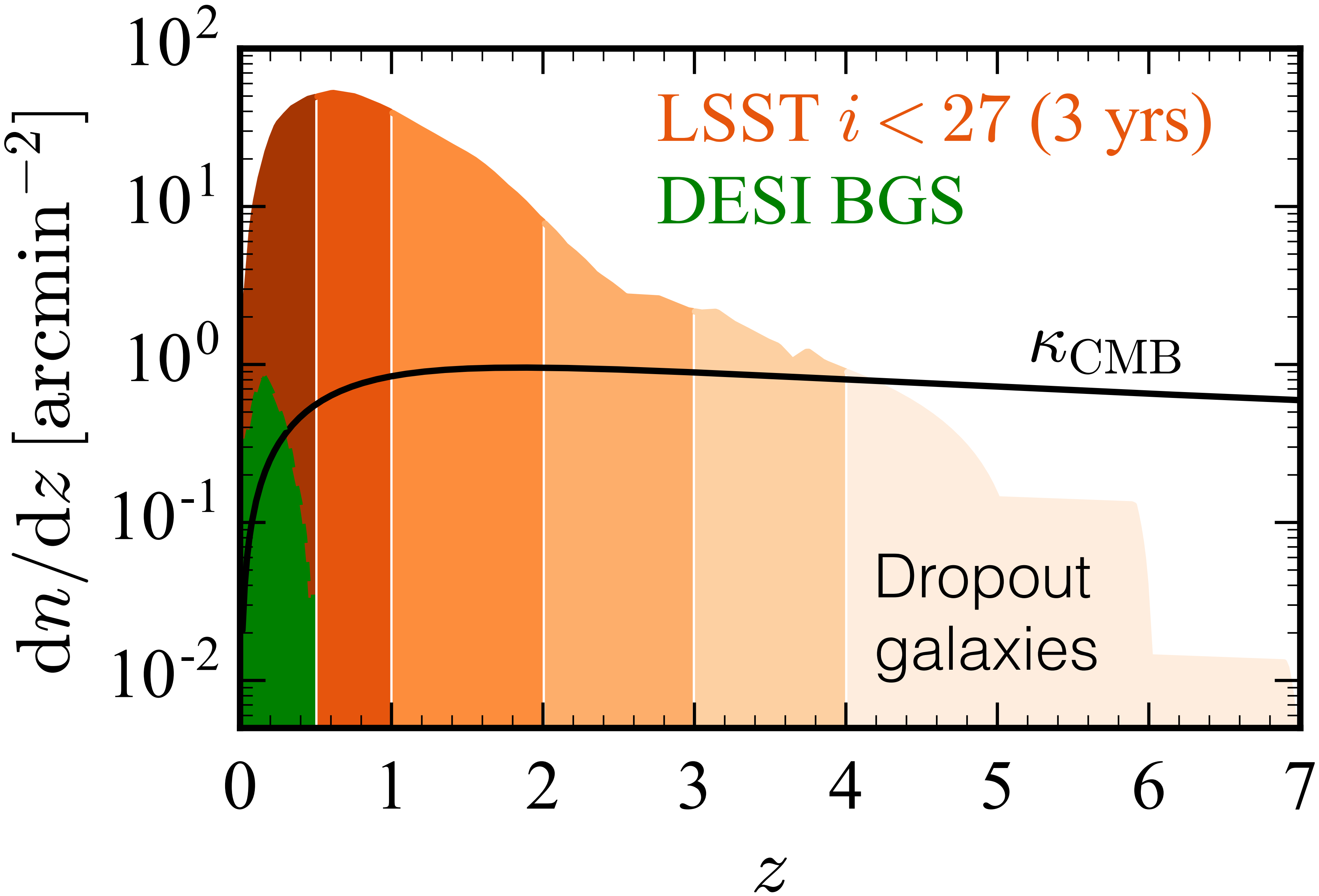
0.5 million z=4-7 LBGs found by HSC/Goldrush in 100 deg^2

Ono, Ouchi+ (2018) 1704.06004

⇒ Expect ~100 million in LSST

Good for CMB lensing Xcorrel

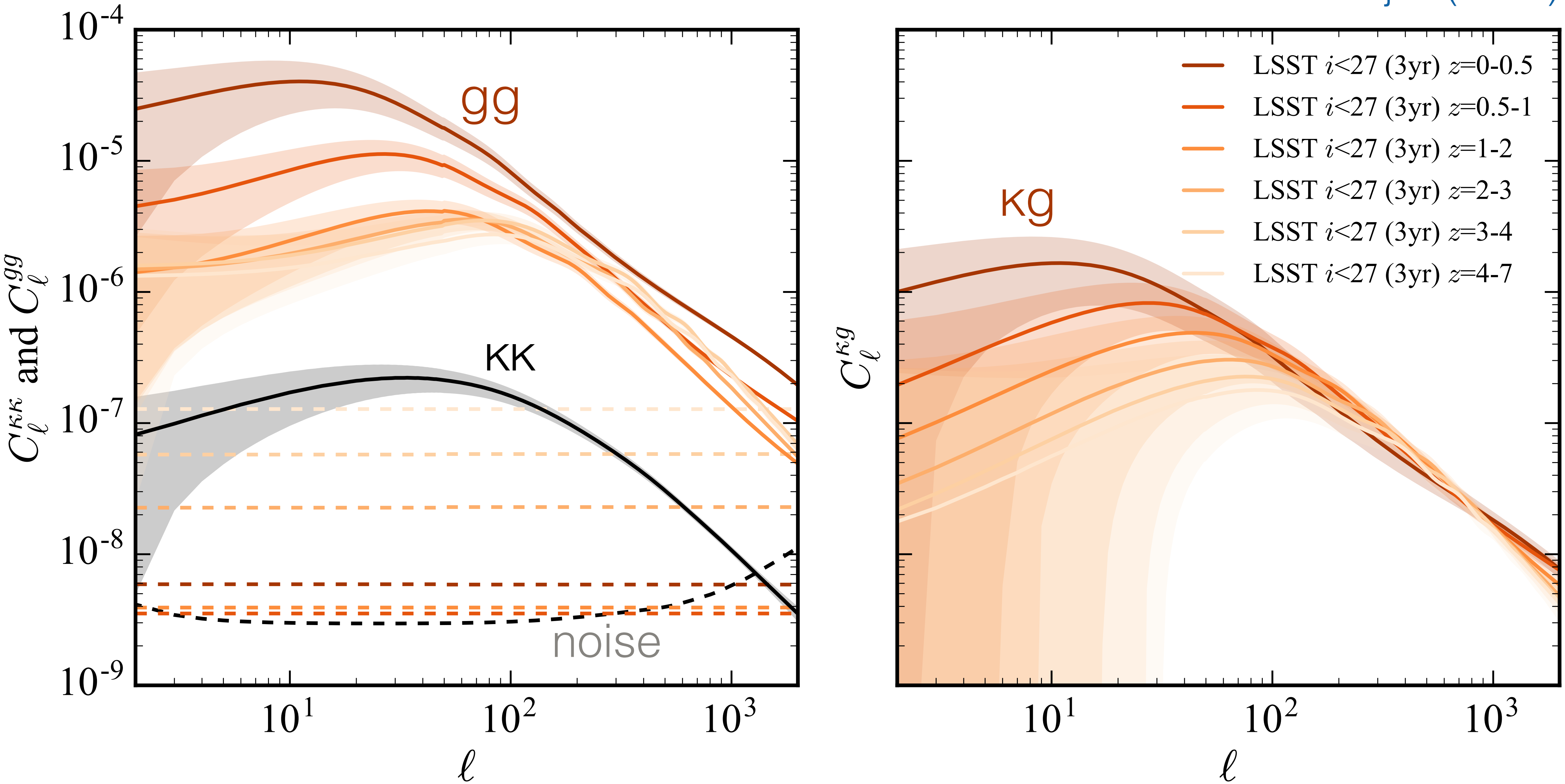
Tomographic redshift bins



MS & Seljak (2018)

Power spectra: CMB-S4 & LSST

MS & Seljak (2018)



Low ell: $C_\ell^{XX'} = \frac{2}{\pi}\int_0^\infty \frac{\mathrm{d}k}{k}\Delta_{X,\ell}(k)\Delta_{X',\ell}(k)k^3 P_{\delta_X\delta_{X'}}(k,z=0),$

$\Delta_{X,\ell}(k) \equiv \int_0^\infty \mathrm{d}\chi\,\bar{W}_X(\chi)j_\ell(k\chi).$

High ell: $C_\ell^{XX'} = \int_z P_{\delta_X\delta_{X'}}(k=\ell/\chi(z),z)$

$\times W_X(z)b_X(z)W_{X'}(z)b_{X'}(z)$

SNR of auto-power spectra

SNR of C^{XX}	$\ell_{\rm max}$ 500	1000	2000
$\kappa_{\rm CMB}$	233	406	539
BOSS LRG z=0-0.9	140	187	230
SDSS $r < 22$ z=0-0.5	247	487	936
SDSS $r < 22$ z=0.5-0.8	247	487	936
DESI BGS z=0-0.5	230	417	665
DESI ELG z=0.6-0.8	158	210	256
DESI ELG z=0.8-1.7	150	194	225
DESI LRG z=0.6-1.2	184	267	349
DESI QSO z=0.6-1.9	44.8	48.8	50.8
LSST $i < 27$ (3yr) z=0-0.5	250	496	982
LSST $i < 27$ (3yr) z=0.5-1	250	496	979
LSST $i < 27$ (3yr) z=1-2	249	492	956
LSST $i < 27$ (3yr) z=2-3	245	469	830
LSST $i < 27$ (3yr) z=3-4	239	444	724
LSST $i < 27$ (3yr) z=4-7	224	387	555

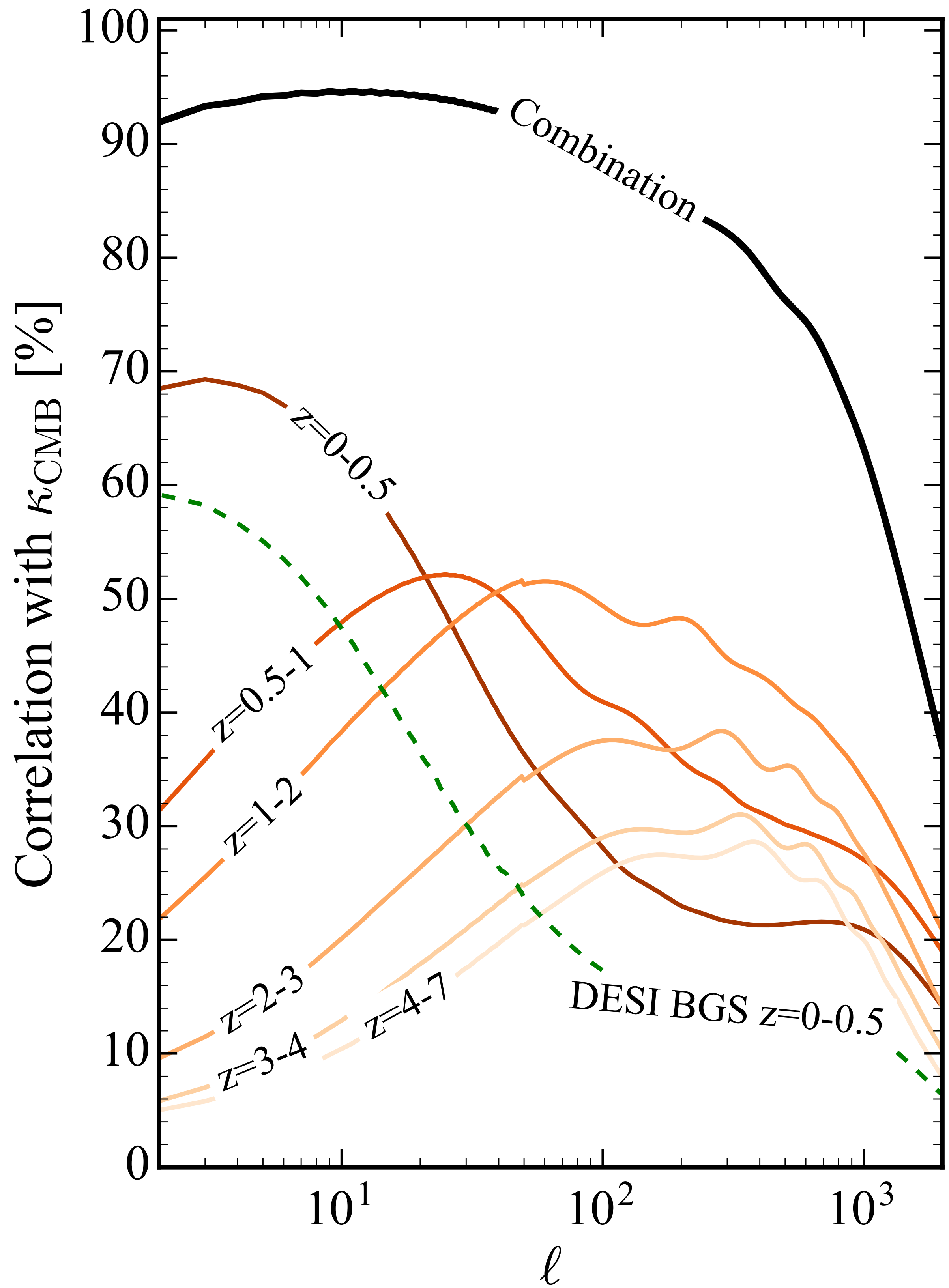
MS & Seljak (2018)

SNR of kg cross-power spectra

	$\ell_{\rm max}$		
SNR of $C^{\kappa_{\rm CMB}X}$	500	1000	2000
BOSS LRG z=0-0.9	77.3	117	159
SDSS $r < 22$ z=0-0.5	88.3	167	284
SDSS $r < 22$ z=0.5-0.8	88.3	167	284
DESI BGS z=0-0.5	50.1	93.5	144
DESI ELG z=0.6-0.8	50.7	73.5	97
DESI ELG z=0.8-1.7	103	148	185
DESI LRG z=0.6-1.2	86.7	133	182
DESI QSO z=0.6-1.9	74.9	94.5	108
LSST $i < 27$ (3yr) z=0-0.5	78.1	150	258
LSST $i < 27$ (3yr) z=0.5-1	112	202	338
LSST $i < 27$ (3yr) z=1-2	144	259	406
LSST $i < 27$ (3yr) z=2-3	121	219	324
LSST $i < 27$ (3yr) z=3-4	101	182	261
LSST $i < 27$ (3yr) z=4-7	94	167	229

MS & Seljak (2018)

Correlation of CMB lensing and galaxies

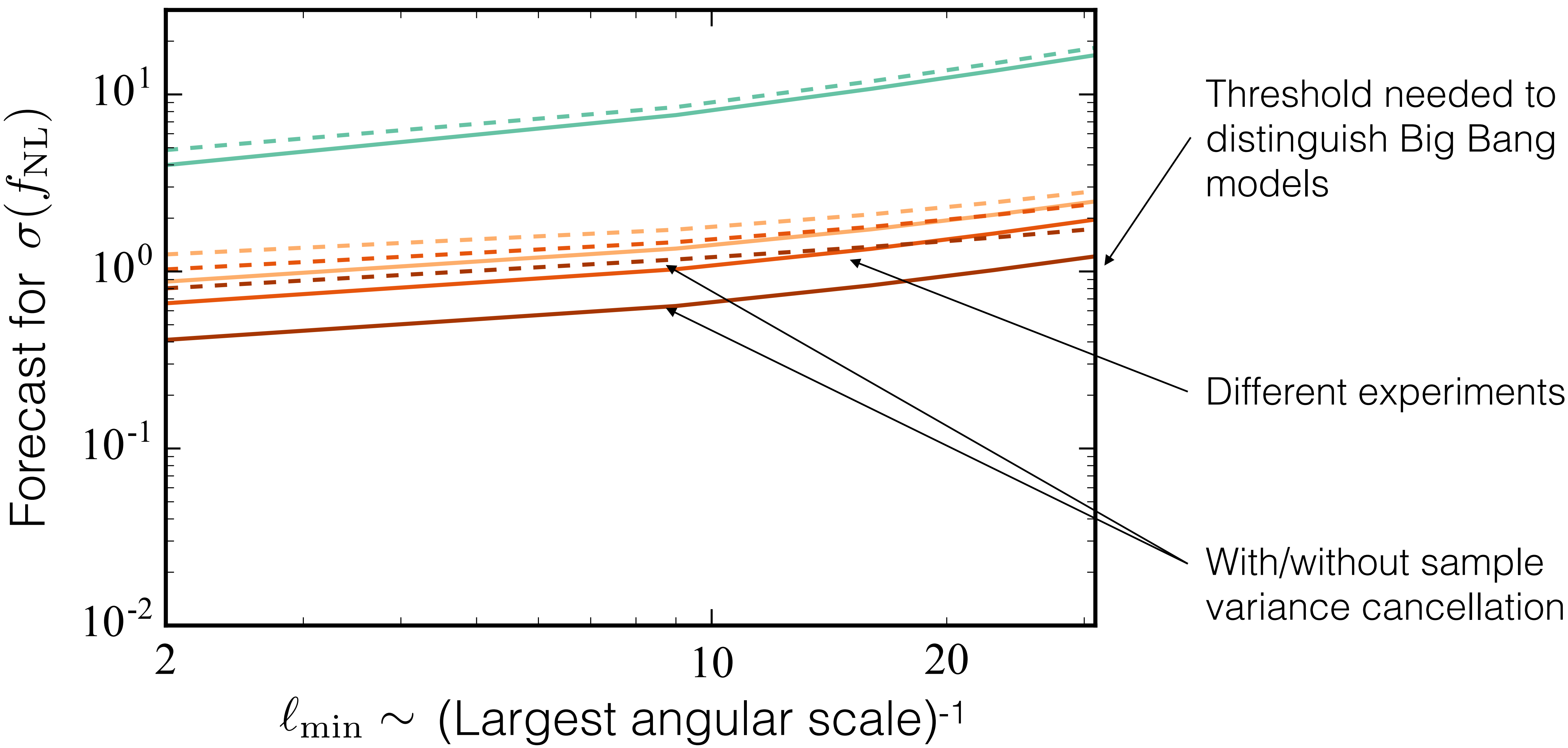


MS & Seljak (2018)

CMB lensing and galaxy maps are up to 95% correlated

$$r_\ell = \frac{C_\ell^{\kappa g}}{\sqrt{\hat{C}_\ell^{\kappa\kappa}\hat{C}_\ell^{gg}}}$$

Forecast for local non-Gaussianity



—> Potential to rule out all Big Bang models driven by a single field

MS & Seljak (2018); Forecast papers for Simons Observatory, CMB-S4, NASA PICO satellite

Other science from cross-correlations

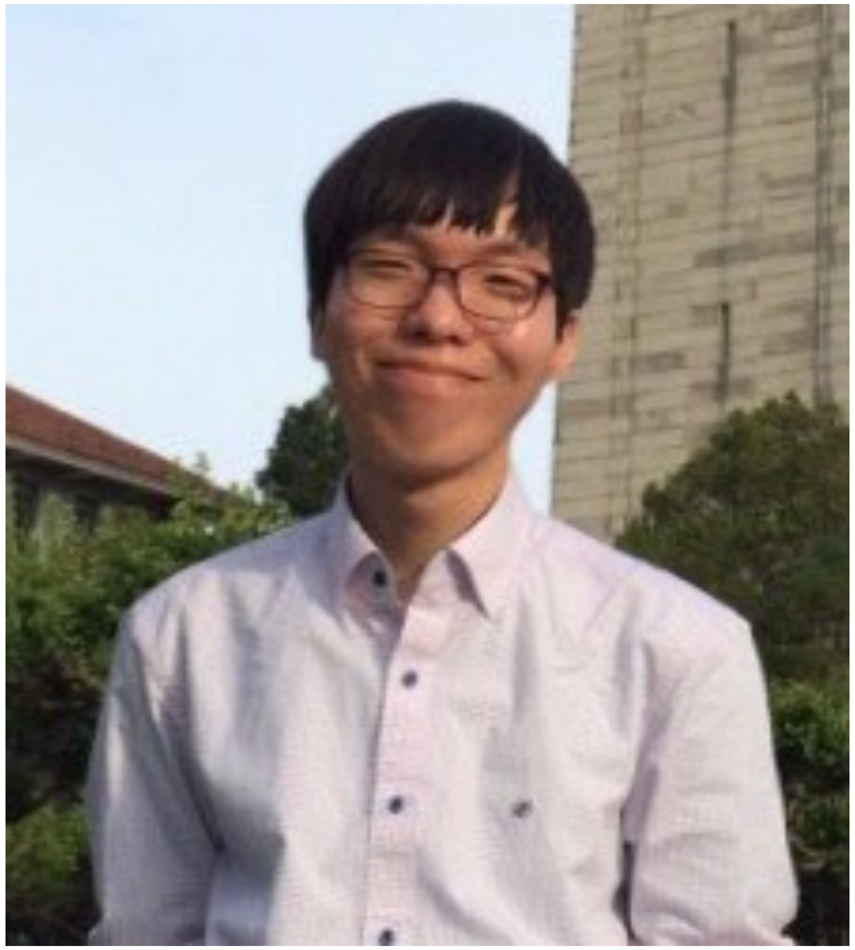
Growth of structure as function of time

Expansion history / geometry / dark energy

Sum of neutrino masses

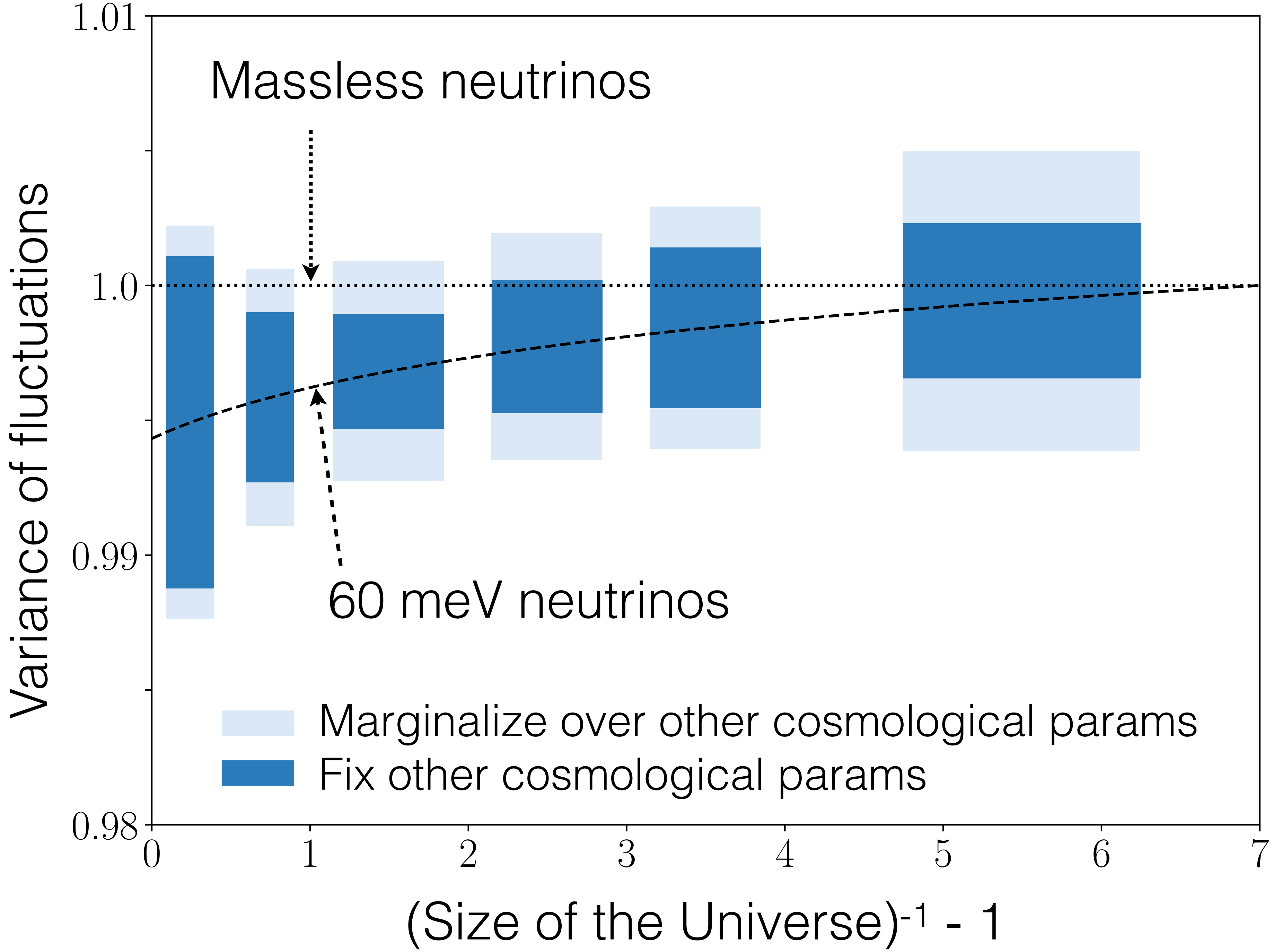
Galaxy formation?

