Gravitational lensing

Martin White
UC Berkeley

Outline

- What and why?
- Background and notation.
- Current status.
- Simulating weak lensing.
- Systematics and theoretical input.
- Clusters and weak lensing.
- Lensing the CMB.

Overview

- *Cosmic shear* is the distortion of the shapes of background galaxies due to the bending of light by the potentials associated with large-scale structure.
- For sources at z_s~1 and structure at $0.1<z<1$ it is a percent level effect which can only be detected statistically.
- Observationally tractable.
- Contains “interesting” information.
- Theoretically clean.

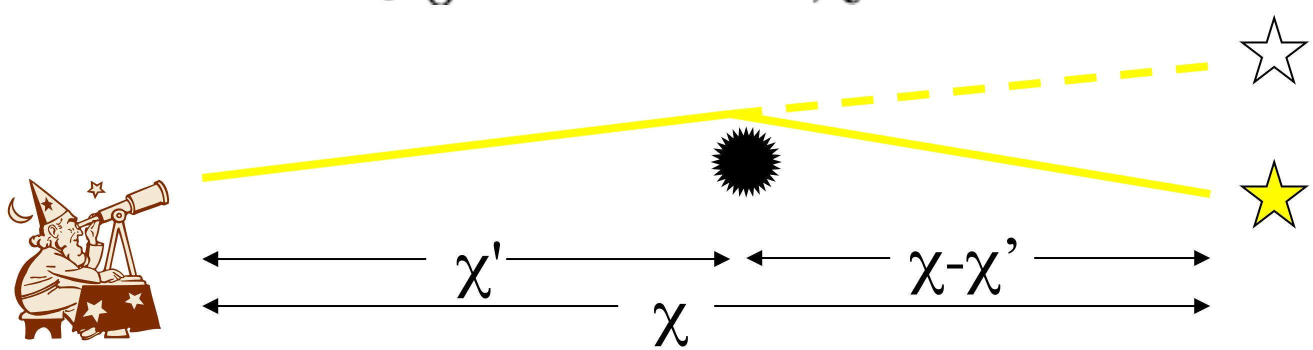
Why Bother?

- Probes large-scale structure
  - Watch structure grow as a function of epoch
  - Observe the formation of objects such as clusters
- Provide estimates of cosmological parameters
  - Measure the mass density (Ω_m)
  - Measure the amplitude of clustering (δ_H, A_s, σ_8)
  - Study Dark Energy (including growth rate!!)
- It is an interesting theoretical problem!

Background

Light from distant sources is deflected by the potentials associated with large-scale structure. Recalling that light deflection goes as the gradient of the potential we can derive the mapping:

$$\vec{\theta}(\chi) = -2\int_0^{\chi} d\chi' \, \frac{\chi - \chi'}{\chi} \nabla_\perp \Phi + \vec{\theta}(0)$$



Unfortunately we do not know *a priori* the positions of the galaxies that we observe, so we need to look at distortions in the shape of galaxies, i.e. the Jacobian of this mapping.

Background (contd)

Thus the "distortion matrix", which describes the how a ray bundle is modified by its transit through the universe is

$$\frac{\partial\theta_i(\chi)}{\partial\theta_j(0)} \equiv \delta_{ij} + A_{ij}$$

where A can be written as the gradient of the mapping or

$$A_{ij} = -2 \int d\chi \; g(\chi) \nabla_i \nabla_j \Phi$$

with

$$g(\chi) \equiv \int_\chi^\infty d\chi_s \; p(\chi_s) \frac{\chi(\chi_s - \chi)}{\chi_s}$$

where $p(\chi_s)$ is the source distribution.

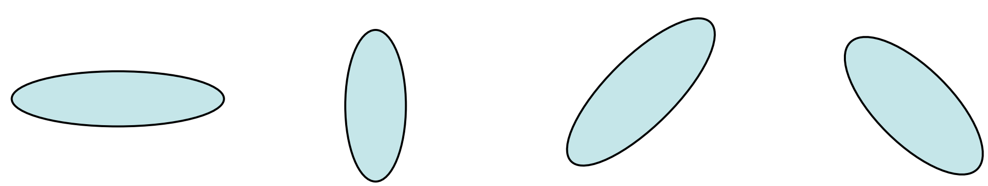
Background (contd)

The distortion matrix ***A*** is conventionally decomposed as

$$(1+A) = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2-\omega \\ -\gamma_2+\omega & 1-\kappa+\gamma_1 \end{pmatrix}$$

where $\kappa \ll 1$ is the convergence and $\gamma \ll 1$ is the shear.
The rotation, ω, only comes from higher order effects and is much smaller than κ or γ.

This maps a circular source to an ellipse:



$\gamma_1>0$ $\gamma_1<0$ $\gamma_2>0$ $\gamma_2<0$

The "Born" approximation

The integral defining **A** should be taken along the perturbed photon path, but the deflection is typically small, so to 1st order we can integrate along a straight line (*Born approximation*).

Then **A** is the second derivative of a projected potential:

$$A_{ij} = -2 \int d\chi \; g(\chi) \nabla_i \nabla_j \Phi \rightarrow \nabla_i \nabla_j \phi$$

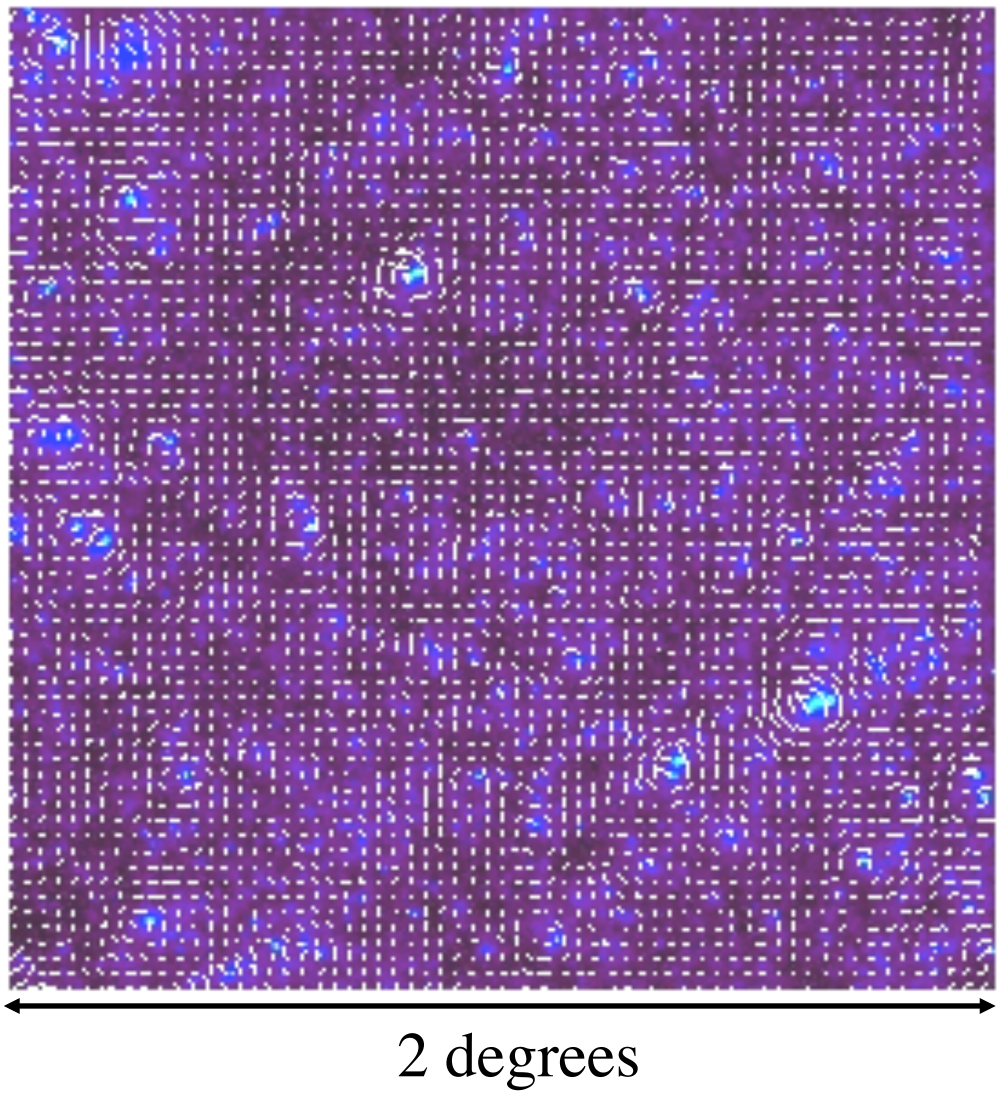
Note κ and γ come from a single potential, ϕ.

If we relate the potential to the density by Poisson's equation, integrate by parts and ignore the surface term

$$\kappa \simeq \frac{3}{2} H_0^2 \Omega_{\mathrm{mat}} \int d\chi \; g(\chi) \frac{\delta}{a}$$

In the Born limit, the convergence is (almost) the projected mass.

A simulated shear field



Obvious non-linear structure, with shear tangential about κ peaks of typical size ~1 arcmin.

Filamentary structure erased by projection.

Shear field sampled (regularly) at about the level achievable observationally from deep space based data.

Measuring Shear

The mapping (1+A) changes the shapes of galaxy images. Thus each galaxy provides a (noisy) measure of the shear at its position:

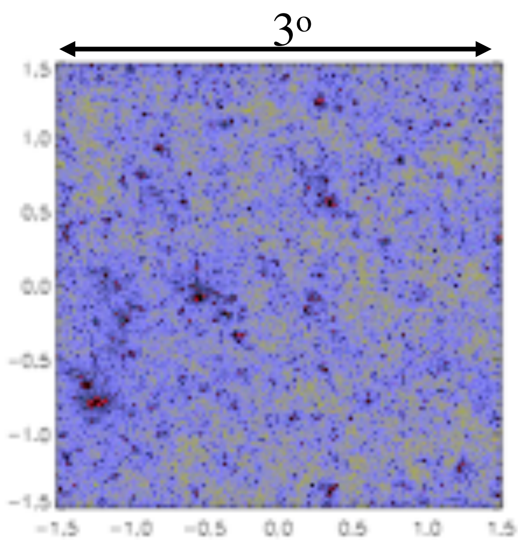
Under the assumption that galaxies are randomly oriented but coherently sheared in some region of the sky, we can simply average the measures of ellipticity to obtain the shear with an error that scales as $e_{rms}/N^{1/2}$ for N galaxies.

$$\gamma_{\rm obs} = \underbrace{\gamma_{\rm intrinsic}}_{0 \pm \sqrt{\frac{\gamma_{\rm int}^2}{N_{\rm gal}}}} + \gamma_{\rm signal} \longrightarrow \text{constant}$$

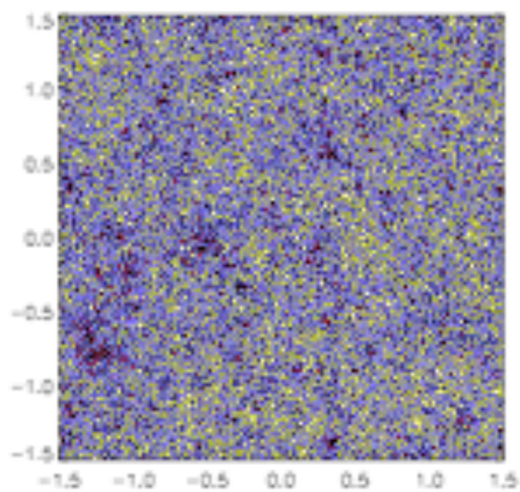
Shot noise

For 10% intrinsic ellipticities and 1% shears we need to average over 100 galaxies to get an estimate of the shear at any position on the sky with S/N~1.

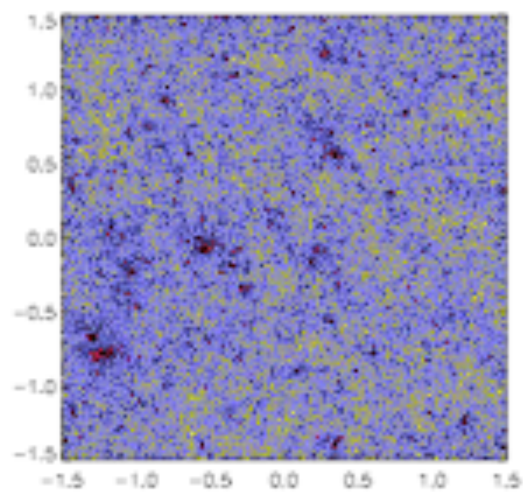
Example: simulated convergence maps with appropriate noise



Input



50 gal/arcmin^2

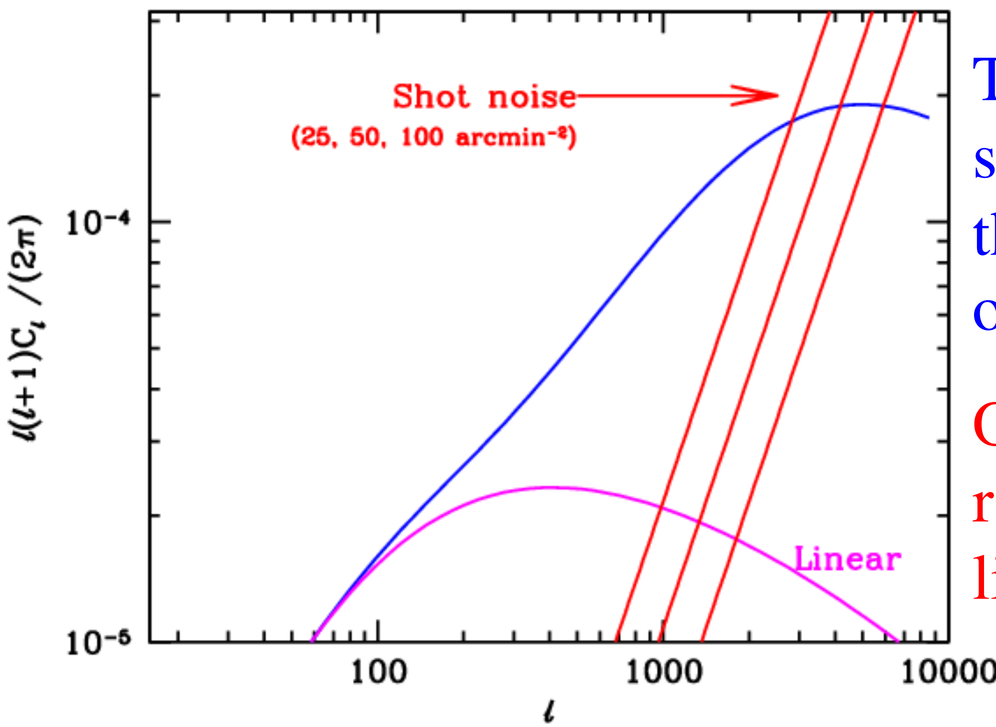


200 gal/arcmin^2

Lensing power spectrum

Within the Born and Limber approximations the shear and convergence power spectra are given by

$$\Delta^2_\kappa(\ell) = \frac{9\pi}{4\ell}\Omega^2_{\rm mat} H_0^4 \int \chi'\, d\chi' \left[\frac{g(\chi')}{a(\chi')}\right]^2 \Delta^2_{\rm m}(k=\frac{\ell}{\chi}, a)$$



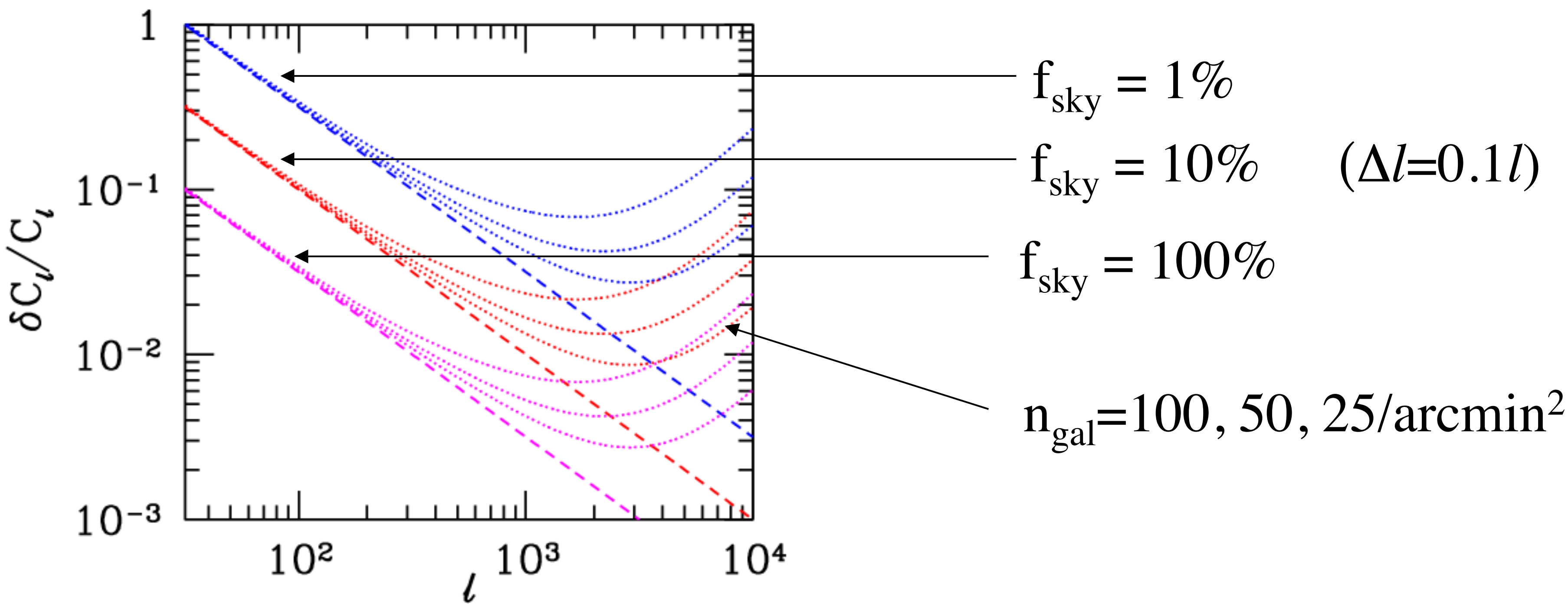
The lensing power spectrum is sensitive to the distance factors, the matter density and the growth of large-scale structure.

