BERKELEY CENTER *for* COSMOLOGICAL PHYSICS

Large Scale Structure: data, simulations, theory

Uroš Seljak
UC Berkeley
Paris, May 24 2016

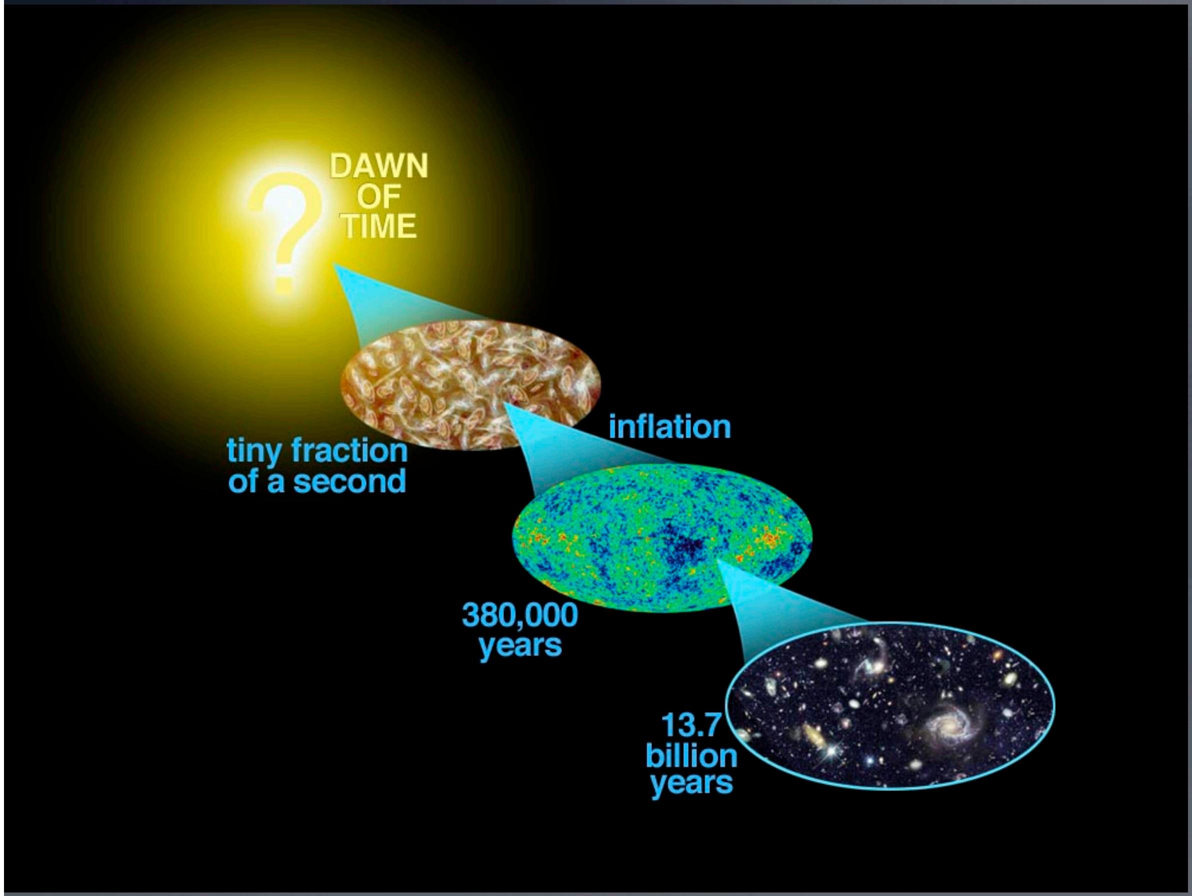
BERKELEY CENTER *for* COSMOLOGICAL PHYSICS

Large Scale Structure: data, simulations, theory (Confesions of a recovering theorist)

Uroš Seljak
UC Berkeley
Paris, May 24 2016

Why doing LSS?

- I: Classical cosmological tests with baryonic acoustic oscillations (in combination with CMB): galaxy 3d surveys
- II: LSS evolves by gravity: clustering of dark matter, which in turn informs us about the origins of structure, state of our universe today, its past evolution…
- types: weak lensing (WL), 3d galaxy clustering (RSD), 2d galaxy clustering (shape of P(k)), Lyman alpha forest..



♦Perturbations can be measured at different epochs:

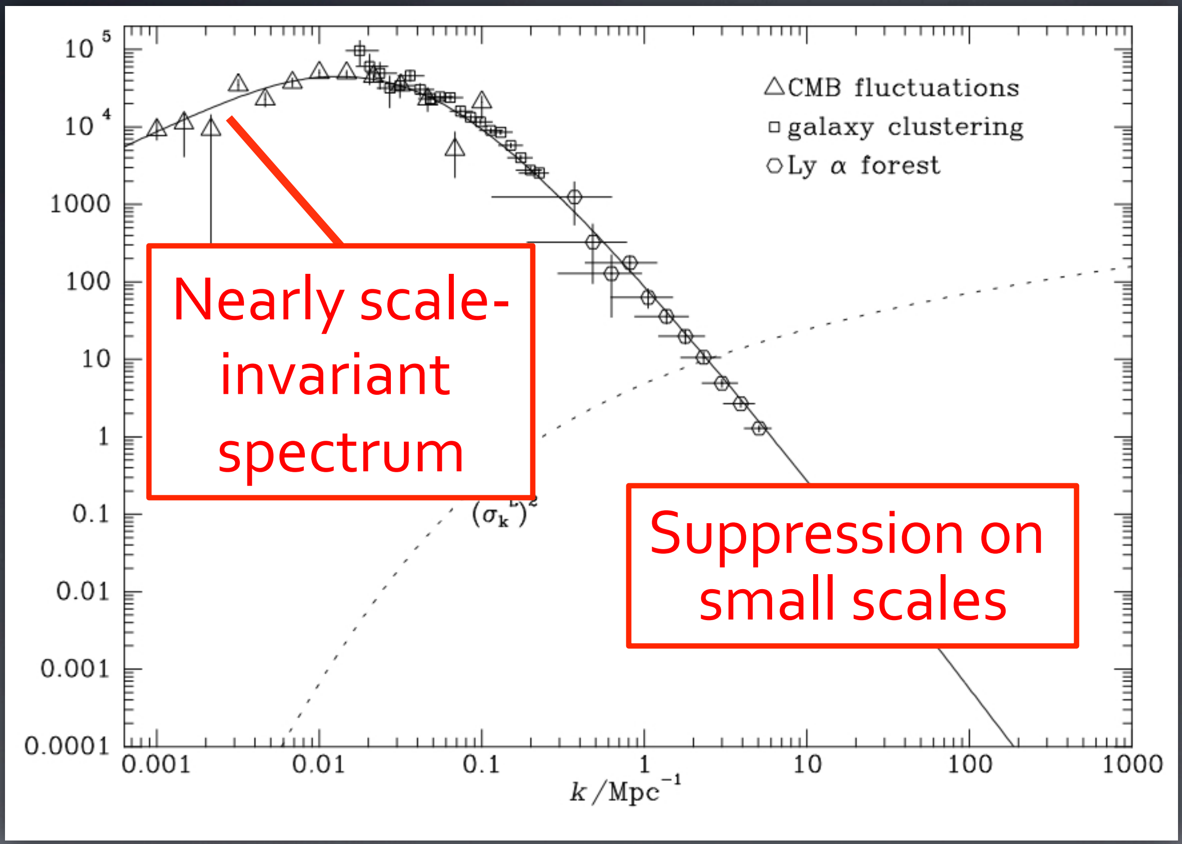
1. CMB z=1000
2. 21cm z=1-10+?
3. Ly-alpha forest z=2-4
4. Weak lensing z=0.3-2
5. Galaxy clustering, clusters z=0-2

Sensitive to matter components, initial conditions…

$$\ddot{\delta} + 2H\dot{\delta} = 4\pi G\bar{\rho}\delta \rightarrow \delta(t)$$

Perturbations: scale dependence

$$\langle \delta(k)\delta^*(k') \rangle = P(k)\delta_D(k-k')$$



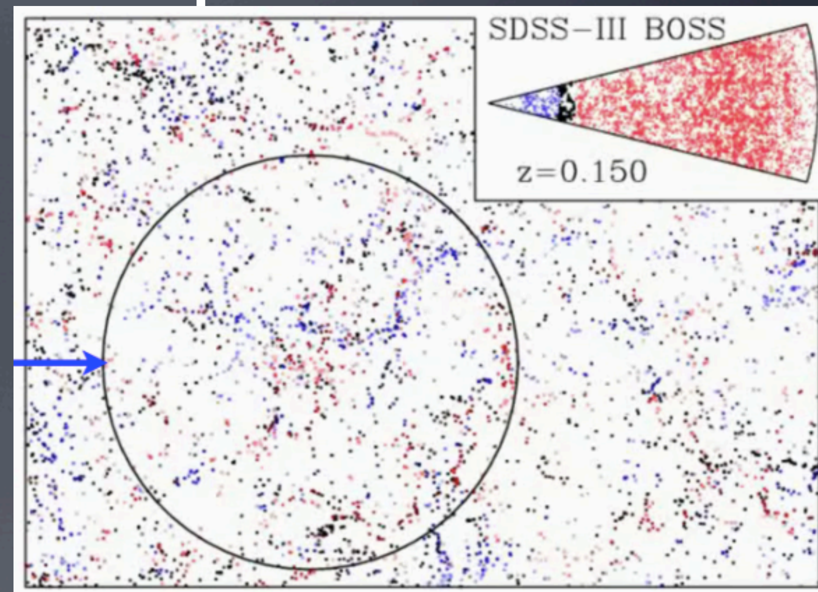
Picture from

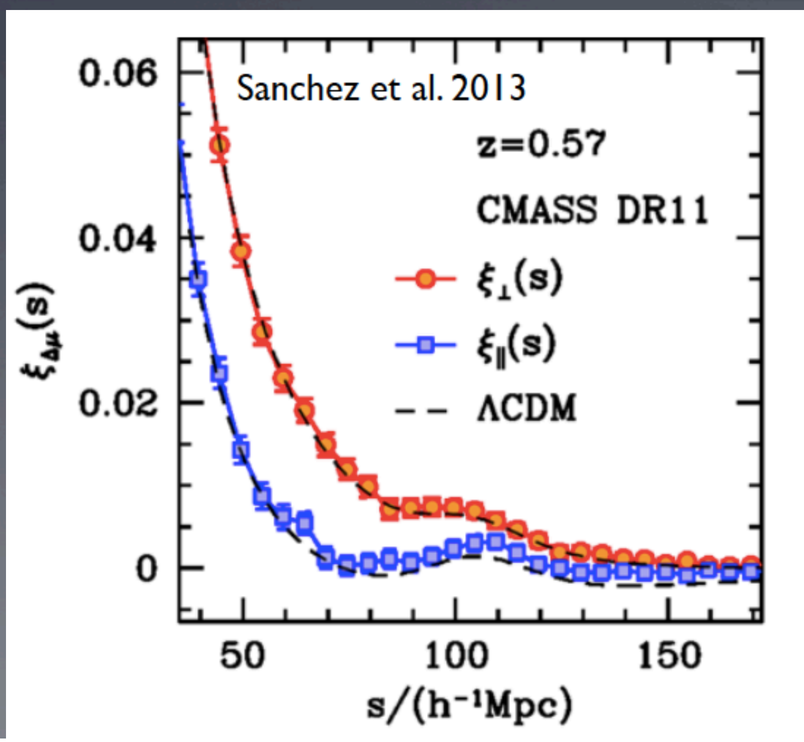
Binney & Tremaine

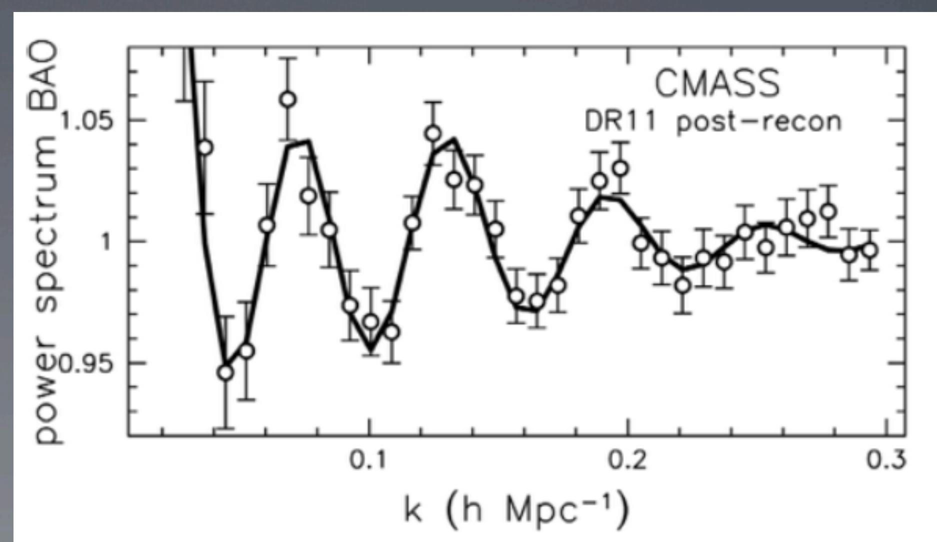
Probes initial conditions

We measure redshifts and angles, and the shape is not a power law, so there are neoclassical tests in broadband measurements as well

BAO: relatively robust in correlation function and power spectrum

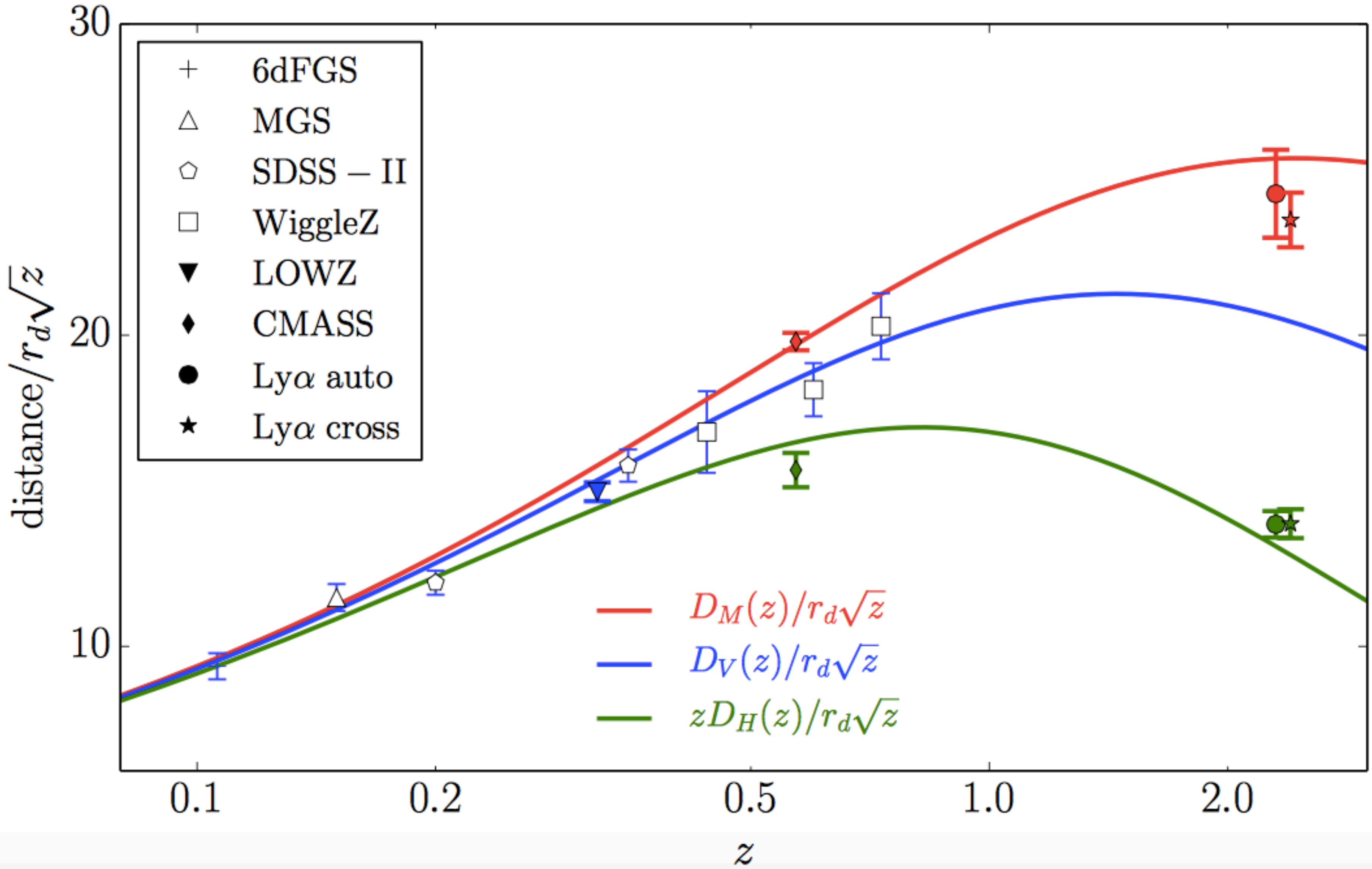






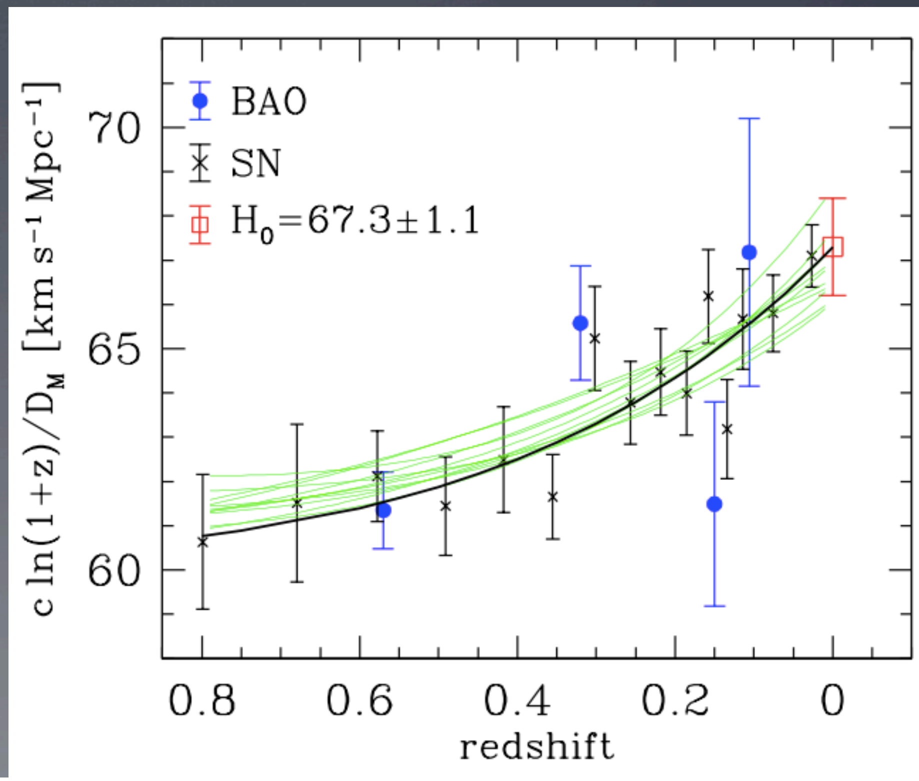
BAO Hubble diagram

Aubourg etal
(BOSS coll)