Example: Sum of neutrino masses



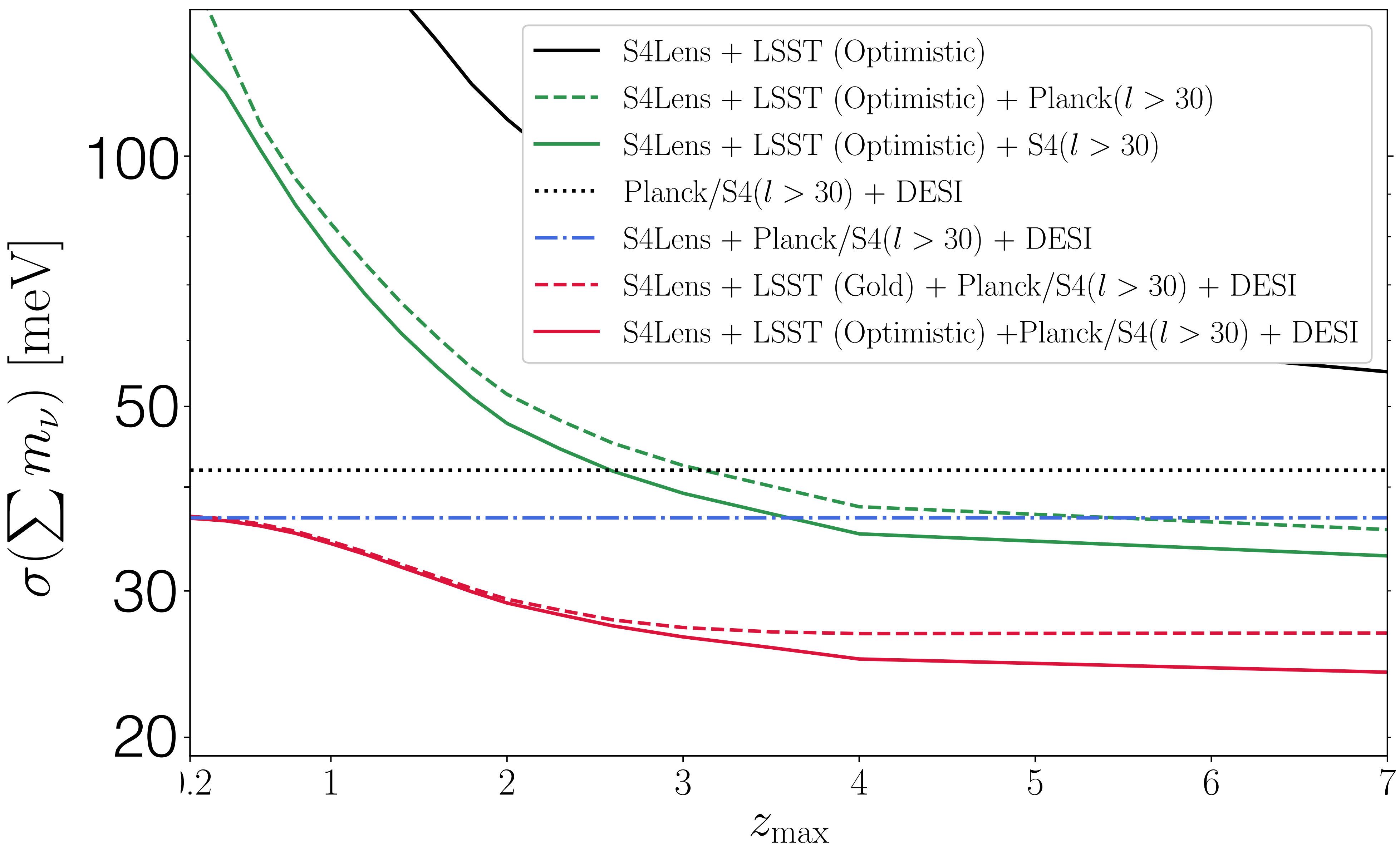
Byeonghee Yu (PhD student @ UC Berkeley)

Example: Sum of neutrino masses



Yu, Knight *et al.* (arXiv:1809.02120)

Example: Sum of neutrino masses



Uses z<7 redshift lever arm

Measures neutrino mass *without optical depth to CMB*

Independent & competitive to usual probes

Yu, Knight, Sherwin+ (1809.02120)

Example: Sum of neutrino masses

	$\sigma(\sum m_\nu)$ [meV] (Gold/Optimistic)		
$k_{\max}$	Lens + LSST	+ Planck/S4 T&P	+ DESI
0.05	307 / 243	94 / 68	32 / 29
0.1	176 / 129	68 / 53	31 / 27
0.2	107 / 71	47 / 38	28 / 25
0.3	84 / 55	40 / 33	27 / 24
0.4	79 / 49	38 / 31	26 / 23

Yu, Knight, Sherwin+ (1809.02120)

Sensitive to $k_{\max}$, but not too much when all data combined

Ignores scale-dependent bias due to neutrinos

LoVerde 2014, LoVerde 2016, Muñoz & Dvorkin 2018, MS & Seljak 2018, Chiang, LoVerde & Villaescusa-Navarro 2019

Challenges for growth measurements

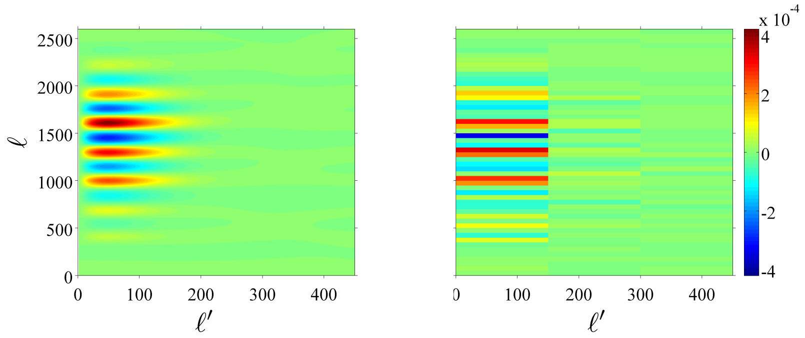
Modeling all power spectra for nonlinear scales (high *L*)

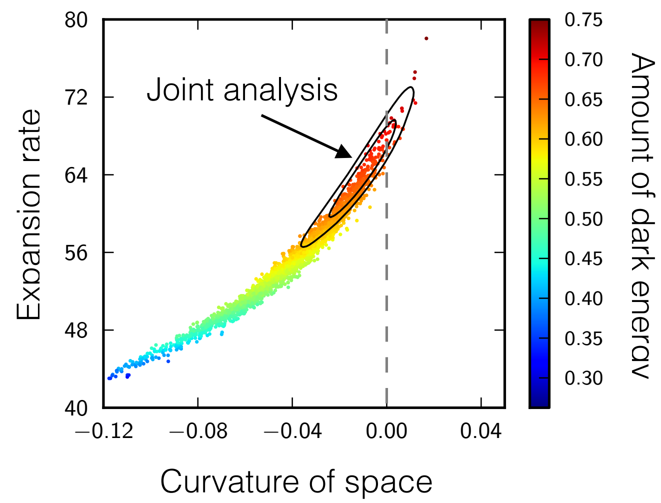
Photometric redshift errors

Relationship between galaxies and dark matter (galaxy bias)

Other cool things to do with CMB lensing

(1) Joint analysis of CMB and CMB magnification, with covariance

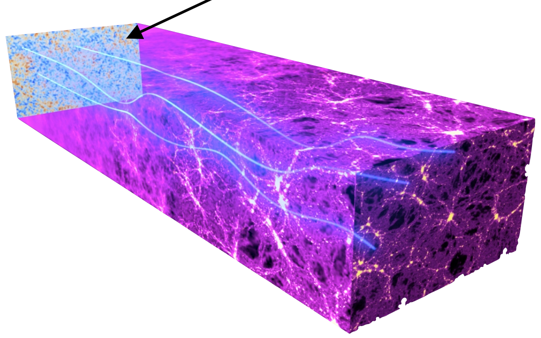




MS, Challinor *et al.* (2013)
Peloton, MS, *et al.* (2017)
Planck collab. (2013, 2018)

(2) Estimate *unlensed CMB*

e.g. Sherwin & MS (2015)

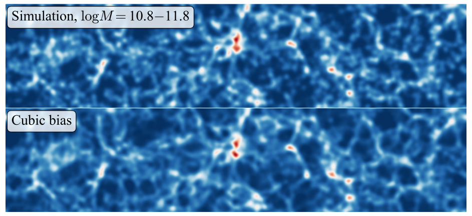


(3) Biases of the magnification estimator

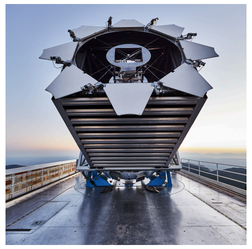
Böhm, MS & Sherwin (2016)
Beck, Fabbian & Errard (2018)
Böhm, Sherwin *et al.* (2018)

Science with galaxy catalogs

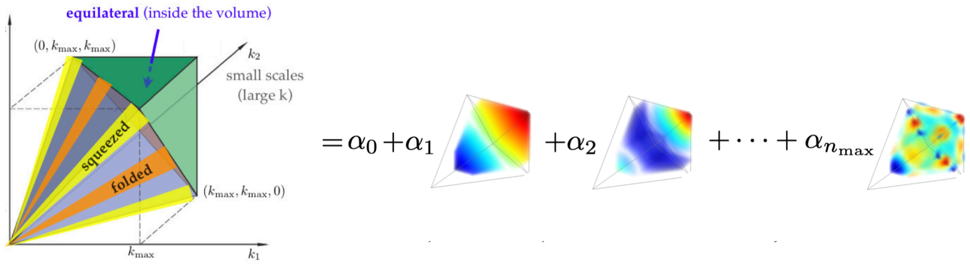
(1) Bias model at the field level



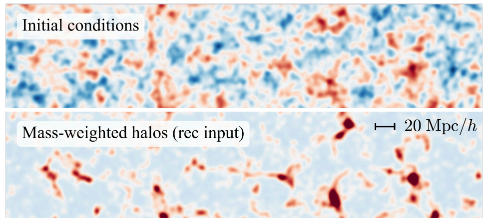
(2) Cosmological parameter analysis



(3) Accounting for skewness

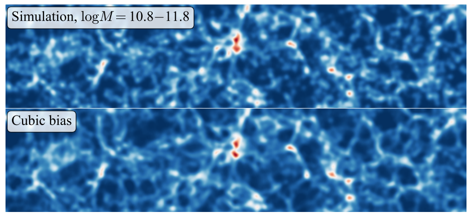


(4) Getting initial from final conditions



Science with galaxy catalogs

(1) Bias model at the field level



(2) Cosmological parameter analysis

(3) Accounting for skewness

(4) Getting initial from final conditions

Overview

We calculate halo density field in PT and compare to simulations

1. How well does perturbative bias expansion work?

2. How correlated is the halo density field with the initial conditions?

3. What are the properties of the noise?

Joint work with Marko Simonović, Valentin Assassi & Matias Zaldarriaga

Overview

These questions have been extensively explored in the past

Desjacques, Jeong, Schmidt: Large-Scale Galaxy Bias

Most of the analyses use *n*-point functions. Disadvantages:

— Cosmic variance, compromise on resolution/size of the box

— At high *k* hard to disentangle different sources of nonlinearities

— Overfitting (smooth curves, many parameters)

— Only a few lowest *n*-point functions explored in practice

— Difficult to isolate and study the noise

Overview

Roth & Porciani (2011)
Baldauf, Schaan, Zaldarriaga (2015)
Lazeyras, Schmit (2017)
Abidi, Baldauf (2018)
McQuinn, D'Aloisio (2018)
Taruya, Nishimichi, Jeong (2018)

Use fields rather than summary statistics

Advantages:

— No cosmic variance, small boxes with high resolution are sufficient

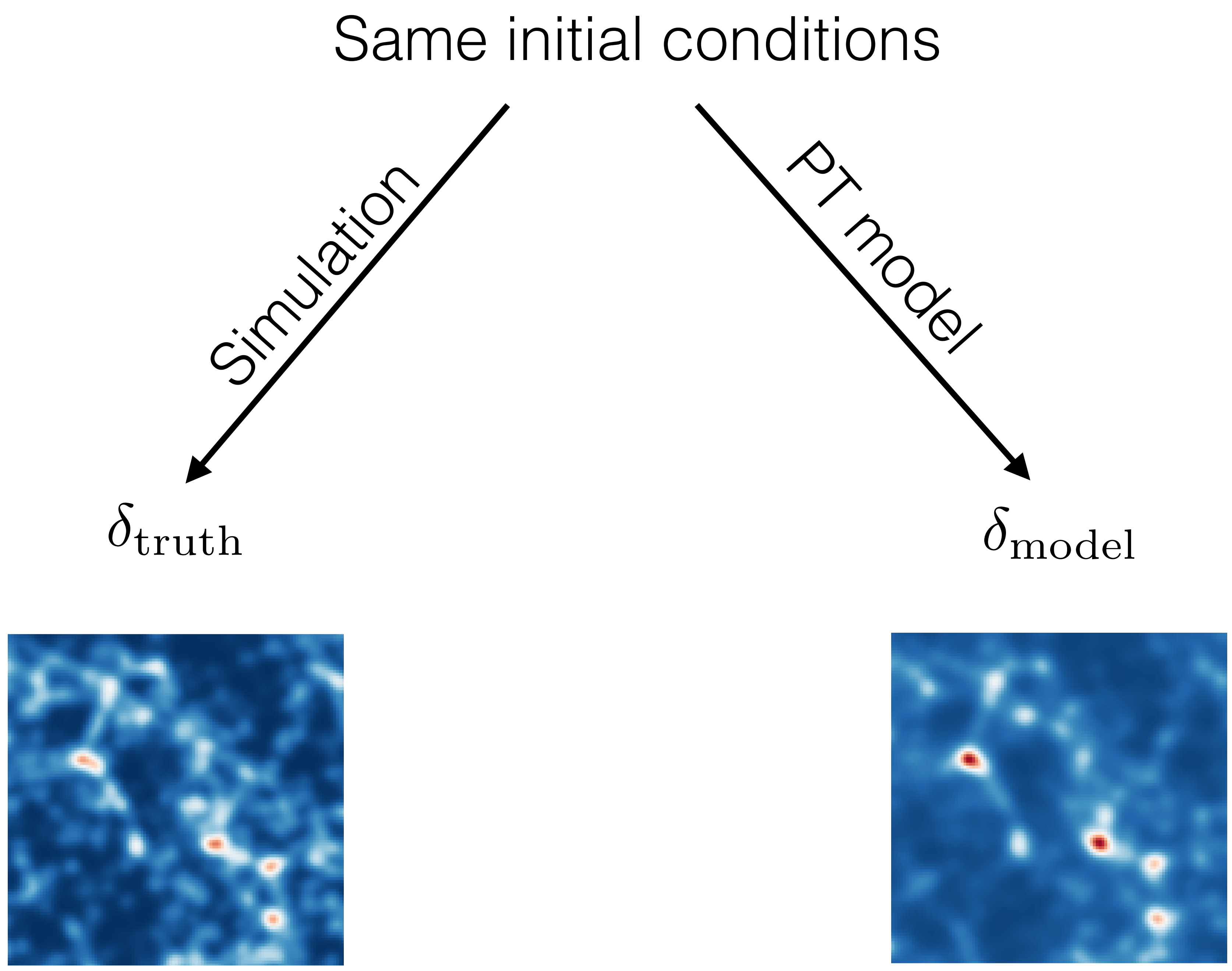
— High S/N at low k, no need to go to the very nonlinear regime

— No overfitting, each Fourier mode (amplitude and phase) is fitted

— "All" n-point functions measured simultaneously

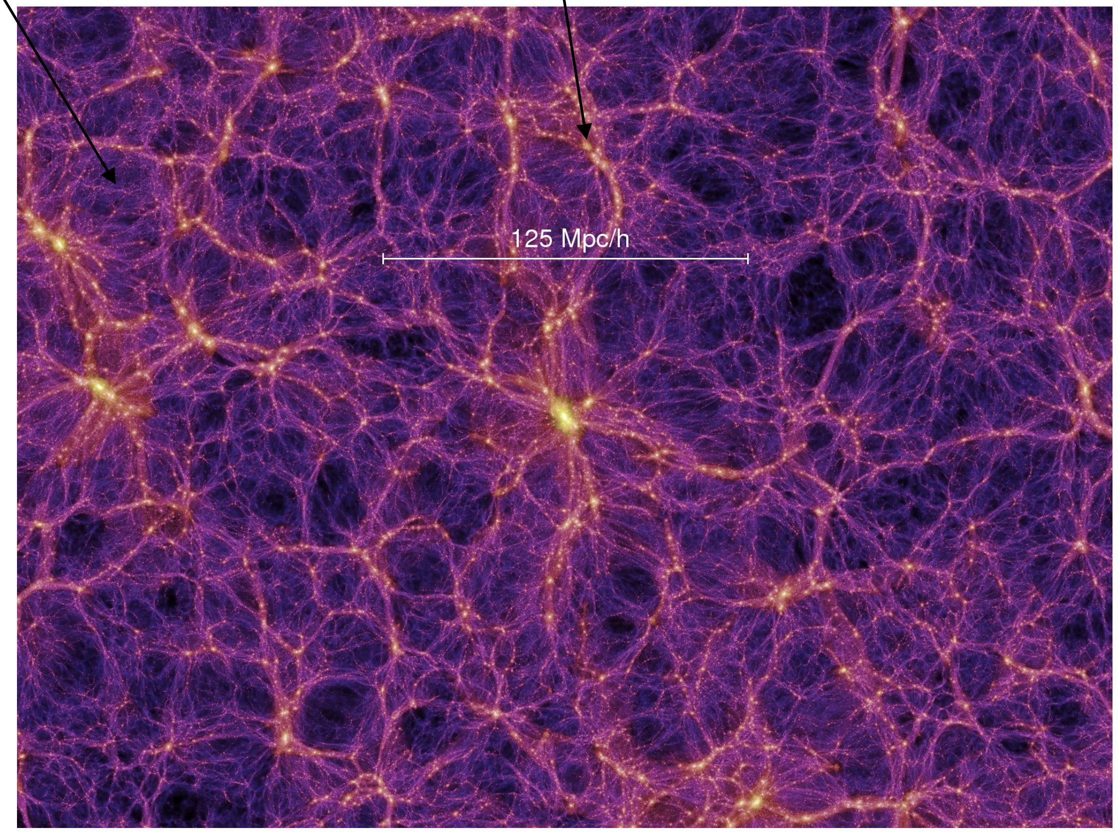
— Easier to isolate and study the noise

Overview


Same initial conditions
Simulation
PT model
$\delta_{\rm truth}$
$\delta_{\rm model}$

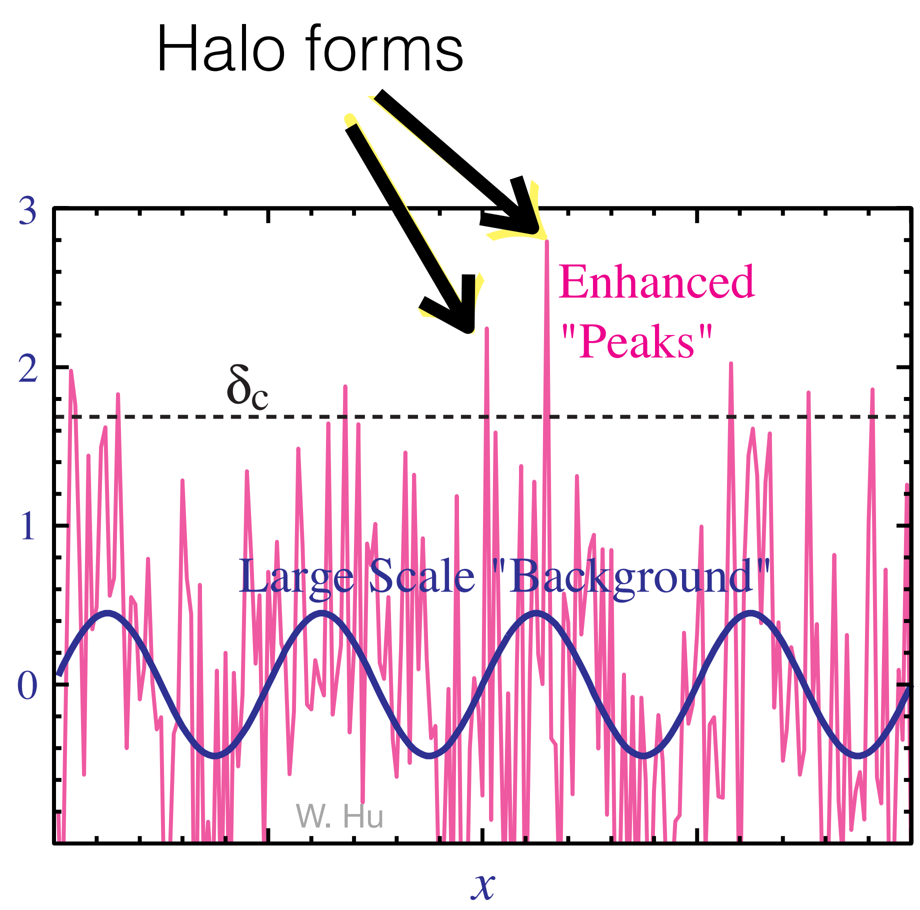
Modeling halo number density

First predict dark matter (purple), then halos (yellow)



Simple peak model

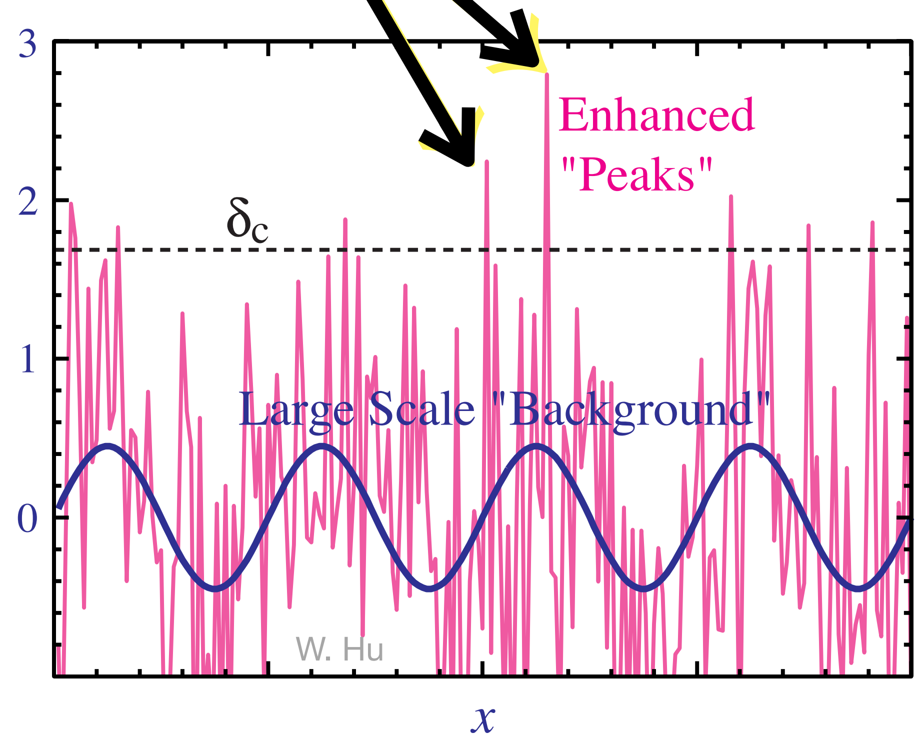
Whenever dark matter density > threshold, a halo forms



Simple peak model

Whenever dark matter density > threshold, a halo forms

Halo forms

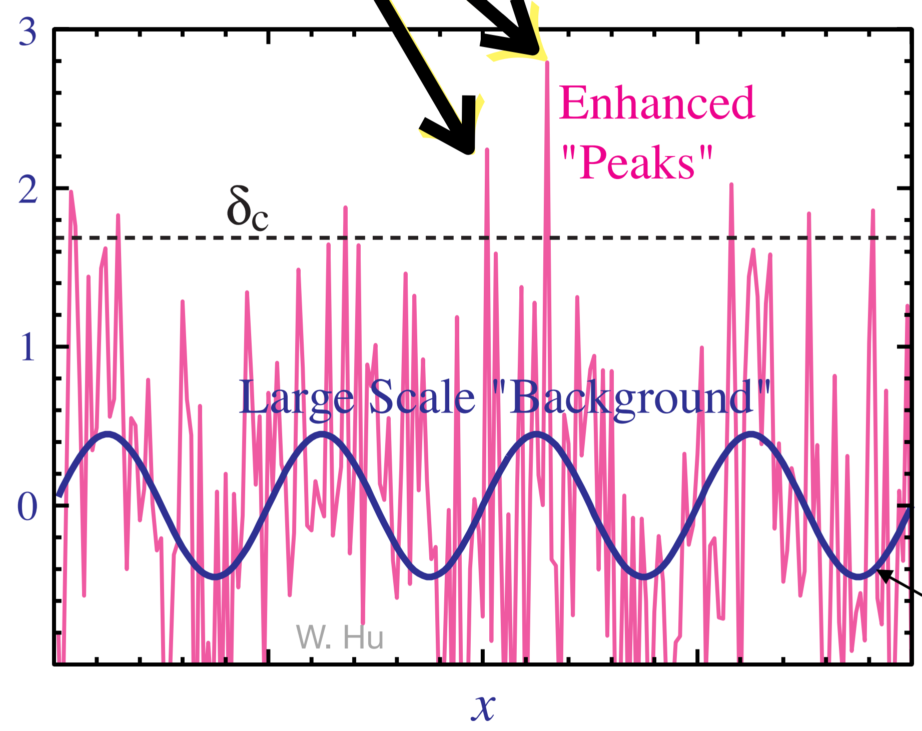


Snow falls

Simple peak model

Whenever dark matter density > threshold, a halo forms

Halo forms



Snow falls

If we zoom out, halos trace the large scale background that modulates the small-scale fluctuations and increases P(>threshold).