Over most of the measurable range it is dominated by non-linear gravitational clustering.

Measuring the power spectrum

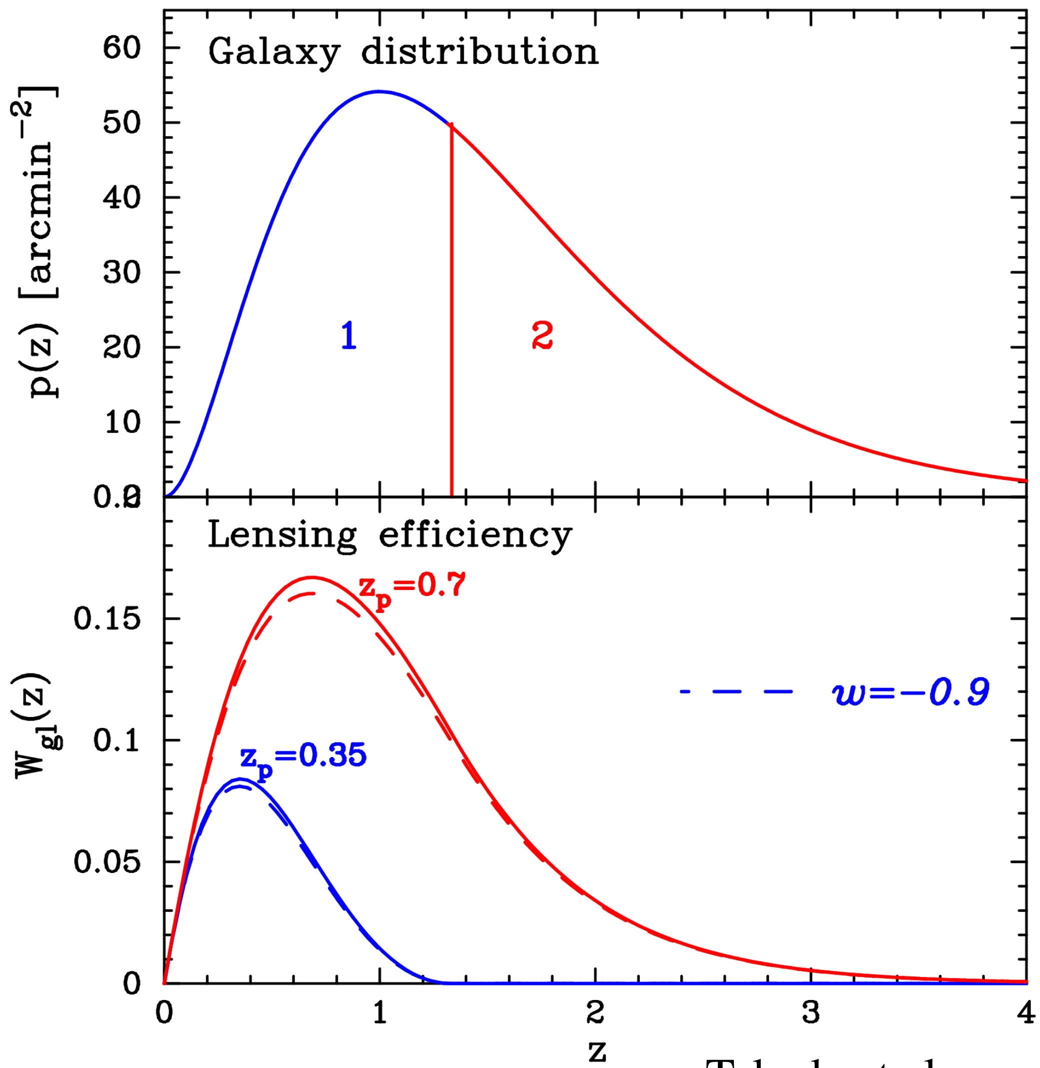
For a Gaussian field measured over f_{sky} of the sky with a finite number of galaxies the error is:

$$\frac{\delta C_\ell}{C_\ell} = \sqrt{\frac{2}{(2\ell+1)f_{\rm sky}}}\left(1 + \frac{\gamma_{\rm rms}^2}{\bar{n}_{\rm gal} C_\ell}\right)$$



Tomography: (2+ε)D surveys

- Tomography refers to the use of information from multiple source redshifts.
- This adds some "depth" information to lensing -- important for evolution studies (Hu 1999).

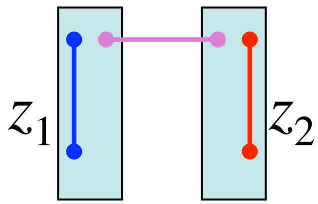


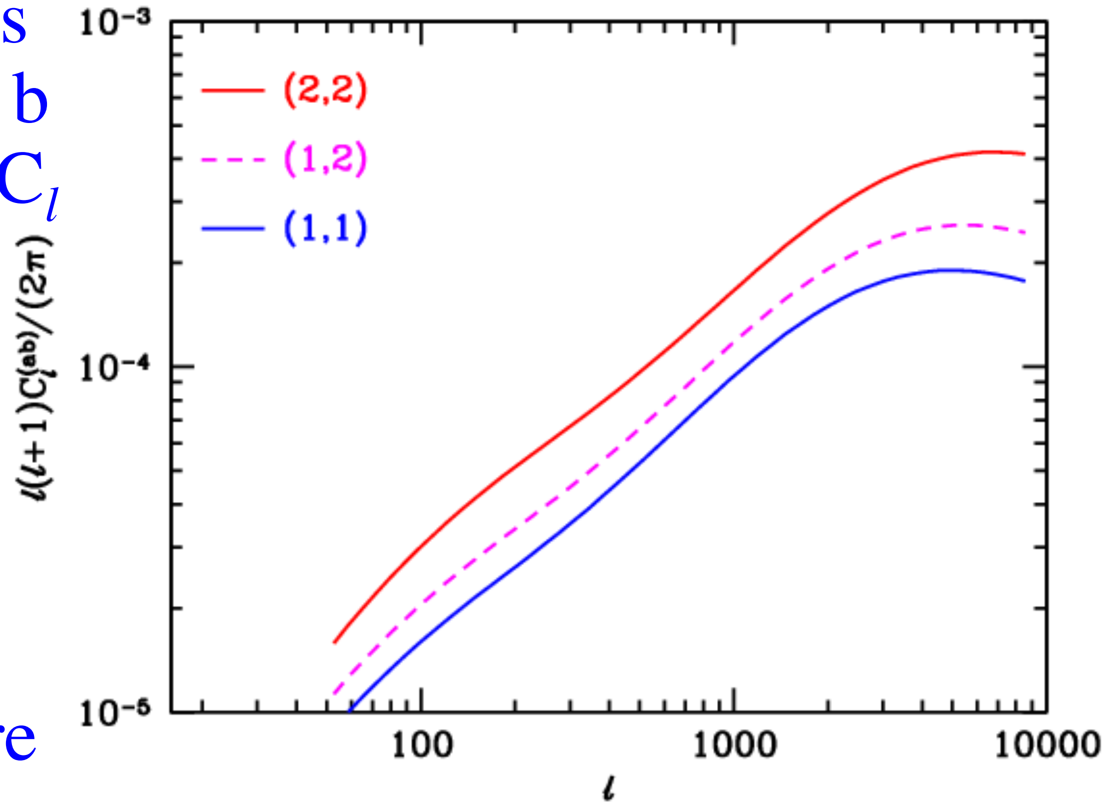
Takada et al.

Tomography (contd)

The generalization is straightforward for any statistic.

If we divide the sources into bins labelled by a, b then we promote C_l to $C_l^{(ab)}$, etc.





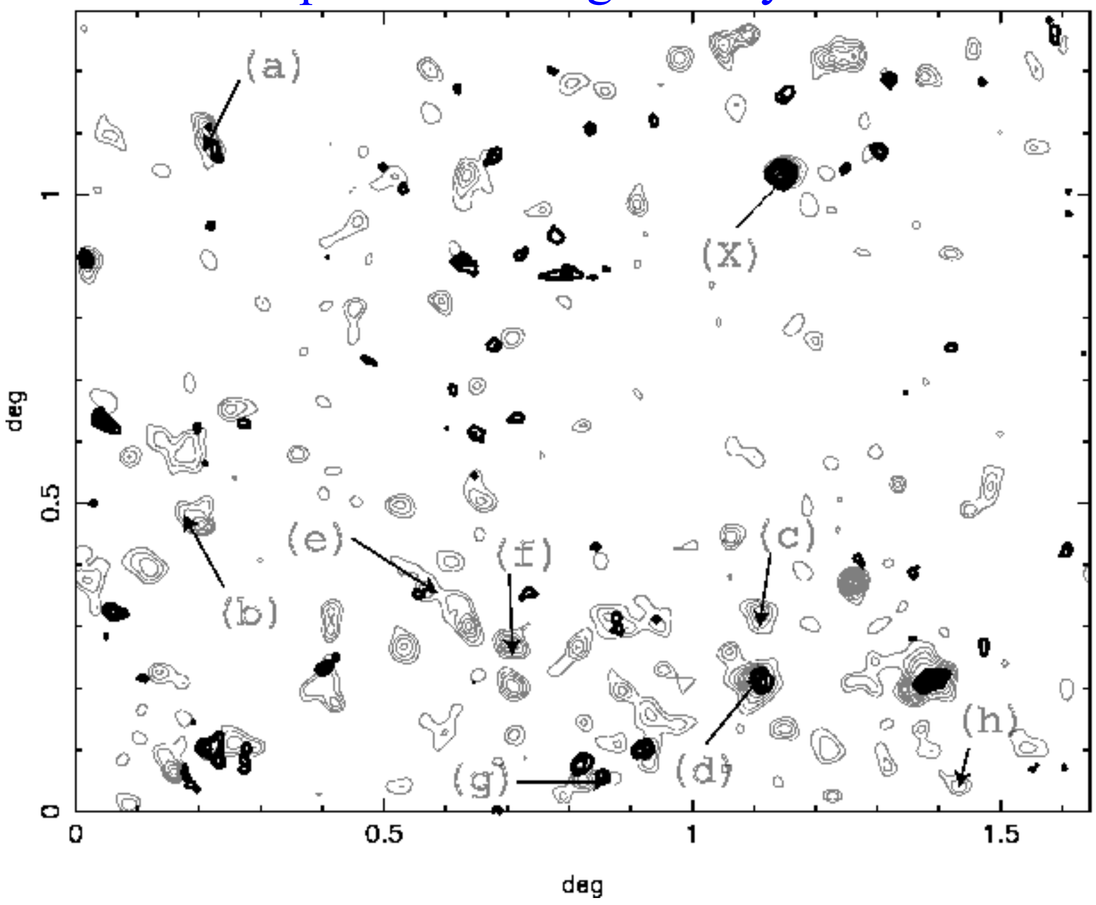
Since $g(\chi)$ is so broad, different source bins are very correlated ($r>0.9$).

Gains saturate quickly!

Observations

First detections of cosmic shear in Spring 2000

Mass map from 2.1 deg^2 survey with Subaru



Miyazaki et al. 2002

Observational status through 2003

Typically tens of galaxies per square arcminute

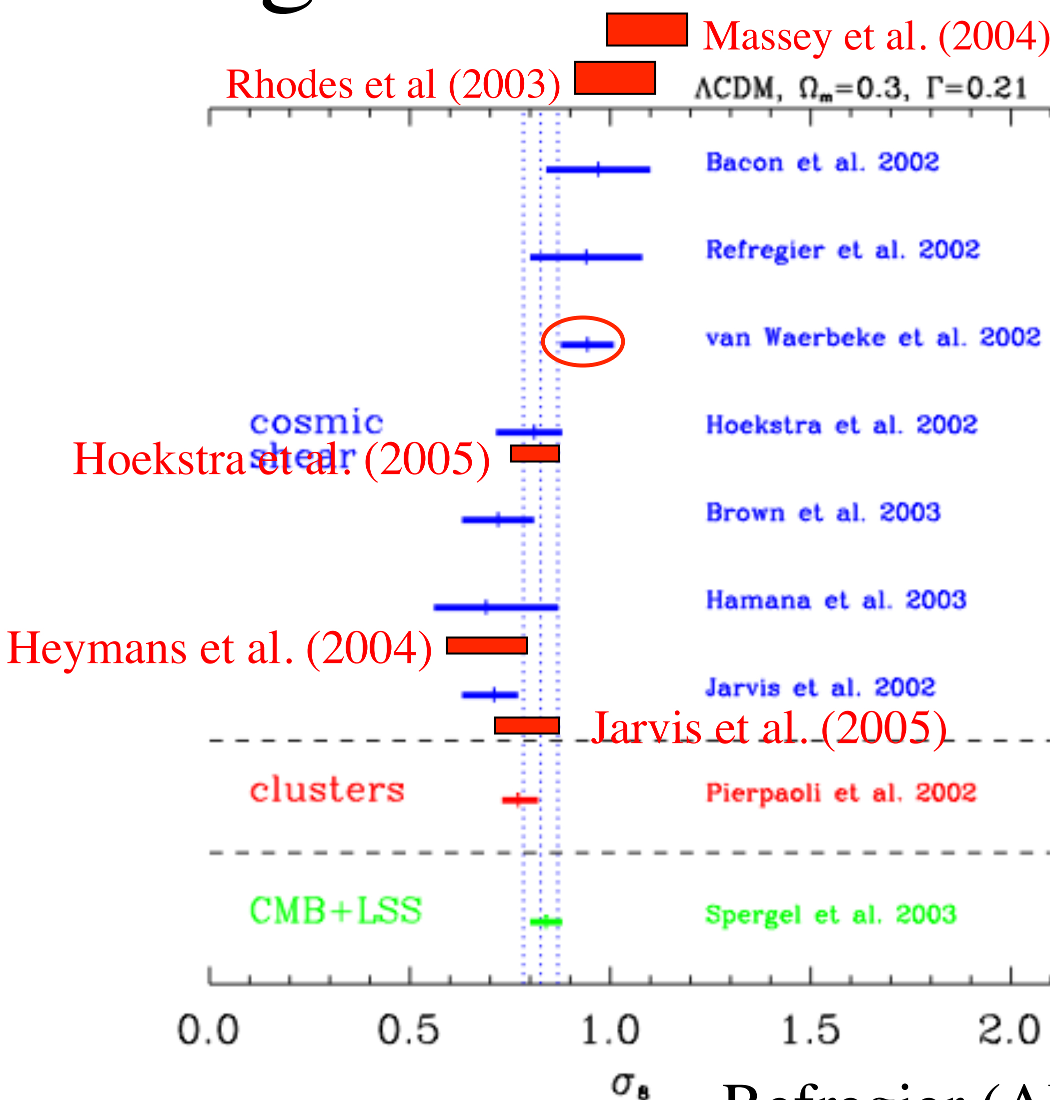
Reference	Year	Telescope	Area (deg^2)	Mag. limit	σ_8
Wittman et al.	2000	CTIO	1.0	$R < 26$	–
van Waerbeke et al.	2000	CFHT	1.7	–	–
Kaiser et al.	2000	CFHT	1.0	$I < 24$	–
Bacon et al.	2000	WHT	0.5	$R < 26$	$1.50^{+0.50}_{-0.50}$
Maoli et al.	2001	VLT	0.7	$I < 25$	$1.03^{+0.03}_{-0.03}$
Rhodes et al.	2001	HST	0.05	$I < 26$	$0.91^{+0.21}_{-0.30}$
van Waerbeke et al.	2001	CFHT	6.5	$I < 25$	$0.88^{+0.02}_{-0.02}$
Hammerle et al.	2002	HST	0.02	–	–
Hoekstra et al.	2002	CFHT	24	$R < 24$	$0.81^{+0.07}_{-0.09}$
van Waerbeke et al.	2002	CFHT	8.5	$I < 25$	$0.98^{+0.06}_{-0.06}$
Refregier et al.	2002	HST	0.4	$I < 24$	$0.94^{+0.14}_{-0.14}$
Bacon et al.	2002	WHT	1.6	$R < 26$	$0.97^{+0.13}_{-0.13}$
Hoekstra et al.	2002	CFHT	53	$R < 24$	$0.86^{+0.04}_{-0.05}$
Jarvis et al.	2002	CTIO	75	$R < 23$	$0.71^{+0.06}_{-0.08}$
Brown et al.	2003	ESO	1.3	$R < 25$	$0.72^{+0.09}_{-0.09}$
Hamana et al.	2003	Subaru	2.1	$R < 26$	$0.69^{+0.18}_{-0.13}$

Recent surveys

New results continue to be published by different groups, mostly on the 2-point function and mostly using single source redshift distributions (with some exceptions).