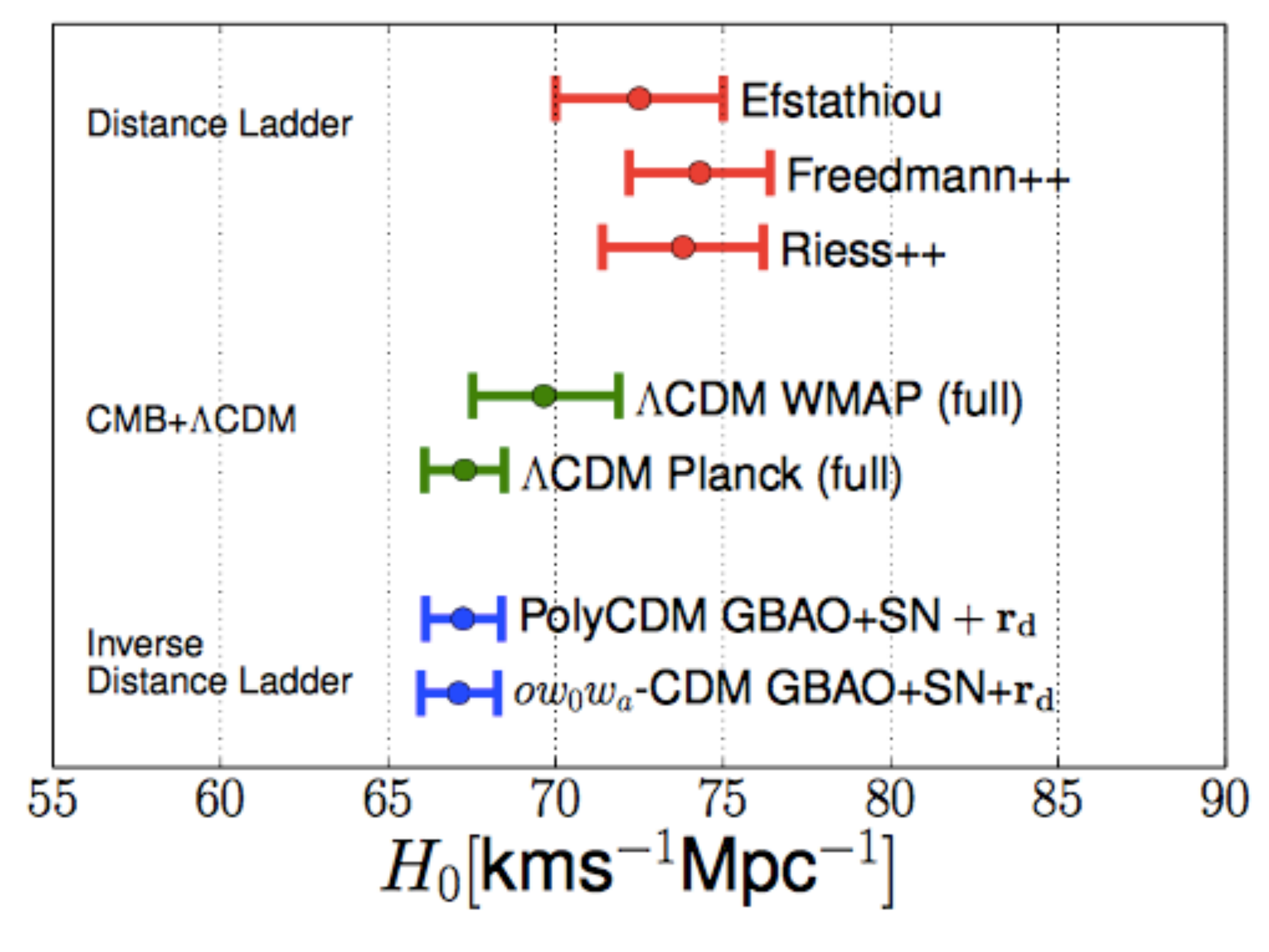
Inverse distance ladder

- BAO is a standard ruler with a known absolute calibration r_d, and at z=0.57 overlaps with SN1A, allowing absolute calibration of SN1A
- Bringing SN1A from z=0.57 down to z=0 gives H_o
- Error 1.5%, dominated by BAO distance error (1%)

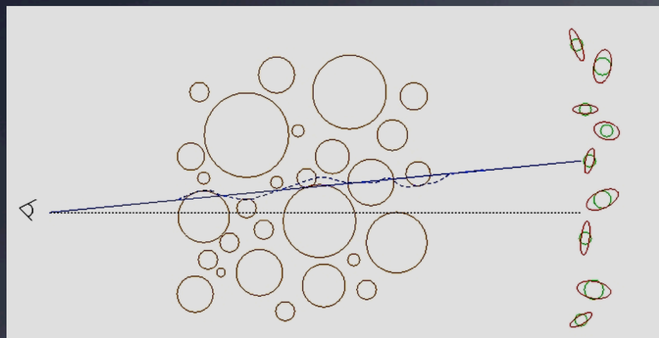


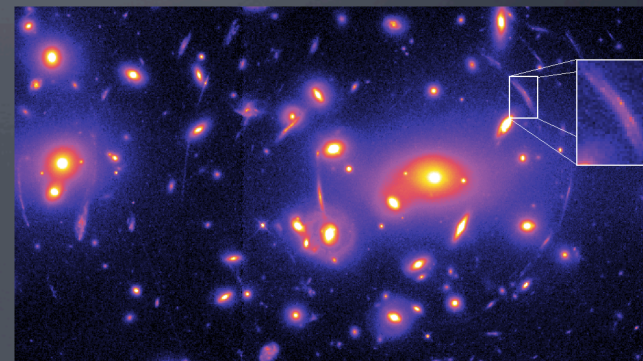
Comparison of H_o

- Perfect agreement with Planck CMB (independent method, relies on 6 parameter ΛCDM model)
- Some discrepancy with direct distance ladder (Riess etal, Freedmann etal): new physics or systematics?

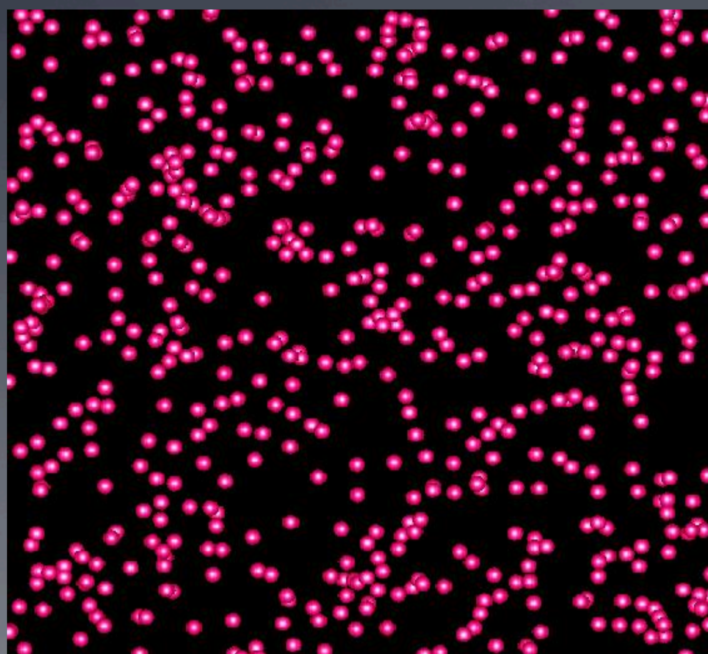


LSS: Weak Gravitational Lensing

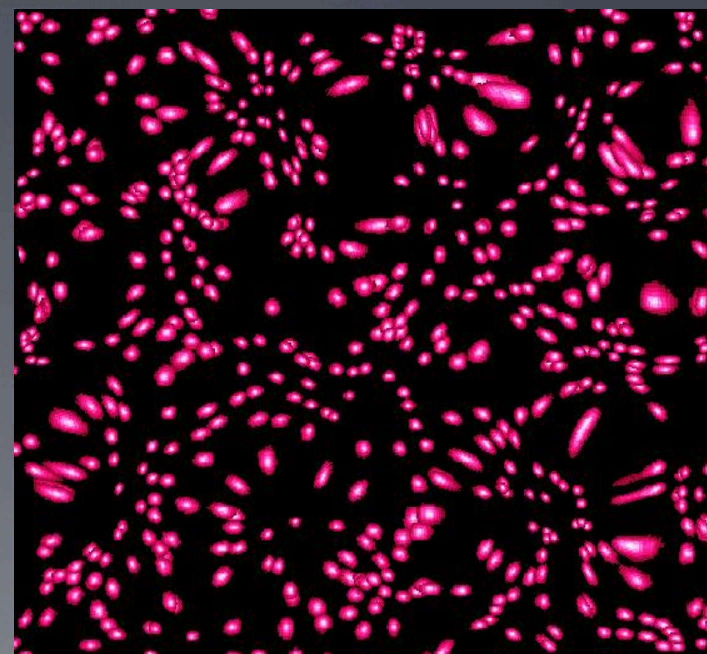




Distortion of background images by foreground matter



Unlensed

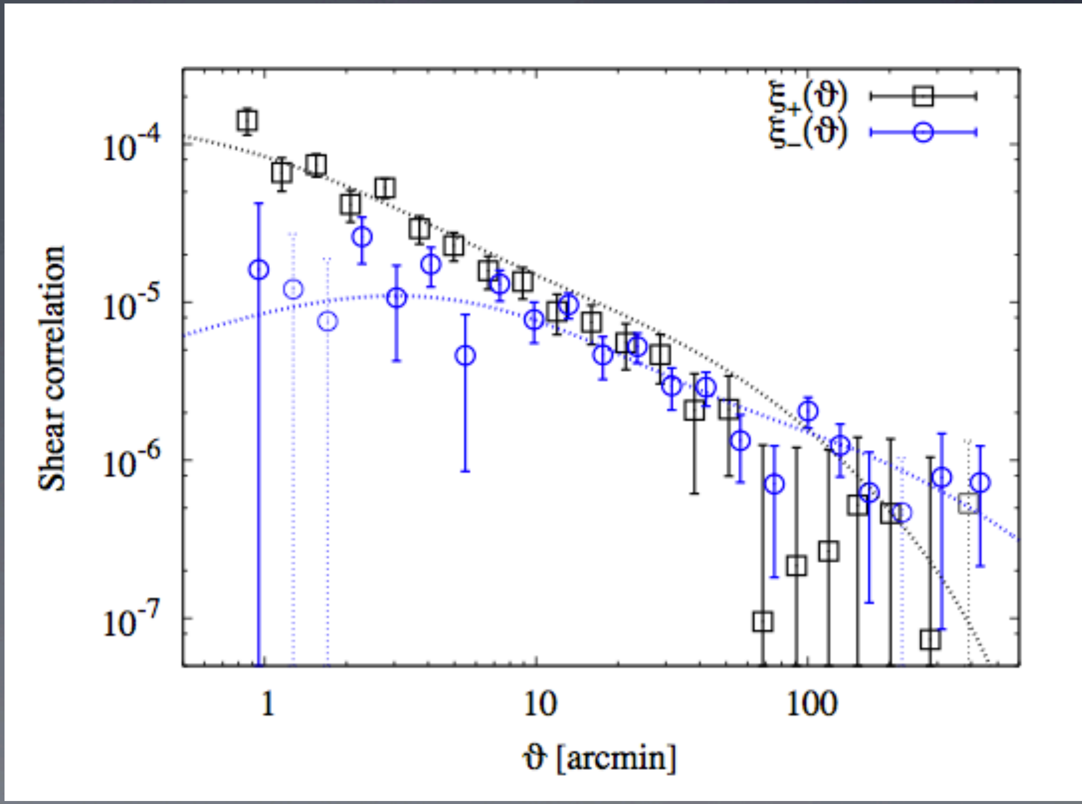


Lensed

WL method I: shear-shear correlations

Dark matter P(k) known (from simulations)

CFHT-LS, Kiblinger et al 2013



Challenges:

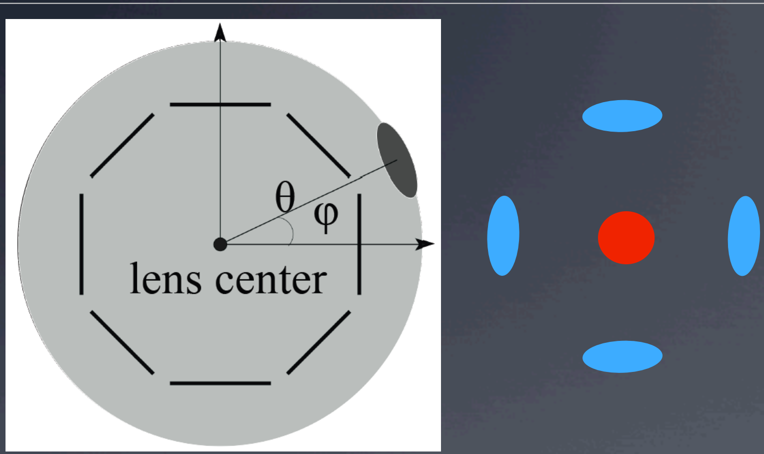
Small scales: could be contaminated by baryonic effects: need to marginalize

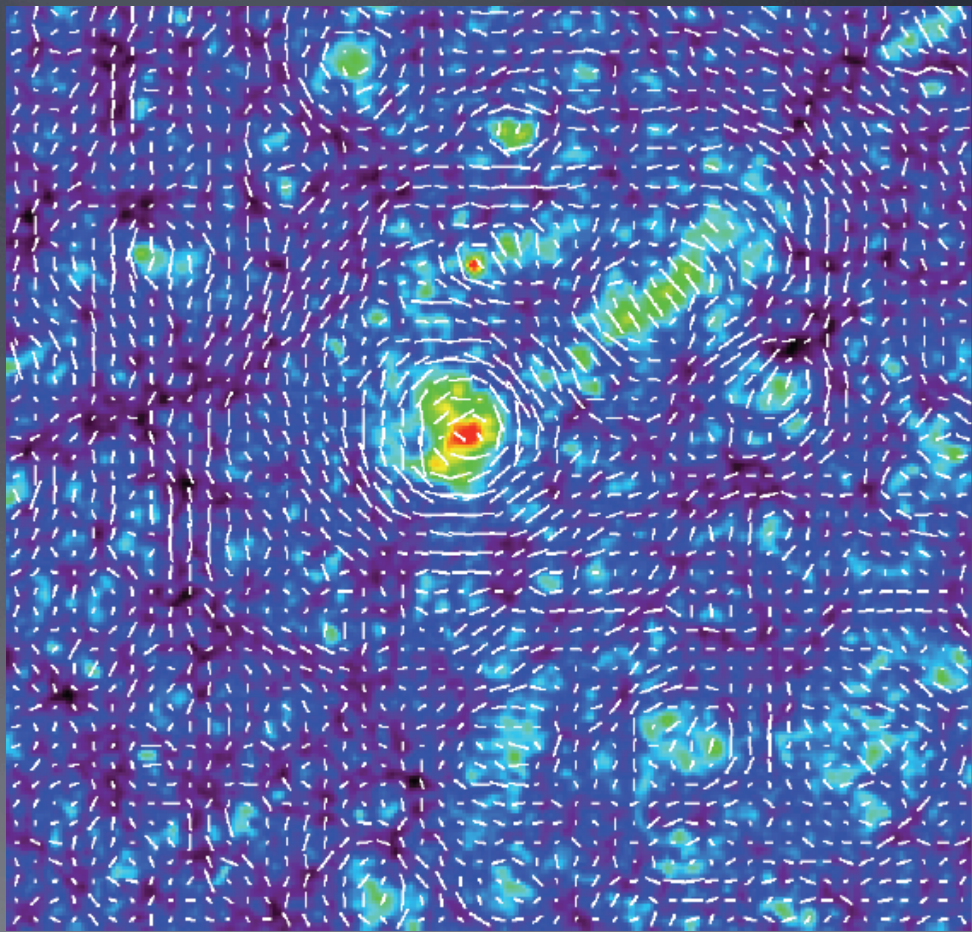
Redshift distributions not completely known