Simple peak model

Tracing the large scale background means halo number density is proportional to matter density, if both smoothed on large scales:

$$n_g(\mathbf{x}) \propto \rho_m(\mathbf{x})$$

Write this as $\boxed{\delta_g(\mathbf{x}) = b_1 \delta_m(\mathbf{x})}$

using fractional deviations from the mean

$$\delta_g(\mathbf{x}) \equiv \frac{n_g(\mathbf{x})}{\langle n_g \rangle} - 1 \qquad \delta_m(\mathbf{x}) \equiv \frac{\rho_m(\mathbf{x})}{\langle \rho_m \rangle} - 1$$

and proportionality constant b_1 (related to how likely a halo forms when changing the threshold density)

Simple peak model

Tracing the large scale background means halo number density is proportional to matter density, if both smoothed on large scales:

$$n_g(\mathbf{x}) \propto \rho_m(\mathbf{x})$$

Write this as $\boxed{\delta_g(\mathbf{x}) = b_1 \delta_m(\mathbf{x})}$ **linear regression**

using fractional deviations from the mean

$$\delta_g(\mathbf{x}) \equiv \frac{n_g(\mathbf{x})}{\langle n_g \rangle} - 1 \qquad \delta_m(\mathbf{x}) \equiv \frac{\rho_m(\mathbf{x})}{\langle \rho_m \rangle} - 1$$

and proportionality constant b_1 (related to how likely a halo forms when changing the threshold density)

Simulations

Ran 5 MP-Gadget[1] DM-only N-body sims with 1536^3 DM particles, 3072^3 mesh for PM forces, L=500 Mpc/h, $m_{\rm ptcle} = 2.9 \times 10^9\, M_\odot/h$

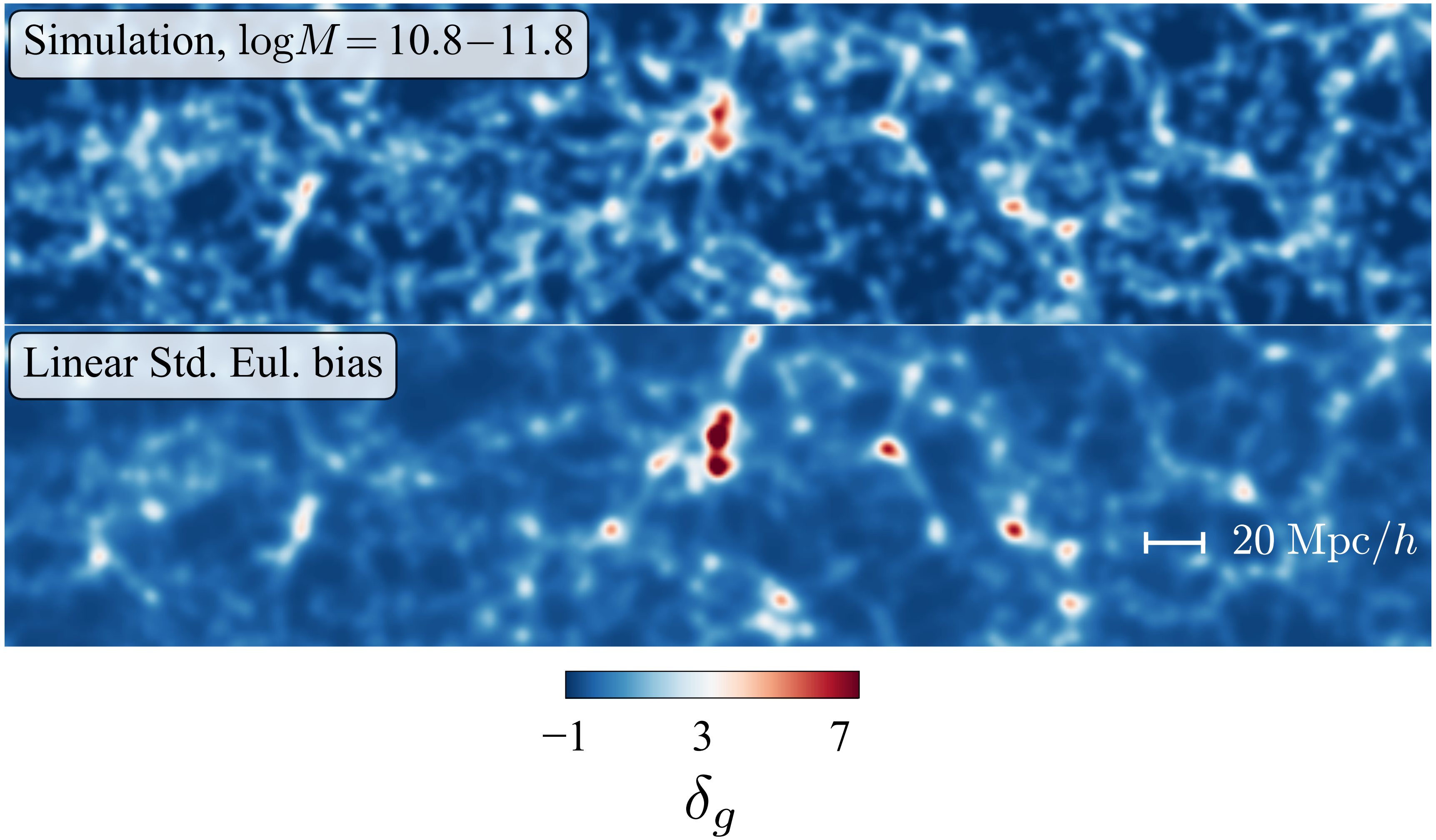
~4000 time steps to evolve z=99 to z=0.6

4 FoF halo mass bins

$\log M[h^{-1}\mathrm{M}_\odot]$	$\bar{n}\,[(h^{-1}\mathrm{Mpc})^{-3}]$	$\bar{n}$ is comparable to
$10.8 - 11.8$	4.3×10^{-2}	LSST [80, 81], Billion Object Apparatus [82]
$11.8 - 12.8$	5.7×10^{-3}	SPHEREx [83, 84]
$12.8 - 13.8$	5.6×10^{-4}	BOSS CMASS [85], DESI [86, 87], Euclid [88–90]
$13.8 - 15.2$	2.6×10^{-5}	Cluster catalogs

[1]Feng et al. https://github.com/bluetides-project/MP-Gadget
[derived from P-Gadget]

Test of linear model



Simulation (= truth)

Model $b_1 \delta_m(\mathbf{x})$

Reasonable prediction on large scales

Missing some structure on small scales

MS, Simonović *et al.* (2019)

Nonlinear model

So far used linear model

$$\delta_g(\mathbf{x}) = b_1 \delta_m(\mathbf{x})$$

Nonlinear model

So far used linear model

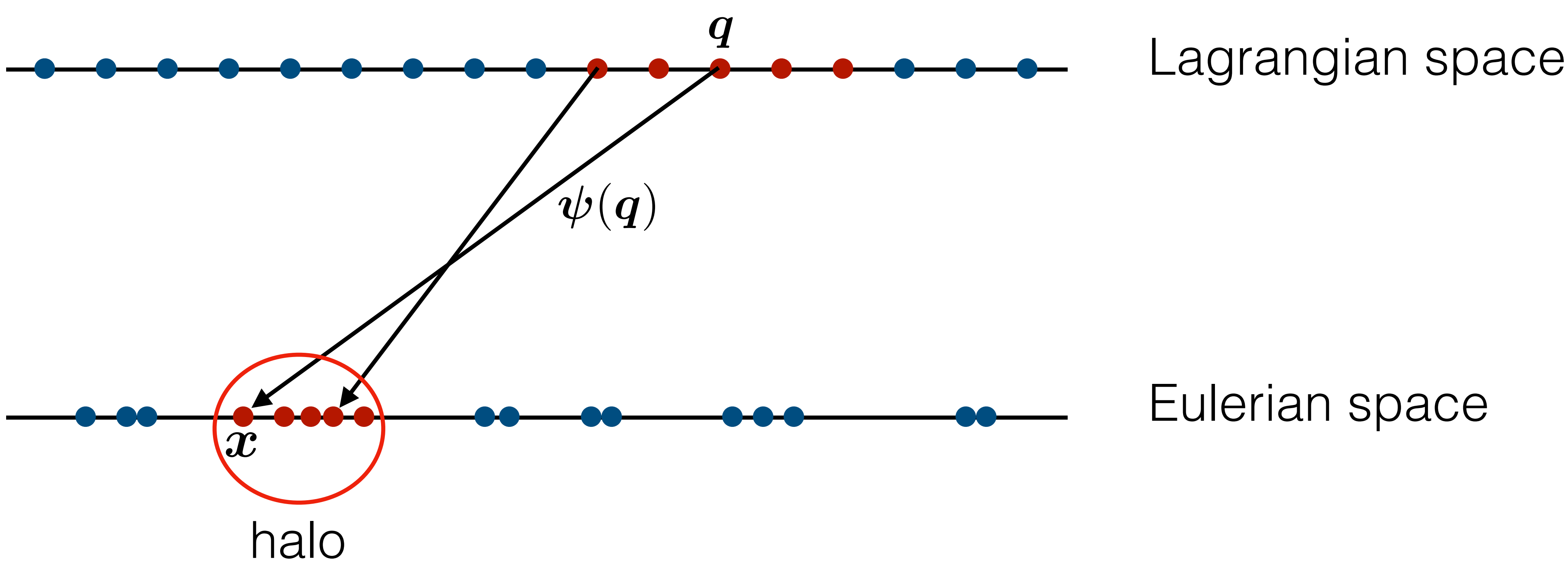
$$\delta_g(\mathbf{x}) = b_1 \delta_m(\mathbf{x}) + b_2 \delta_m^2(\mathbf{x}) + \text{tidal term} + b_3 \delta_m^3(\mathbf{x}) + \cdots$$

Include all nonlinear terms allowed by symmetries (EFT)

Fit parameters b_i by minimizing mean-squared error (least-squares 'polynomial' regression)

MS, Simonović *et al.* (2019)

Bulk flows



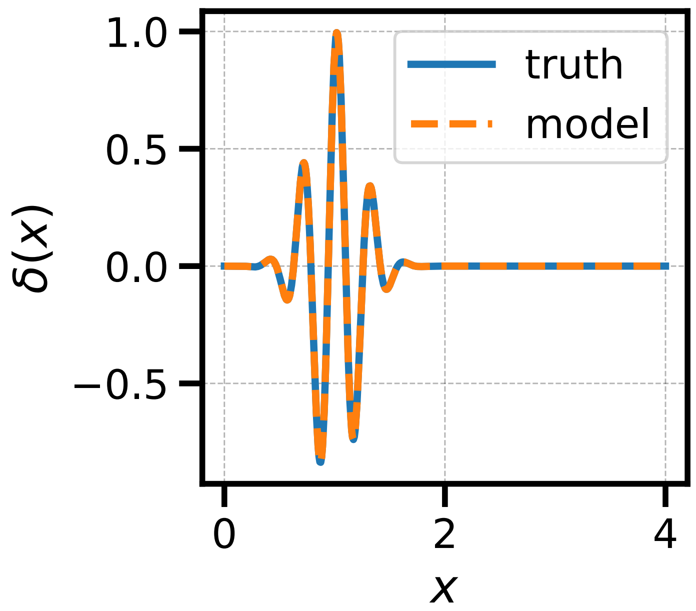
$$\boldsymbol{\psi}_1(\boldsymbol{q}) = \int_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{q}}\, \frac{i\boldsymbol{k}}{k^2}\, \delta_1(\boldsymbol{k})$$

linear displacement is large

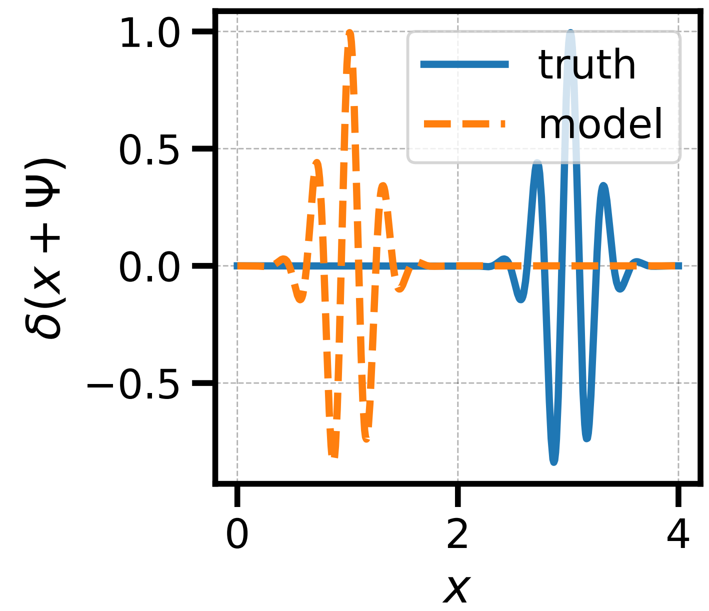
Bulk flows

Good model of δ_{truth}

Displace δ_{truth} by a shift Ψ



$$x \rightarrow x + \Psi$$



Large bulk flows lead to large model error $\epsilon(x) \equiv \delta_{\mathrm{truth}} - \delta_{\mathrm{model}}$ because fields are incoherent

Need a model that takes into account large bulk flows

Including bulk flows in the model

$$\delta_h^{\mathrm{L}}(\boldsymbol{q}) = b_1^{\mathrm{L}}\,\delta_1(\boldsymbol{q}) \,+\, b_2^{\mathrm{L}}\,(\delta_1^2(\boldsymbol{q}) - \sigma_1^2) \,+\, b_{\mathcal{G}_2}^{\mathrm{L}}\mathcal{G}_2(\boldsymbol{q}) + \cdots$$

$$\sigma_1^2 = \langle \delta_1^2(\boldsymbol{q}) \rangle = \int_0^\infty \frac{\mathrm{d}k}{2\pi^2}\,k^2 P_{11}(k)$$

$$\delta_h(\boldsymbol{k}) \equiv \int \mathrm{d}^3\boldsymbol{x}\,(1 + \delta_h(\boldsymbol{x}))\,e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} = \int \mathrm{d}^3\boldsymbol{q}\,(1 + \delta_h(\boldsymbol{q}))\,e^{-i\boldsymbol{k}\cdot(\boldsymbol{q}+\boldsymbol{\psi}(\boldsymbol{q}))}$$

$$\delta_h(\boldsymbol{k}) = \int \mathrm{d}^3\boldsymbol{q}\,\Big(1 + b_1^{\mathrm{L}}\,\delta_1(\boldsymbol{q}) \,+\, b_2^{\mathrm{L}}\,(\delta_1^2(\boldsymbol{q}) - \sigma_1^2) \,+\, b_{\mathcal{G}_2}^{\mathrm{L}}\mathcal{G}_2(\boldsymbol{q}) + \cdots \\ - i\boldsymbol{k}\cdot\boldsymbol{\psi}_2(\boldsymbol{q}) + \cdots\Big)\,e^{-i\boldsymbol{k}\cdot(\boldsymbol{q}+\boldsymbol{\psi}_1(\boldsymbol{q}))}$$

Usual approximation in (C)LPT

for example: Vlah, Castorina, White (2016)

Shifted operators

Motivates bias expansion in "shifted" operators (incl. bulk flows)

$$\tilde{\mathcal{O}}(\boldsymbol{k}) \equiv \int \mathrm{d}^3\boldsymbol{q}\ \mathcal{O}(\boldsymbol{q})\, e^{-i\boldsymbol{k}\cdot(\boldsymbol{q}+\boldsymbol{\psi}_1(\boldsymbol{q}))}$$

$\delta_h(\boldsymbol{k}) = b_1\,\tilde{\delta}_1(\boldsymbol{k}) + b_2\,\tilde{\delta}_2(\boldsymbol{k}) + b_{\mathcal{G}_2}\,\tilde{\mathcal{G}}_2(\boldsymbol{k}) + \cdots$ + noise

Shifted operators

Motivates bias expansion in "shifted" operators (incl. bulk flows)

$$\tilde{\mathcal{O}}(\boldsymbol{k}) \equiv \int \mathrm{d}^3\boldsymbol{q}\ \mathcal{O}(\boldsymbol{q})\, e^{-i\boldsymbol{k}\cdot(\boldsymbol{q}+\boldsymbol{\psi}_1(\boldsymbol{q}))}$$

$$\delta_h(\boldsymbol{k}) = b_1\,\tilde{\delta}_1(\boldsymbol{k}) + b_2\,\tilde{\delta}_2(\boldsymbol{k}) + b_{\mathcal{G}_2}\,\tilde{\mathcal{G}}_2(\boldsymbol{k}) + \cdots \text{ + noise}$$

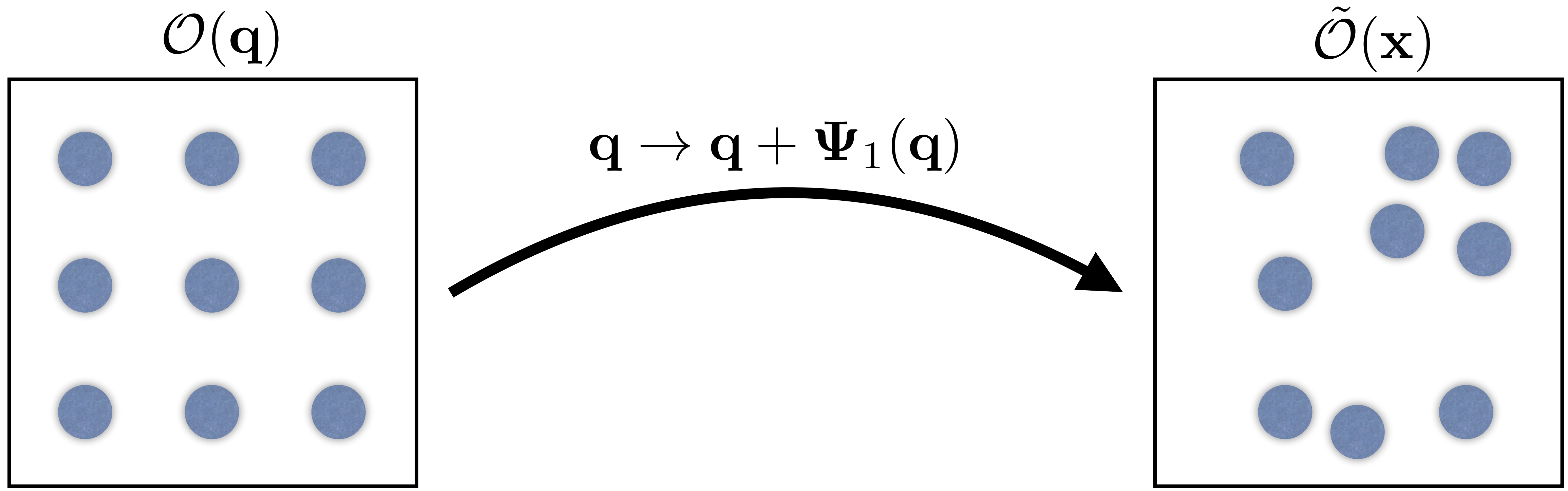
Result is in Eulerian space so easy to compare against simulations

Has IR resummation, giving correct halo positions and BAO spread

Model is perturbative, only linear fields used in the construction

Power spectrum agrees with resummed 1-loop PT

Model on the grid



Distribute 1536^3 particles on regular grid $\mathbf{q}$

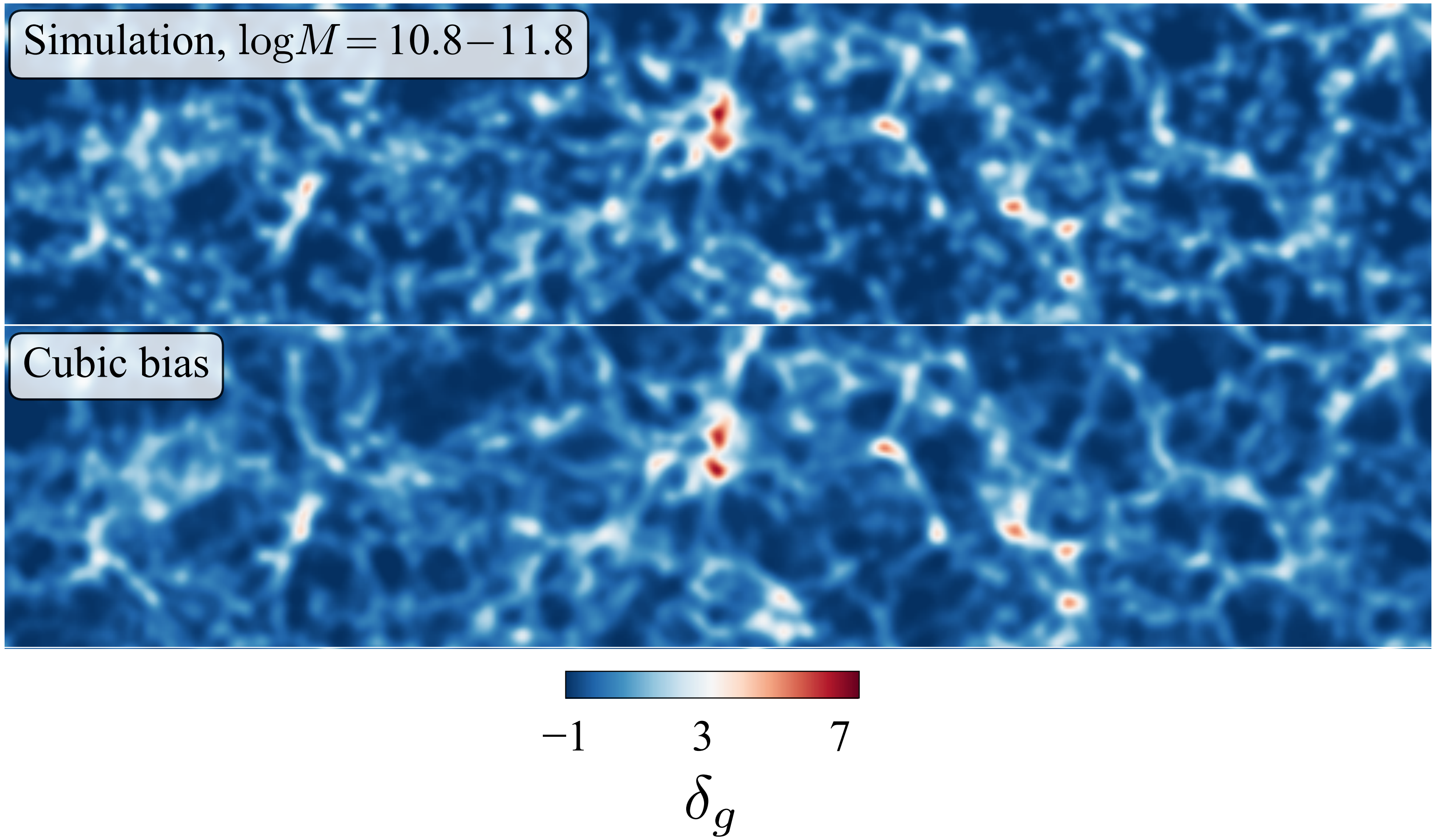
Assign artificial particle masses $m_i = \mathcal{O}(\mathbf{q}_i)$