Reference	Year	Area	z	σ_8
Rhodes et al.	2004	0.25	1.0	1.02 ± 0.16
van Waerbeke et al.	2005	8.5	$0.8 - 1.0$	0.83 ± 0.07
Jarvis et al.	2005	75	0.6 ± 0.1	0.72 ± 0.08
Massey et al.	2005	4.5	0.8 ± 0.08	1.02 ± 0.15
Heymans et al.	2005	0.22	1.0 ± 0.1	0.68 ± 0.13

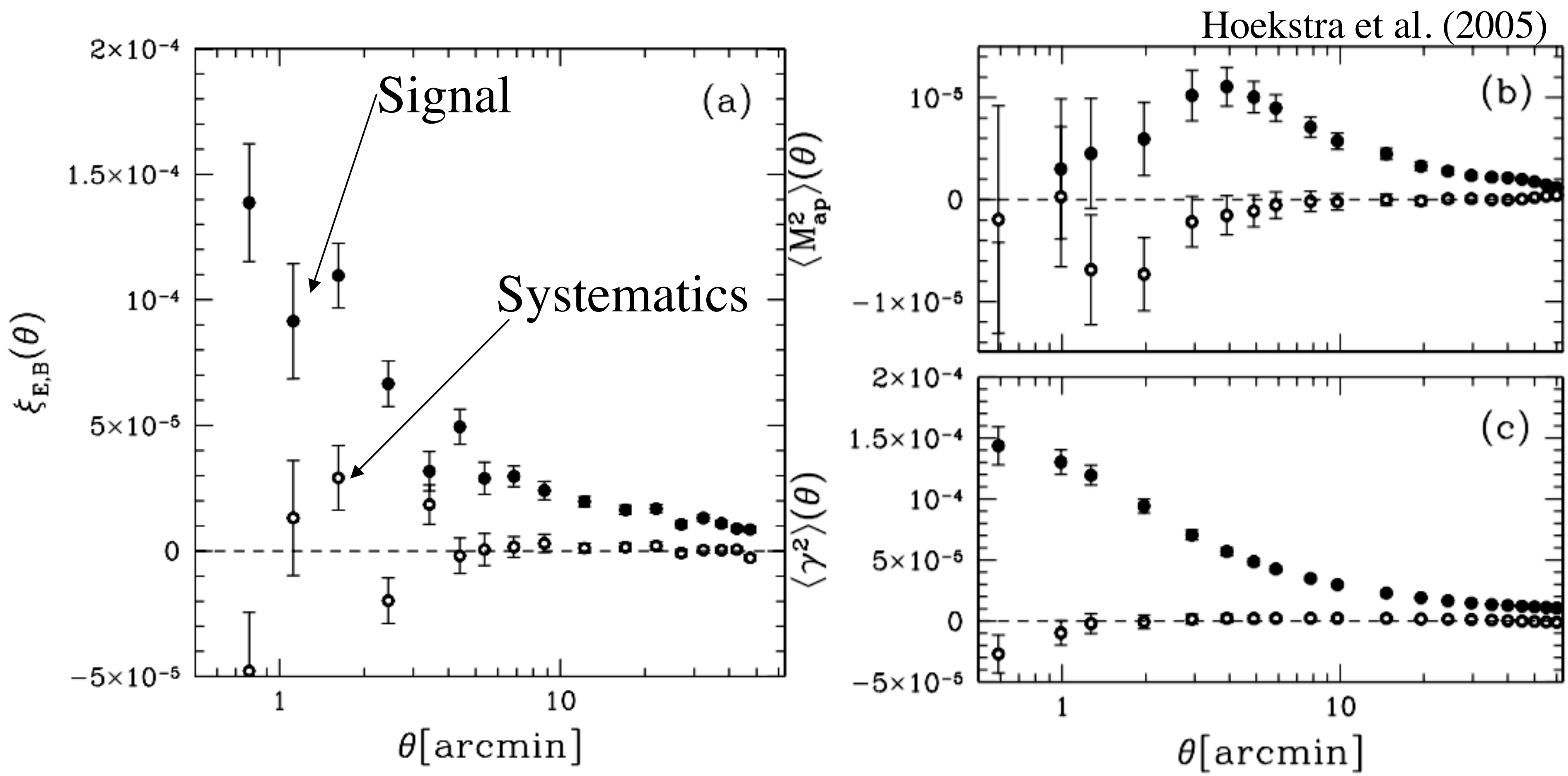
Agreement isn't bad, but …



Refregier (ARAA, 2003)

Need the equivalent of 1% precision in σ_8 to be able to measure dark energy *w*.

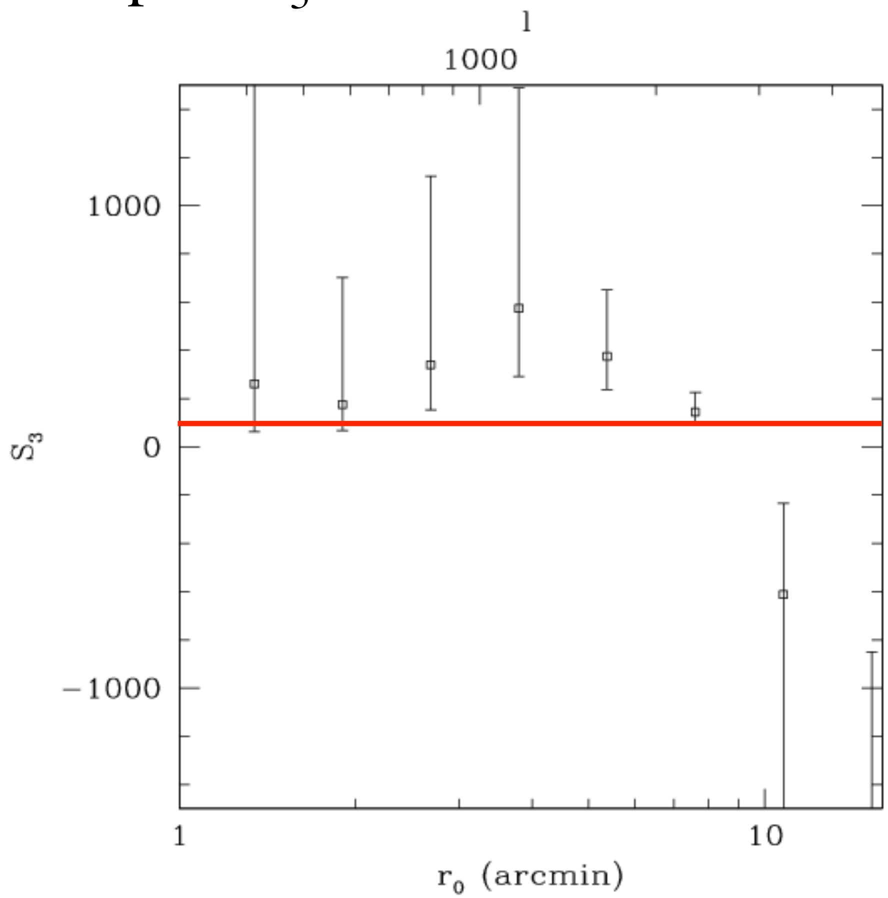
The 2-point function: state of the art



We are beginning to measure the power spectrum. B-modes gone!

The skewness

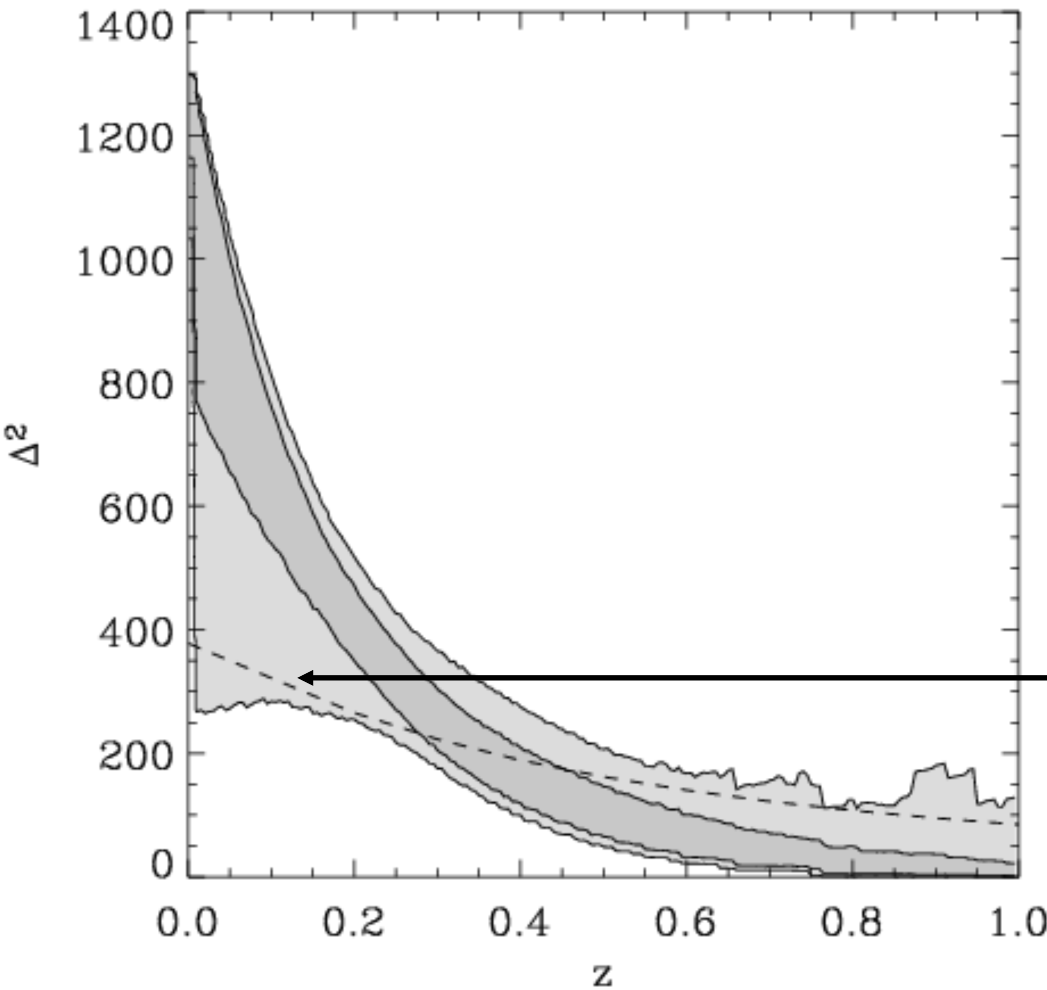
By measuring the 2- and 3-point functions of the shear, the VIRMOS-DESCARTES group (Pen et al. 2003) were able to compute $S_3 = \langle\kappa^3\rangle / \langle\kappa^2\rangle^2$ over a range of scales.



The errors include an allowance for the non-zero *B*-mode they found during this earlier analysis.

Re-analysis in progress.

Structure grows!?



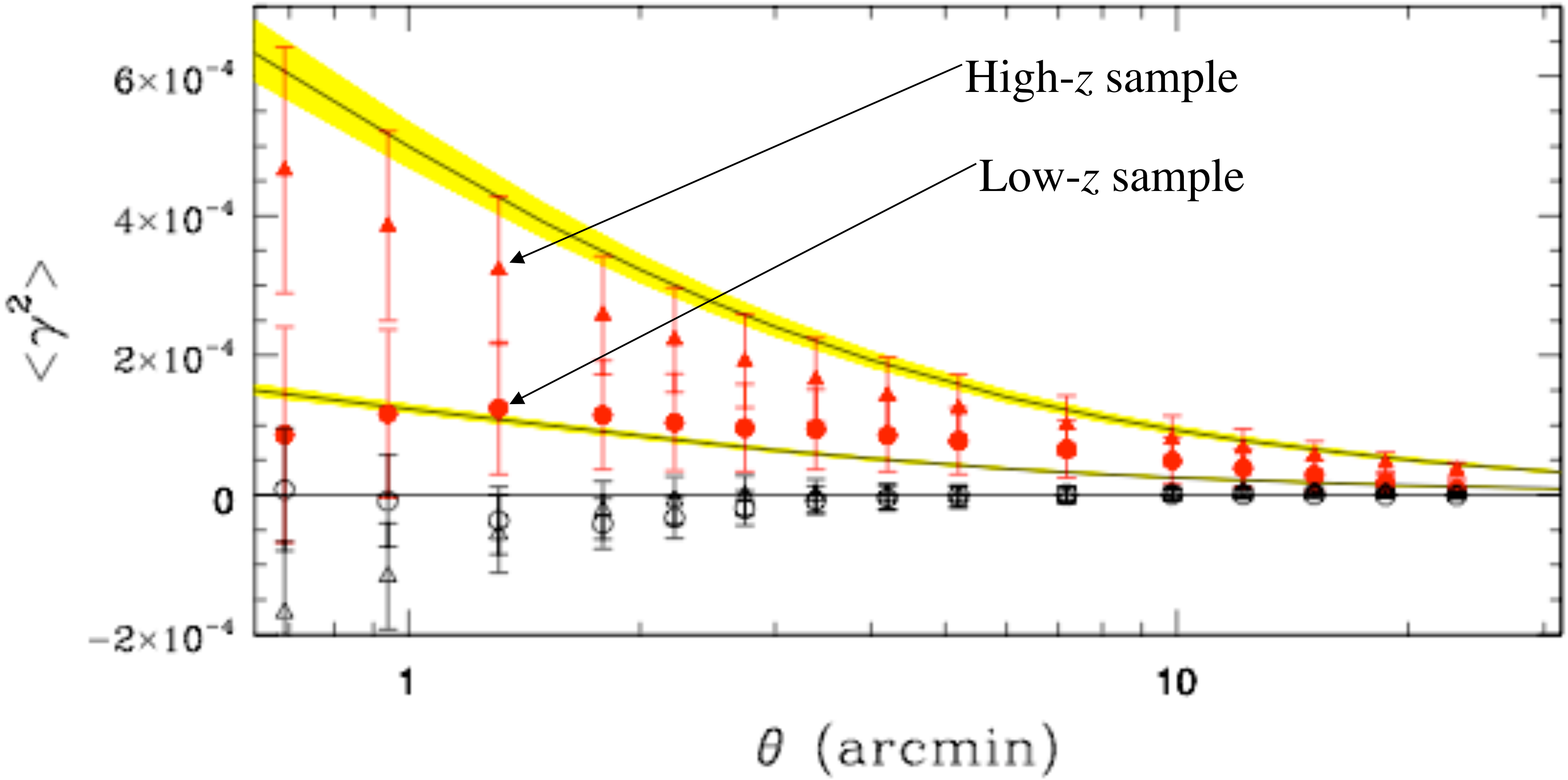
Evolution of power, Δ^2 (14/Mpc), from the COMBO-17 survey.

(Bacon *et al*. 2004)

ΛCDM with σ_8=0.7

Tomography demonstrated!

Semboloni et al. (2005)



Future projects

The 2nd generation of surveys will use "good" telescopes and tested observational techniques: expect dramatic improvement.

Survey	FoV ($\deg^2$)	Area ($\deg^2$)	Start	
DLS	2×0.3	28	1999	
CFHT-LS	1	~ 100	2003	
VST	1	$x00$	2005	?
VISTA	2	$10,000$	2007	??
Pan-Starrs	4×4	$31,000$	2008	
DES	2.1	5,000	200X	
LSST	7	31,000	201X	
SNAP	0.7	1,000	201X	

Computing weak lensing

Theory or simulation?

All lensing "theory" is simulation based … whether it uses fits to halo profiles, halo mass functions and N-body power spectra (semi-"analytic") or direct simulation.

Types and uses of simulations

Lensing lends itself to numerical simulation …

We need numerical simulations to refine and calibrate algorithms and analytic approximations, and potentially serve as templates when the data become available.