Intrinsic alignments

Various systematics

WL Method II: galaxy-shear correlations



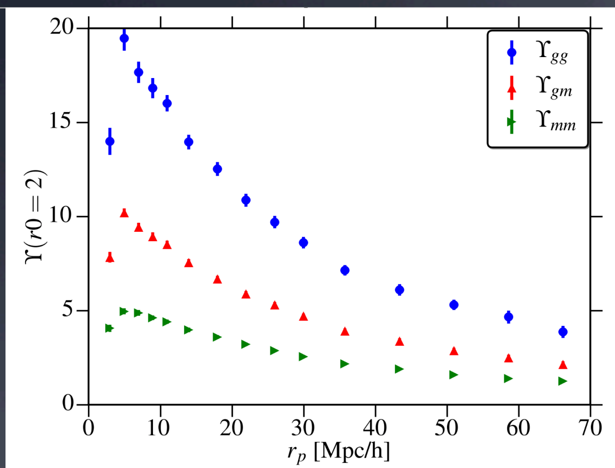


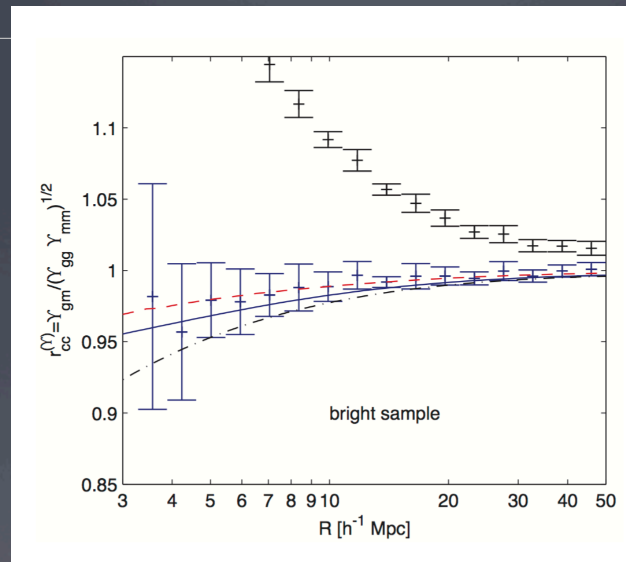
On small scales: galaxy-halo (mass, profile) connection

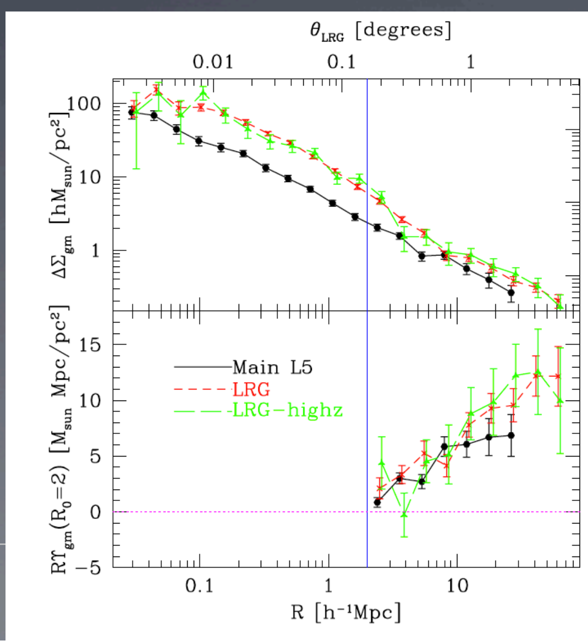
On large scales: cross-correlation proportional to bias b

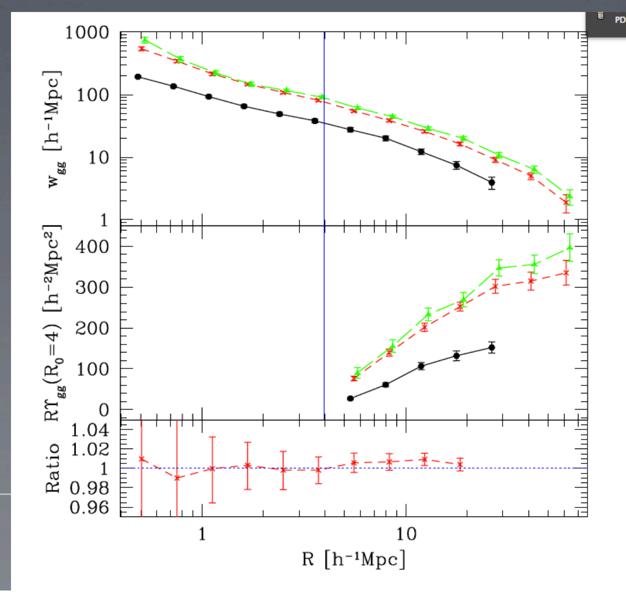
Galaxy auto-correlation proportional to b^2

Combining g-shear and g-g clustering









Relies on r_{cc} being known

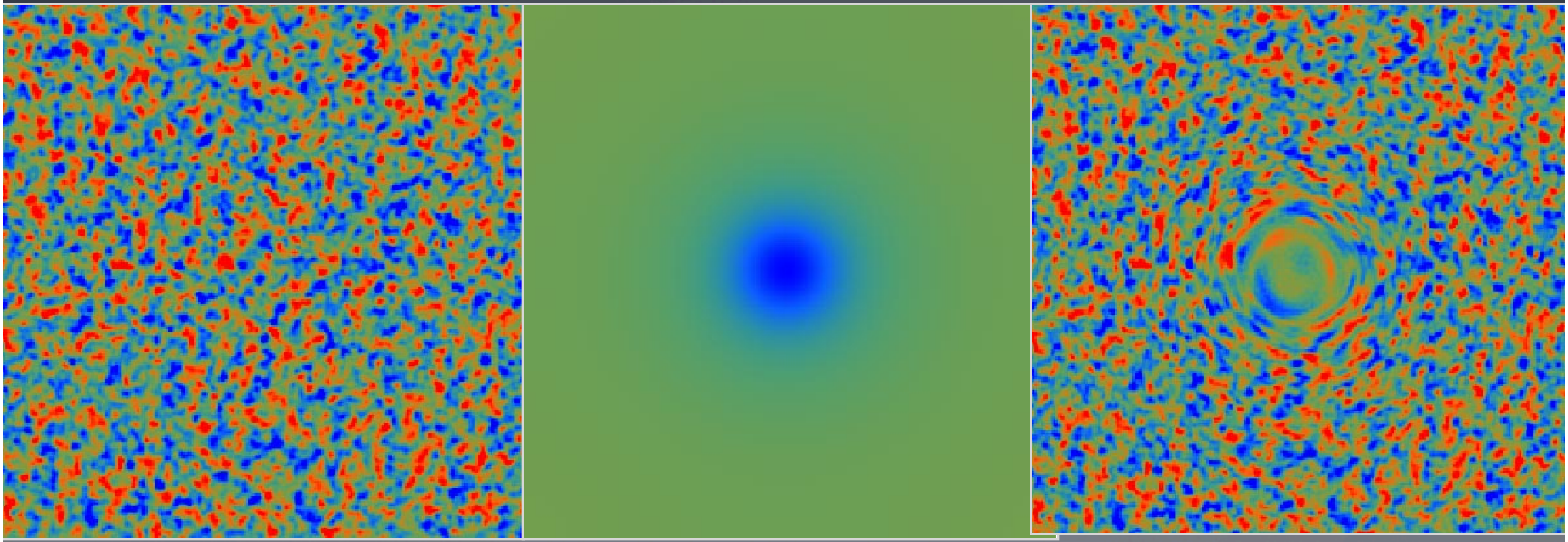
Eliminating small scale from lensing makes r_{cc} close to 1

Mandelbaum, Baldauf, US etal, 2013

LSS gravitational lensing on CMB

$$T_{lensed}(\vec{\mathbf{n}}) = T_{unlensed}(\vec{\mathbf{n}} + \vec{\mathbf{d}}) \qquad \vec{\mathbf{d}} = -2\vec{\nabla}\nabla^{-2}\kappa$$

- Here κ is the convergence and is a projection of the matter density perturbation.



Okamoto and Hu 2002

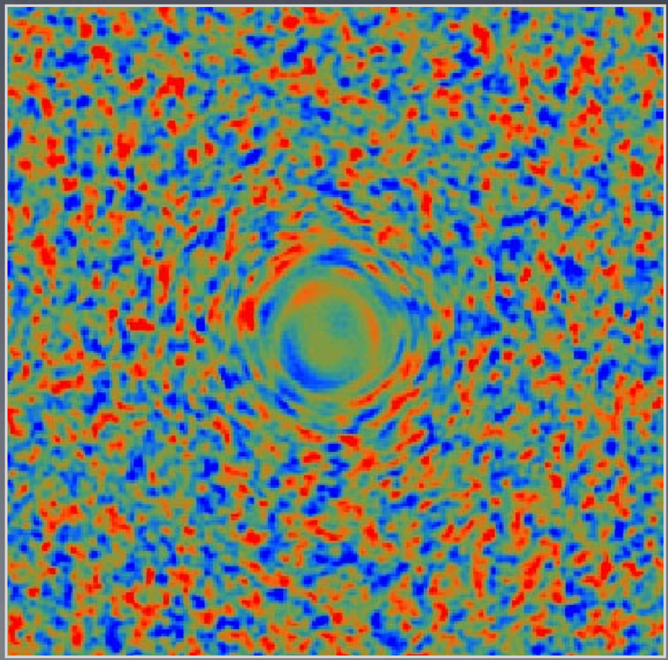
Reconstruction of lensing from CMB

Zaldarriaga and US 1998

Local estimate of typical patch size or shape

Compare to global average

Must be quadratic in T, hence 4 point function <TTTT> for κ^2



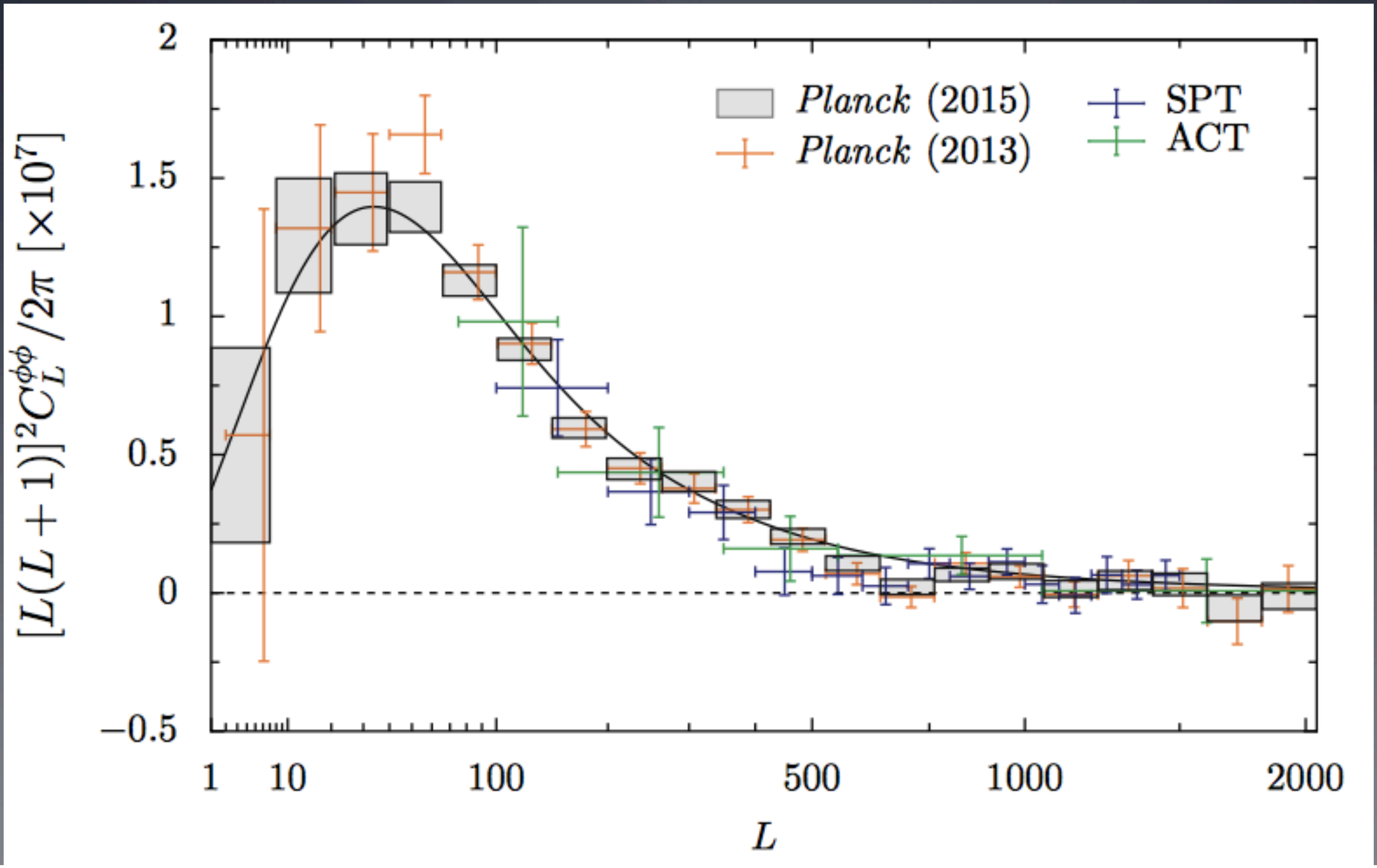
$$\kappa \propto (\nabla_x T)^2 + (\nabla_y T)^2$$

$$\gamma_1 \propto (\nabla_x T)^2 - (\nabla_y T)^2$$

$$\gamma_2 \propto 2(\nabla_x T)(\nabla_y T)$$

CMB lensing: Planck

40 sigma in Planck 2015, 200+ sigma with S3, S4
2-d projection of matter distribution



Covariance matrix challenge

- simulations have a hard time converging on covariance matrix, its inverse is "hard": e.g. 12,000 simulations in Blot et al. 2014
- Disconnected part: "gaussian" is easy: we should compute it analytically using window functions (note: this is not done currently)
- Connected part: smooth response to long wavelength modes

$$\mathrm{Cov}(P(k_i), P(k_j)) = P(k_i)P(k_j)V^{-1}\left(\frac{4\pi^2}{k_i^2 \Delta k}\delta_{ij} + \delta_{A_0}^2\right)$$

Mohammed & US 2014

PT approach to Covariance

- Modes from outside the survey (do not average to zero): tree level effects from survey window function very important (supersample variance), easy to calculate, depend on whether the mean density is computed from within the survey or not (Li, Takada, Hu 2014)

$$\delta \ln P(k) = \left(\frac{47}{21} - \frac{1}{3}\frac{d\ln P}{d\ln k}\right)\delta_b = \left(\frac{68}{21} - \frac{1}{3}\frac{d\ln(k^3 P)}{d\ln k}\right)\delta_b$$

- Use 26/21 instead of 68/21 for local mean density
- Modes inside the survey (average to zero): use PT trispectrum