Displace by linear displacement $\mathbf{\Psi}_1(\mathbf{q})$

Interpolate to Eulerian grid using CIC weighted by particle masses

[Very similar to generating N-body initial conds./Zeldovich density]

Test of nonlinear model



Simulation (= truth)

Nonlinear model

Much better agreement than previous model

MS, Simonović *et al.* (2019)

Measures of success

Error power spectrum (= MSE per wavenumber *k*)

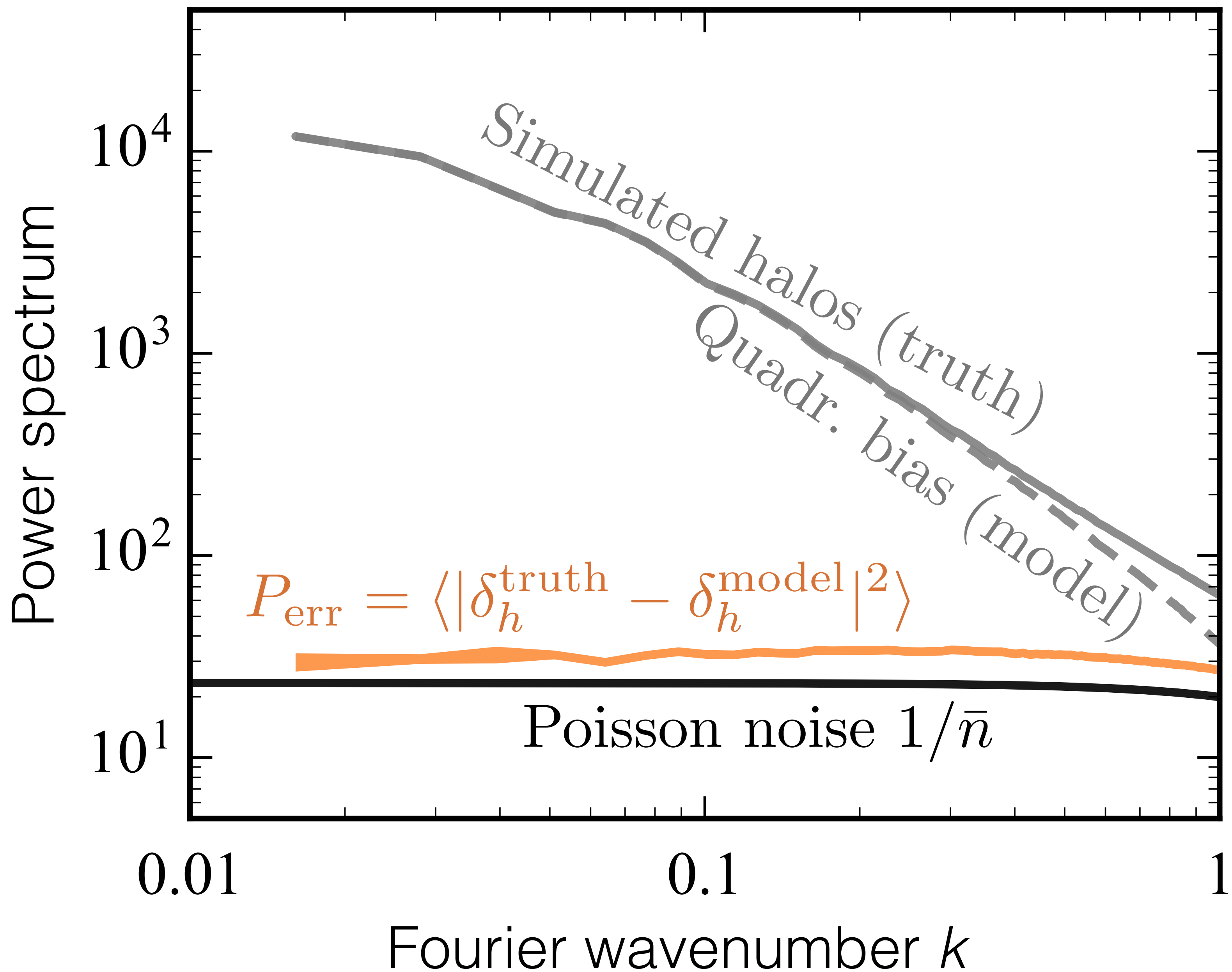
$$P_{\rm err}(k) \equiv \langle |\delta_{\rm truth}(\mathbf{k}) - \delta_{\rm model}(\mathbf{k})|^2 \rangle$$

Cross-correlation between model and truth

$$r_{cc}(k) = \frac{\langle \delta_{\rm truth}(\mathbf{k}) \delta^*_{\rm model}(\mathbf{k}) \rangle}{\sqrt{P_{\rm truth}(k) P_{\rm model}(k)}}$$

For best-fit model, $P_{\rm err} = P_{\rm truth}(1 - r_{cc}^2)$, so focus on $P_{\rm err}$ here

Error power spectrum (= MSE)

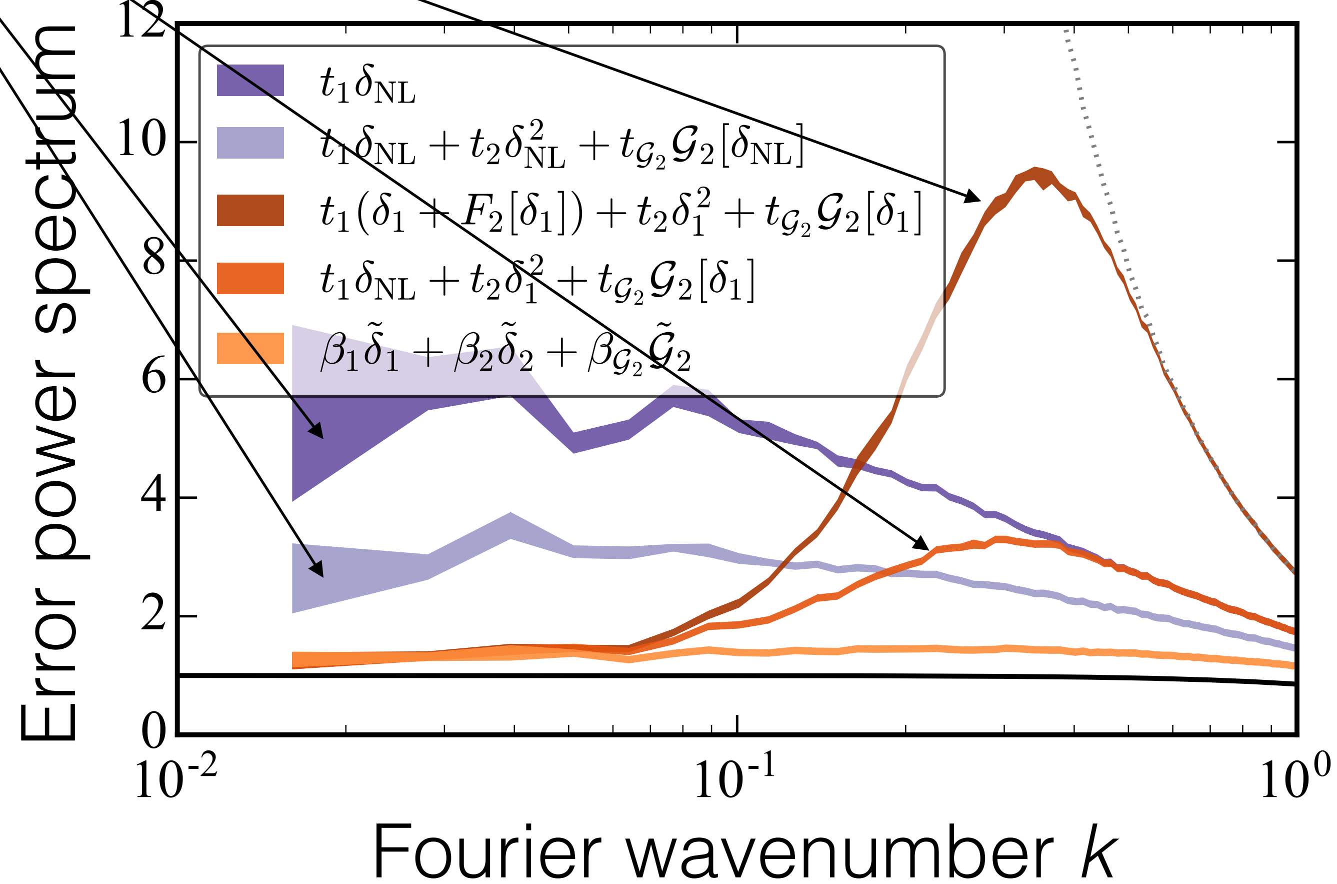


Error is quite white, similar to Poisson shot noise (discrete sampling)

MS, Simonović *et al.* (2019)

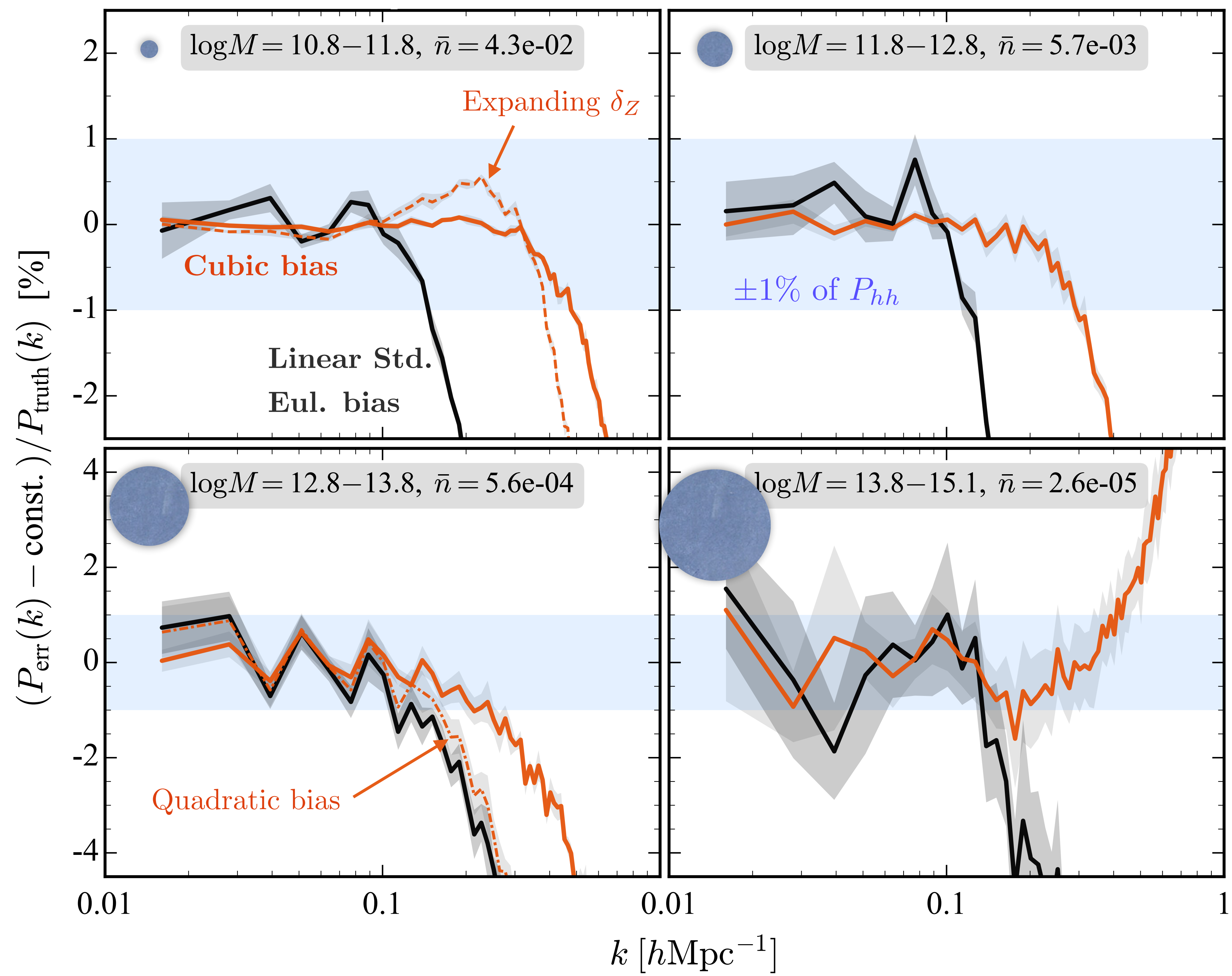
Tried many other nonlinear bias operators

Give few times larger model error



MS, Simonović *et al.* (2019)

Scale dependence of the error



Scale dependence of the noise important around the nonlinear scale

Potentially dangerous because can bias cosmological parameters

k_{max} up to which noise is constant

$k_{\rm max}$ when scale dependence of $P_{\rm err}$ is detectable with 1σ in V=10 h^{-3}Gpc^3 volume (or V=0.5 in brackets):

		$k_{\rm max}\,[h\mathrm{Mpc}^{-1}]$	
$\log M\,[h^{-1}\mathrm{M}_\odot]$	$\bar{n}\,[(h^{-1}\mathrm{Mpc})^{-3}]$	Lin. Std. Eul.	Cubic
$10.8 - 11.8$	4.3×10^{-2}	0.1 (0.14)	0.3 (0.37)
$11.8 - 12.8$	5.7×10^{-3}	0.08 (0.1)	0.18 (0.24)
$12.8 - 13.8$	5.6×10^{-4}	0.07 (0.1)	0.13 (0.18)
$13.8 - 15.2$	2.6×10^{-5}	0.1 (0.14)	0.24 (0.32)

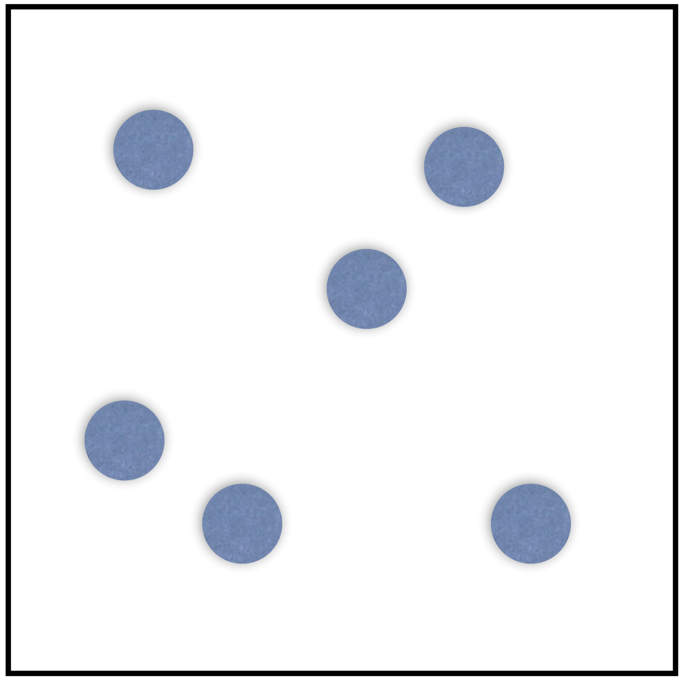
Nonlinear bias has 2-3x higher $k_{\rm max}$

⇒ 8-30x more Fourier modes

⇒ 4-5x smaller error bars (in principle; also have more params!)

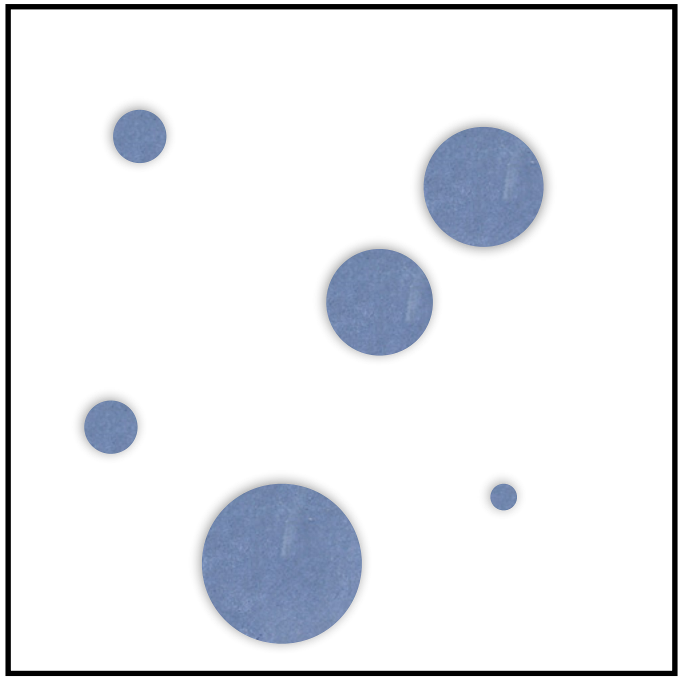
Weighting halos by their mass

Halo *number* density
(how many halos per cell)



used so far

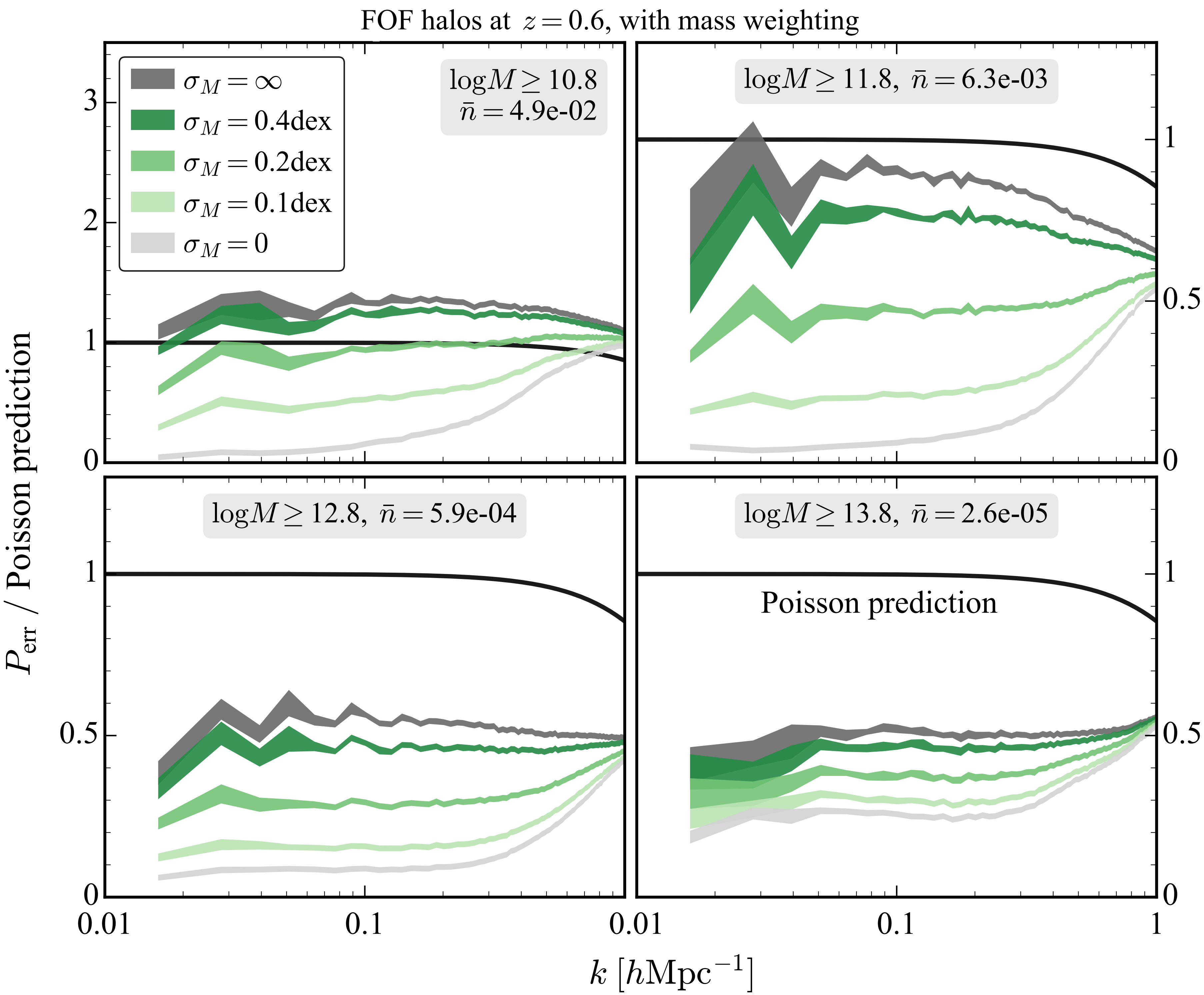
Halo *mass* density
(how much halo mass per cell)



more similar to dark matter
⇒ smaller shot noise

Seljak, Hamaus & Desjacques (2009)
Hamaus, Seljak & Desjacques (2010, 2011, 2012)
Cai, Bernstein & Sheth (2011)

Weighting halos by their mass



Shot noise (squared model error) 17x lower for light halos, 2-7x lower for heavy halos

With 60% halo mass scatter (green), still get factor few

How well can we do observationally?

Mass weighting questions

How well can halo masses be measured (e.g. BOSS, DESI)?

What observable properties of galaxies can we use? What sims?

New ideas to get halo masses?

For shot noise limited applications, gain may be large

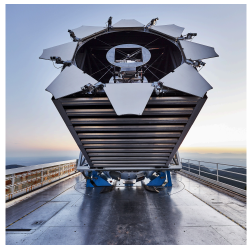
What if mass estimates are biased?

Use for BAO reconstruction? (Suffers from high shot noise)

Science with galaxy catalogs

(1) Bias model at the field level

(2) Cosmological parameter analysis



(3) Accounting for skewness

(4) Getting initial from final conditions

(2) Cosmological parameter analysis

Compare model with galaxy catalogs to infer cosmological parameters θ: Expansion rate, amount of dark matter, curvature, …

Work with power spectrum $\hat{P}_{gg}(k)$ of the galaxy number density (variance of fluctuations as a function of scale).

Monte-Carlo sample θ to get posterior

$$P(\theta|\hat{P}_{gg}) = \frac{\mathcal{L}(\hat{P}_{gg}|\theta)\,P(\theta)}{\int \mathrm{d}\theta' P(\hat{P}_{gg}|\theta')P(\theta')}$$

(2) Cosmological parameter analysis

Compare model with galaxy catalogs to infer cosmological parameters θ: Expansion rate, amount of dark matter, curvature, …

Work with power spectrum $\hat{P}_{gg}(k)$ of the galaxy number density (variance of fluctuations as a function of scale).