Simulations can be used to extract:

- Halo abundances and shapes
- Mass power spectra (and covariance matrices)
- Projected mass maps
- Ray tracing maps
- Mock galaxy catalogues

We have implemented all of these approaches…

Ray tracing: the MLP algorithm

- The gold standard of simulation algorithms is the "multiple lens plane" algorithm, where we trace ray bundles through the evolving mass distribution in an N-body simulation.
- The lensing equations are discretized and the integrals turned into sums:

$$\vec{\theta}_n = -\sum_{p=1}^{n-1} \frac{r(\chi_n - \chi_p)}{r(\chi_n)} \nabla_\perp \psi_p + \vec{\theta}_1$$

$$U_{ij} \equiv \frac{\partial^2 \psi_p}{\partial x_i \partial x_j}$$

$$\mathbf{A}_n = \mathbf{I} - \sum_{p=1}^{n-1} g(\chi_p, \chi_n) \mathbf{U}_p \mathbf{A}_p$$

Tests of the MLP

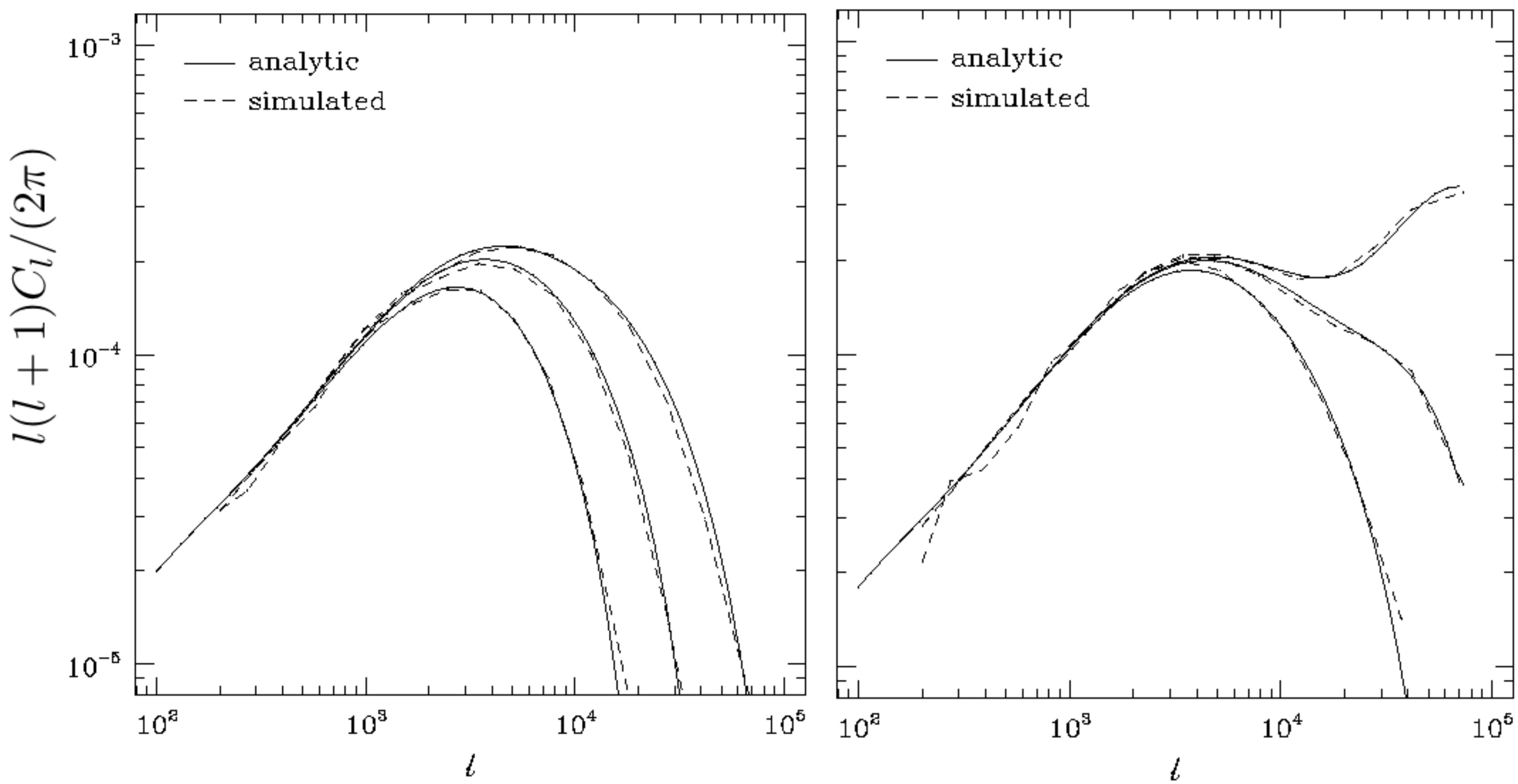
With Chris Vale we have made extensive tests of the MLPA and its convergence properties:

- The effect of border discontinuities
- The "ray-plane perpendicular" approximation
- The first fully 3-d ray tracing protocol
- Time evolution effects
- Number of lens planes necessary
- Numerical resolution issues — Vale & White (2003)
- Test common analytic approximations

Bottom line: MLP is good to at least a few percent in the power spectrum; the limiting computational cost is the generation of N-body simulations.

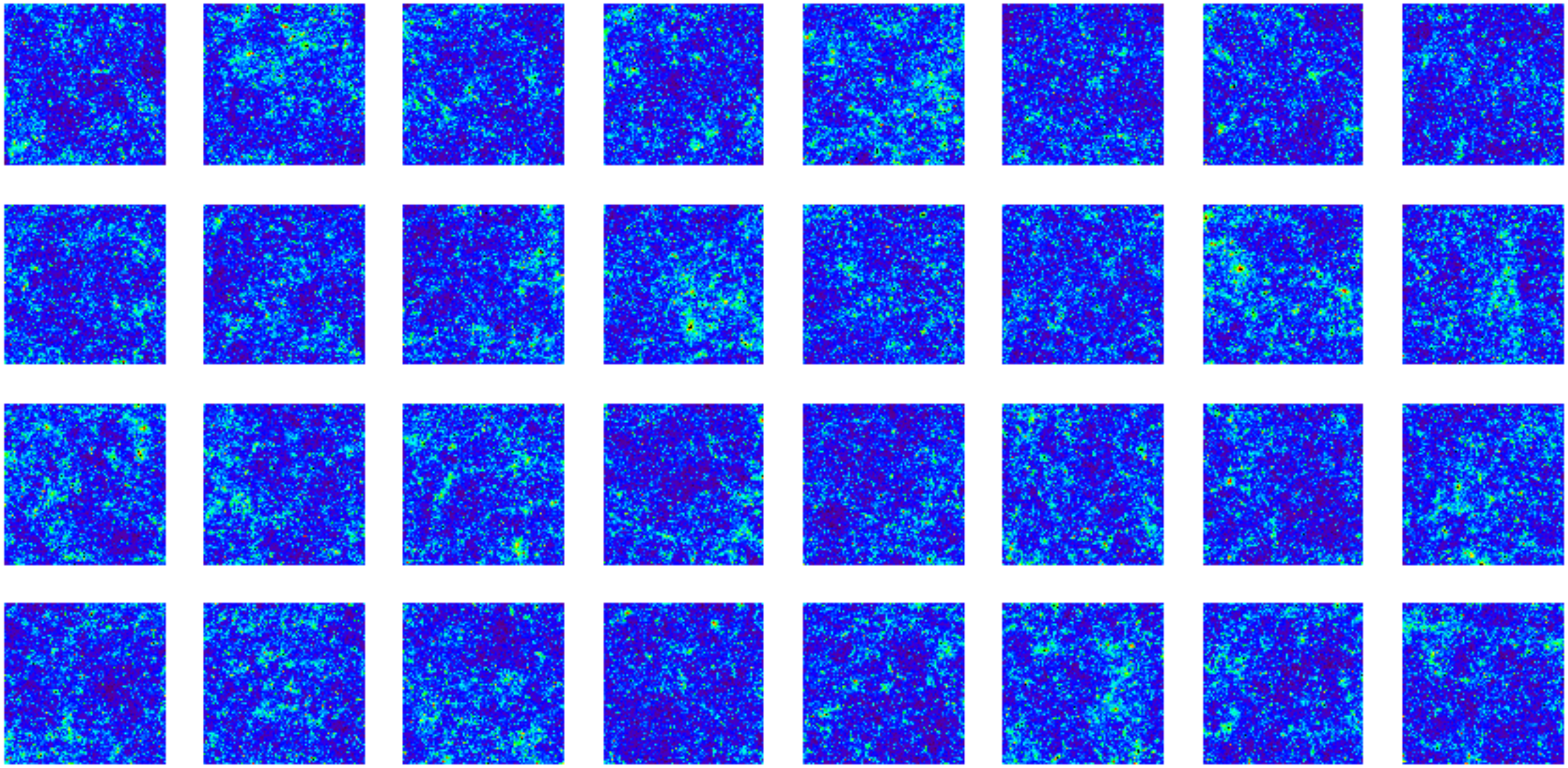
Understanding Our Simulations

(Vale & White 2003)



$\Omega_m = 0.357 \quad \omega = -0.8 \quad h = 0.64 \quad n = 1.00 \quad \sigma_8 = 0.88 \quad \tau = 0.15$

(with Chris Vale)



32 convergence maps, 3^o on a side

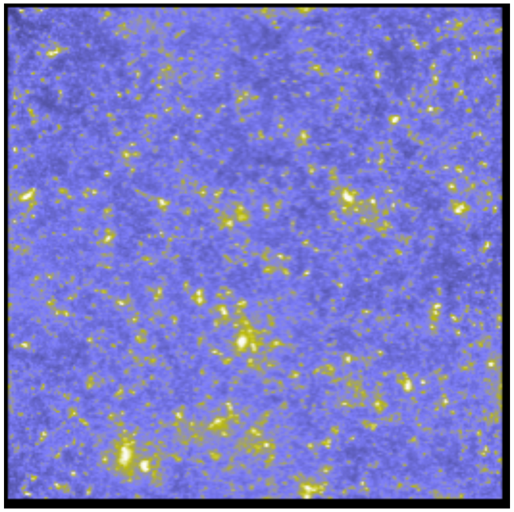
http://mwhite.berkeley.edu/Lensing/

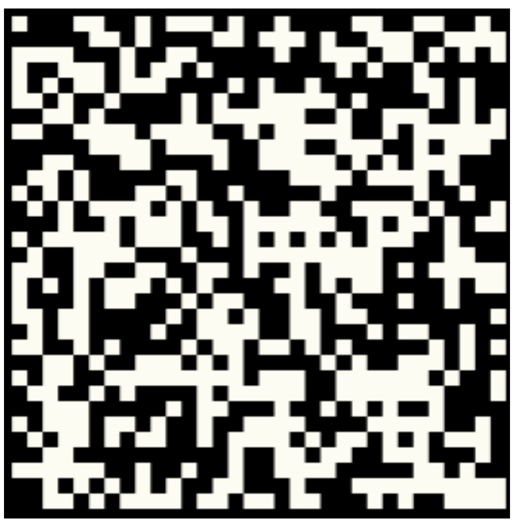
Theory & Analysis

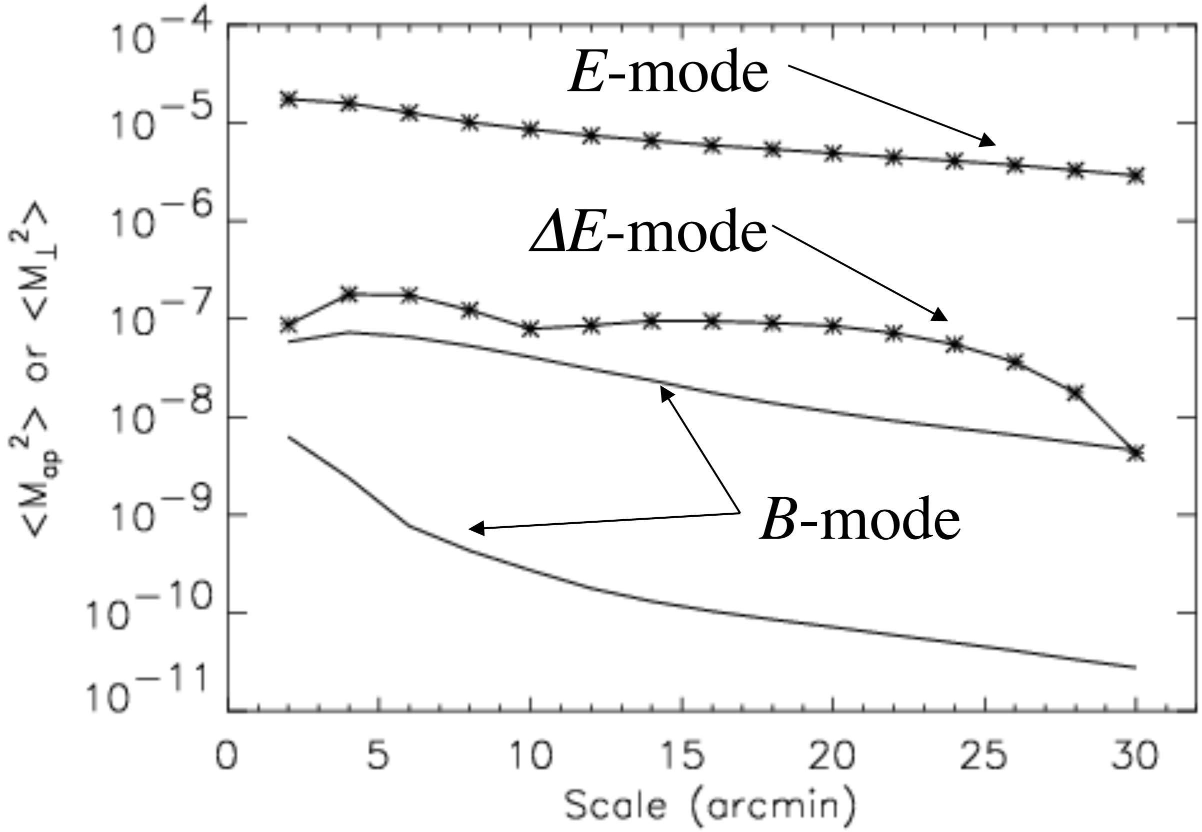
- These maps are very useful for investigating higher order functions and higher order effects.
- The maps make good tests of algorithms.
- The maps can be used to model systematic errors and their removal or estimate error bars from sample variance.
- Available
  - Convergence and shear maps [different $p(z_s)$]
  - Halo catalogs
  - Sheared "galaxy" catalogs
  - Power spectra, …

Example: systematics

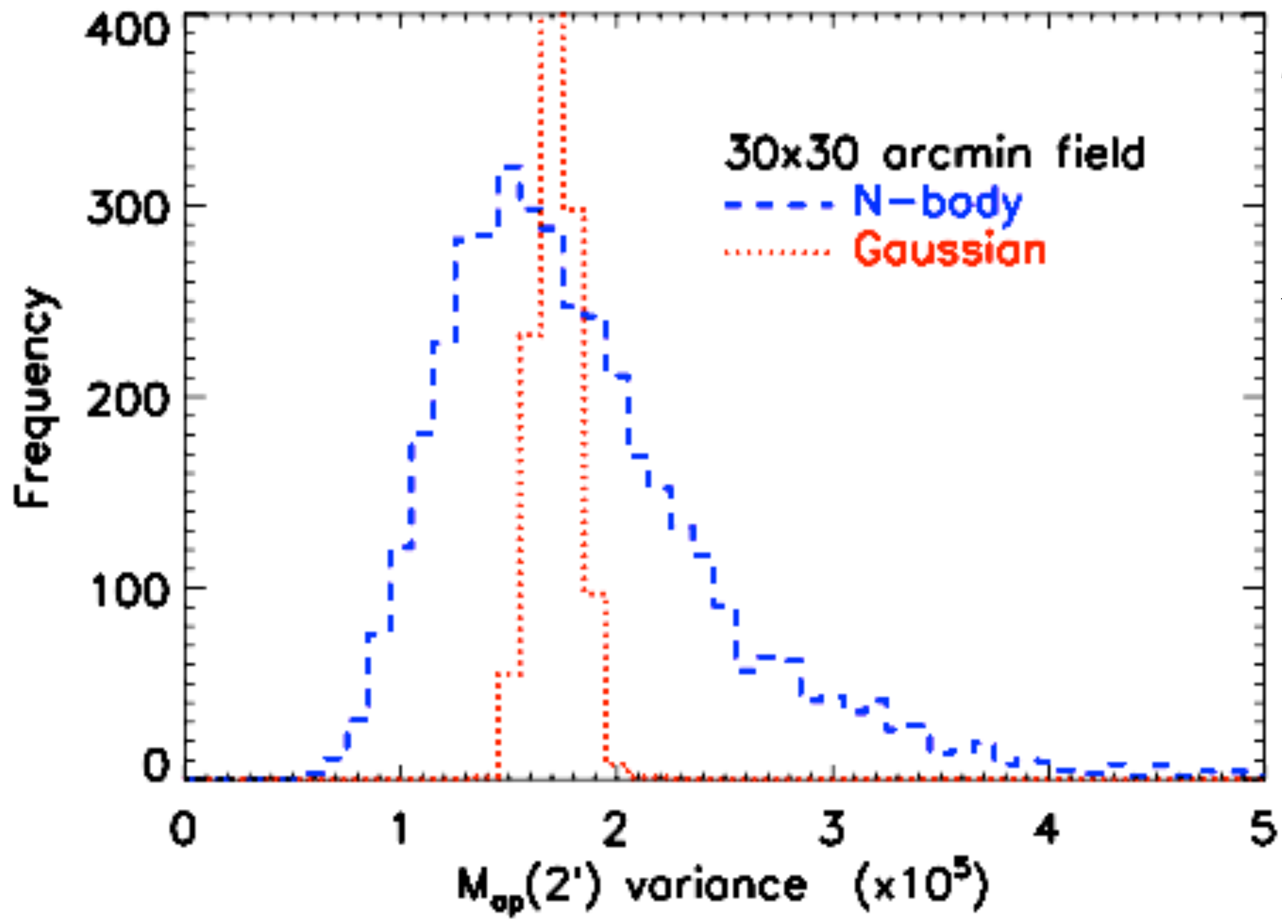
(Vale, Hoekstra, van Waerbeke & White 2004)







Non-Gaussianity & Sample variance



The distribution of variances is not well approximated by a Gaussian on small scales. Sample variance is a larger effect than a naïve calculation would indicate.