PT trispectrum

- Tree-level calculation (Scoccimarro etal 1999)

$$C_{ij} \equiv \langle \hat{P}(k_i)\hat{P}(k_j)\rangle - \langle \hat{P}(k_i)\rangle\langle \hat{P}(k_j)\rangle = V_f \, [\frac{2P_i^2}{V_s(k_i)}\delta_{ij} + \bar{T}(k_i,k_j)]$$

$$\bar{T}(k_i,k_j) = \int_{k_i}\frac{d^3\boldsymbol{k}_1}{V_s(k_i)}\int_{k_j}\frac{d^3\boldsymbol{k}_2}{V_s(k_j)}\, T(\boldsymbol{k}_1,-\boldsymbol{k}_1,\boldsymbol{k}_2,-\boldsymbol{k}_2)$$

$$\bar{T}(k_i,k_j) = \int_{k_i}\frac{d^3\boldsymbol{k}_1}{V_s(k_i)}\int_{k_j}\frac{d^3\boldsymbol{k}_2}{V_s(k_j)}\Big[12F_3(\boldsymbol{k}_1,-\boldsymbol{k}_1,\boldsymbol{k}_2)P_1^2P_2 + 8F_2^2(\boldsymbol{k}_1-\boldsymbol{k}_2,\boldsymbol{k}_2)P(|\boldsymbol{k}_1-\boldsymbol{k}_2|)P_2^2$$
$$+ \quad 16F_2(\boldsymbol{k}_1-\boldsymbol{k}_2,\boldsymbol{k}_2)F_2(\boldsymbol{k}_2-\boldsymbol{k}_1,\boldsymbol{k}_1)P_1P_2P(|\boldsymbol{k}_1-\boldsymbol{k}_2|) + (\boldsymbol{k}_1 \leftrightarrow \boldsymbol{k}_2)\Big]$$

- 1-loop terms: sample variance of low k modes

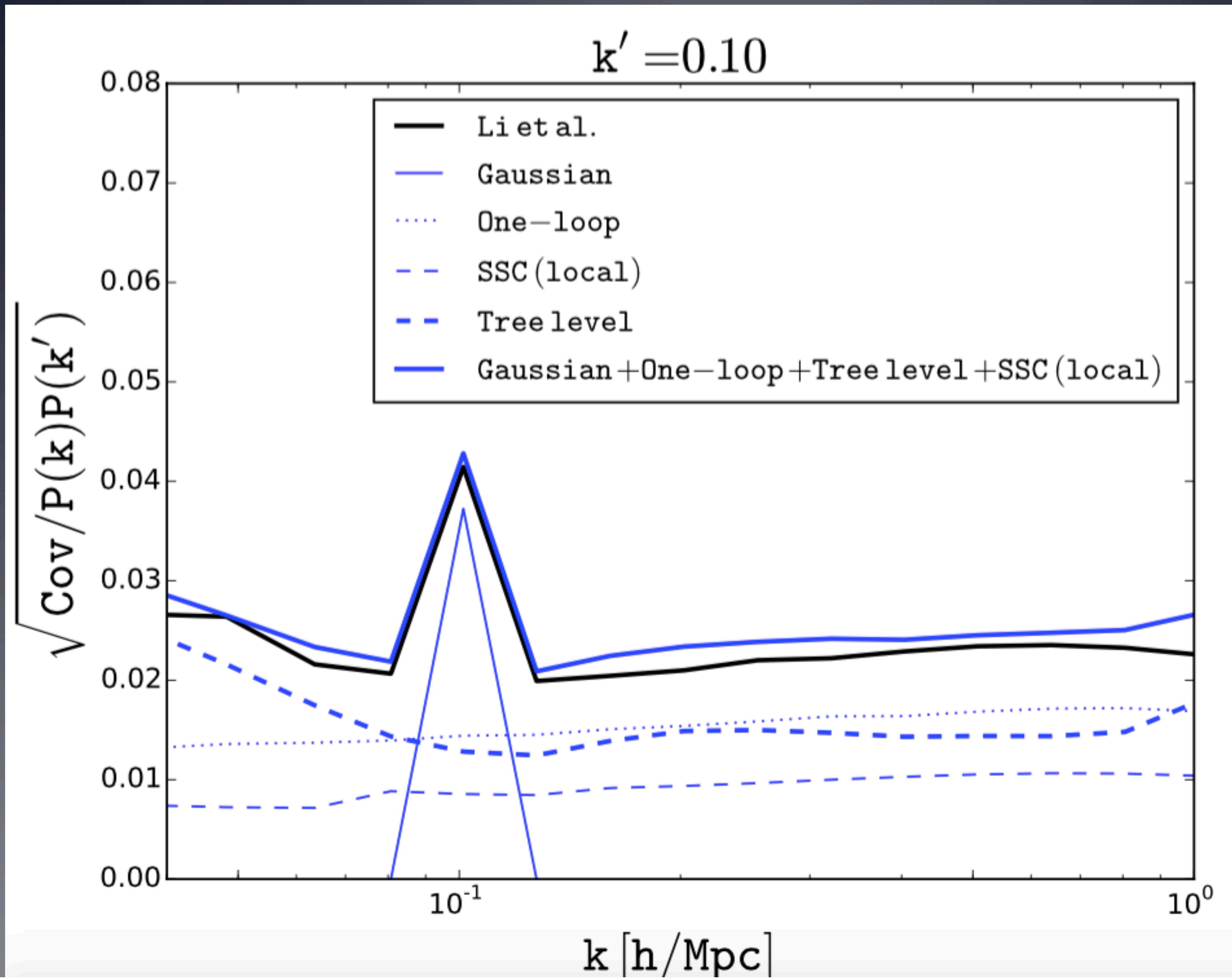
$$P = P_{22} + P_{13} = \{\frac{2519}{2205}P_{S0}(k) - \frac{47}{105}kP'_{S0}(k) + \frac{1}{10}k^2P''_{S0}(k)\}\langle\delta_L^2\rangle$$

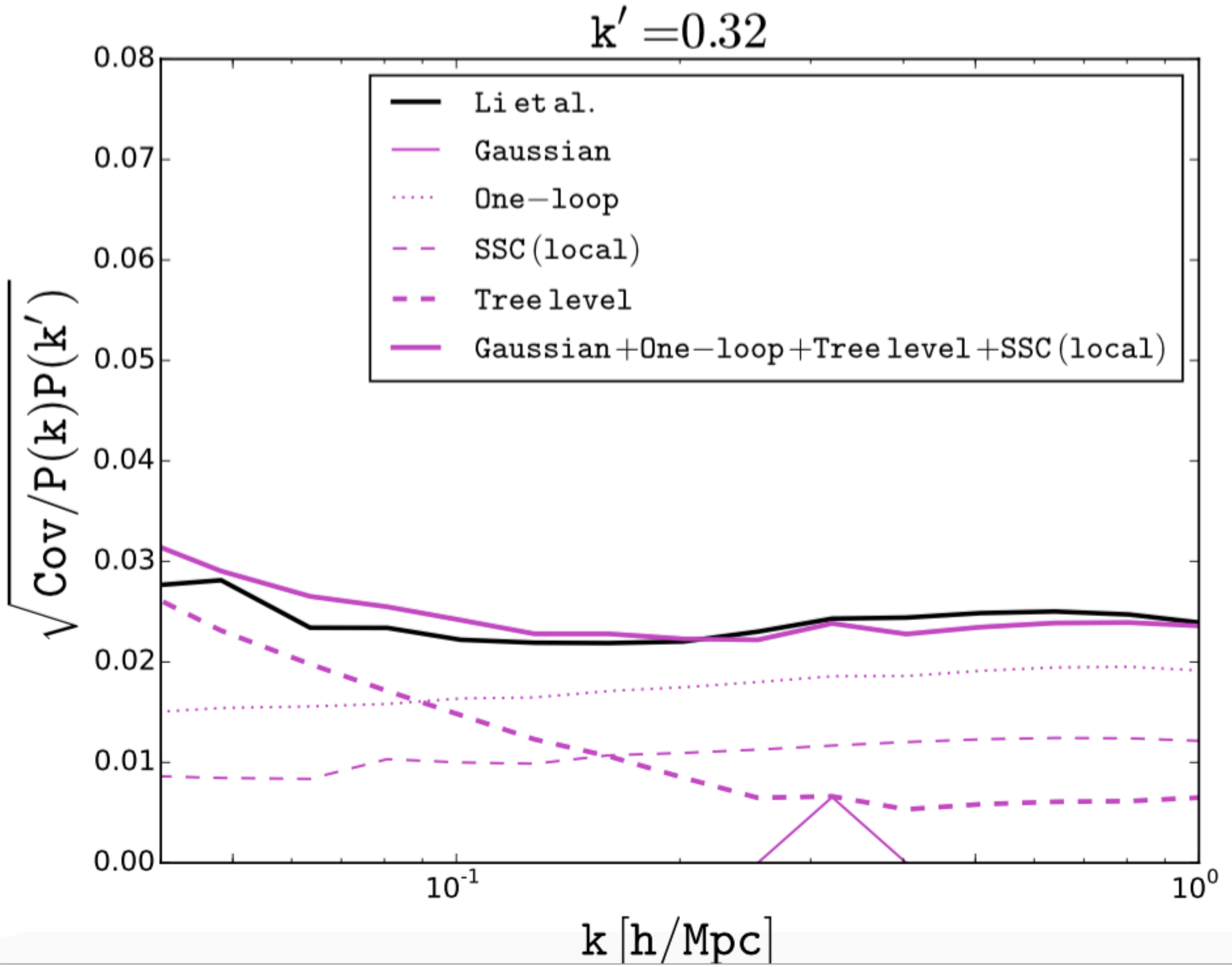
$$\frac{\mathbf{Cov}_{ij}}{P(k_i)P(k_j)} = \left(\frac{1}{\pi^2}\int P_{\rm Lin}^2(k)k^3 d\ln k\right)\mathbf{W}_i\mathbf{W}_j$$

Large contribution from low k, hence large volume needed

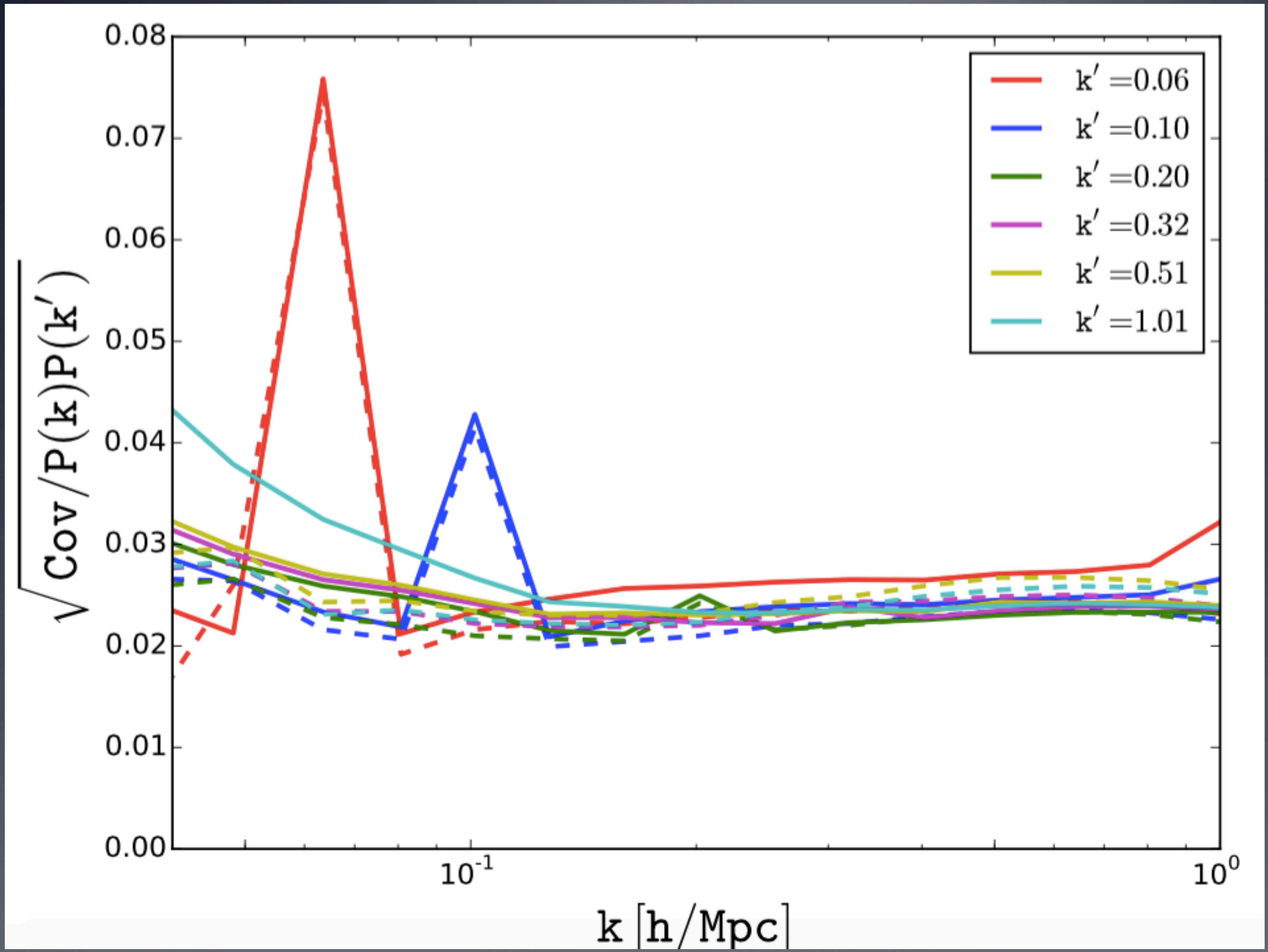
$$\mathbf{W}_i = \frac{2519}{2205}E_2(k_i) - \frac{47}{105}\frac{d\ln P(k_i)}{d\ln k_i} + \frac{1}{10}\frac{d^2\ln P(k_i)}{d\ln k_i^2}$$

PT vs simulations





k′ =0.32
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
$\sqrt{Cov/P(k)P(k')}$
Li et al.
Gaussian
One−loop
SSC (local)
Tree level
Gaussian + One−loop + Tree level + SSC (local)
10^{-1}
10^{0}
k [h/Mpc]

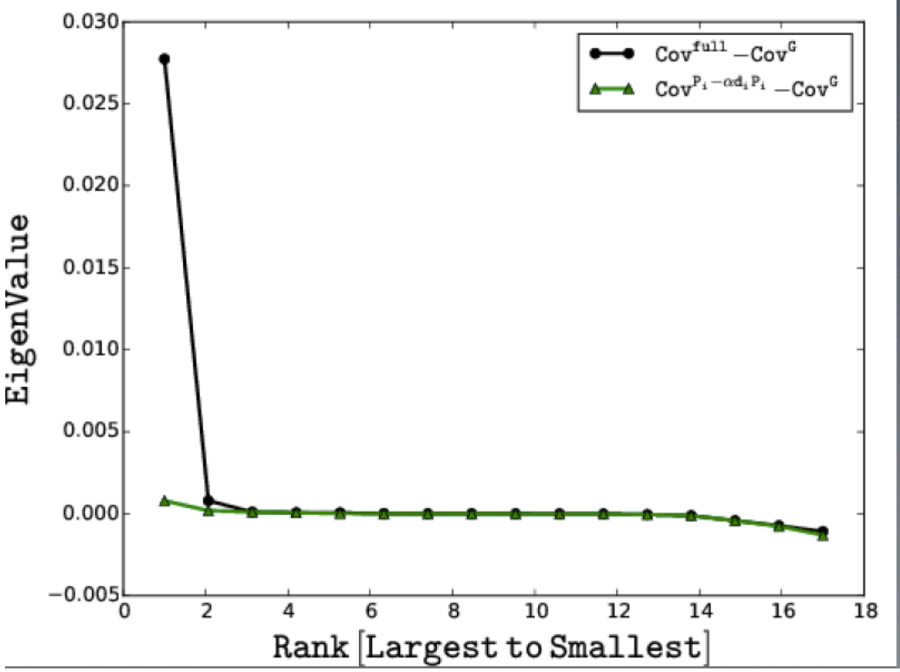


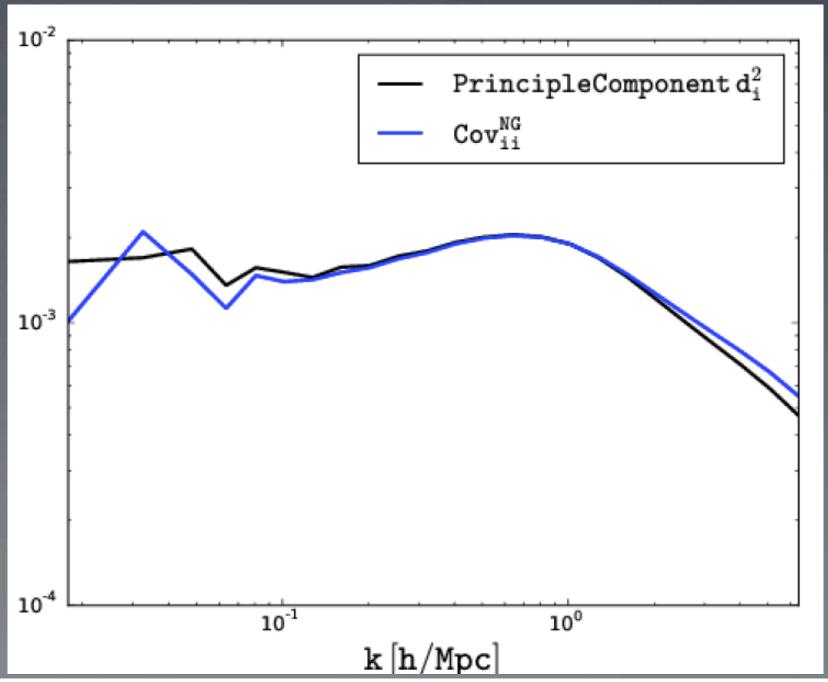
Lessons: connected part of covariance is very correlated
For k>0.2h/Mpc all the modes are strongly correlated
Correlations generated by long wavelength modes:
beware of jackknife/bootstrap methods