Monte-Carlo sample θ to get posterior

$$P(\theta|\hat{P}_{gg}) = \frac{\mathcal{L}(\hat{P}_{gg}|\theta)\, P(\theta)}{\int \mathrm{d}\theta' P(\hat{P}_{gg}|\theta') P(\theta')}$$

shifted bias operators

$$\mathcal{L}(\hat{P}_{gg}|\theta) \propto \exp\left[-\frac{1}{2}\,\frac{\left(\hat{P}_{gg} - P_{gg}^{\mathrm{model}}(\theta)\right)^2}{\mathrm{var}(\hat{P}_{gg})}\right]$$

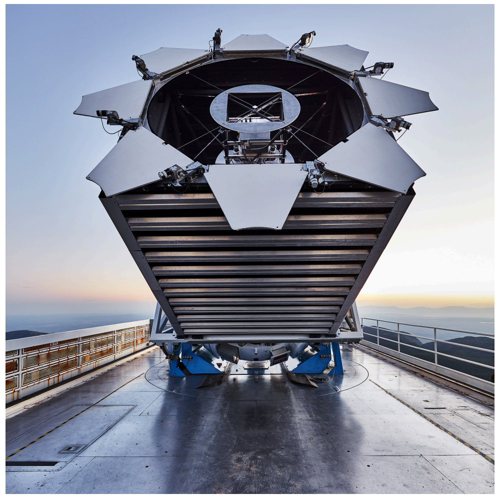
(2) Cosmological parameter analysis

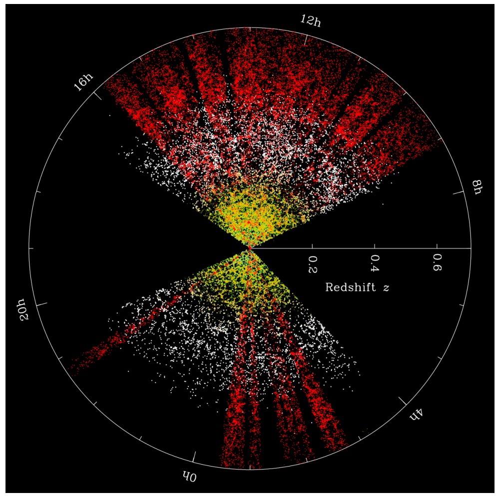
For sampling, need fast evaluation of the model power spectrum (reduce 2D integrals to 1D FFTs)

McEwen+ (2016); MS+ (2016); Cataneo+ (2017); Simonović+ (2018)

Recently applied to Sloan Digital Sky Survey (SDSS)

D'Amico, Gleyzes+; Ivanov, Simonović & Zaldarriaga; Tröster, Sanchez+

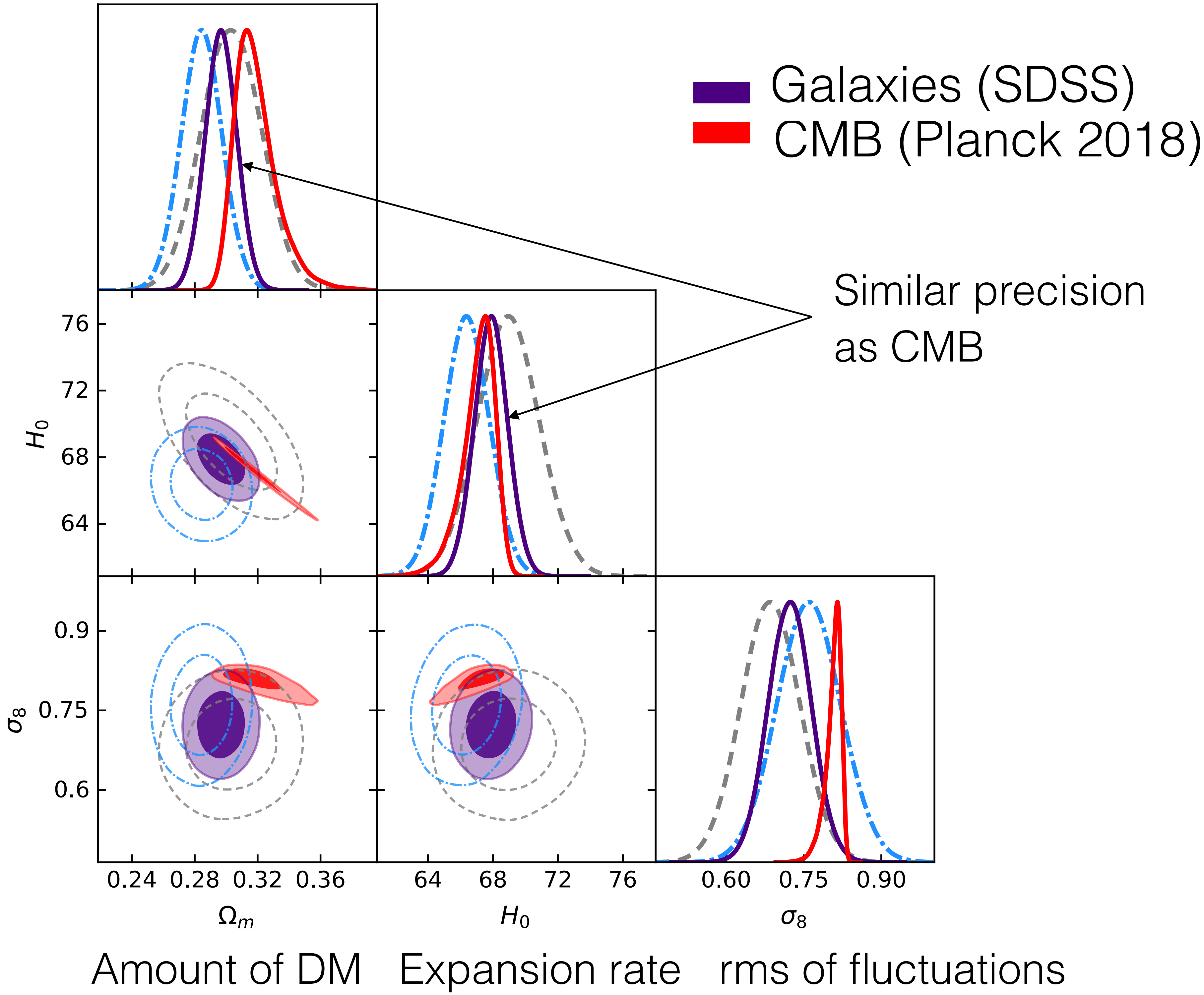




SDSS

Validated using Masahiro's challenge simulations

(2) Cosmological parameter analysis



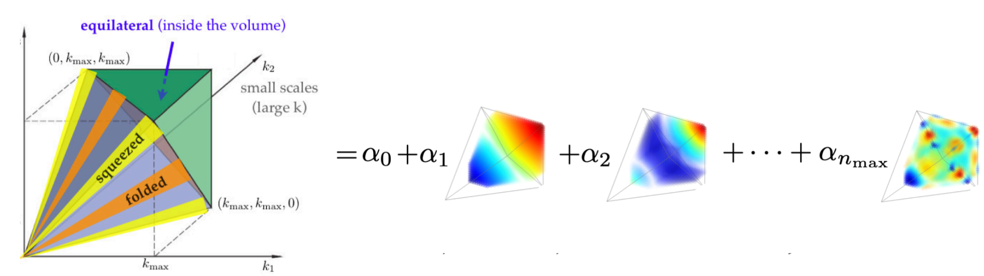
Ivanov+ (arXiv:1909.05277)

Science with galaxy catalogs

(1) Bias model at the field level

(2) Cosmological parameter analysis

(3) Accounting for skewness

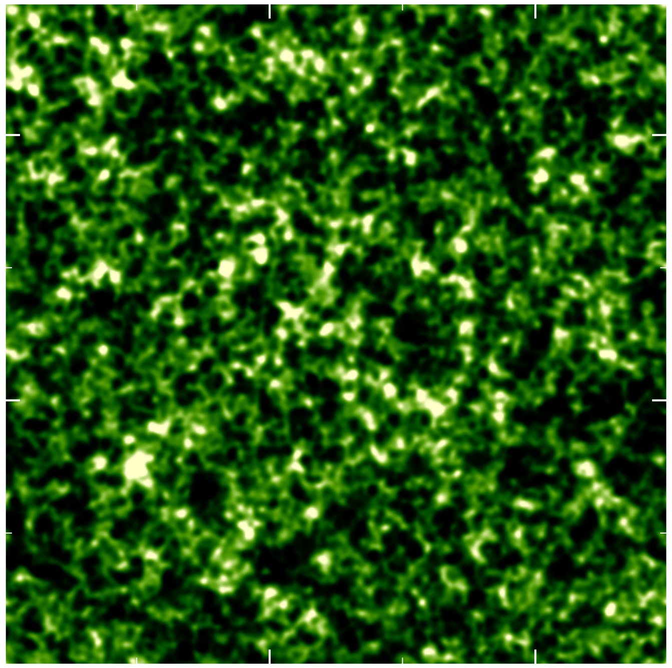


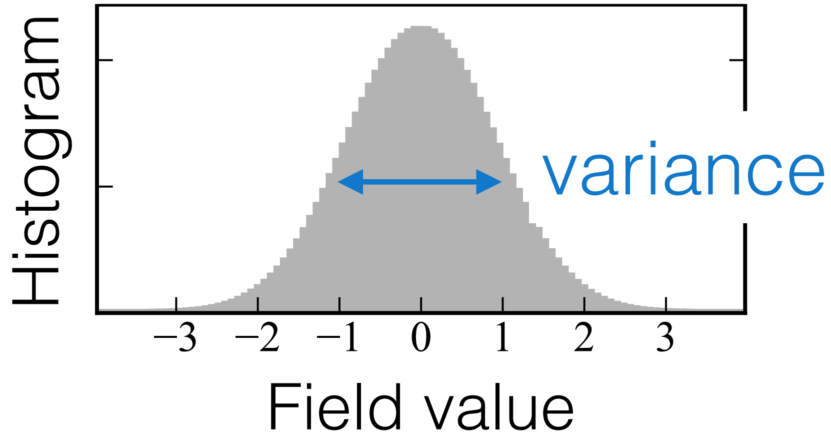
(4) Getting initial from final conditions

Normal distribution

Only used power spectrum, i.e. sample variance of each scale

Gaussian random field



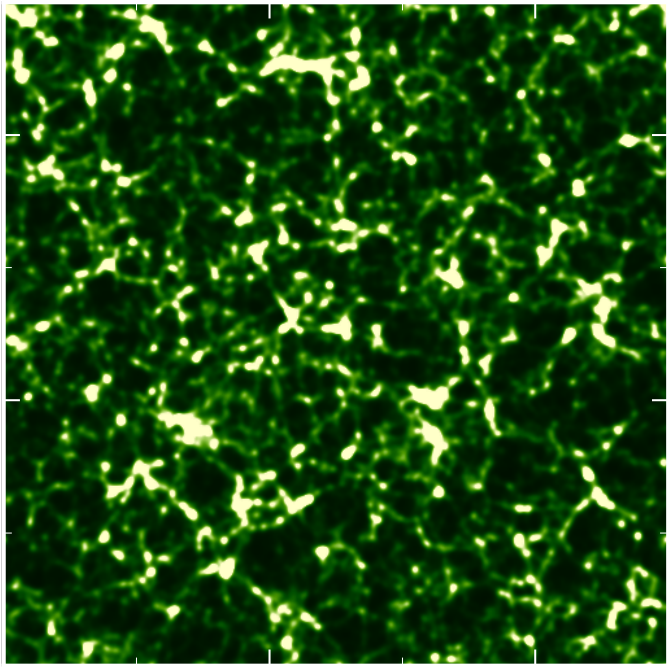


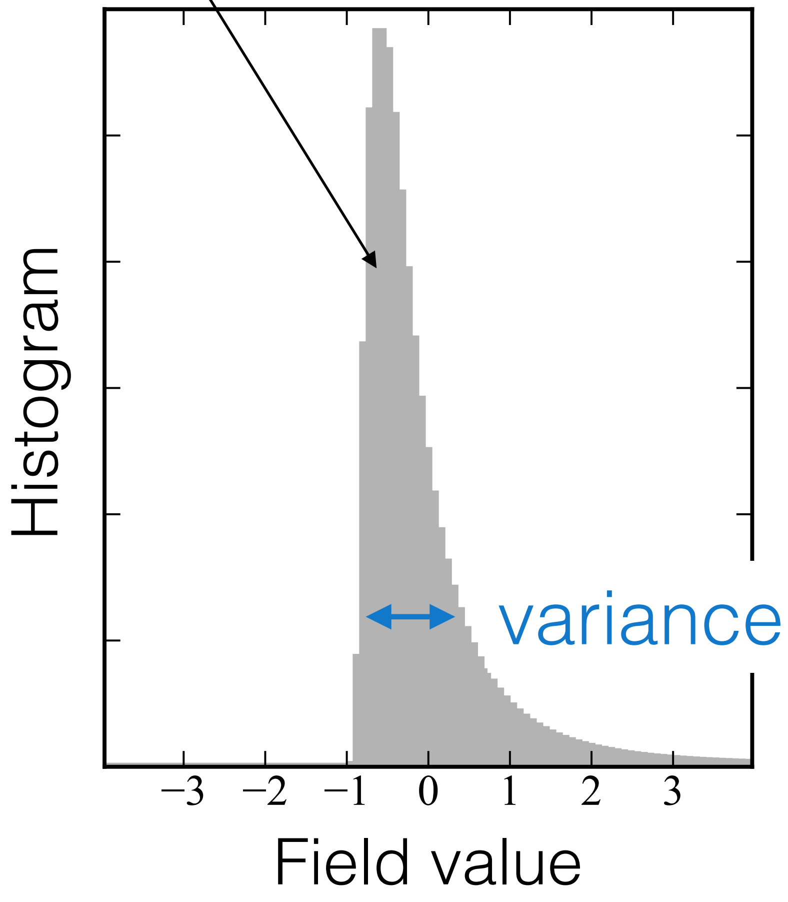
Sufficient if halo/galaxy distribution were a Gaussian random field

Halos

Not normally distributed, pdf is highly skewed

Halo number density





Information in the tails can improve precision of parameter estimates

How to extract information from the tails?

Challenges:

(a) Distribution has no simple analytical form

(b) What summary statistics should we use?

(c) How can we make sure that we extract all the information?

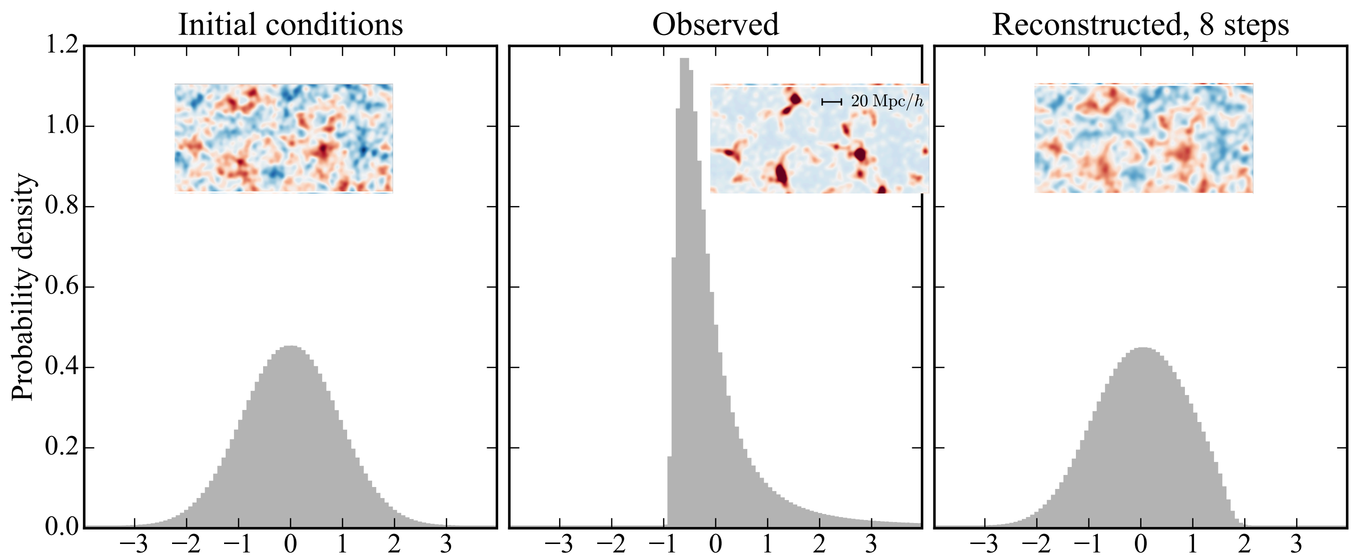
How to extract information from the tails?

Some options:

(a) Count outliers

(b) Measure higher-order moments: Skewness, kurtosis, etc

(c) Reconstruct initial conditions (normally distributed) & measure their sample variance

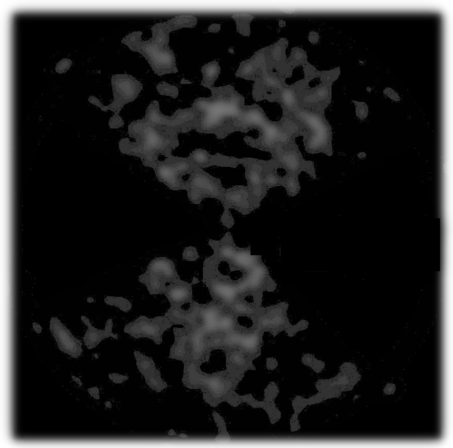


Skewness

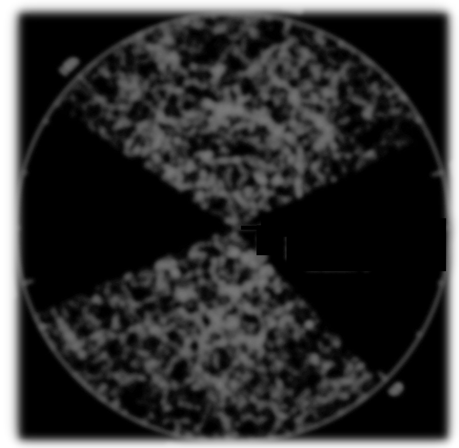
Need scale-dependent generalization of skewness

Correlation of 3 Fourier modes

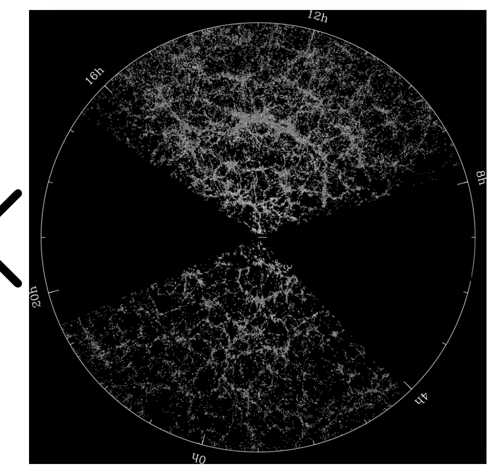
$$\langle\, \delta(\mathbf{k}_1)\, \delta(\mathbf{k}_2)\, \delta(\mathbf{k}_3)\, \rangle \;=$$



$\times$



$\times$

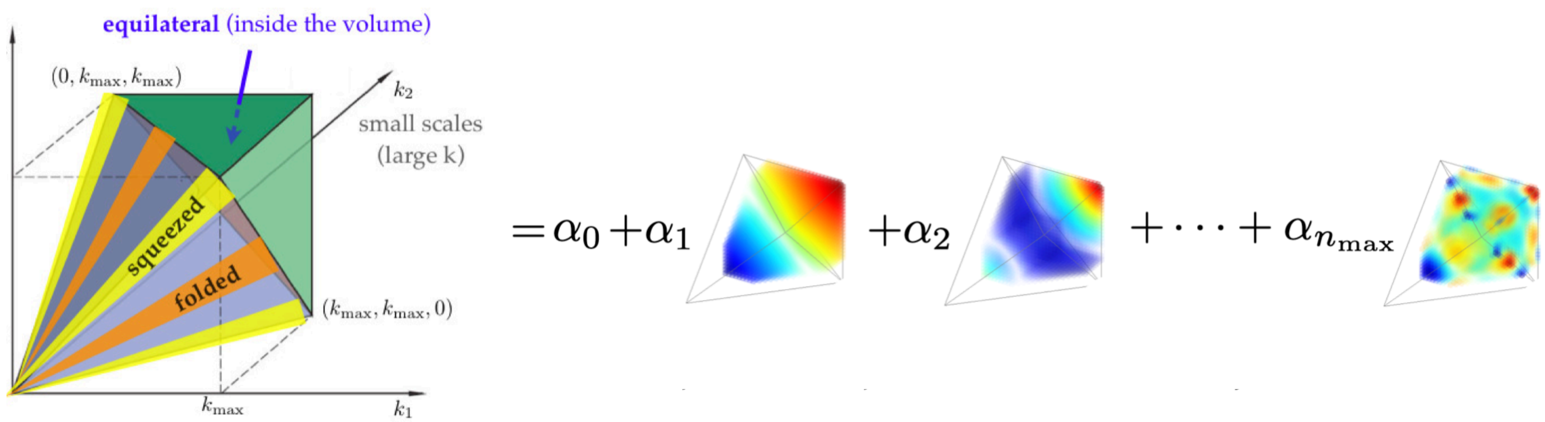


Contains info about Big Bang, relation between galaxies and dark matter, rms amplitude of fluctuations

Measuring skewness

Challenging: Can form too many Fourier mode triplets

Solution 1: In space of all triplets, use simple basis functions to capture most of the information



Solution 2: Given signal of interest, can compute its maximum likelihood estimator (matched filter). This reduces to a sum over all triplets which can be computed using a few 3D FFTs

MS, Regan & Shellard (2013)
MS, Baldauf & Seljak (2015)