Reduced shear

- Unless we have a measurement of the intrinsic size or magnification of a galaxy we cannot measure γ but only $g=\gamma/(1-\kappa)$

$$\frac{\partial\theta^{\mathrm{src}}}{\partial\theta^{\mathrm{img}}} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} = (1-\kappa)\begin{pmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{pmatrix}$$

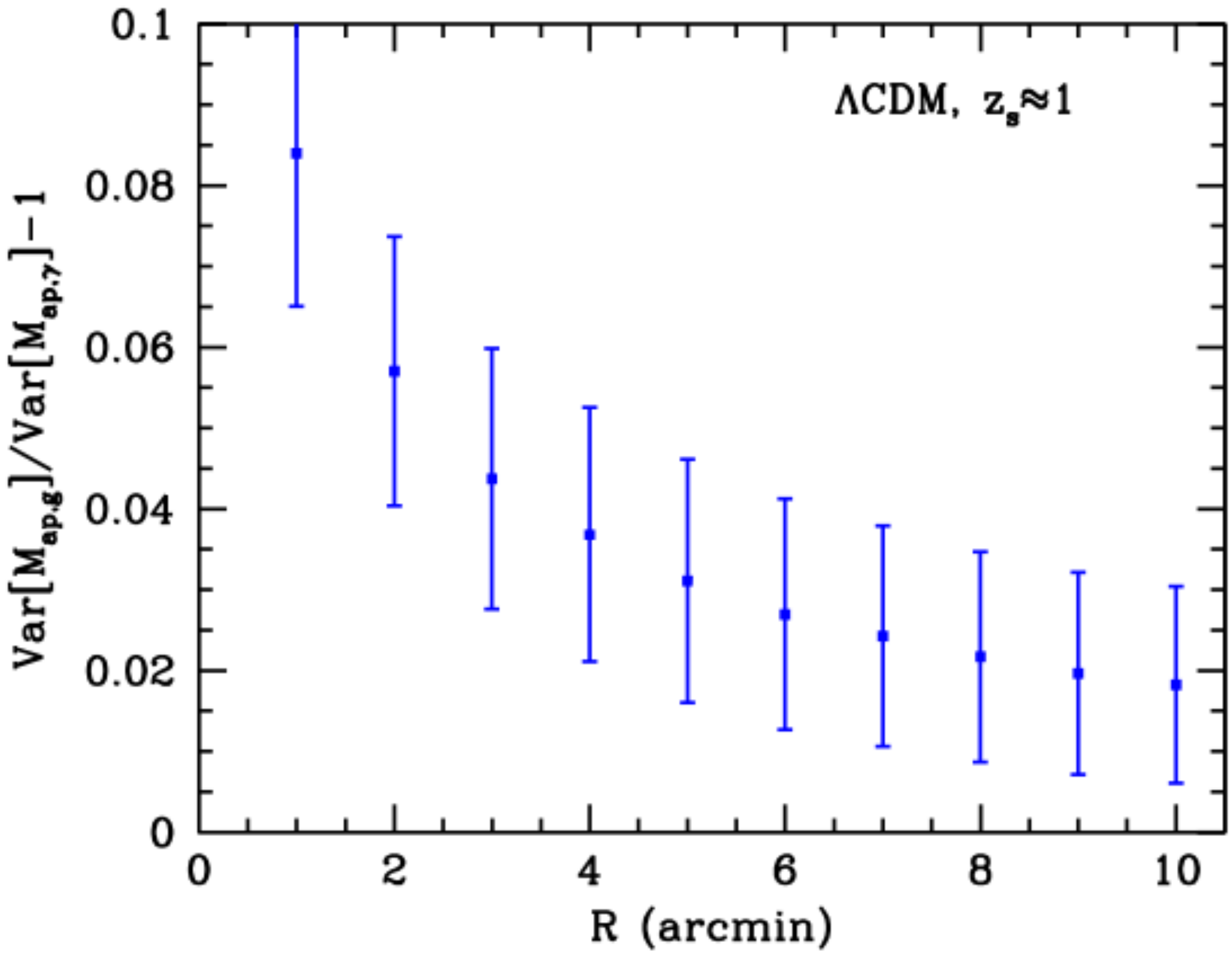
- Since γ and κ are usually small this difference is often neglected (except around clusters).
- Can be a few percent effect on arcminute scales!

Reducing shear enhances shear

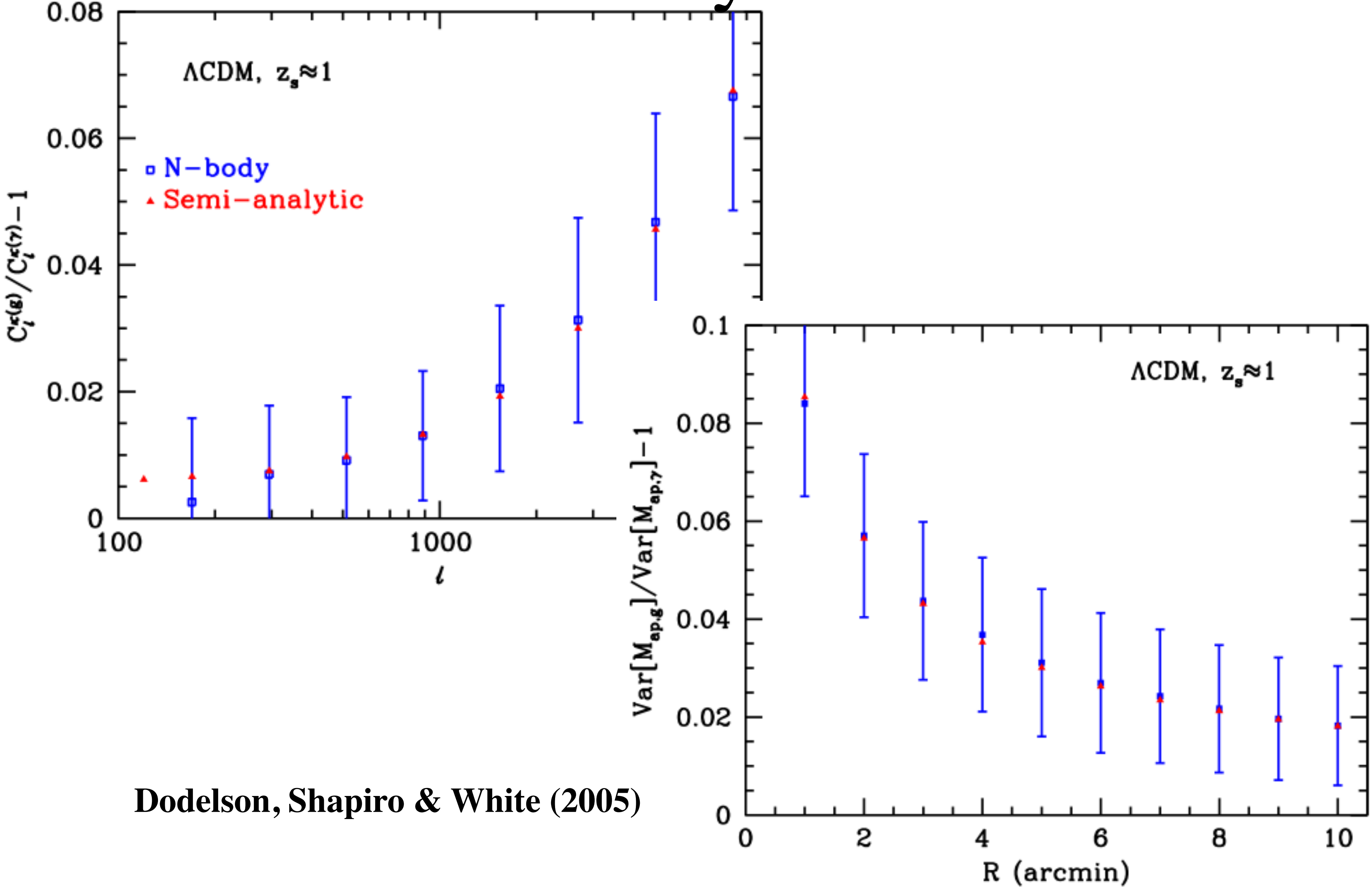
- On small scales κ can be quite large, and spatial smoothing does not commute with the "reducing" operation.
- Generally g has larger fluctuations than γ because κ is skew positive.
  - Excess small-scale power compared to naïve predictions.
- The effect is different for different estimators
  - A signal of "reduced shear" vs. e.g. intrinsic alignments or systematics.
- The effect is non-linear
  - Provides cross-check on shear calibration

Reduced shear

We don't measure the shear, γ, but the reduced shear $g=\gamma/(1-\kappa)$

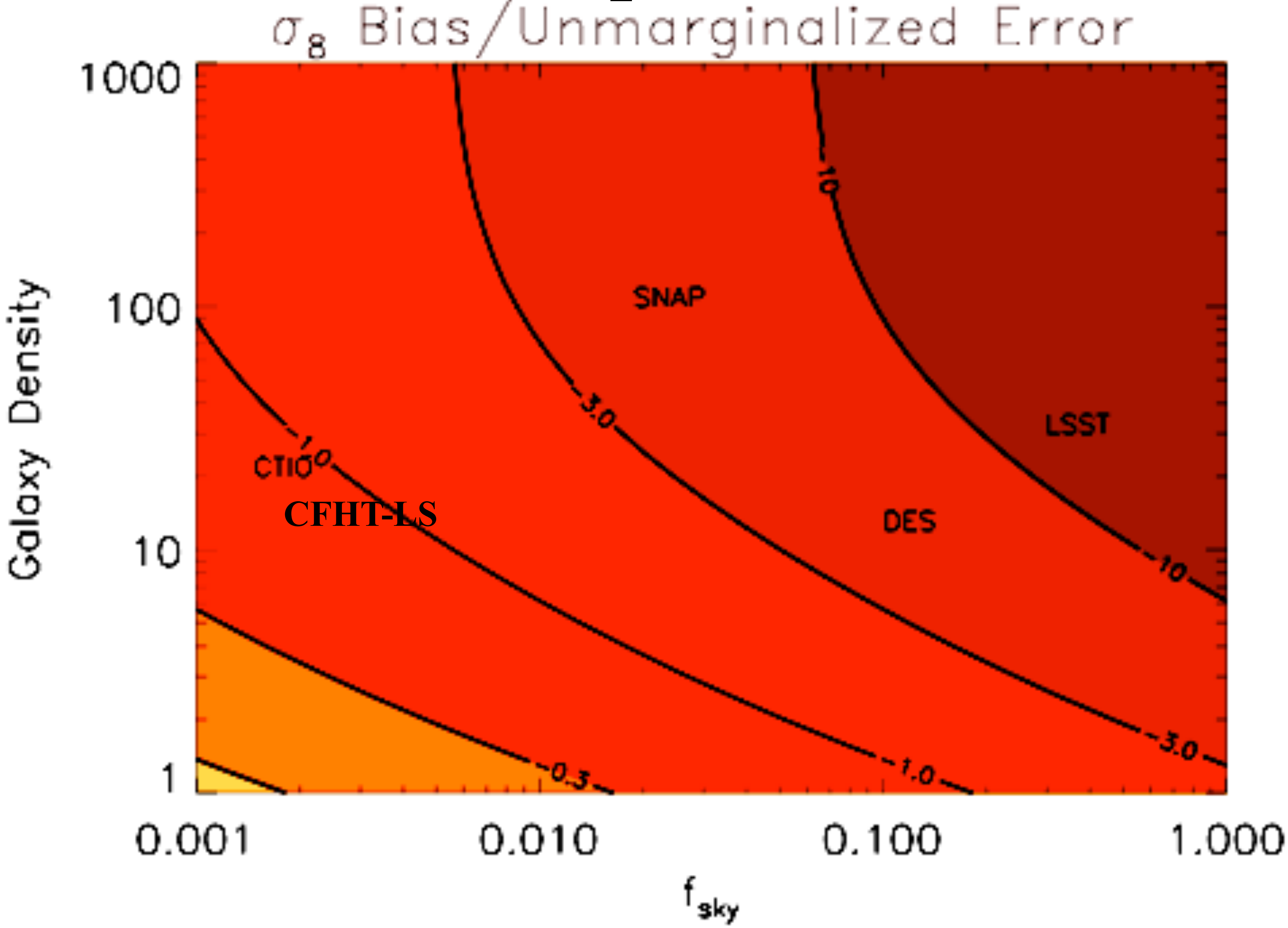


A semi-analytic model

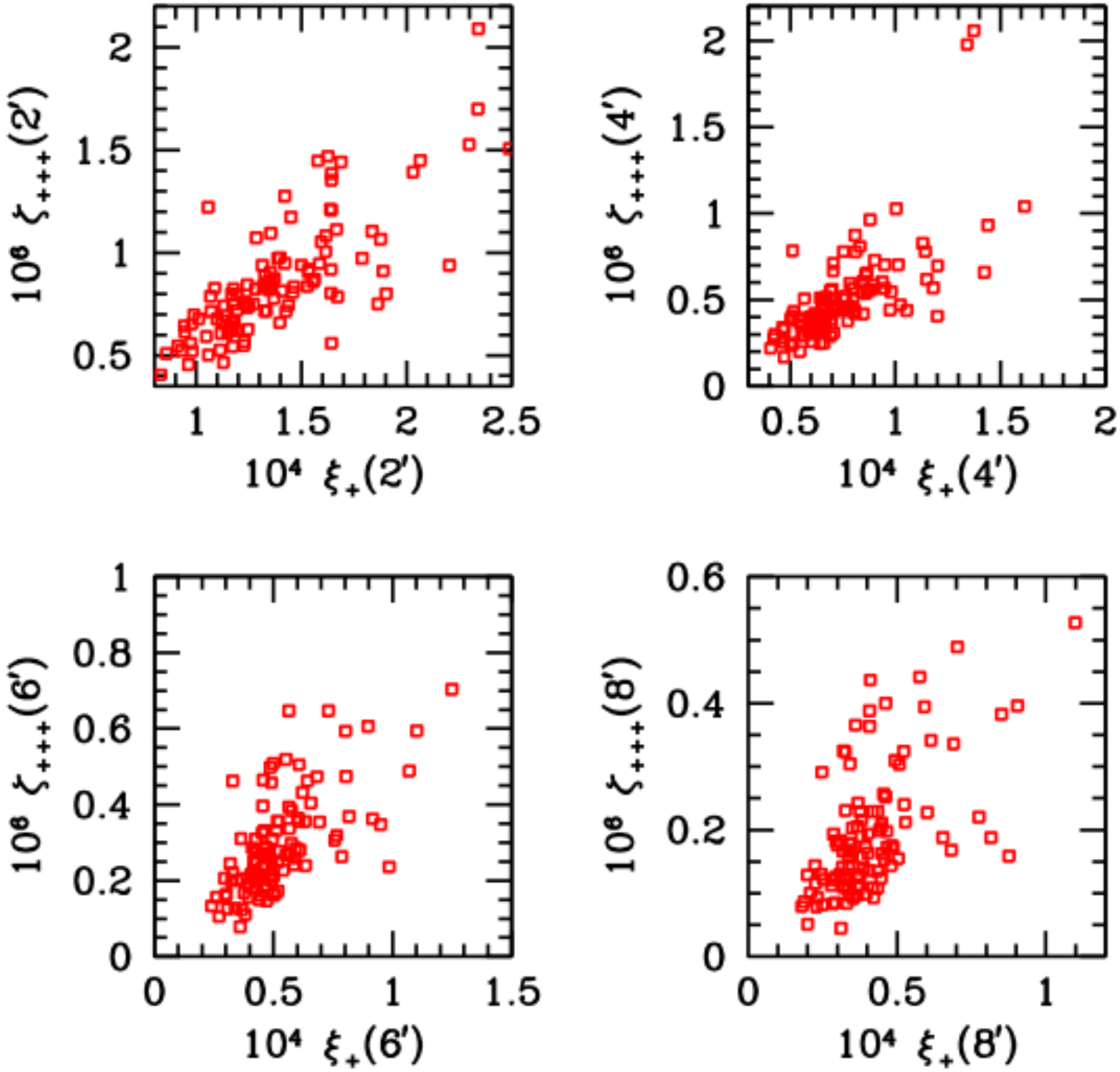


Dodelson, Shapiro & White (2005)

Bias in parameters



Correlations in clustering

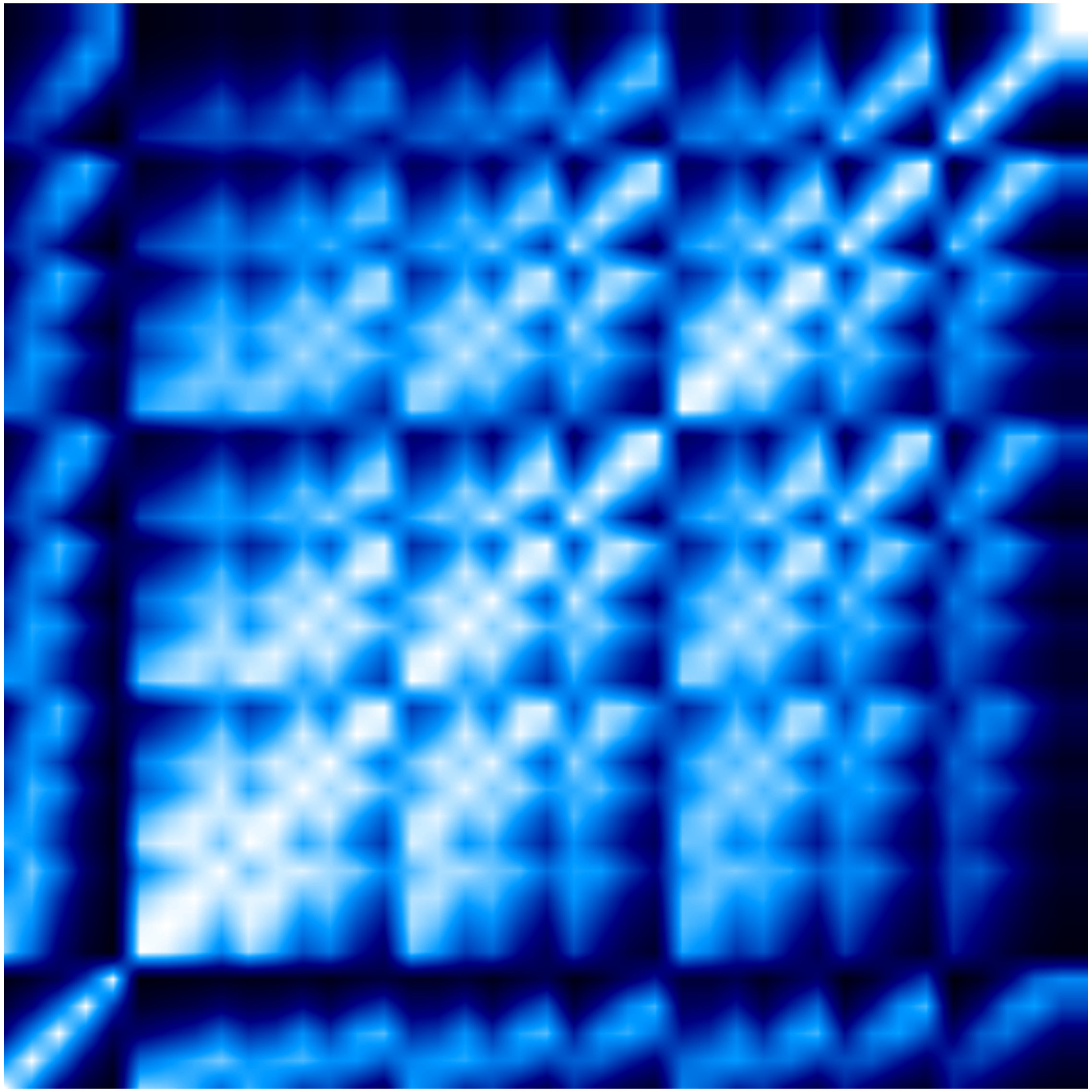


Find that the 2-point and 3-point functions are highly correlated on small scales.