Covariance matrix as an external parameter

- Most of the connected covariance comes from a small scale response to long wavelength modes

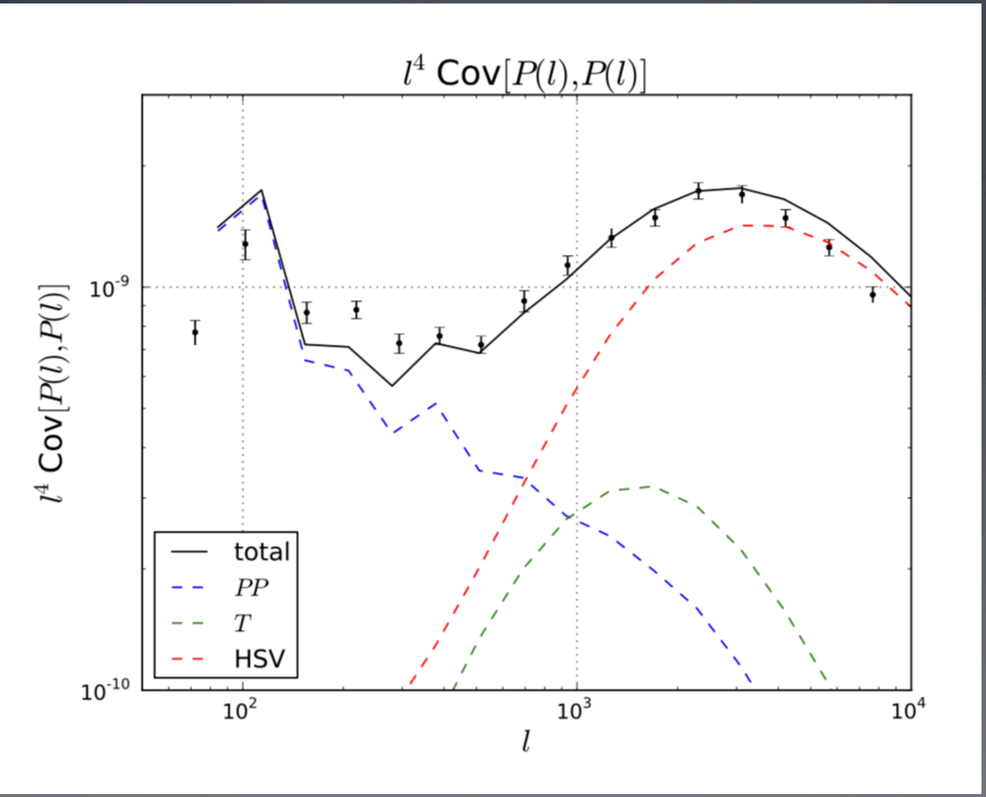
- The connected part can be written as a single eigenmode $C_{ij}=Ad_id_j$, where i represents k_i amplitude and d_i is a response at that k_i





For WL need to project trispectrum

Done in the context of halo model (Schaan etal), not yet PT (work in progress)



Remaining challenges: combining baryonic effects and covariance matrix, both limit information from high l

RSD: PT model for galaxies

- Need to develop nonlinear models that are sufficiently general to allow for any reasonable nonlinear effects present in the data, while preserving as much of cosmological information as possible
- Some can be modeled by perturbation theory (PT)+biasing
- There is always more information on small scales, but same of it is hopelessly corrupted by nonlinear effects that cannot be modeled in PT
- One needs to model our ignorance obeys all symmetries (e.g. $k^2 P_L(k)$ at low k) and all physics (the biasing parameters are physical, e.g. FoG is determined by halo mass…)
- In recent years a workhorse has been the halo model+biasing+PT

local bias

non-local bias

linear: can be measured via $P^{hm}(k)$ at large scales

2nd-order: measured via $B^{hmm}(k)$ at large scales se

3rd-order

$$\begin{aligned}\delta_{\rm h}(\boldsymbol{x}) =& c_\delta \delta_{\rm m}(\boldsymbol{x}) \\ &+\frac{1}{2}c_{\delta^2}\delta_{\rm m}(\boldsymbol{x})^2 + \frac{1}{2}c_{s^2}s(\boldsymbol{x})^2 \\ &+\frac{1}{3!}c_{\delta^3}\delta_{\rm m}(\boldsymbol{x})^3 + \frac{1}{2}c_{\delta s^2}\delta_{\rm m}(\boldsymbol{x})s(\boldsymbol{x})^2 + c_\psi\psi(\boldsymbol{x}) + c_{st}s(\boldsymbol{x})t(\boldsymbol{x}) + \frac{1}{3!}c_{s^3}s(\boldsymbol{x})^3 \\ &+c_\epsilon\epsilon + \ldots,\end{aligned}$$

where

$$s_{ij}(\boldsymbol{x}) \equiv \partial_i\partial_j\phi(\boldsymbol{x}) - \frac{1}{3}\delta^{\rm K}_{ij}\delta_{\rm m}(\boldsymbol{x}) = \left[\partial_i\partial_j\partial^{-2} - \frac{1}{3}\delta^{\rm K}_{ij}\right]\delta_{\rm m}(\boldsymbol{x}),$$ tidal field

$$t_{ij}(\boldsymbol{x}) \equiv \partial_i v_j - \frac{1}{3}\delta^{\rm K}_{ij}\theta_{\rm m}(\boldsymbol{x}) - s_{ij}(\boldsymbol{x}) = \left[\partial_i\partial_j\partial^{-2} - \frac{1}{3}\delta^{\rm K}_{ij}\right][\theta(\boldsymbol{x}) - \delta_{\rm m}(\boldsymbol{x})],$$

$$\psi(\boldsymbol{x}) \equiv [\theta(\boldsymbol{x}) - \delta_{\rm m}(\boldsymbol{x})] - \frac{2}{7}s(\boldsymbol{x})^2 + \frac{4}{21}\delta_{\rm m}(\boldsymbol{x})^2.$$

(halo density)-(matter density) McDonald & Roy (2010)

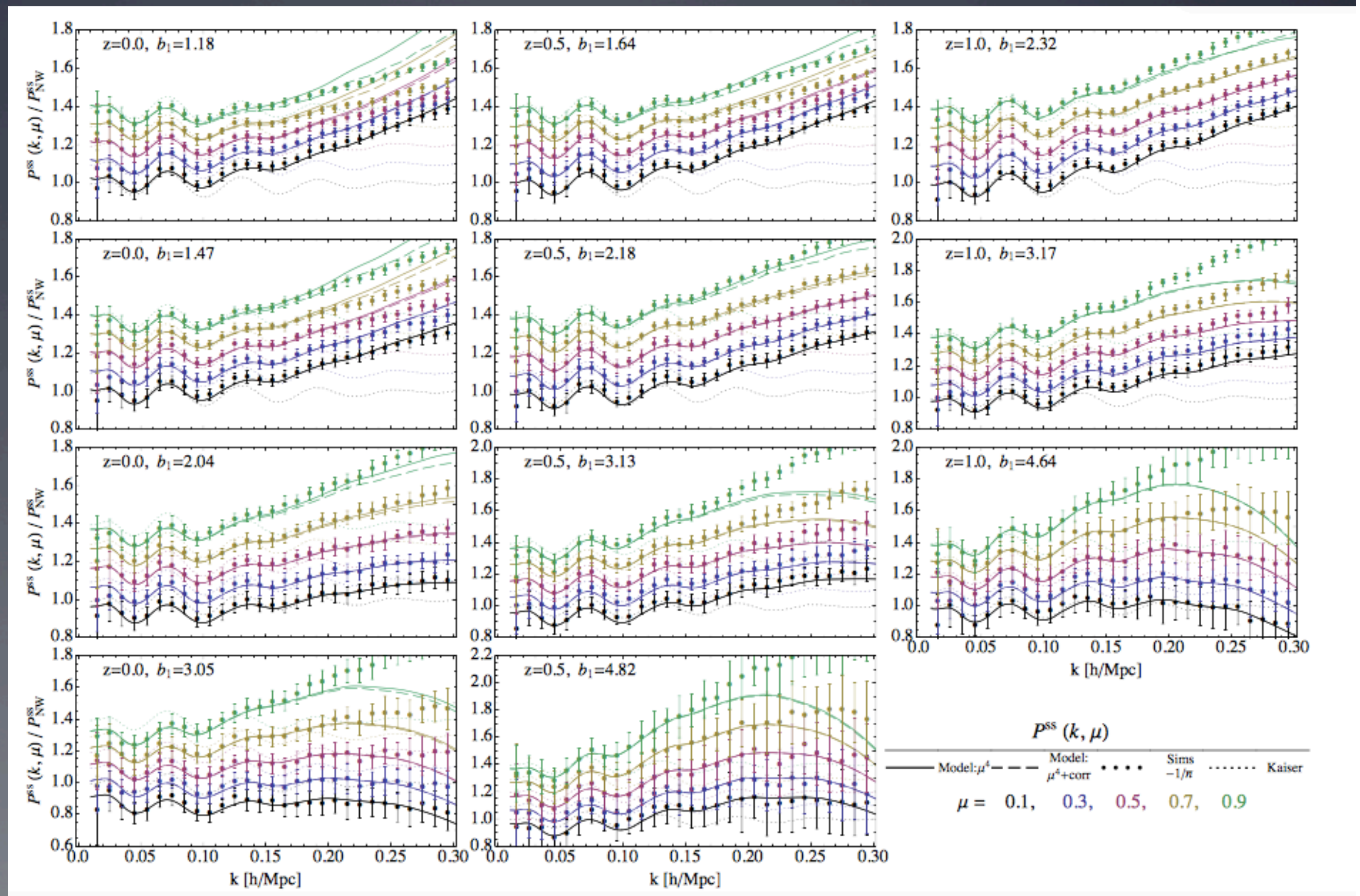
$$\begin{aligned}P^{\rm hm}_{00}(k) =& \left(c_\delta + \frac{34}{21}c_{\delta^2}\sigma^2 + \frac{1}{2}c_{\delta^3}\sigma^2 + \frac{1}{3}c_{\delta s^2}\sigma^2 + \frac{1}{2}c_{\delta\epsilon^2}\sigma_\epsilon^2 + \frac{68}{63}c_{s^2}\sigma^2 - \frac{16}{63}c_{st}\sigma^2\right)P^{\rm NL}_{\delta\delta}(k) \\ &+c_{\delta^2}\int\frac{d^3q}{(2\pi)^3}P(q)P(|\boldsymbol{k}-\boldsymbol{q}|)F^{(2)}_{\rm S}(\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}) \\ &+c_{s^2}\int\frac{d^3q}{(2\pi)^3}P(q)P(|\boldsymbol{k}-\boldsymbol{q}|)F^{(2)}_{\rm S}(\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q})S^{(2)}(\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}) \\ &+\left(-\frac{16}{21}c_{s^2} + \frac{32}{105}c_{st} + \frac{512}{2205}c_\psi\right)\sigma_3^2(k)P(k) \\ =& b_1P^{\rm NL}_{\delta\delta}(k) + b_2P_{b2,\delta}(k) + b_{s^2}P_{bs2,\delta}(k) + b_{3\rm nl}\,\sigma_3^2(k)P(k),\end{aligned}$$

origin: (1)x(1) or (1)x(3)→linear bias

PT+halo model for halos

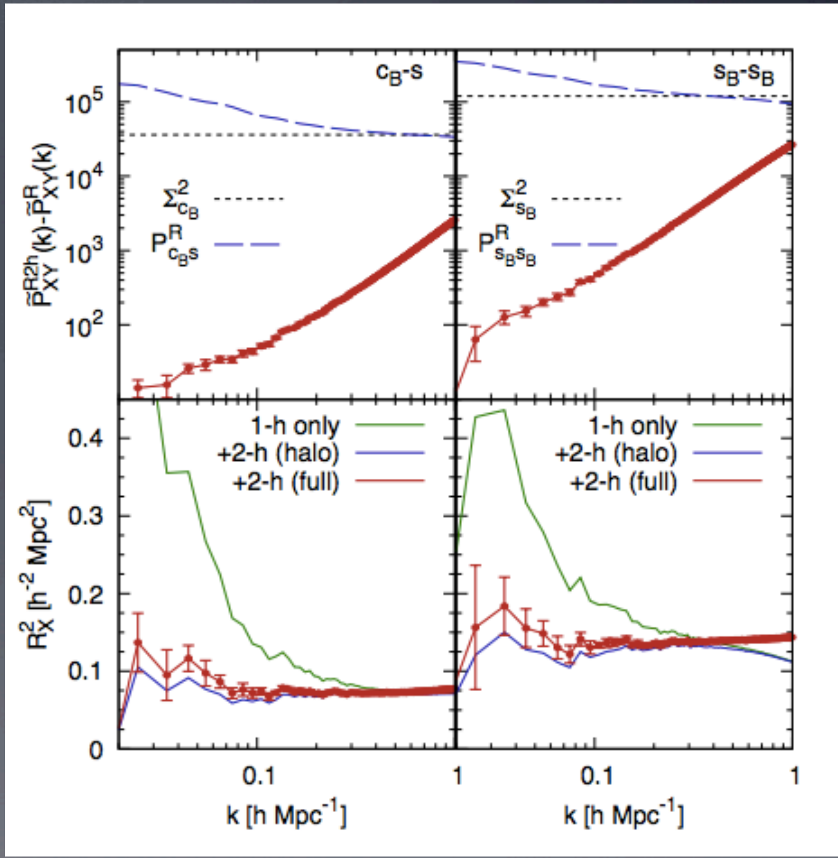
- Use PT to model halos, account for all bias terms
- Biasing: local, non-local, k², stochastic…
- General principle: everything that is allowed by symmetry is also present in reality (many biasing terms)
- Example: 1 loop SPT modeling of RSD for halos (Vlah etal 2013)

RSD is never linear for k>0.1h/Mpc



Beyond PT: effect of satellites in real space

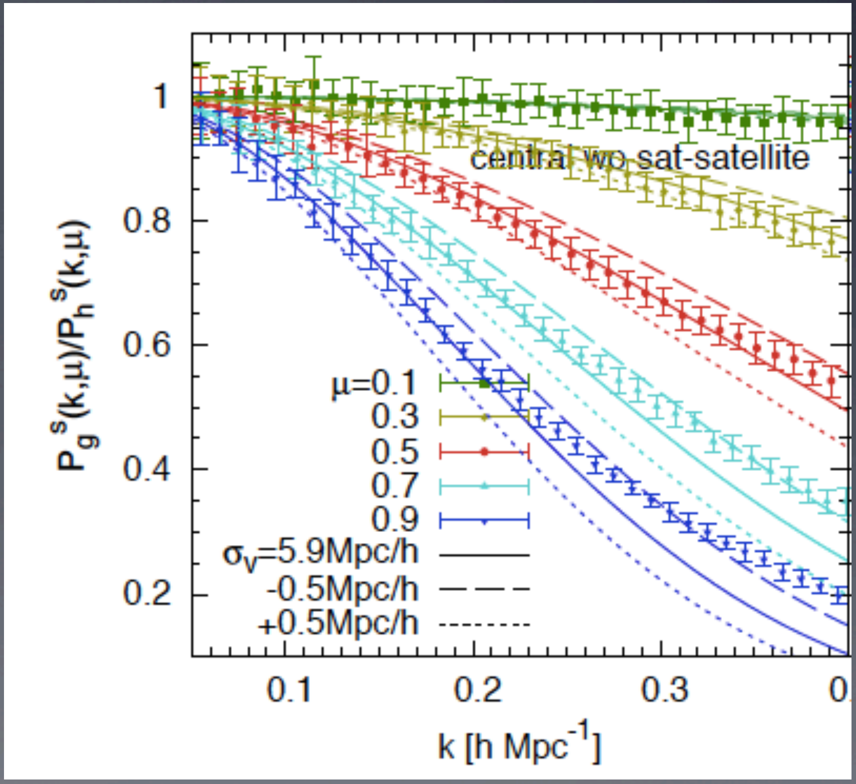
- Satellite-central pairs and satellite-satellite pairs inside halos create additional 1 halo term
- Leading term at low k: additional Poisson shot noise amplitude $\Sigma^2 = V/N_{cs}$ given by number of central-satellite pairs
- Leading correction due to radial distribution $\Sigma^2(1-k^2R^2_{vir})$, same also for 2-halo term
- Must vanish at high k to give 0 or 1/n



Okumura et al 2015

Beyond PT: redshift space distortions

- Supplemented by satellite velocities inside the halos (Fingers of God), inducing 2 halo term and 1 halo term
- FoG term is large: virial theorem $\sigma_{vir}^2 = 50R_{vir}^2$, must use resummed version, e.g. $\exp(-k^2\mu^2\sigma_{vir}^2)$
- Leading term: 2 halo correlation between central galaxies and satellites in different halos
- FoG also applies to 1 halo term