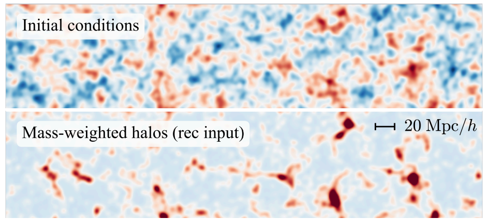
Science with galaxy catalogs

(1) Bias model at the field level

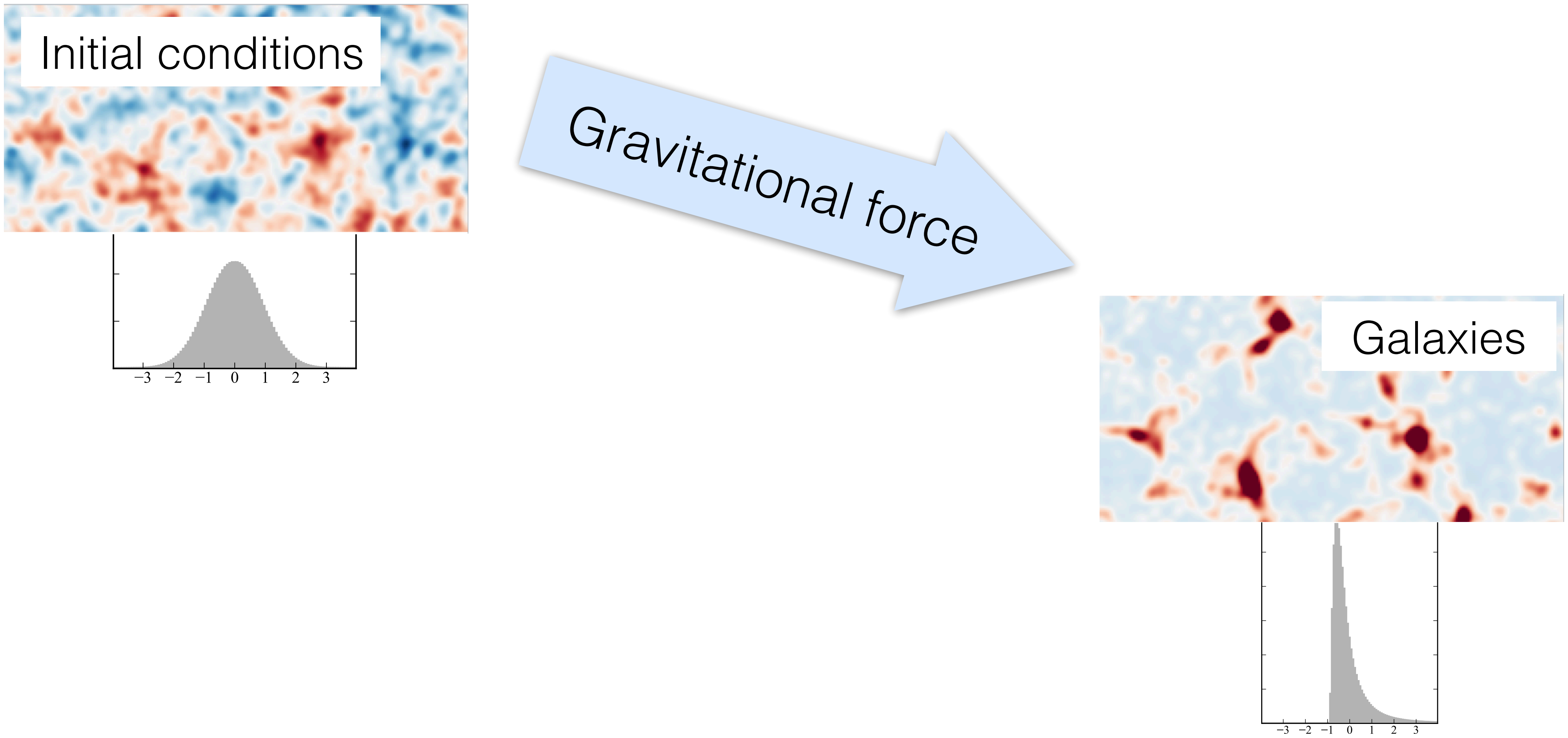
(2) Cosmological parameter analysis

(3) Accounting for skewness

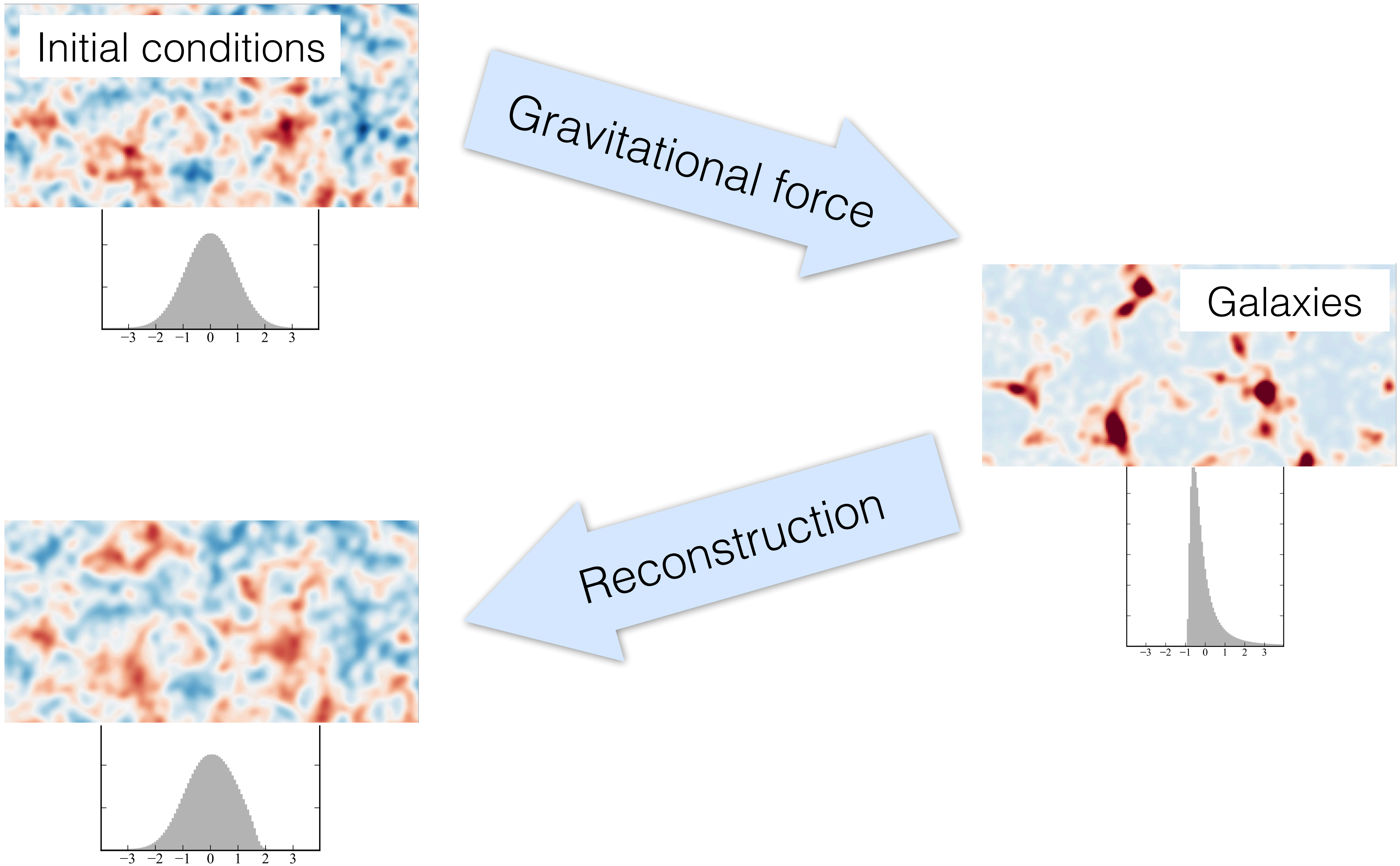
(4) Getting initial from final conditions



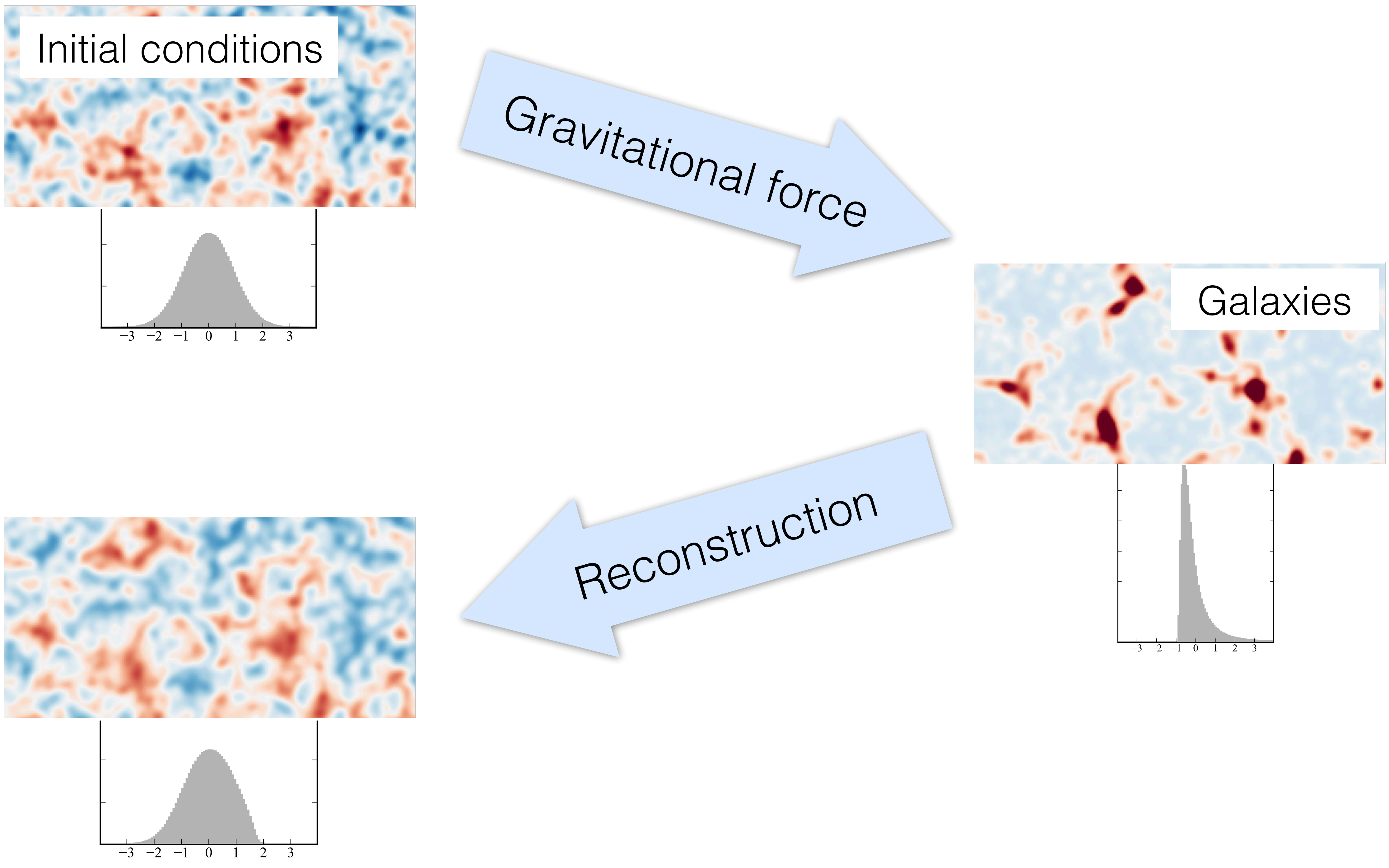
Getting initial from final conditions



Getting initial from final conditions



Getting initial from final conditions



Goal: Reconstruct initial conditions & measure their sample variance

Getting initial from final conditions

Challenges:

(a) Forward evolution is highly nonlinear, so how to invert?

(b) Not injective: Multiple initial conds. can give same final conds.

(c) Algorithm must be fast to be applicable to data

Getting initial from final conditions

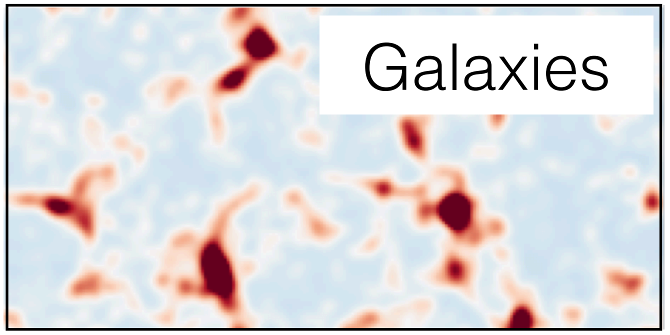
Long history — recently quite active:
Zhu, Yu+ ('17)
Wang, Yu+ ('17)
MS, Baldauf+ ('17)
Seljak, Aslanyan+ ('17)
Modi, Feng+ ('18), Shi+ ('18), Hada+ ('18), Modi, White+ ('19), Sarpa+ ('19), Schmidt+ ('19), Elsner+ ('19), Yu & Zhu ('19), Zhu, White+ ('19)

Also sampling (Jasche, Kitaura, Lavaux, Wandelt)

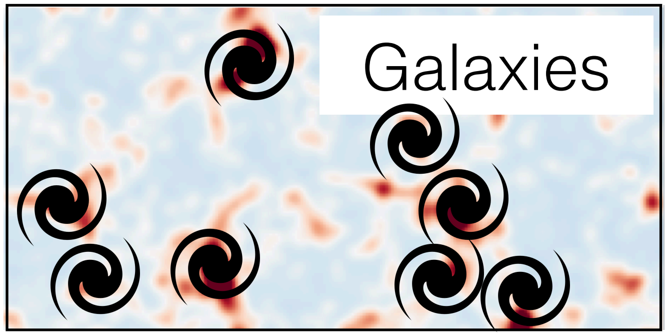
machine learning (Ho, Li, …)

theory (e.g. MS+ ('15), Hikage+ ('17), Wang+ ('18), Sherwin+ ('18))