This is not too surprising when thought of from an "object" perspective but is not often assumed.

Correlations contd.

- Correlation matrix for 2nd and 3rd order M_{ap} statistics (computed from κ maps).
- Uses Mexican hat filter with scales 1, 2, 4, 8 & 16 arcmin (40 measures: 5x 2-pt and 35x 3-pt).



Beyond gravity

- Non-gravitational physics becomes important on small scales, becoming dominant beyond l~3000.
  - White (2005), Zhan & Knox (2005)
- Dramatic progress in modeling extra physics!
  - Expect small # of simulations including relevant physics will be available within 5-10 years.
  - Can mock up some of the physics in gravity-only simulations
    - Put gas in hydrostatic equilibrium with known DM potential.
    - Apply adiabatic contraction to halos where gas would have cooled.
- Use photo-z to apply "nulling tomography".
  - Huterer & White (2005)

Beyond the 2-point function

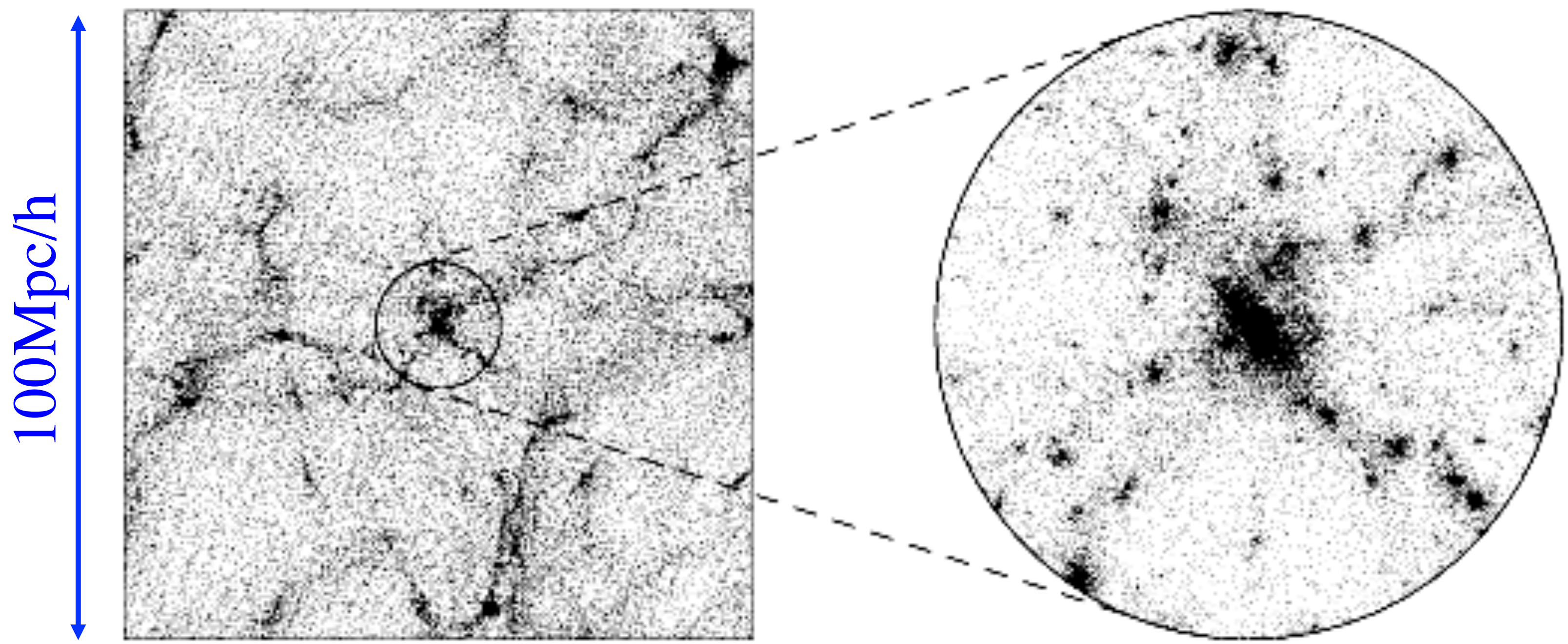
Non-gaussianity as blessing or curse?

Finding clusters with weak lensing

- Lensing maps are obviously non-Gaussian.
  - You can point to structures in the maps
- Higher order functions or peak statistics contain additional information beyond the 2-pnt function.
- The obvious extension of non-Gaussian thinking is to look at the extrema of the maps
  - Finding clusters.
- Unfortunately lensing measures the *projected* mass along the line of sight.
- Projection effects can be severe and need to be modeled.

Projection effects lead to scatter in the shear-mass relation

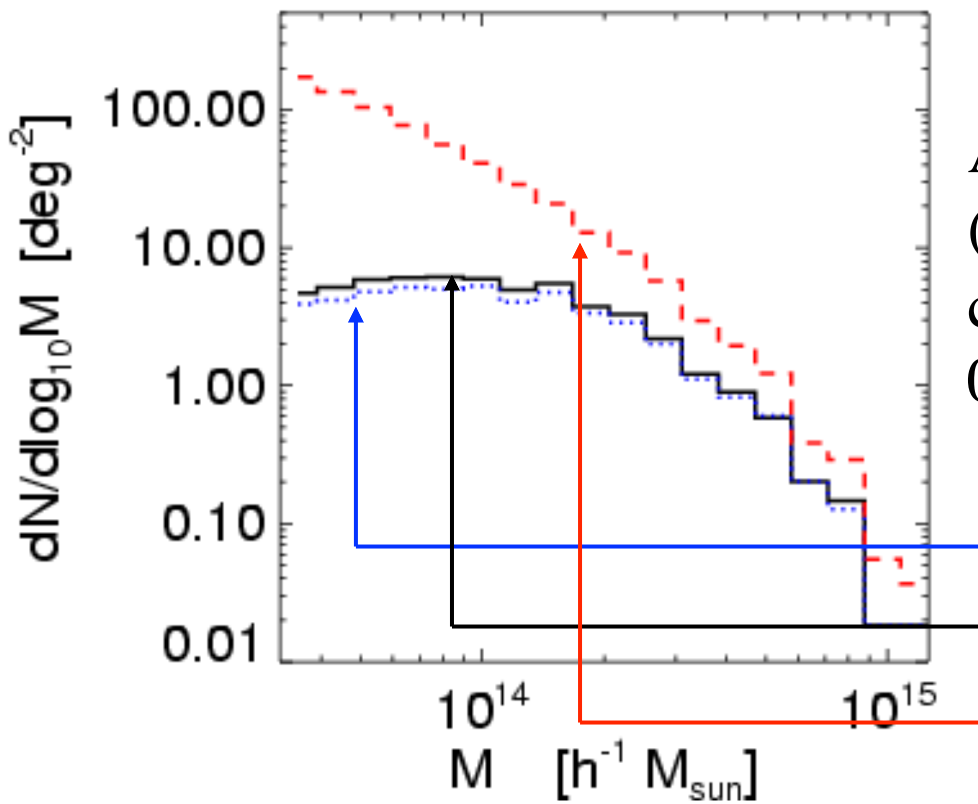
Scatter in the shear-mass relation means lensing does *not* produce a mass selected sample, but a shear selected sample! This has implications for doing cosmology.



Metzler, White & Loken

Projection effects can be severe

Hennawi & Spergel have used a tomographic matched filter algorithm to find clusters and determine their redshifts. The tomographic information helps reduce projection effects, but cannot eliminate them entirely.



Assuming 60% efficiency (3.5σ) and searching for clusters in the range 0.2<z<0.8

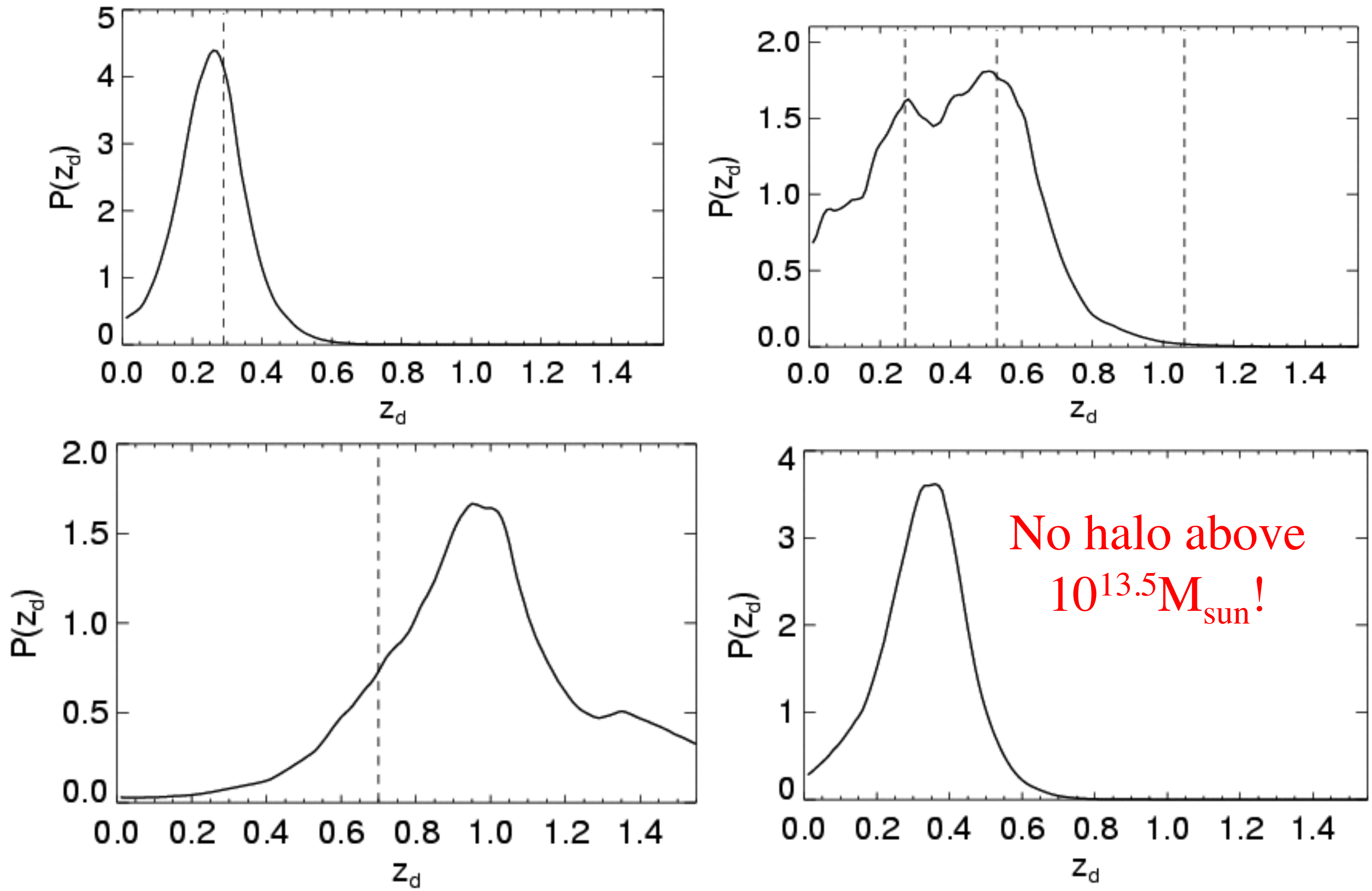
No tomography

Tomography

Truth

Tomographic (MF) redshifts

(Hennawi and Spergel 2004)



What does lensing add?

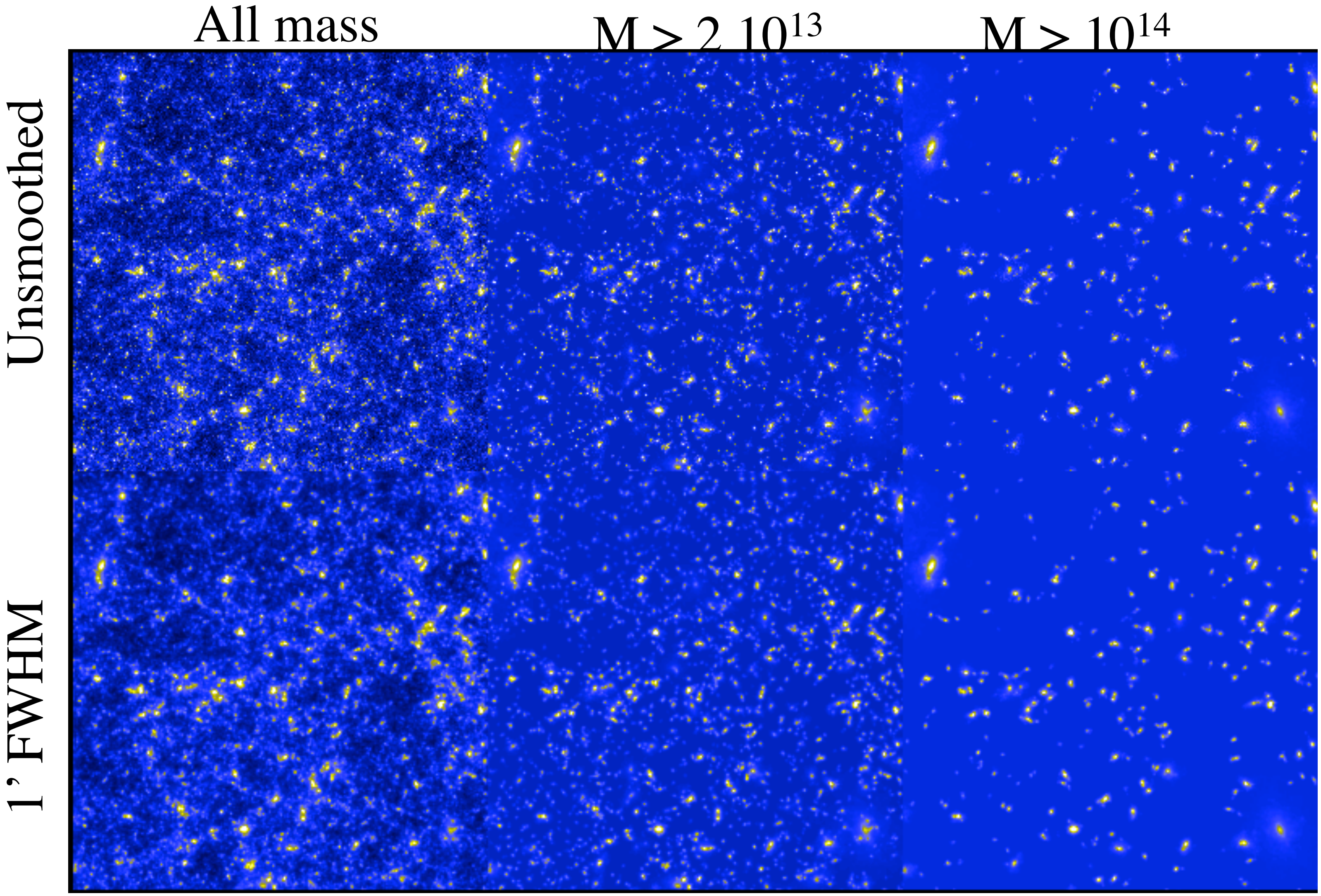
- While lensing offers an advantage in being independent of the dynamical state of the cluster or the luminosity of the material, it suffers a huge disadvantage due to projection.
  - In principle lensing can be accurately modeled, and measurements compared to predictions.
- Lensing may not be an efficient way to find clusters.
  - But perhaps it doesn't need to be!
- Clusters can probably be found by other means, e.g. as galaxy overdensities in deep optical images.