Okumura etal 2015

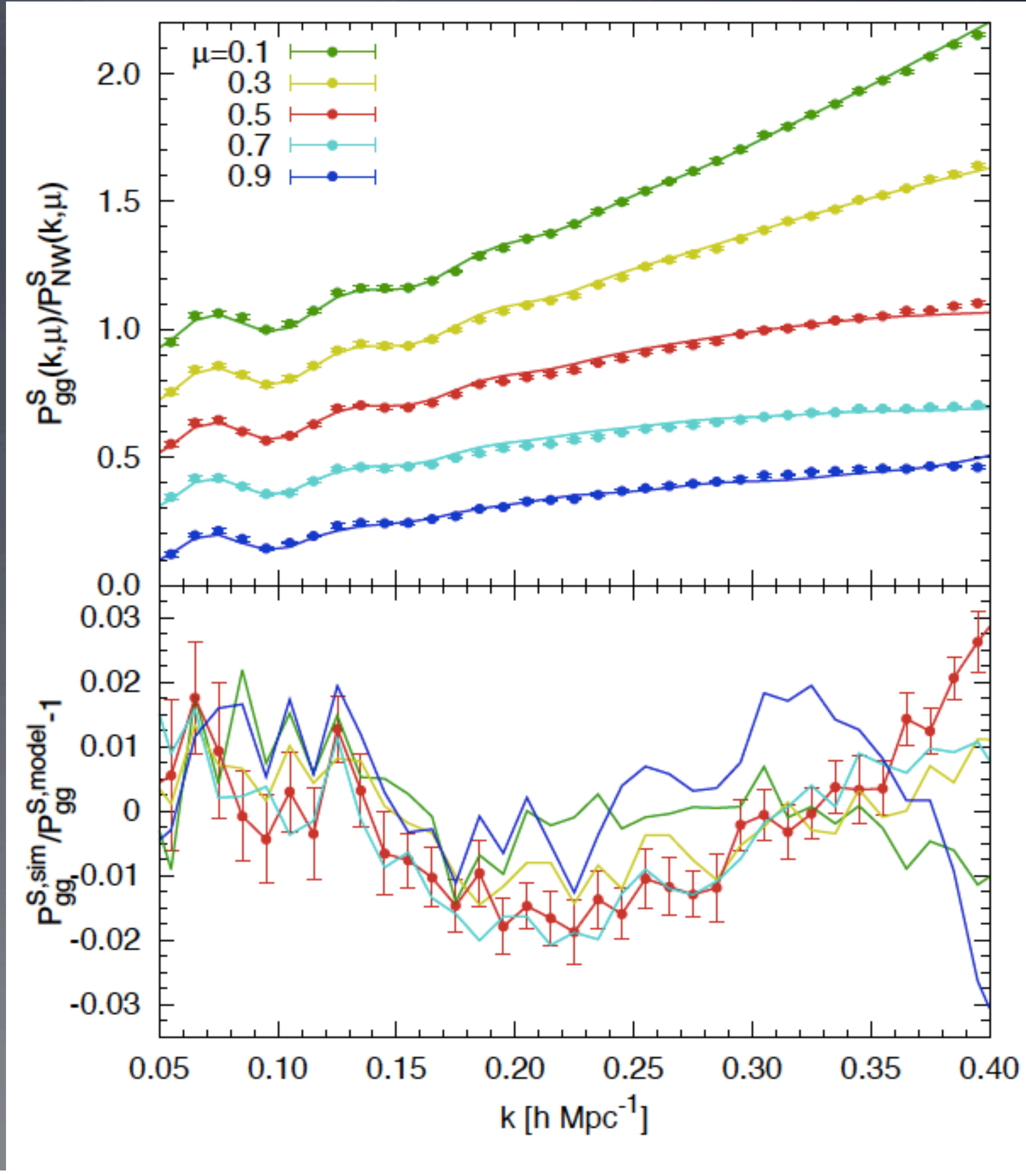
Example: modeling of CMASS in simulations

Okumura etal 2015

Linear Kaiser never a good model

With SPT, 1 shot noise term and 1 FoG term one can go to k=0.2h/Mpc (current state of the art: e.g. Beutler et al, based on models by Saito, Taruya, Scoccimarro)

To go beyond one needs more PT biasing parameters. New NL model achieves 1% to k=0.4h/Mpc by introducing many physical parameters in the PT model: central and satellite galaxies, each with a bias and FoG, 1-halo contribution from central-satellite pairs, halo exclusion…



Gains with scale

	$k_{\max} = 0.2\ h/\text{Mpc}$	$k_{\max} = 0.3\ h/\text{Mpc}$	$k_{\max} = 0.4\ h/\text{Mpc}$
$b_1\sigma_8$	$1.266^{+0.004}_{-0.004}$	$1.267^{+0.004}_{-0.004}$	$1.270^{+0.004}_{-0.004}$
$f\sigma_8$	$0.470^{+0.012}_{-0.009}$	$0.481^{+0.007}_{-0.008}$	$0.480^{+0.007}_{-0.007}$

Hand etal, in prep

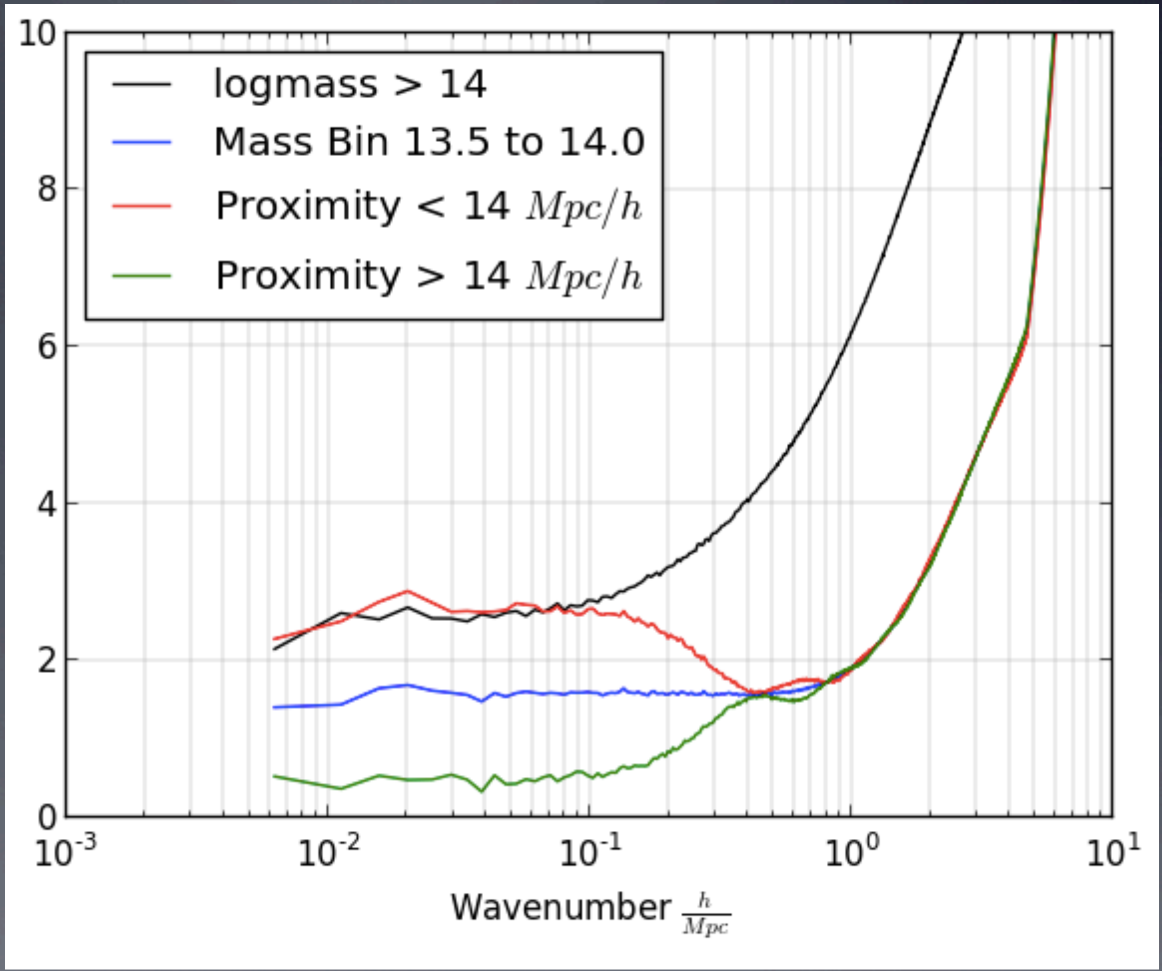
- Increasing k_{max} does not reduce the errors by $k_{max}^{3/2}$
- Errors significantly larger than predicted by Fisher matrix calculations: price one pays for marginalizations
- Current errors from BOSS using k_{max} =0.2h/Mpc at 5-10% level on $f\sigma_8$ (Beutler etal, Sanchez etal, Grieb etal, Gil-Marin etal…): not competitive with CMB lensing
- Reid etal (2014) claim 2.5% error from BOSS!?

Challenges: bias a function of more than halo mass?

- "Assembly" bias defined as bias depending on more than just halo mass.
- It has been for a long time: identify halos with linking length of 0.1 (SO 1000), or local density maxima
- Some of these "halos" embedded inside larger halos with larger bias: subhalo satellites inside a large halo
- HOD models attempt to model environment dependence of "halos" by splitting them into centrals and satellites etc
- There is nothing special about linking length of 0.2 (SO 200)

Environmental and assembly bias

- Take halos of equal mass and split them by proximity to larger halos
- They pick up bias of larger halo on large scales, but bias of small halo on small scales
- Huge bias difference
- Strong scale dependence of the bias
- Lesson: these halos can be thought as being satellites in larger halos
- Is this a disaster for LSS?



The scale dependence is a potential issue

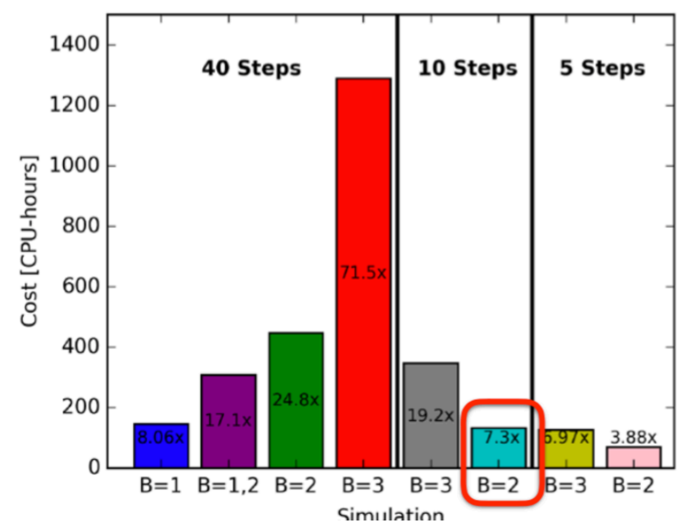
Future directions: emulators

- Key challenge: accurate small scale modeling to extract full information from 2pt function
- Solution: giving up PT. A fully numerical marginalization over all nuisance parameters
- Need simulation based RSD model that is general and fast
- General: HOD modeling, including assembly bias etc
- Fast: FastPM code (Yu etal 2016)
- Learn as you go emulator (Aslanyan 2015): simultaneously samples the posterior and emulates the nonlinear P(k)
- Evaluates the proximity to already evaluated points in parameter space, if too far calls a full simulation
- Fast (HOD) and slow (cosmology) parameters

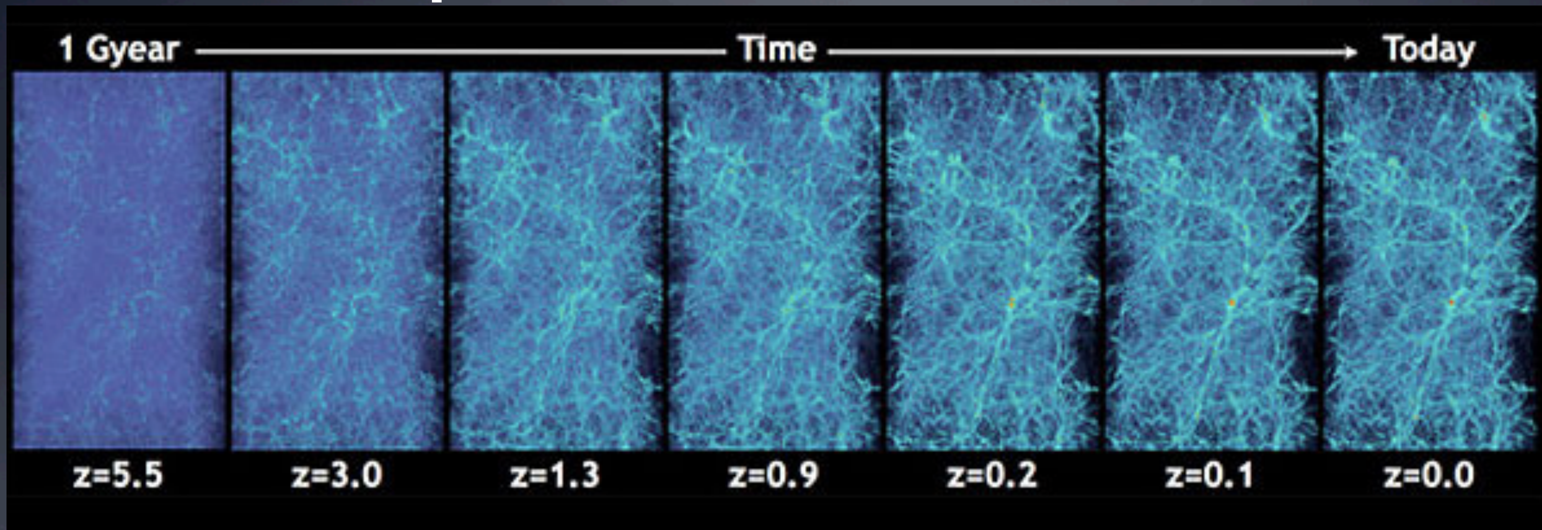
Fast modeling: FastPM

- Need a fast simulation that predicts the data sufficiently well
- FastPM: PM which enforces correct evolution on large scales even with few time steps (typical simulation 1000+ steps)
- Kick-Drift scheme is exact on Zeldovich (different from usual PM)
- Strong scaling tested to 10^4 cores (pencil FFTs)
- 5-10 time steps already give very good results: 100 CPU hours (minutes) for 10^{10} particles
- 4-7 times cost of 2LPT
- Force just PM (unlike COLA)