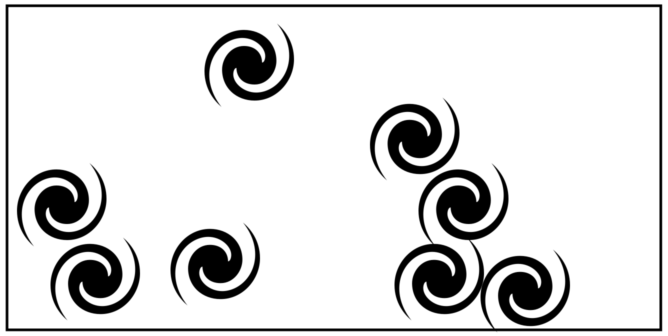
Getting initial from final conditions



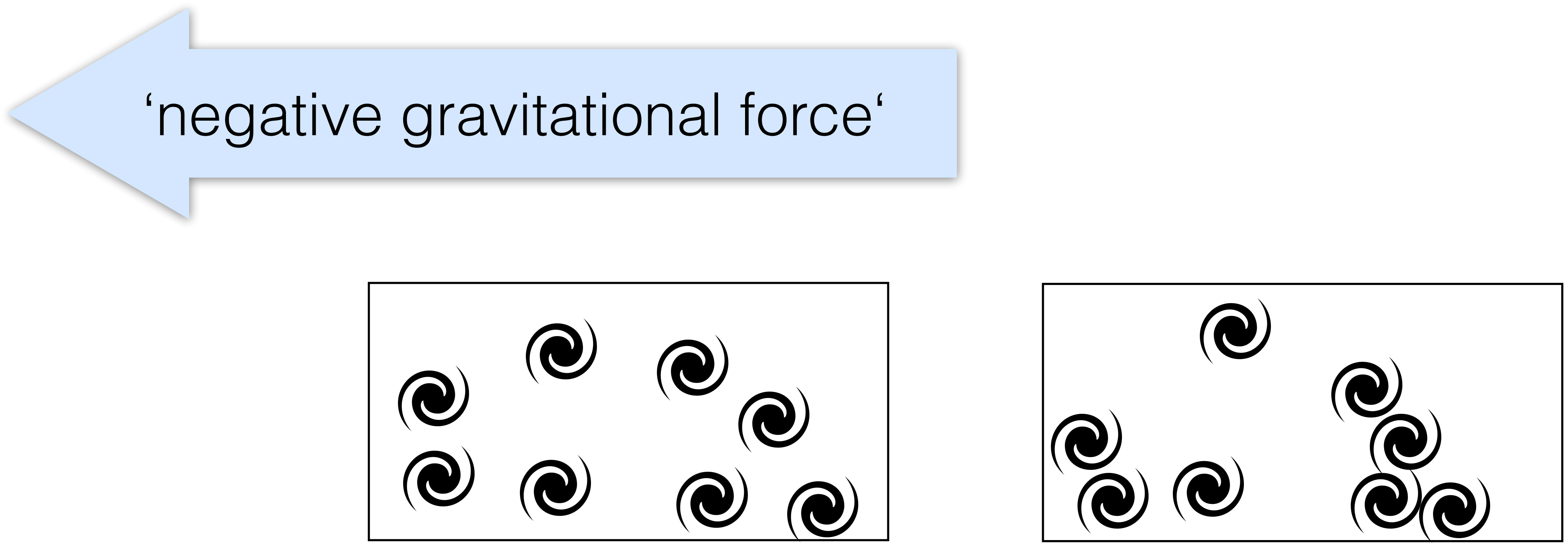
Getting initial from final conditions



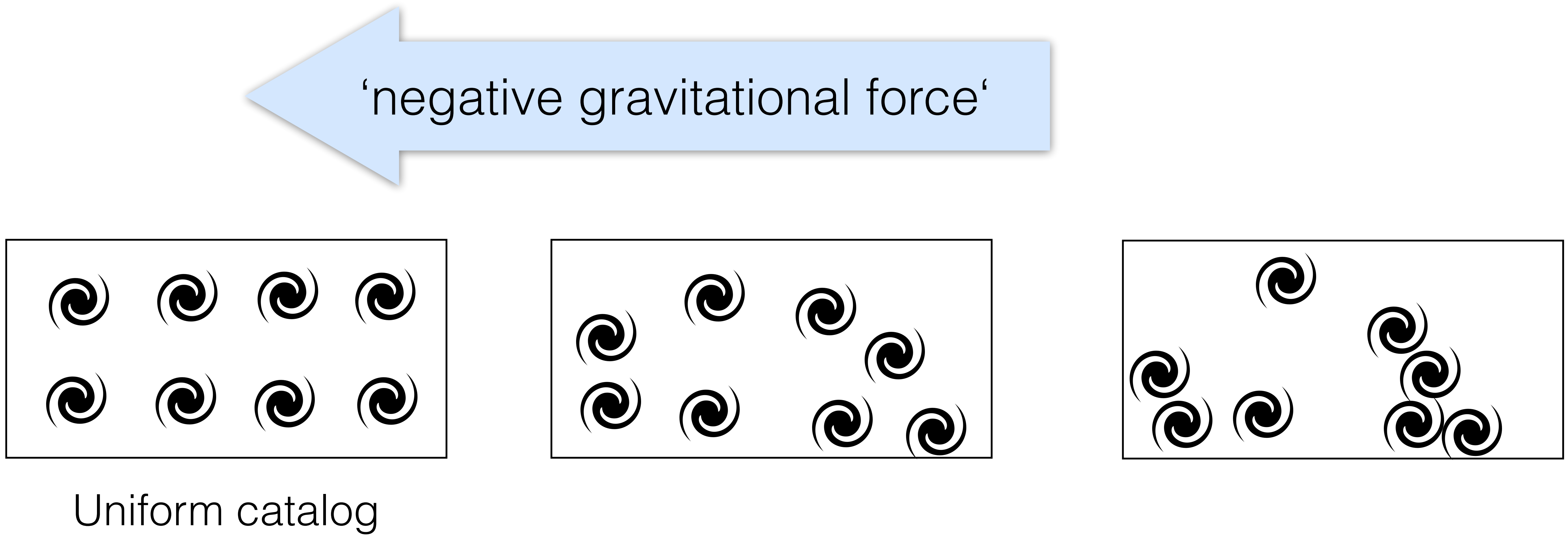
Getting initial from final conditions



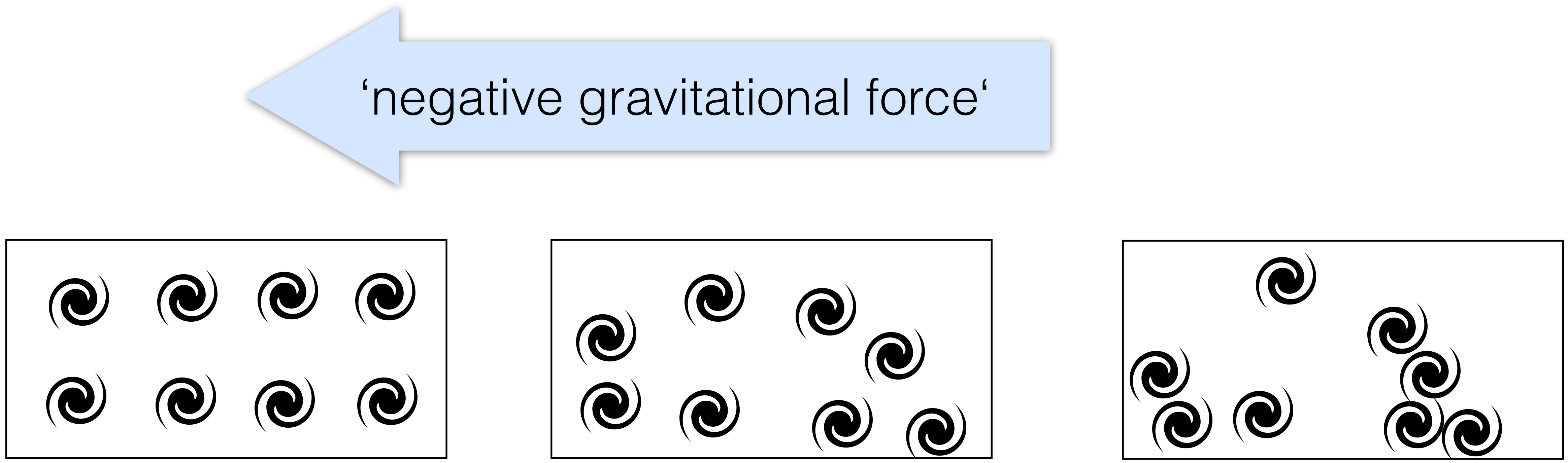
Getting initial from final conditions



Getting initial from final conditions



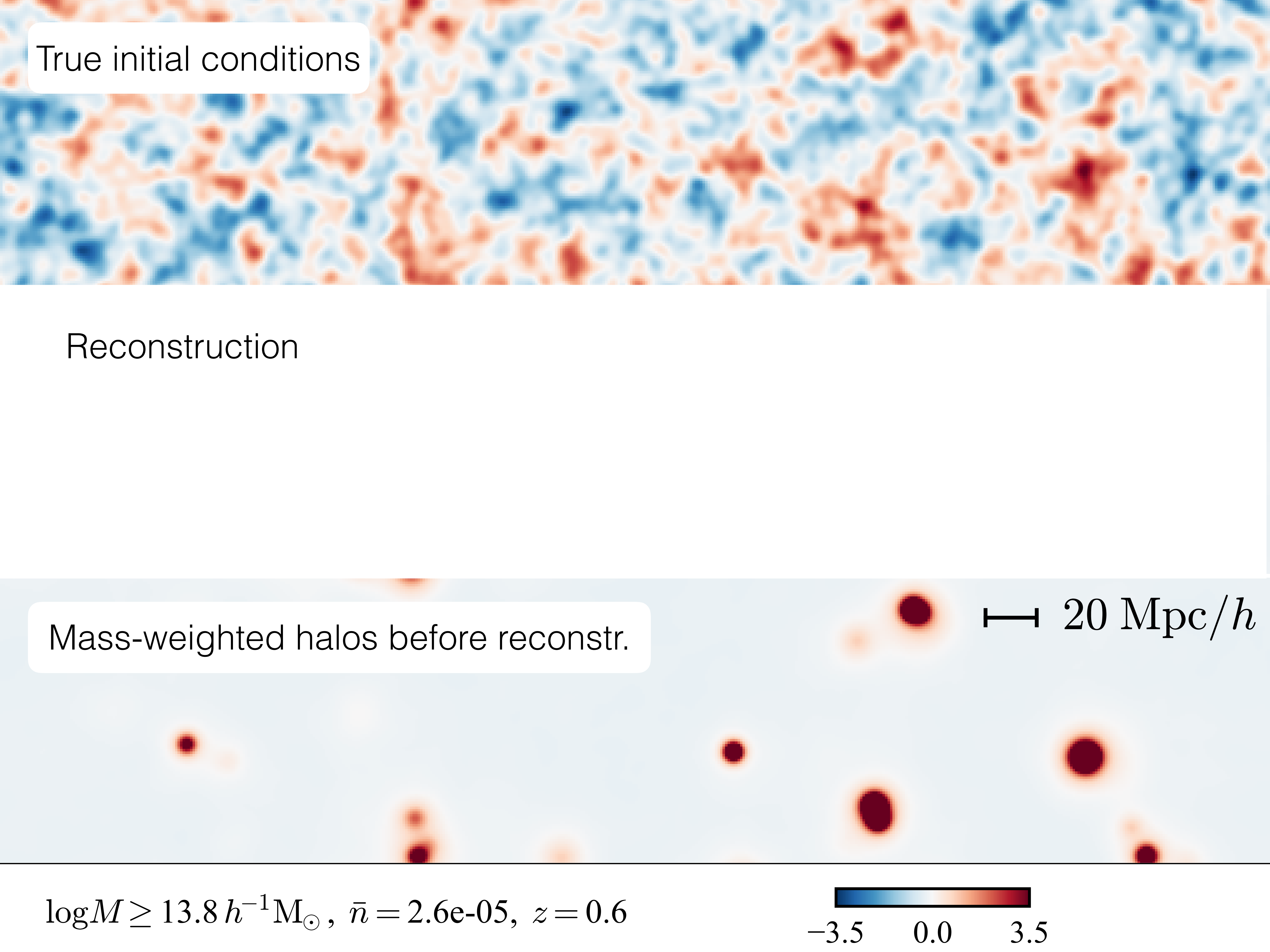
Getting initial from final conditions



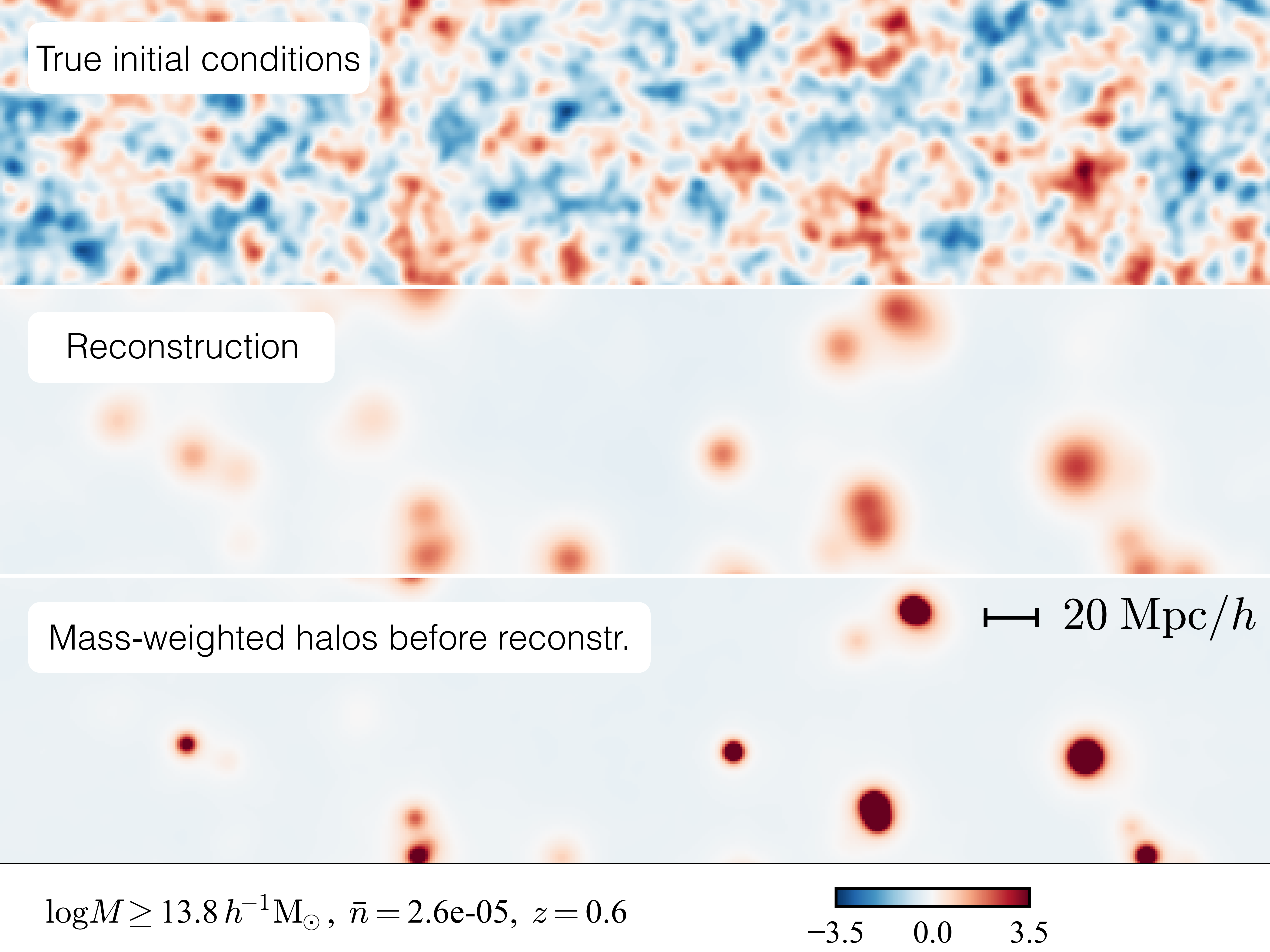
Uniform catalog

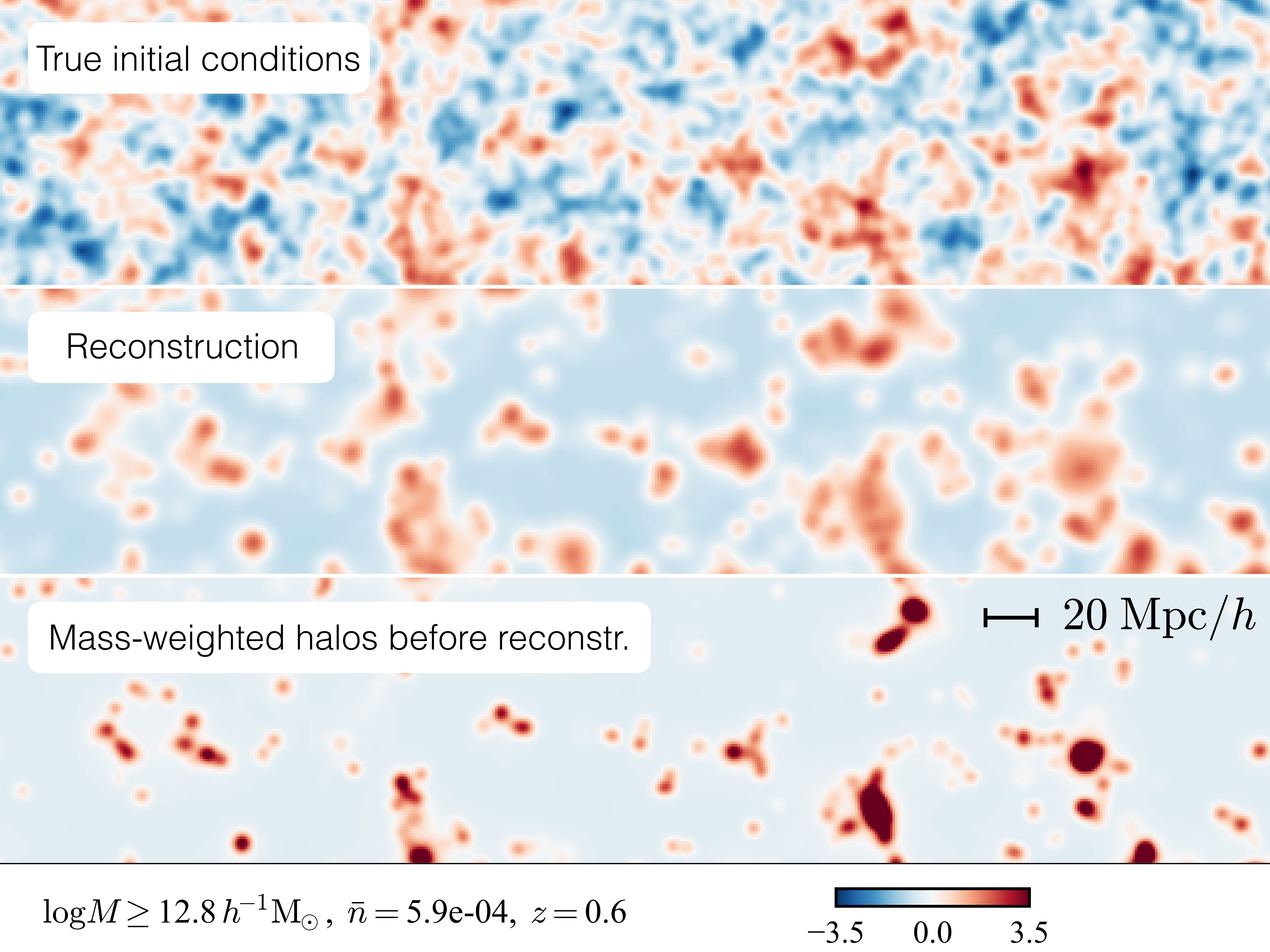
From path of each galaxy, can estimate initial conditions

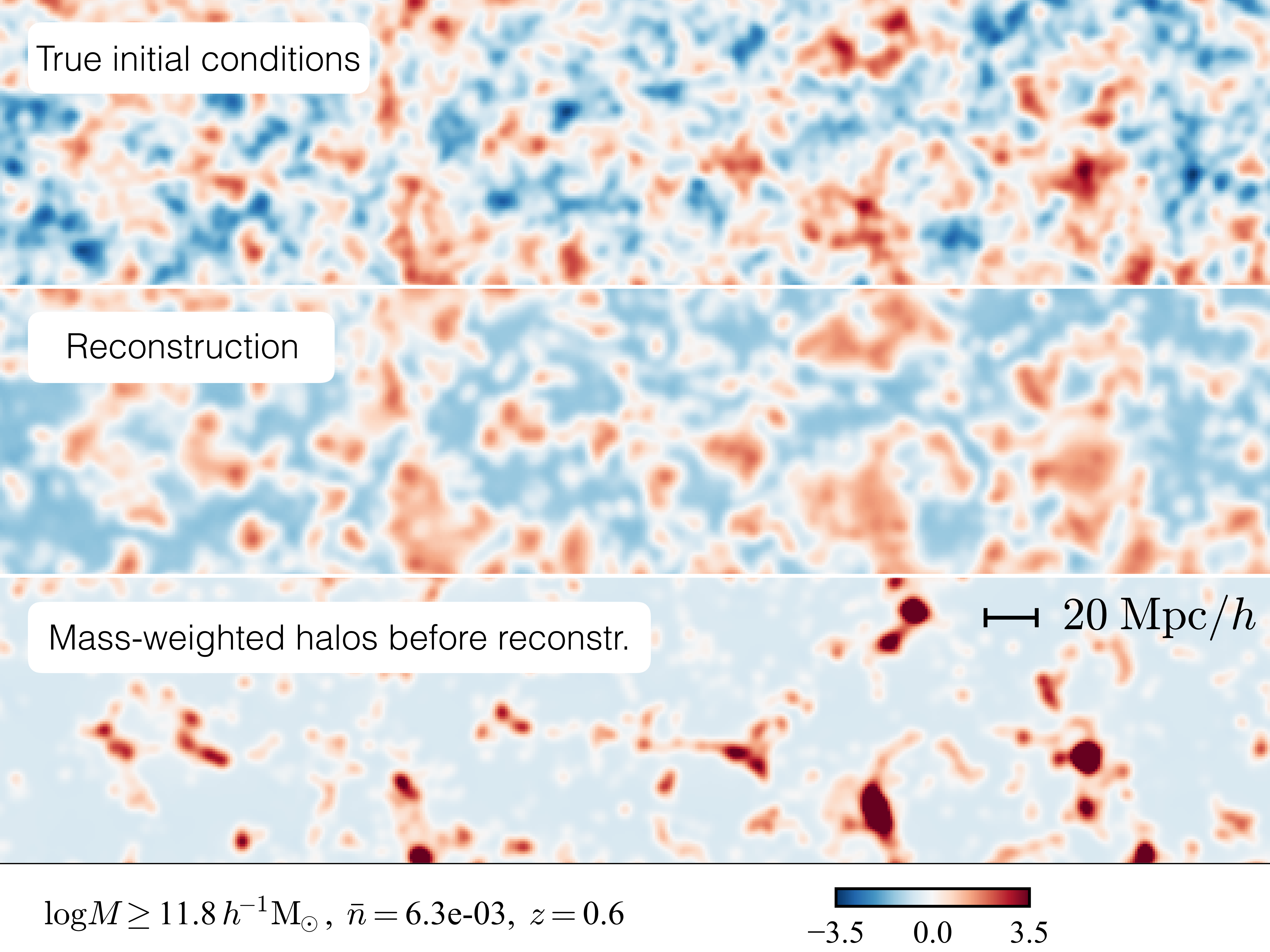
MS, Baldauf & Zaldarriaga (2017)

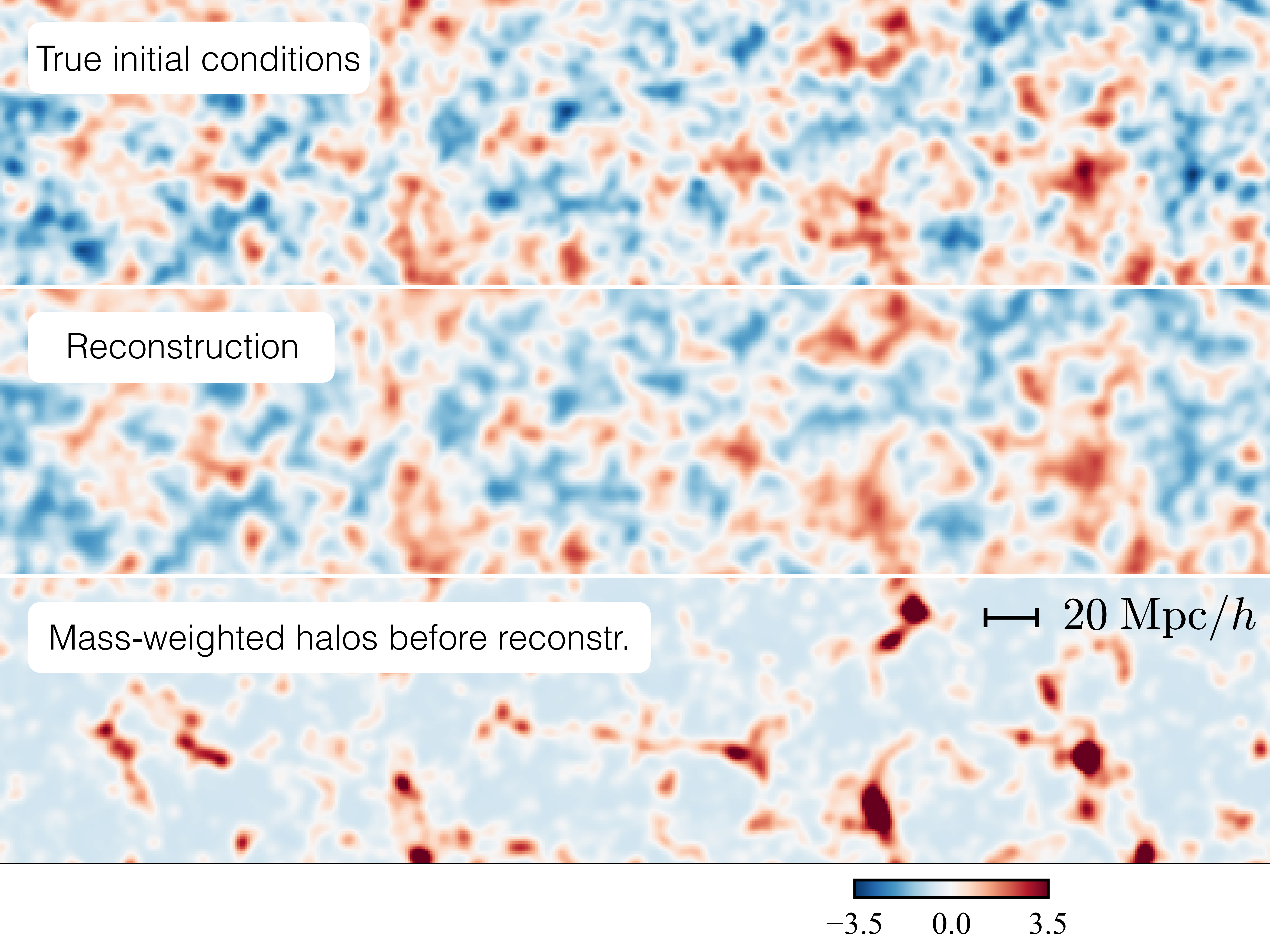


True initial conditions
Reconstruction
Mass-weighted halos before reconstr.
20 Mpc/h
$\log M \geq 13.8\, h^{-1}\mathrm{M}_\odot$, $\bar{n} = 2.6\text{e-}05$, $z = 0.6$
-3.5
0.0
3.5


True initial conditions
Reconstruction
Mass-weighted halos before reconstr.
20 Mpc/h
$\log M \geq 13.8\, h^{-1}\mathrm{M}_{\odot}$, $\bar{n} = 2.6\text{e-}05$, $z = 0.6$
−3.5
0.0
3.5


True initial conditions
Reconstruction
Mass-weighted halos before reconstr.
20 Mpc/h
$\log M \geq 12.8\,h^{-1}\mathrm{M}_{\odot}$, $\bar{n} = 5.9\text{e-}04$, $z = 0.6$
−3.5
0.0
3.5


True initial conditions
Reconstruction
Mass-weighted halos before reconstr.
20 Mpc/h
log M ≥ 11.8 h^-1 M_⊙ , n̄ = 6.3e-03, z = 0.6
−3.5
0.0
3.5



True initial conditions
Reconstruction
Mass-weighted halos before reconstr.
20 Mpc/h
−3.5
0.0
3.5

Measures of success

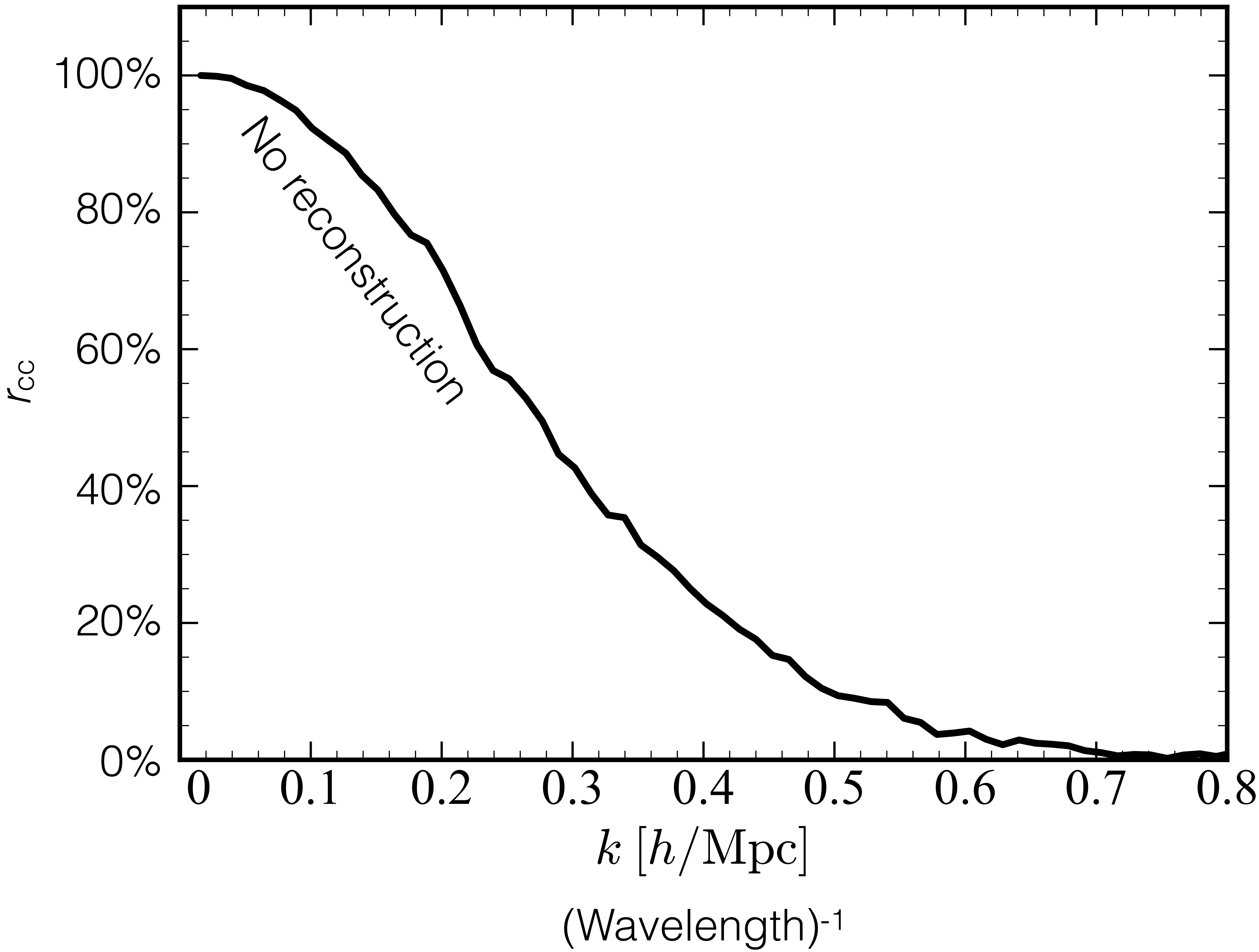
Error power spectrum (= MSE per wavenumber *k*)

$$P_{\rm err}(k) \equiv \langle |\delta_{\rm true\,ICs}(\mathbf{k}) - \delta_{\rm rec}(\mathbf{k})|^2 \rangle$$

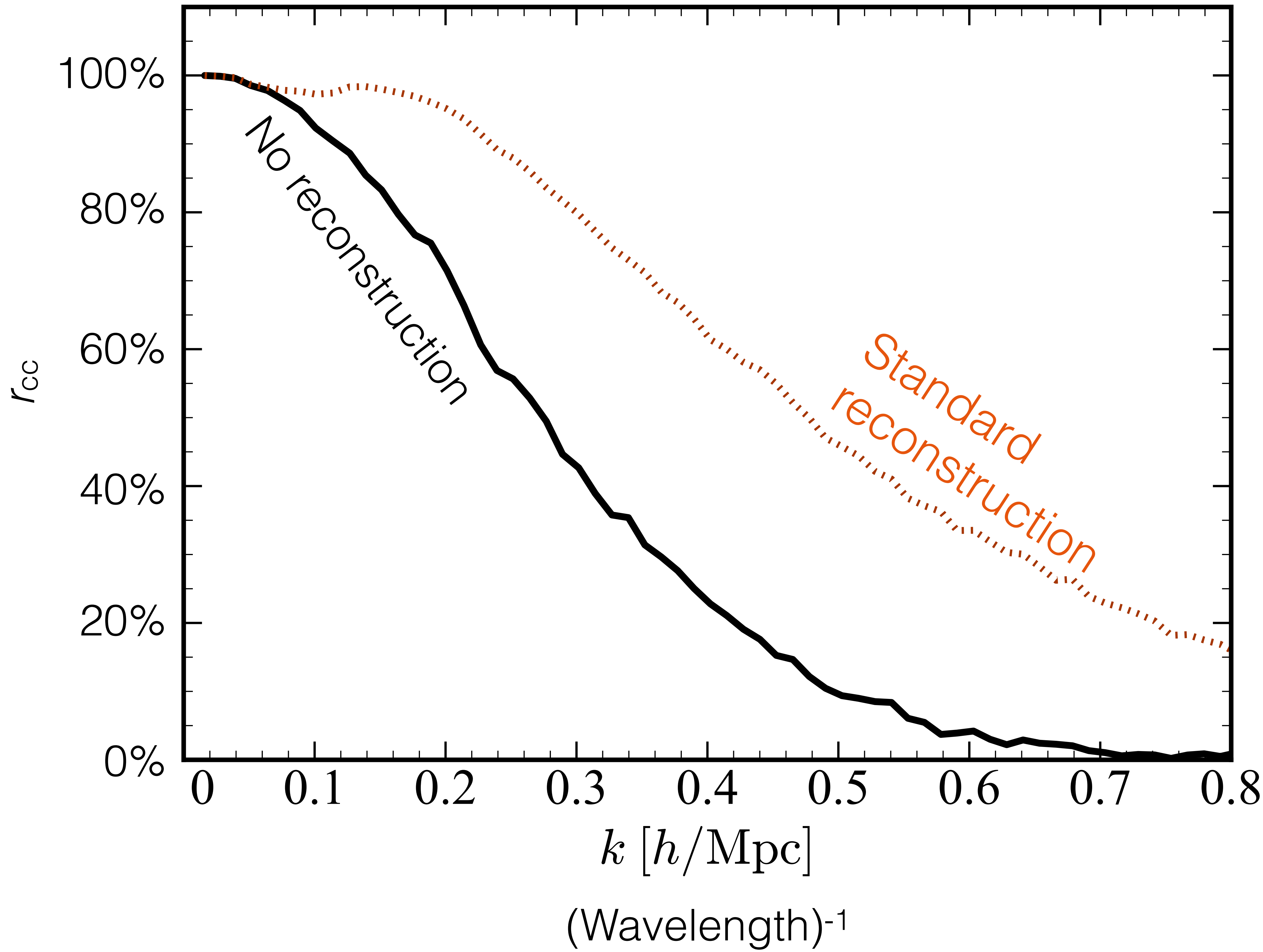
Correlation of reconstruction with true initial conditions

$$r_{cc}(k) = \frac{\langle \delta_{\rm true\,ICs}(\mathbf{k}) \delta^*_{\rm rec}(\mathbf{k}) \rangle}{\sqrt{P_{\rm true\,ICs}(k) P_{\rm rec}(k)}}$$