What does lensing add to the 3D galaxy distribution?

A calibration sample

- It has been emphasized (Majumdar & Mohr) that even a small sample of clusters where the observable-M relation has been calibrated would dramatically improve cluster counting constraints on cosmological parameters (e.g. DE).
- Can lensing be used to calibrate the *O*-M relation?
- Suppose I had a cluster catalog with a richness estimate for each halo *and* a weak lensing map of the same region.
  - Can I use the galaxy data to remove the line-of-sight projection on a halo-by-halo basis?
  - If not, can I rely on simulations to *calibrate* the shear-mass relation (bias and scatter)?

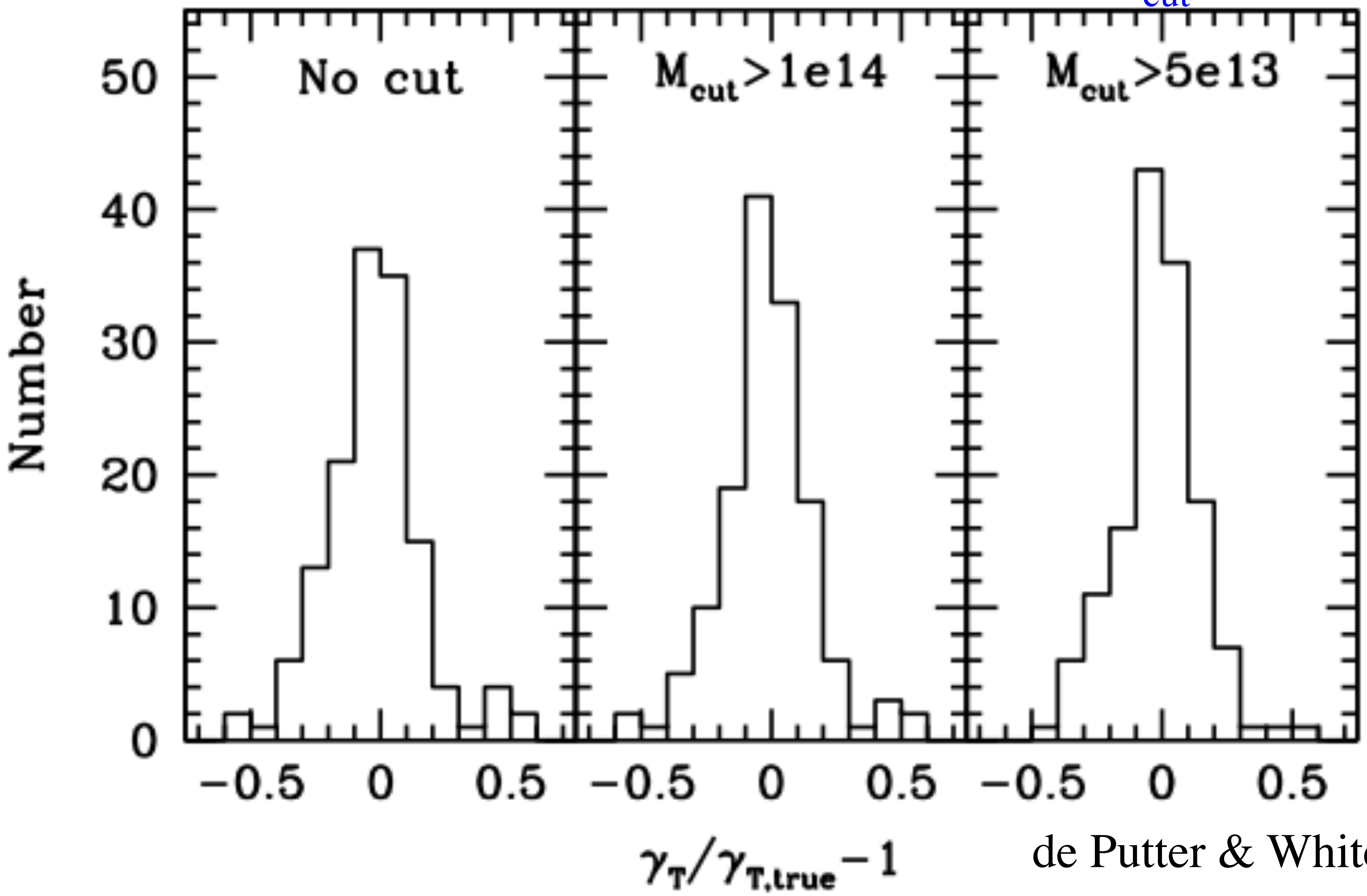
Halos and lensing



Modeling the line-of-sight

Around efficient lensing halos with $M > 3 \times 10^{14}\ h^{-1} M_{sun}$

subtract the effect of halos with $M > M_{cut}$

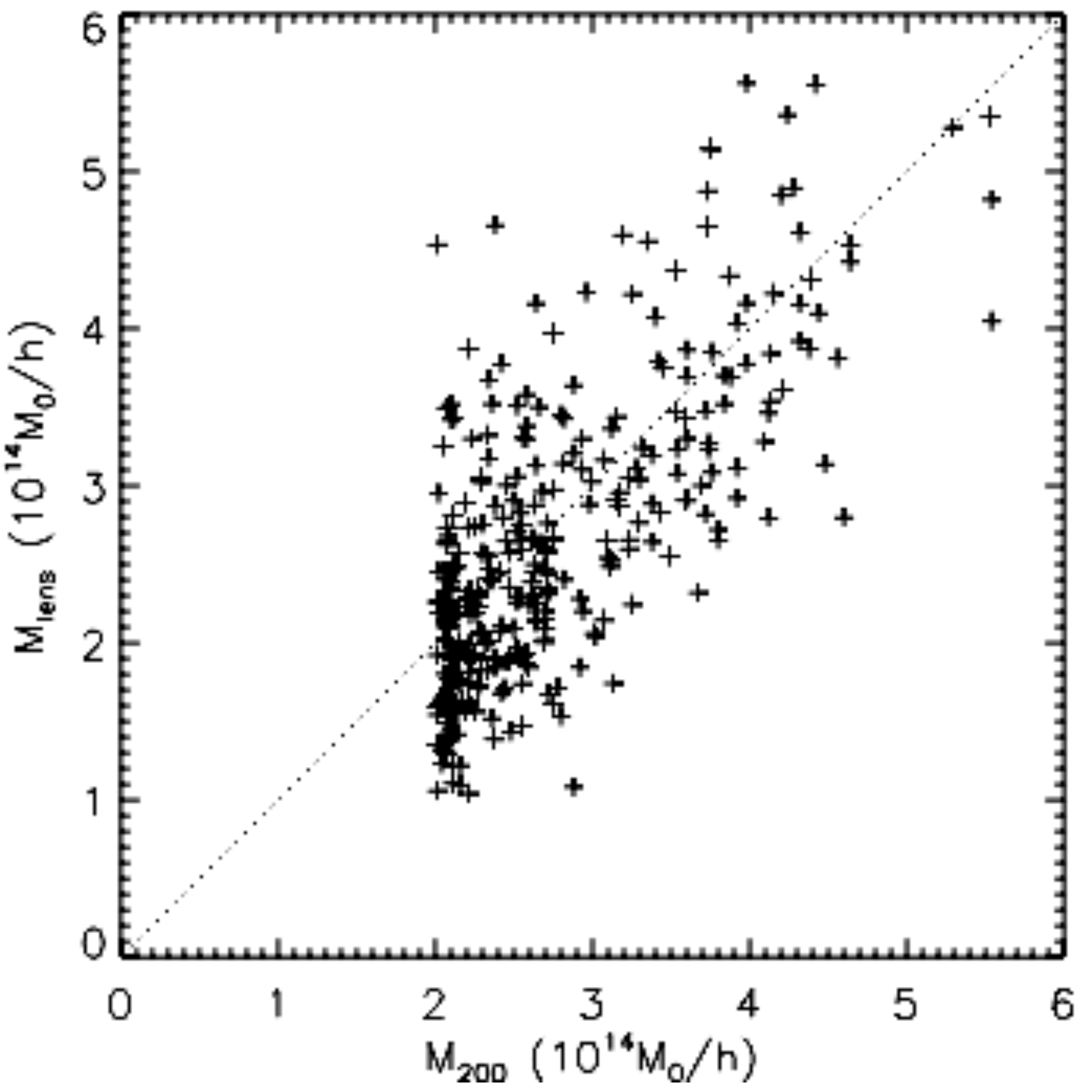


de Putter & White (2005)

Mass function too steep

- Convergence is very slow, because of the nature of the halo mass function.
- It appears that (unless one can go to extremely low mass halos) halo-by-halo mass estimation won't work.
- We need to use a statistical method, where we calibrate an observable-M relation using a large sample of clusters.

Scatter in mass estimator

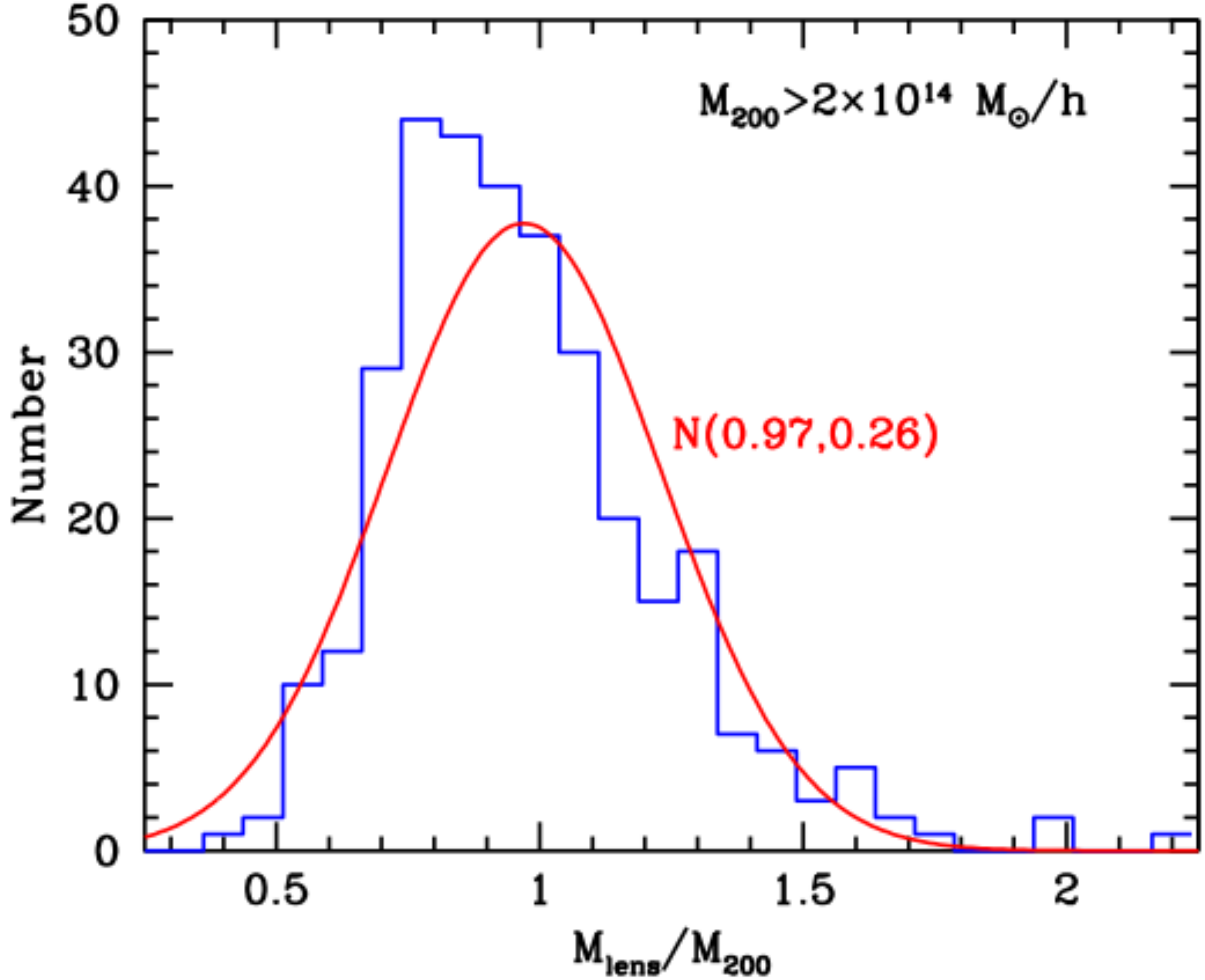


Correction for line-of-sight contamination by fitting a projected NFW profile to κ and computing M_{200} from the fit.

Bias is a few percent.

Scatter is ~25%.

Distribution of errors



Calibration vs testing

- Used simulations to *test* line-of-sight correction procedure.
- Can I rely on simulations to *calibrate* the shear-mass relation (bias and scatter)?
  - Should be tractable in principle, but never demonstrated in even approximately realistic conditions.
  - If I can trust the simulations “perfectly”, what do I gain by using peaks rather than the whole map?
  - Is this easier than calibrating the richness-mass relation given advances in theoretical understanding of galaxies (halo models)?

Lensing of the CMB

Of course galaxies aren't the only source of (lensed) light in the universe. Any screen will do. The CMB is the furthest screen!

Large-scale structure will lens the CMB anisotropy.

Since we don't know the "shape" of the CMB *a priori* we need to use more statistical information.

Seljak (1996)

Zaldarriaga & Seljak (1999)

Zaldarriaga (2000)

Seljak & Zaldarriaga (2000)

Hu (2001)

Hu & Okamoto (2002)

Okamoto & Hu (2002, 2003)

Cooray & Kesden (2003)

Hirata & Seljak (2003)

Amblard, Vale & MW (2004)

Lensing of the CMB (contd)

Consider the CMB, lensed

$$T(\hat{n}) = \widetilde{T}(\hat{n} - \nabla\Phi) \simeq \widetilde{T}(\hat{n}) - \nabla\Phi \cdot \nabla\widetilde{T}$$

The correlation function will depend on Φ, allowing us to make a quadratic estimator assuming everything is Gaussian and the deflection angle is small.

$$\hat{\kappa} \sim \nabla \cdot \widehat{\nabla\Phi} \sim \nabla \cdot [T_w \nabla T_w]$$

(Hu; Hirata & Seljak)

We should be able to detect this effect with upcoming experiments (e.g. APEX-SZ, SPT, ACT)!

But how well do these estimators work, and how sensitive are they to observational strategy, foregrounds and systematics?

Lensing of CMB by LSS

- Worry about violations of assumptions:
  - Potential field is non-Gaussian.
  - Deflection angle is not infinitesimal.
- Since estimator is looking for small amounts of (lensing induced) non-Gaussianity on top of the Gaussian CMB, worry about the effect of foregrounds,
  - IR sources
  - kSZ and
  - O-V at hi-z
  - etc..

Numerical study

Want to simulate some of these issues and investigate whether APEX-SZ, SPT and ACT could see lensing.

- Simulations include
  - Primary CMB map
  - Gaussian and non-Gaussian lensing fields
  - Idealized detector noise
  - Kinetic SZ signal (optional)
- Primary configuration
  - 30x30 degrees
  - 0.8' FWHM resolution
  - 2μK/arcmin (white) noise or more
- Make lensed maps, apply quadratic estimator, apply corrections …

Assessing the results

It is non-trivial to assess the numerical results. We use visual inspection and two power spectra as our metrics:

- Cross-spectrum $\left\langle \kappa^{\rm true} \kappa^{\rm est} \right\rangle$
  - Not measurable
  - Looks for bias in the method or a mis-estimated normalization for κ
- Auto-spectrum $\left\langle \kappa^{\rm est} \kappa^{\rm est} \right\rangle - \mathcal{N}$
  - Measurable
  - Can be both multiplicatively and additively biased due to misestimates of the noise terms.