Feng et al. 2016

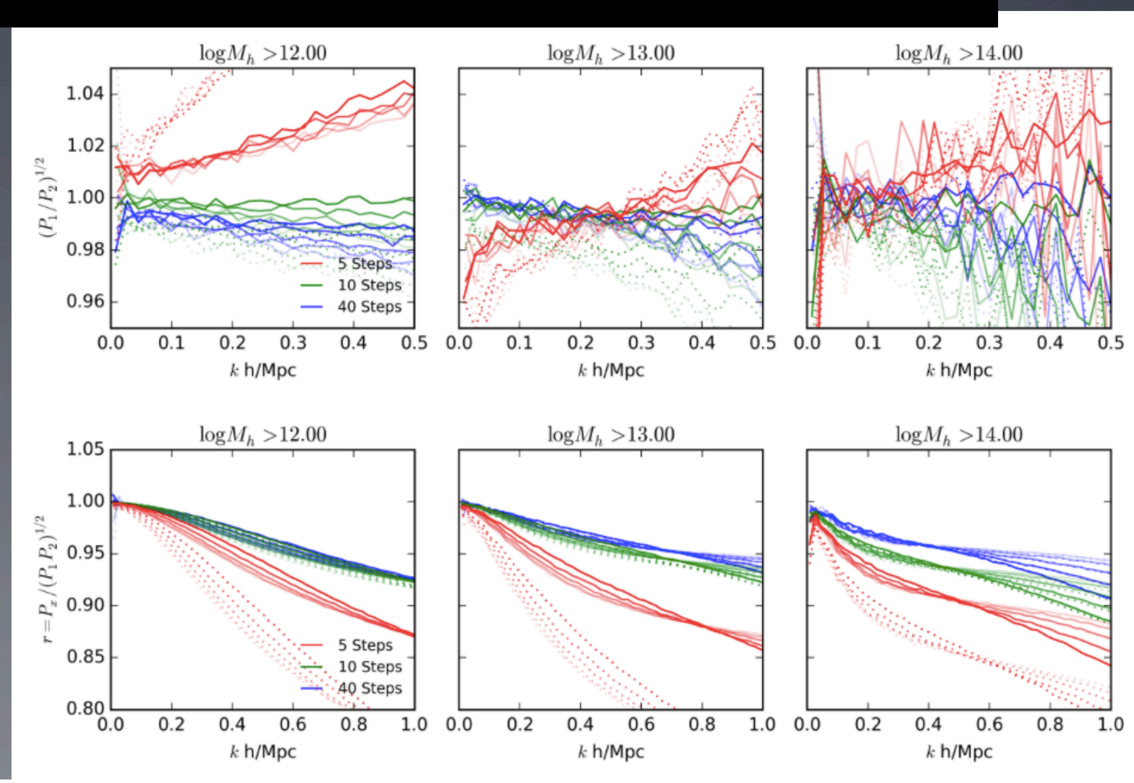


FastPM performance on halos



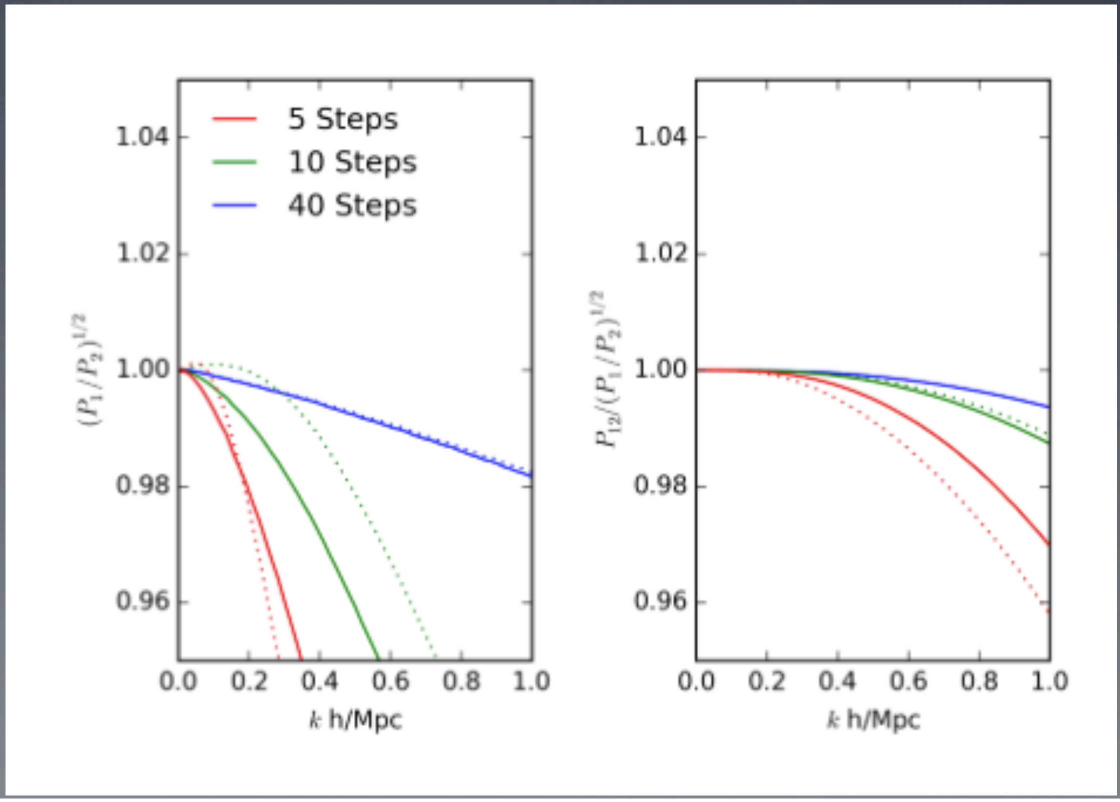
Comparison against very high resolution simulation: 1-2% accurate for 10 time steps

Stochasticity small compared to scatter



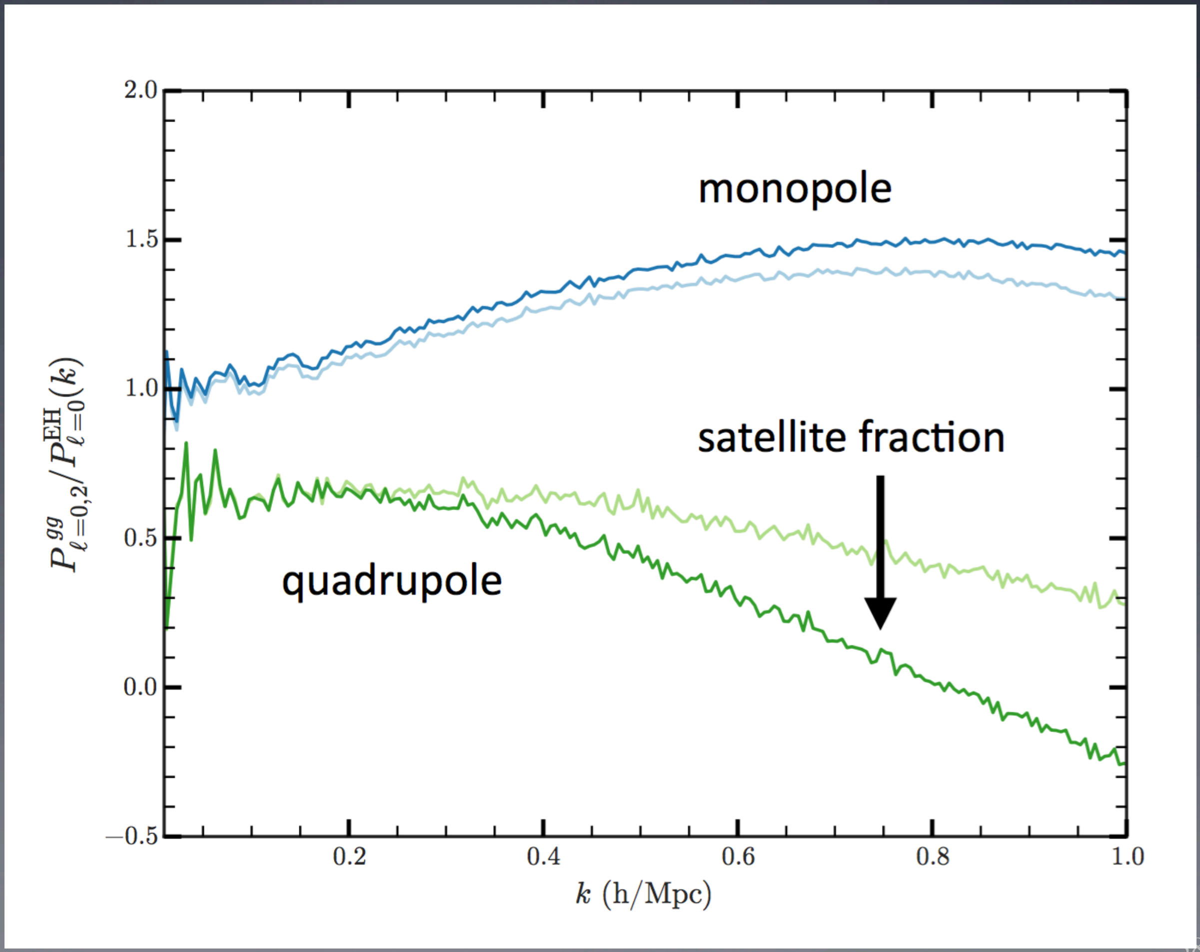
FastPM performance on dark matter

- DM cross-correlation coefficient 0.99 at k=1h/Mpc for 10 steps
- transfer functions less accurate, but can be calibrated on higher resolution sims
- Reason: we do not resolve internal structures of halos with 10 steps



Emulator implementation

- FastPM
- Halo finder using Nbodykit code
- HOD
- P(k) using Nbodykit code
- Example: vary satellite fraction
- Work in progress



What is the ultimate reach of LSS?

- We only know how to connect galaxies to dark matter on large scales ($k<0.1h$/Mpc). Number of modes scales as k^3: huge gains can be achieved on small scales, but we need to get to initial modes
- Currently only crude statistical analyses applied to data: we do not use higher order correlations (bispectrum and higher), or other types of analysis (peaks, voids…), we do not know where information is
- Even if we know the reach of our models in the final power spectrum, we still have little idea of the amount of information that has cascaded to higher order correlations that may be retrievable from the data
- maybe some quiet corner of the universe, far from violent events like massive clusters, can tell us about initial modes on very small scales?

"Theorem": to extract all information need minimum variance map

- There is a close connection between minimum variance map and minimum variance P(k). Need the former to get the latter. Proven for linear case (Seljak 1998)
- Can be generalized to the nonlinear case (Seljak etal 2016, in prep): need to build Fisher matrix and noise bias
- To get the nonlinear minimum variance map: solve the optimization problem

$$L(\boldsymbol{s}|\boldsymbol{d}) = (2\pi)^{-(M+N)/2}\det(\boldsymbol{S})^{-1/2}\det(\boldsymbol{N})^{-1/2}\exp\left(-\frac{1}{2}\left\{\boldsymbol{s}^\dagger\boldsymbol{S}^{-1}\boldsymbol{s} + [\boldsymbol{d}-\boldsymbol{F}(\boldsymbol{s})]^\dagger\boldsymbol{N}^{-1}[\boldsymbol{d}-\boldsymbol{F}(\boldsymbol{s})]\right\}\right)$$

- How to predict data F(s) given initial modes s: run a simulation for each configuration of s
- How to find the maximum posterior for millions of modes s? Curse of dimensionality: $V=2^N$, $N>10^6$, means we cannot search blindly

How to find a minimum variance map in 10^6++ parameter space?

- Maximize likelihood to find minimum variance map

$$L(\boldsymbol{s}|\boldsymbol{d}) = (2\pi)^{-(M+N)/2}\det(\boldsymbol{S})^{-1/2}\det(\boldsymbol{N})^{-1/2}\exp\left(-\frac{1}{2}\left\{\boldsymbol{s}^\dagger\boldsymbol{S}^{-1}\boldsymbol{s} + [\boldsymbol{d}-\boldsymbol{F}(\boldsymbol{s})]^\dagger\boldsymbol{N}^{-1}[\boldsymbol{d}-\boldsymbol{F}(\boldsymbol{s})]\right\}\right)$$

$$\boldsymbol{g} = \frac{1}{2}\frac{\partial\chi^2}{\partial\boldsymbol{s}} = \frac{\boldsymbol{s}_m}{S} - \boldsymbol{R}^\dagger\boldsymbol{N}^{-1}[\boldsymbol{d}-\boldsymbol{F}(\boldsymbol{s}_m)]$$

$$R_{ij} = \frac{\partial F(\boldsymbol{s}_m)_i}{\partial s_j}$$

gradient

$$\boldsymbol{D} = \frac{1}{2}\frac{\partial\chi^2}{\partial\boldsymbol{s}\partial\boldsymbol{s}} = \boldsymbol{S}^{-1} + \boldsymbol{R}^\dagger\boldsymbol{N}^{-1}\boldsymbol{R} + \boldsymbol{F}''[\boldsymbol{d}-\boldsymbol{F}(\boldsymbol{s}_m)]$$

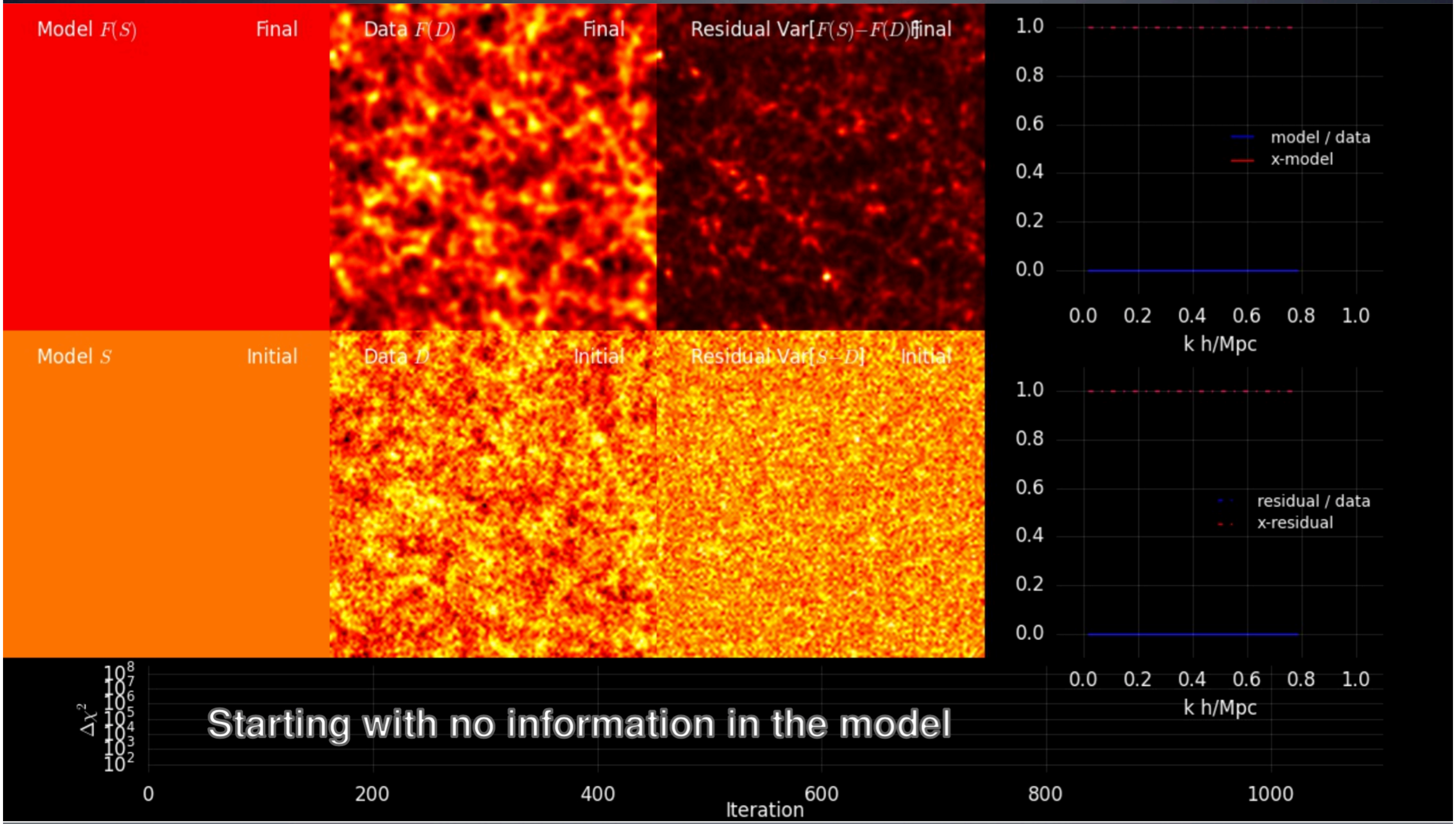
Hessian

$$\delta\boldsymbol{s} = -\boldsymbol{D}^{-1}\boldsymbol{g}.$$

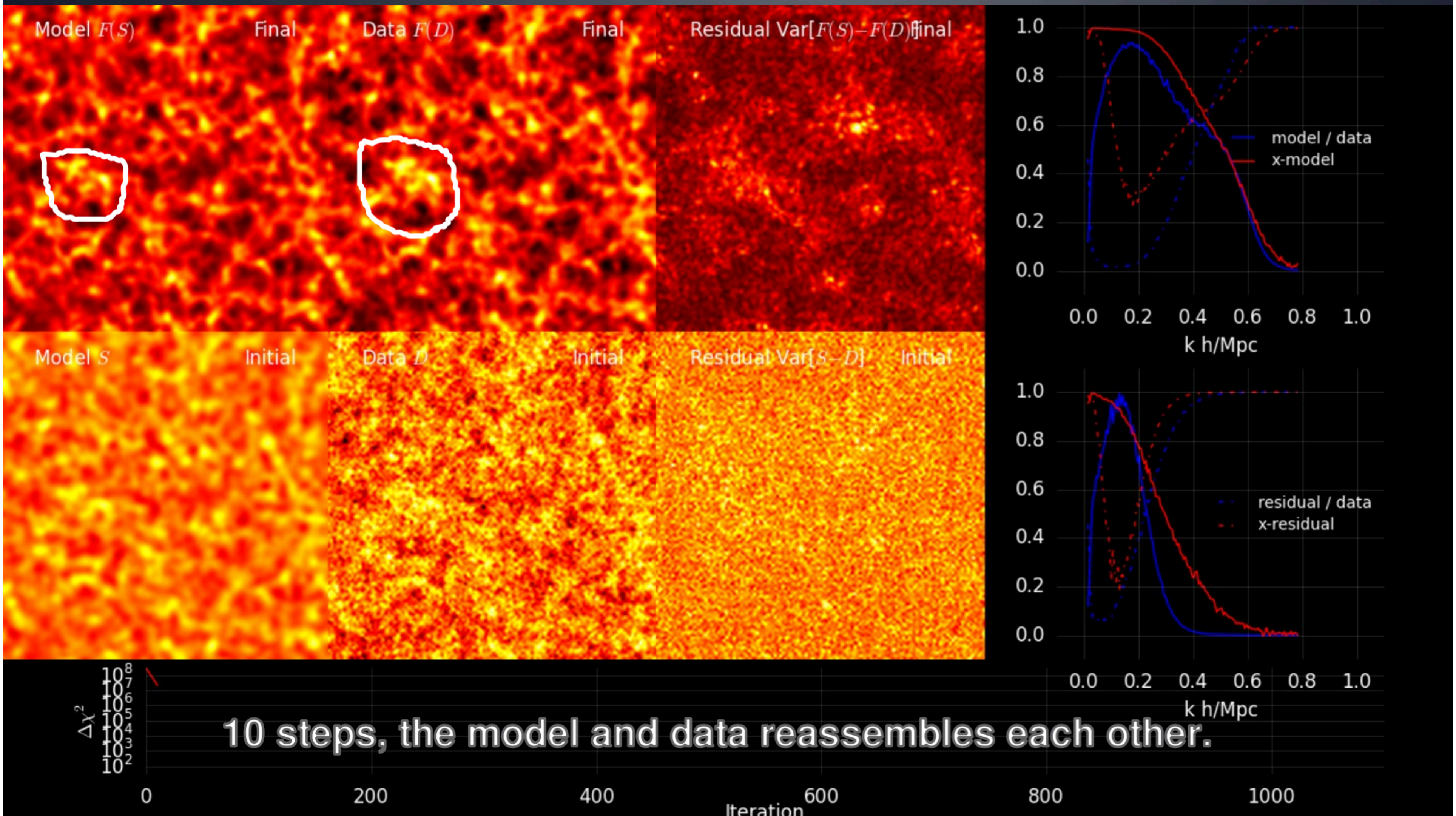
Newton's method

* Need a gradient R_{ij}: derivative of a full simulation wrt all initial modes s
* Also need nonlinear model F(s): a full simulation