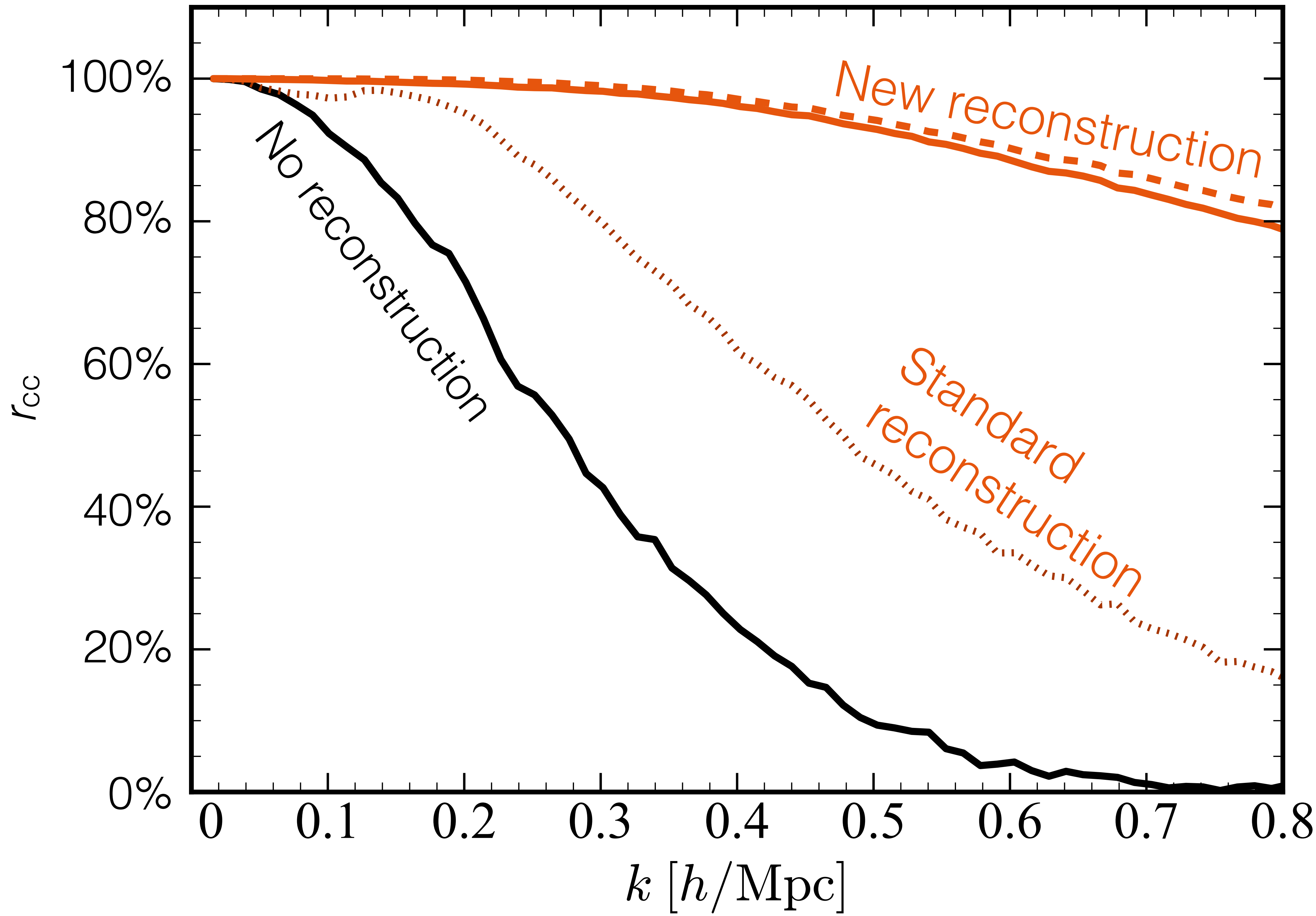
Correlation with true initial conditions



Correlation with true initial conditions

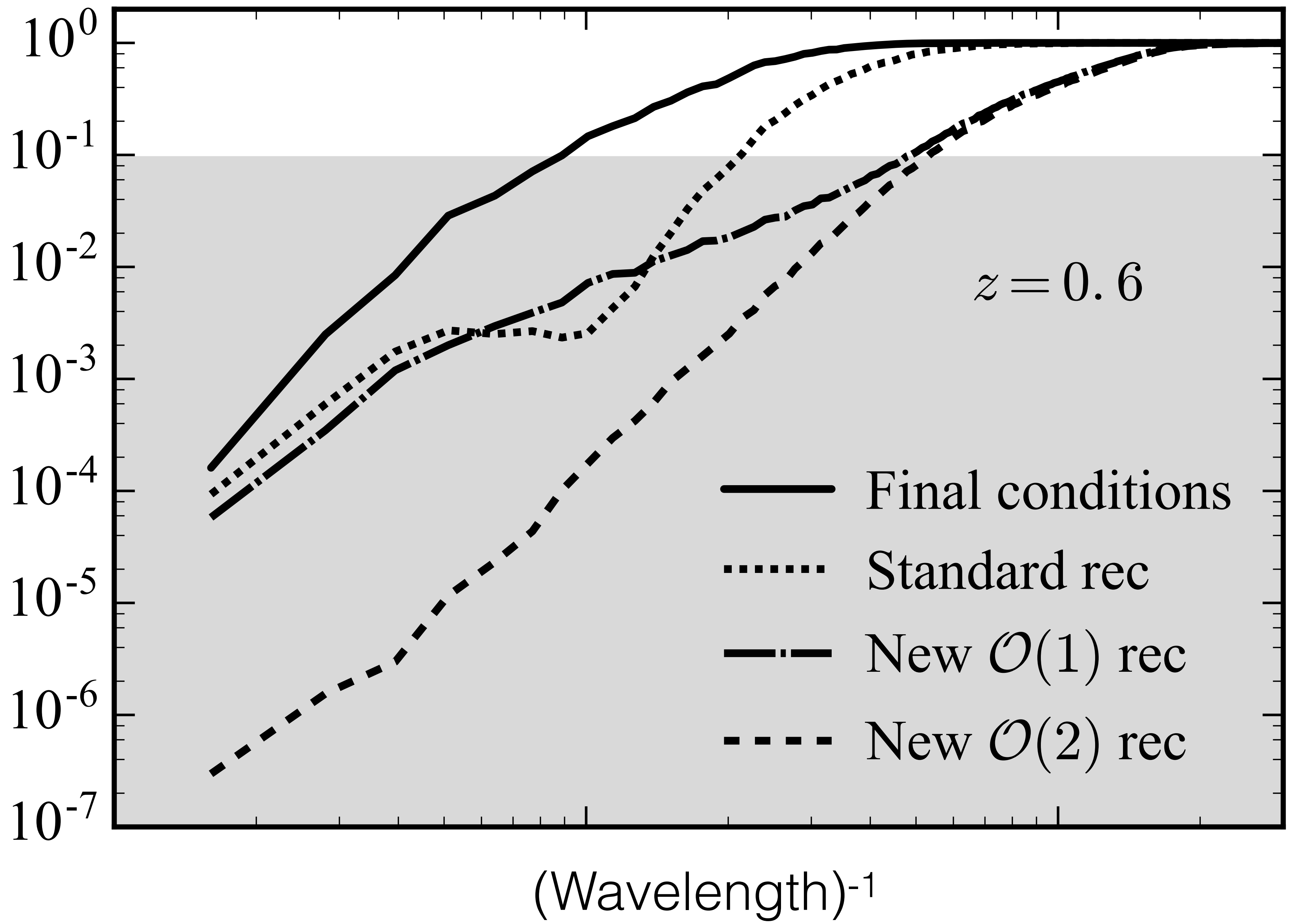


Correlation with true initial conditions



MS, Baldauf & Zaldarriaga (2017), similar to Zhu, Yu+ (2017); noise-free DM

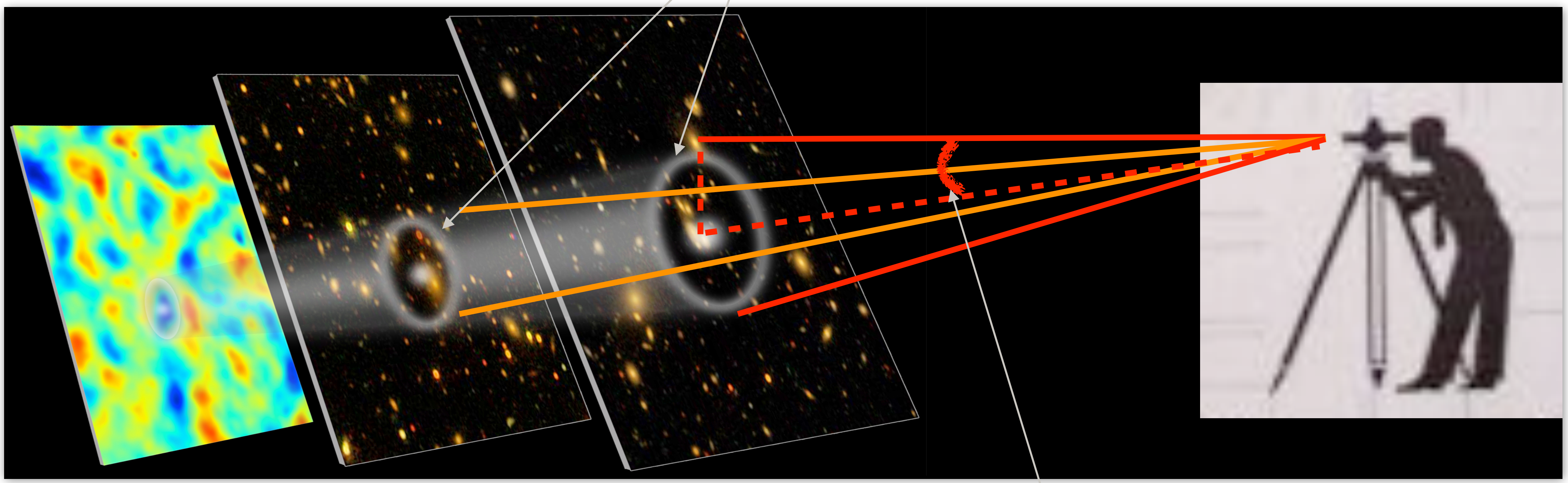
[Mean-square fractional error]



MS, Baldauf & Zaldarriaga (2017)

Expansion rate

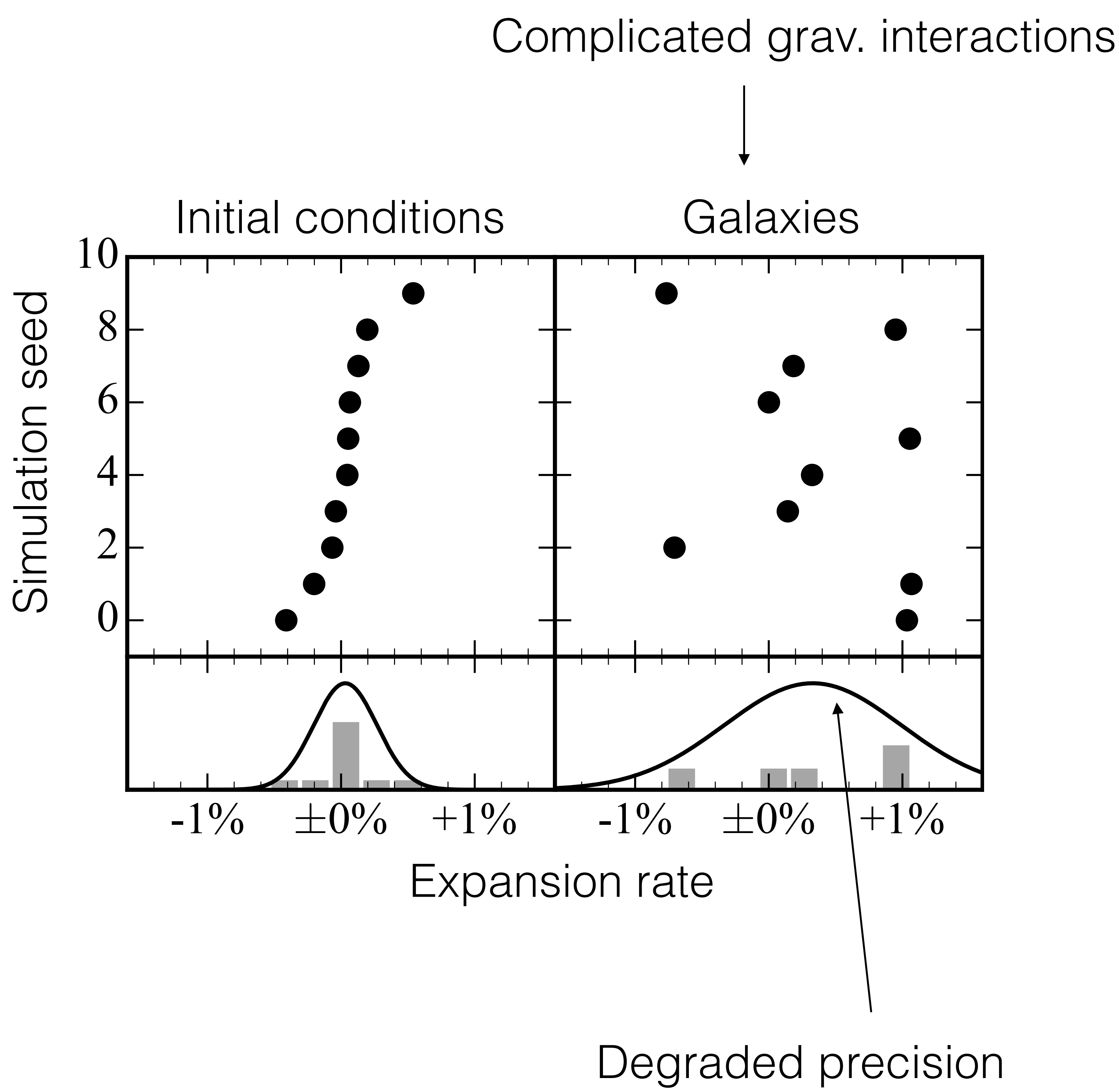
Initial conditions have a clear feature



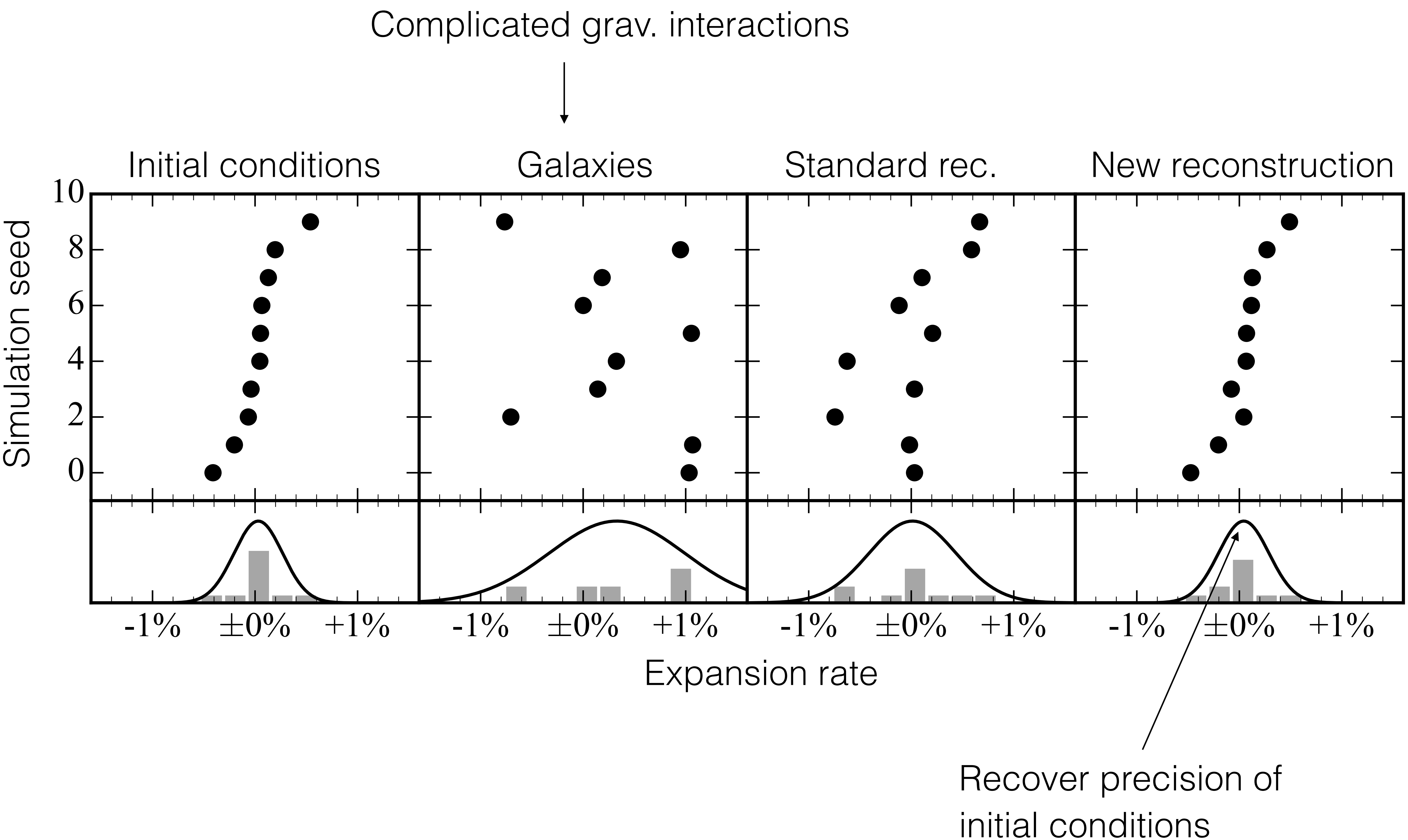
Can estimate expansion rate by measuring angular extent

More precise in initial conditions than observed galaxies

Expansion rate in 10 simulations



Expansion rate in 10 simulations



MS, Baldauf & Zaldarriaga (2017)

Alternative reconstruction approach

Previous algorithm works well, but based on intuition & heuristics

Alternative: Use gradient descent to maximize posterior of initial conditions (1M+ parameters)

From simulation or polynomial regression

$$P(\delta_{\rm IC}|\delta_g) = \frac{\mathcal{L}(\delta_g|\delta_{\rm IC})P(\delta_{\rm IC})}{P(\delta_g)}$$

Normal distribution (Gaussian ICs)

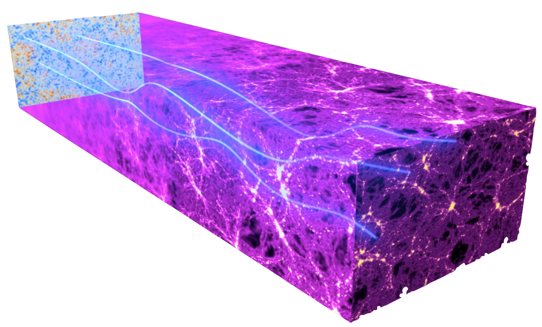
Fast because shifted operators model has easy gradients

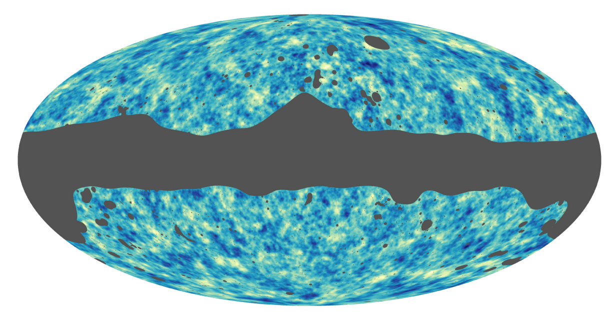
Seljak, Aslanyan *et al.* (2017)
Schmidt, Elsner *et al.* (2019)
Modi, White *et al.* (arXiv:1907.02330)

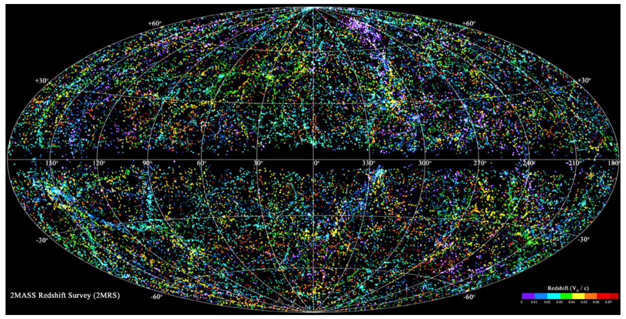
Gravitational lensing of the CMB

(0) Joint analysis of lensed CMB and magnification

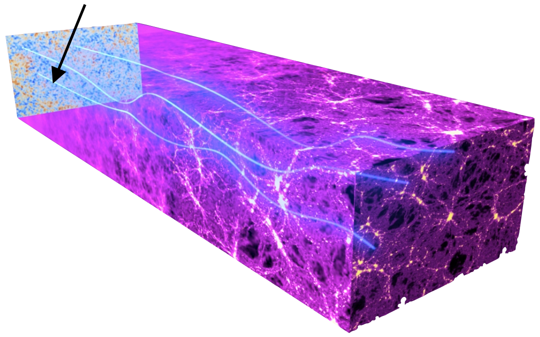
(1) Cross-correlate magnification with galaxy catalogs







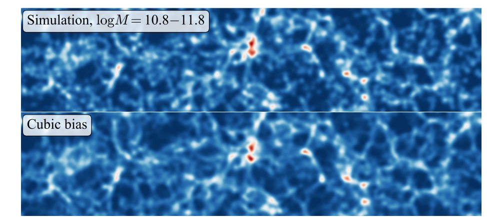
(2) Estimate *unlensed CMB*



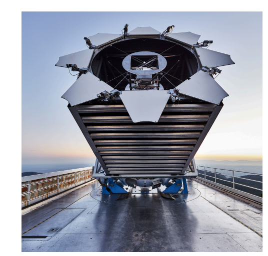
(3) Biases of the magnification estimator

Science with galaxy catalogs

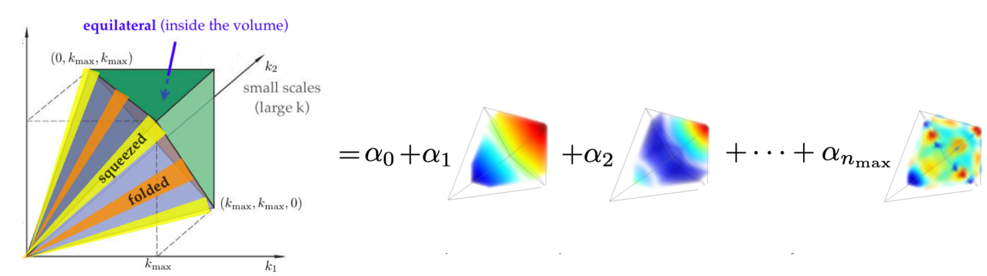
(1) Bias model at the field level



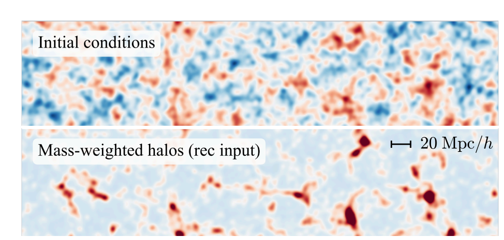
(2) Cosmological parameter analysis



(3) Accounting for skewness



(4) Getting initial from final conditions



Thanks