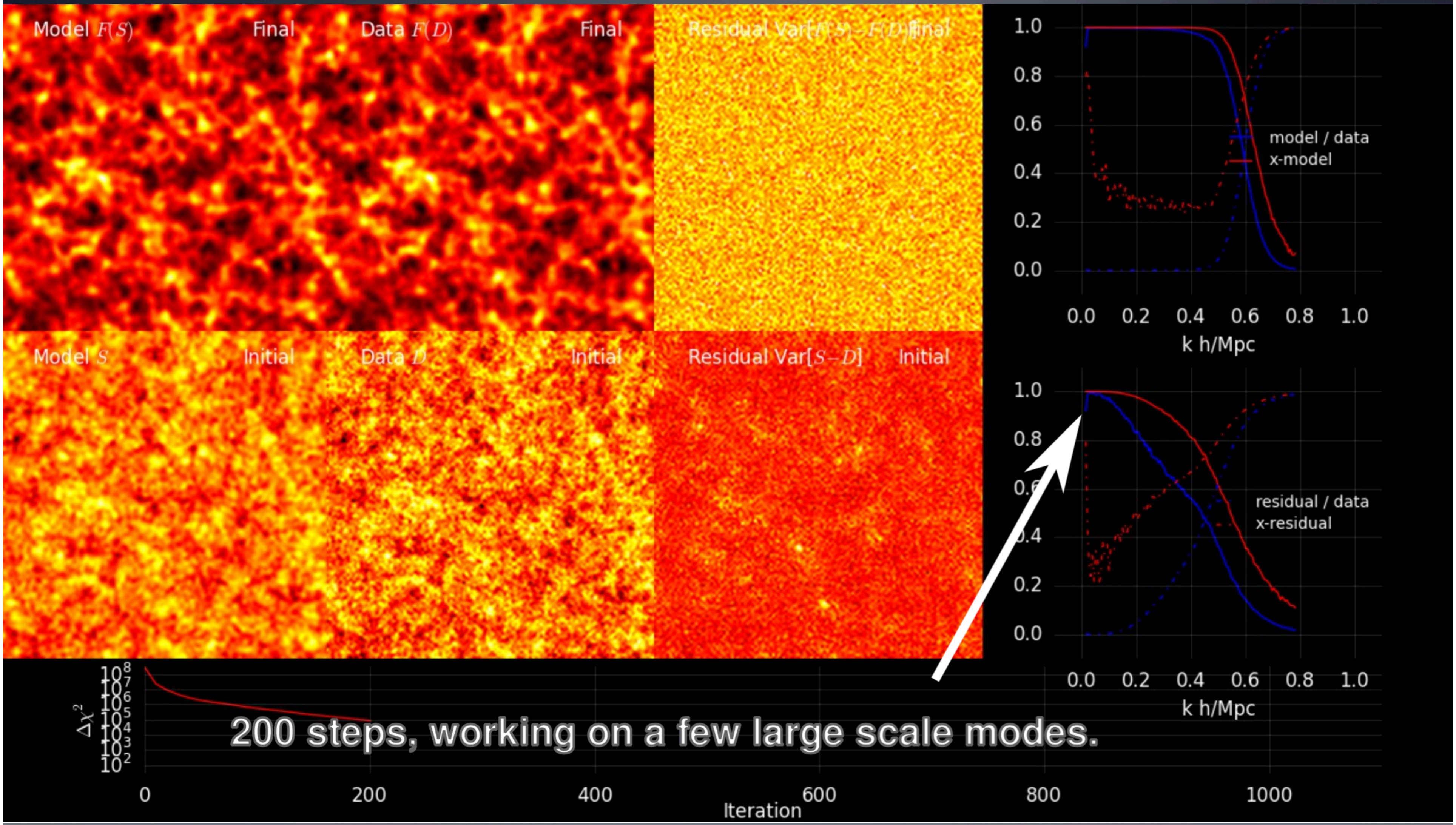
- Need to compute fast F(s) and its gradient
- Our approach: L-BFGS
- Other approaches: HMC (Wang etal, Jasche & Wandelt), slower in our tests

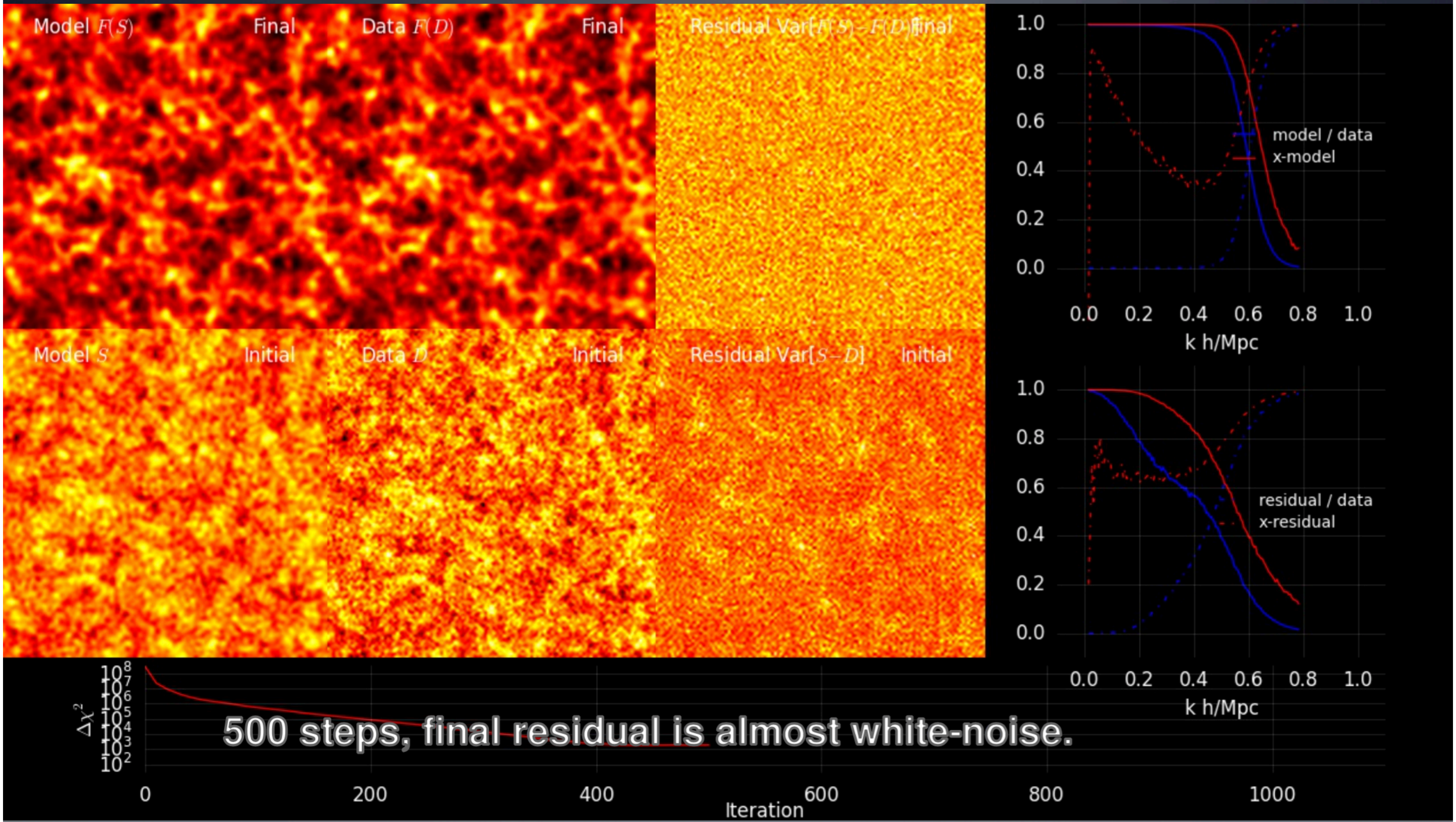


Starting with no information in the model

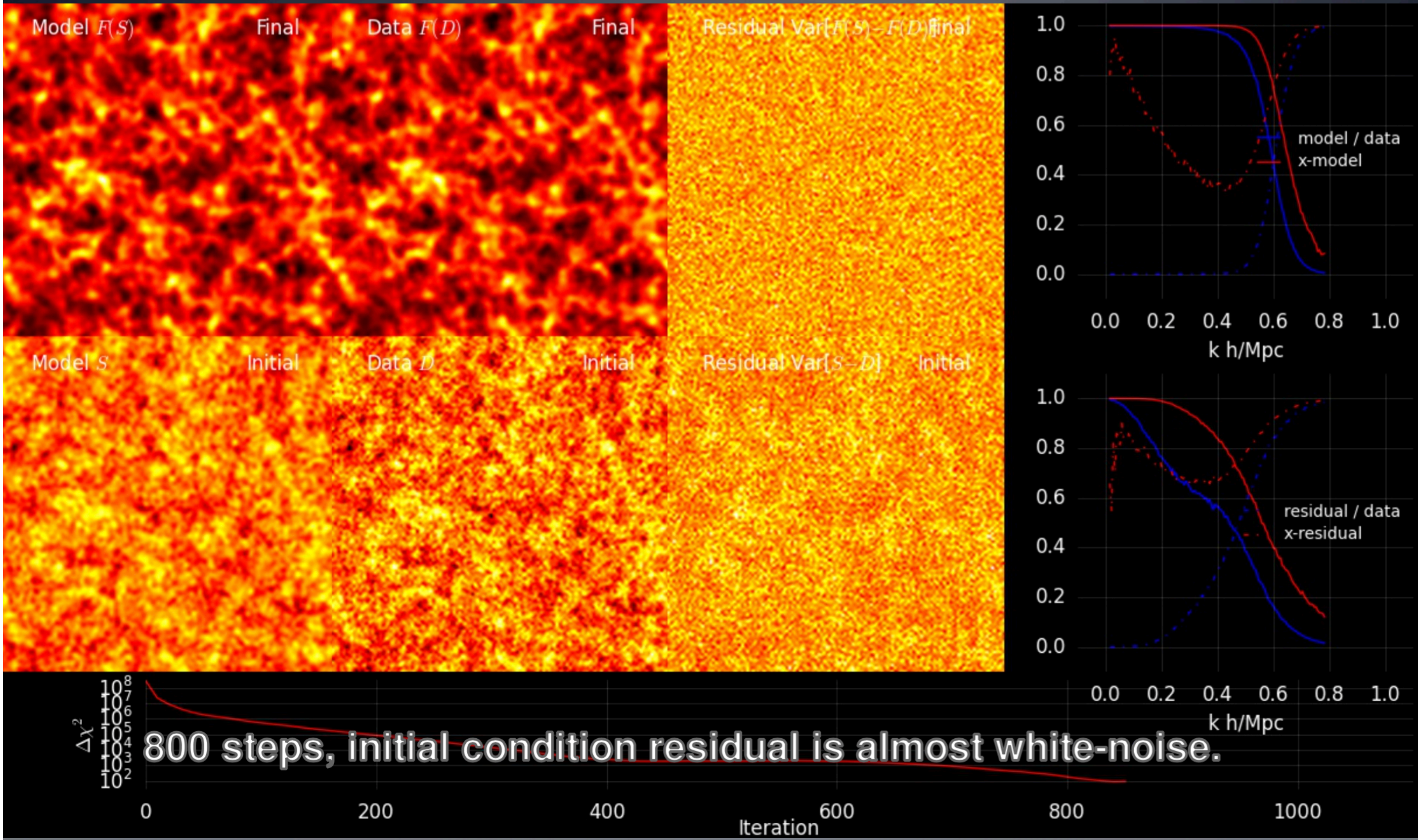


10 steps, the model and data reassembles each other.


Model $F(S)$ Final
Data $F(D)$ Final
Residual Var$[F(S)-F(D)]$ Final
model / data
x-model
Model S Initial
Data D Initial
Residual Var$[S-D]$ Initial
residual / data
x-residual
k h/Mpc
$\Delta\chi^2$
Iteration
200 steps, working on a few large scale modes.



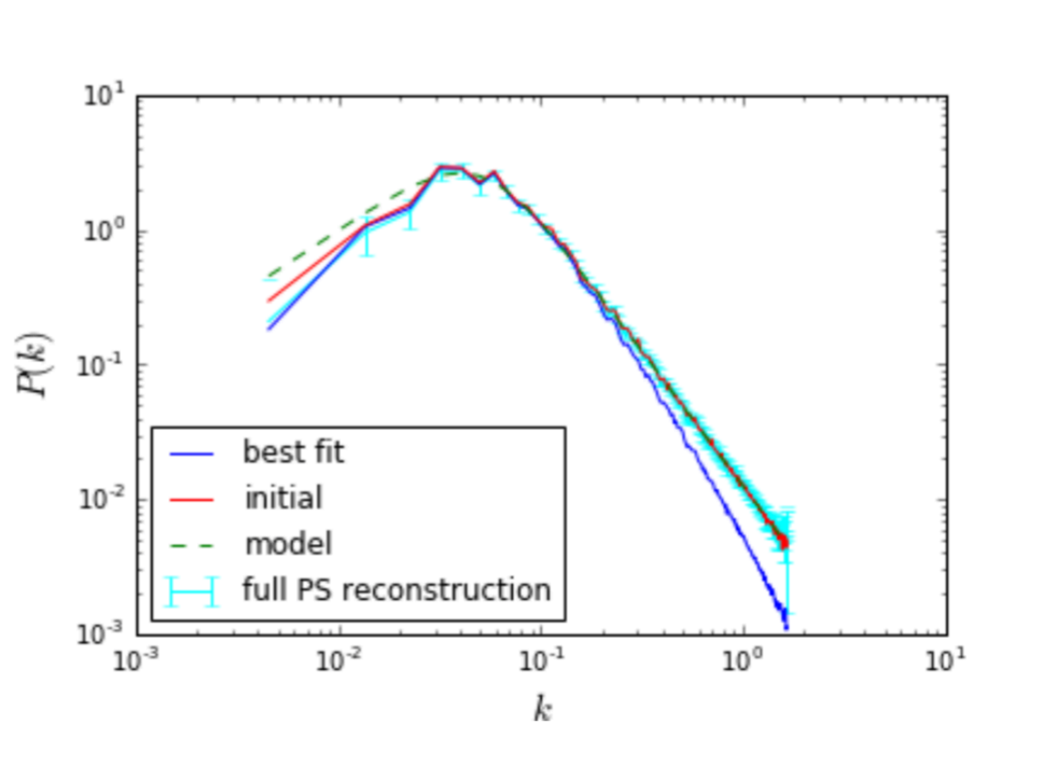
Model F(S) Final
Data F(D) Final
Residual Var[F(S) − F(D)] Final
model / data
x-model
Model S Initial
Data D Initial
Residual Var[S − D] Initial
residual / data
x-residual
k h/Mpc
Iteration
500 steps, final residual is almost white-noise.



800 steps, initial condition residual is almost white-noise.

Optimal initial power spectrum reconstruction

- Need to construct Fisher matrix and noise bias
- Never tried before: first attempt (work in progress, with Yu Feng, Grigor Aslanyan and Chirag Modi)



Implications

- We have a way to formally extract all information and prove it: work in progress, but path is clear
- "Theorem" says one has to make a map of our universe first: hence reduced nonlinear statistics (power spectrum, bispectrum etc) are not needed
- Conversely, to make a map need to solve equations at a field level: discretize, solve for force and displacements... This is exactly what a simulation does, and it does this (nearly) exactly
- No point in doing approximate field calculations (e.g. 1,2,3LPT), as full N-body not much more expensive
- Analytic calculations need to be replaced with numerical methods if all the information is to be extracted

Lessons

- Nonlinear regime of LSS is challenging and rich
- PT is hard and falls apart just when it gets interesting: we have no idea of its reach
- If we want to extract all information need to go numerical