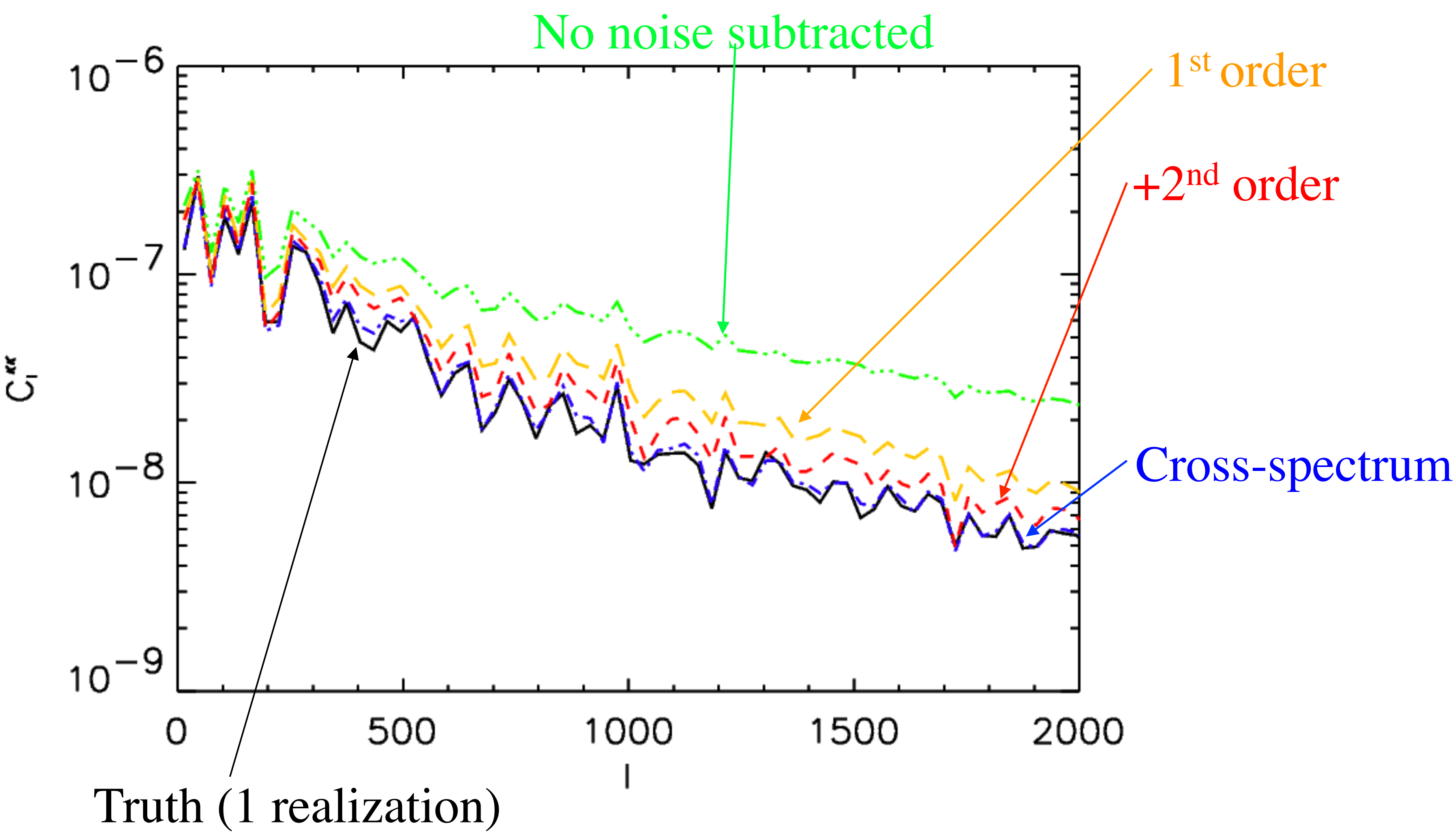
Biases

- Even in the absence of foregrounds we find that the quadratic estimator is both multiplicatively and additively biased.
- The bias depends on the level of signal.
- The additive bias comes from
  - Higher order terms in the noise
  - Non-Gaussianity in the lensing field.
- The multiplicative bias is due to non-linearity

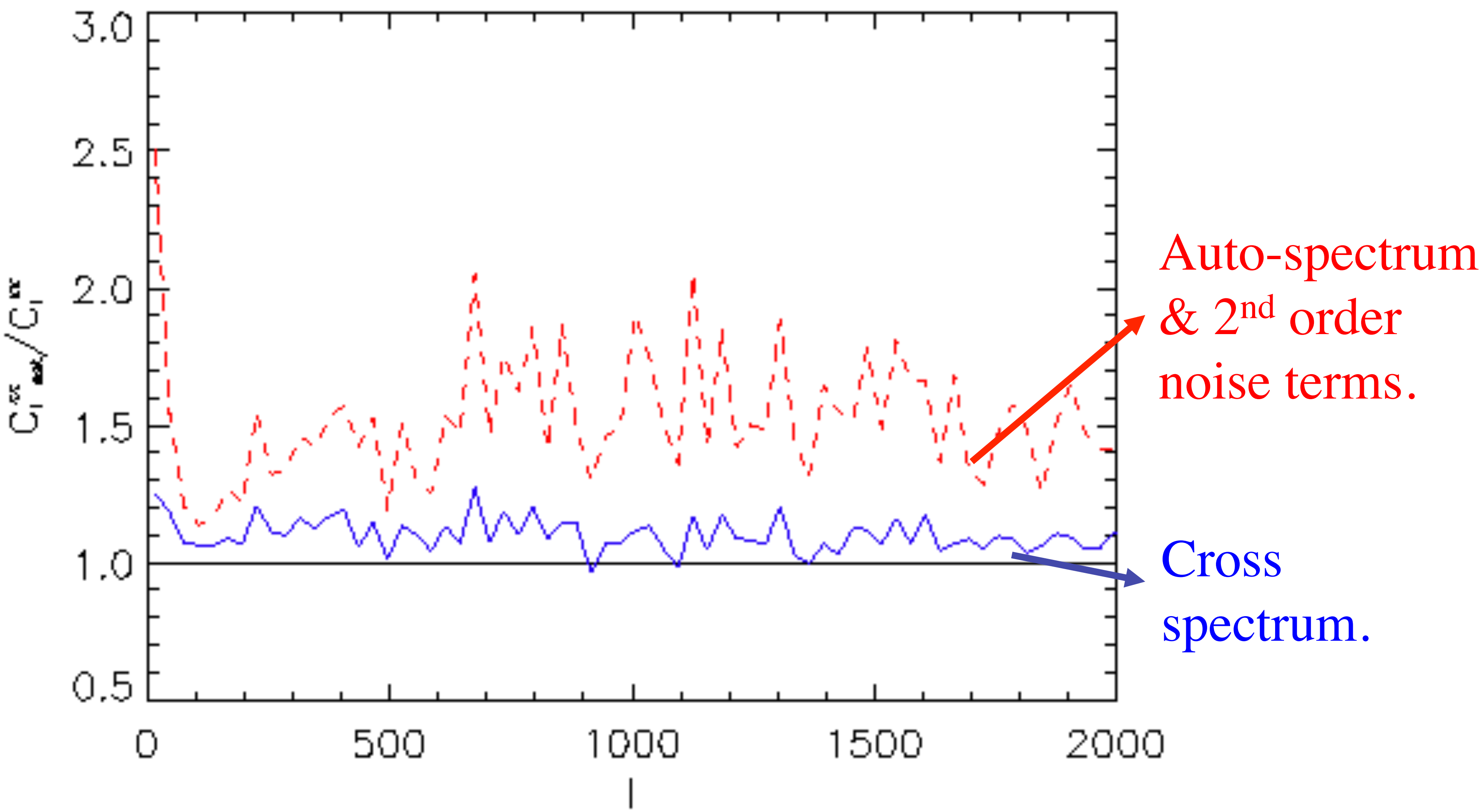
Quoted errors are $\mathcal{O}(1\%)$!

Assumption	C_l Error
Quad Est	70%
2nd order	25%
NG	20-30%

Additive bias for Gaussian maps



Adding non-Gaussianity



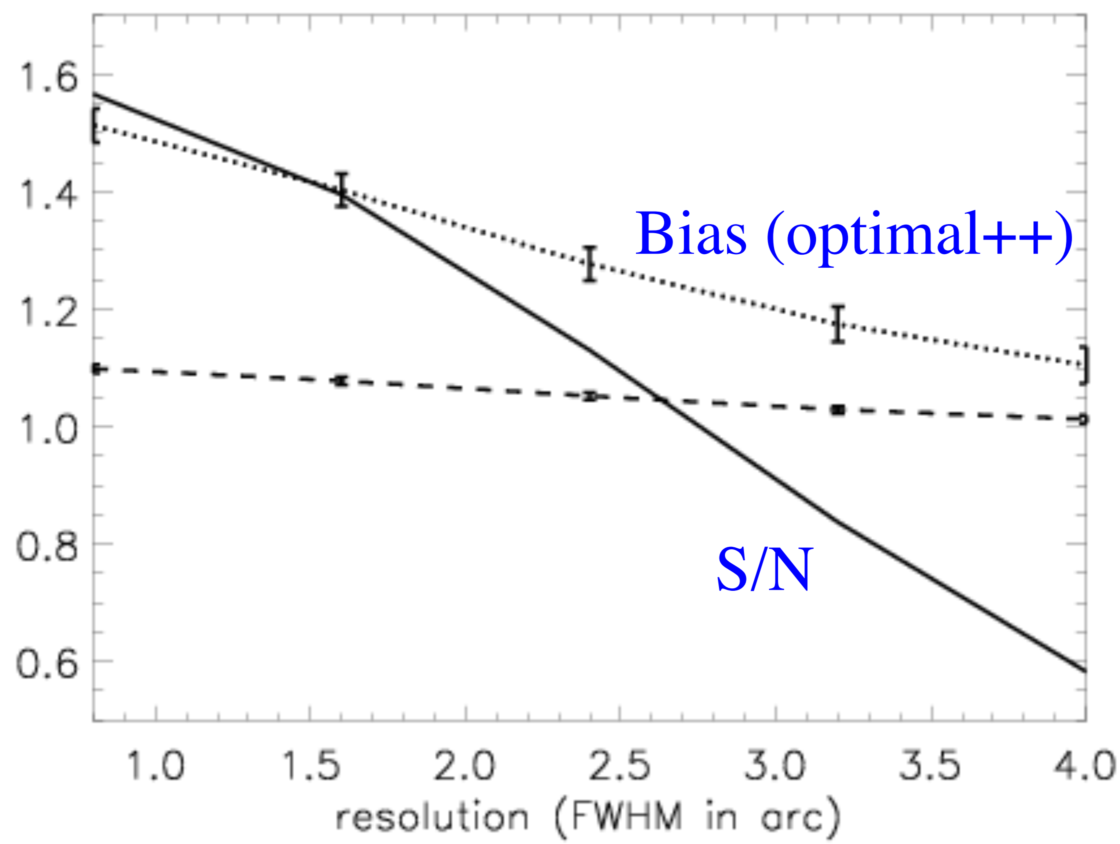
Now have non-vanishing 3-point function, etc., so Gaussian estimate for normalization is insufficient.

Effects of resolution

Working at higher resolution increases the S/N.

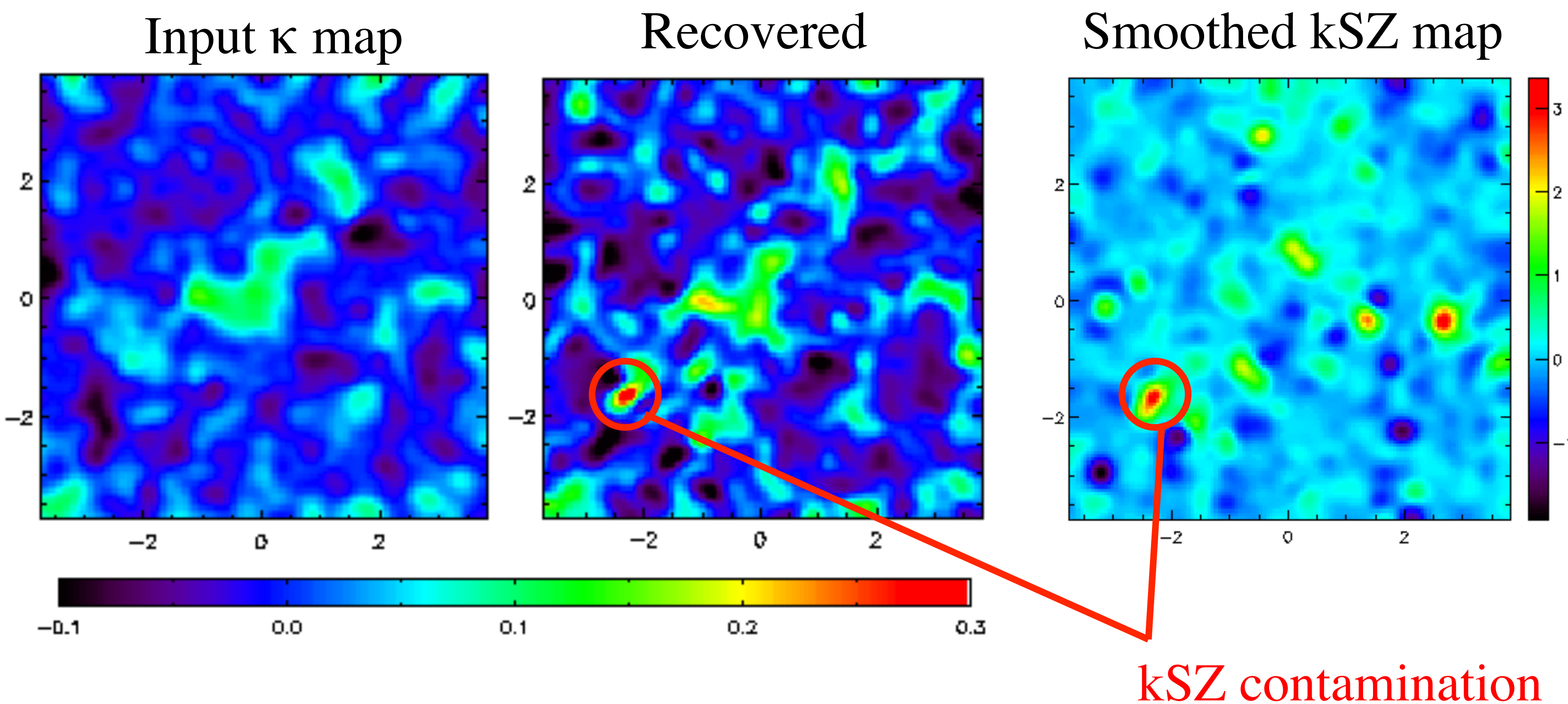
Higher resolution brings more signal, but additional theory uncertainty:

- LSS is non-Gaussian
- Deflection angle not small



Similar effects for polarization (Amblard, in prep)

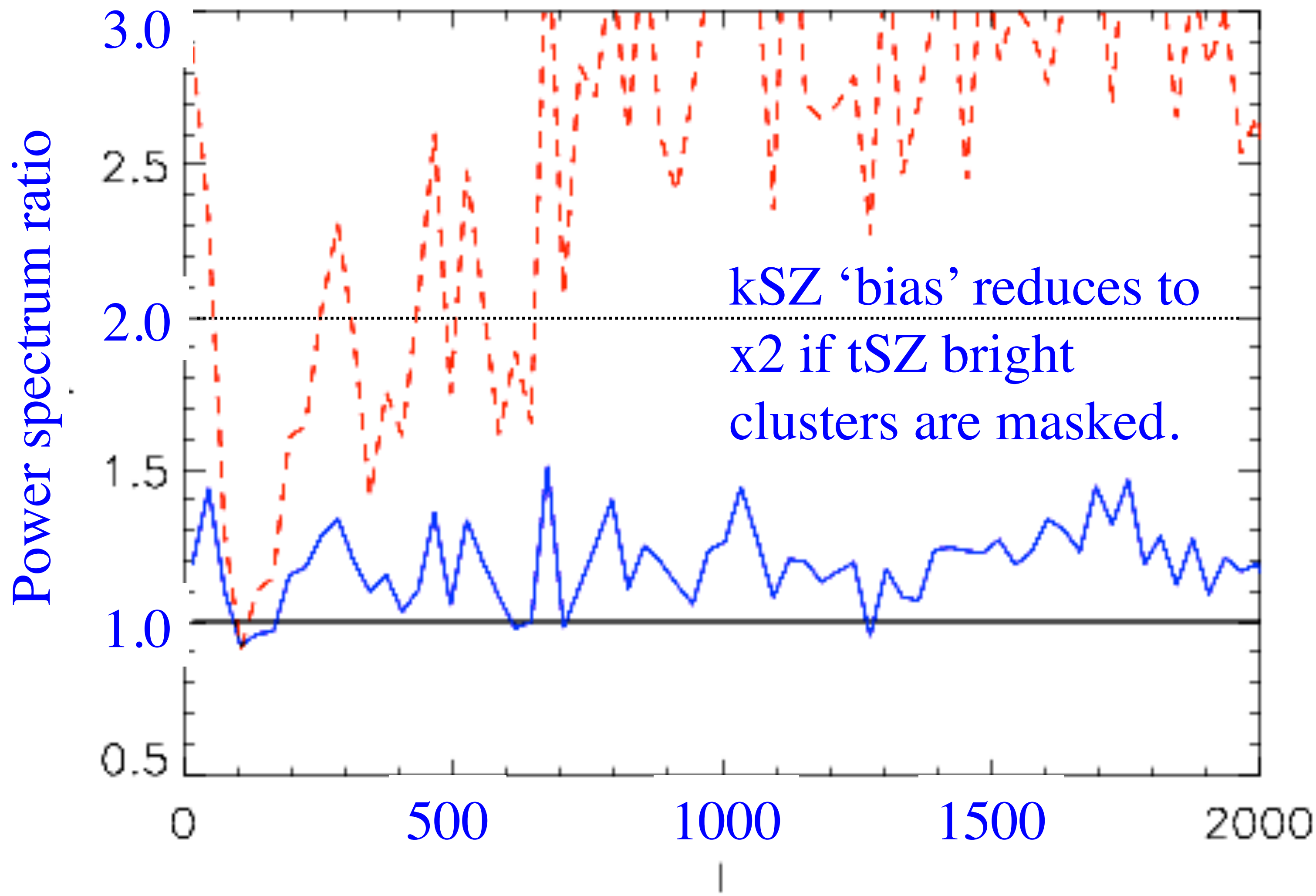
Adding foregrounds: kSZ



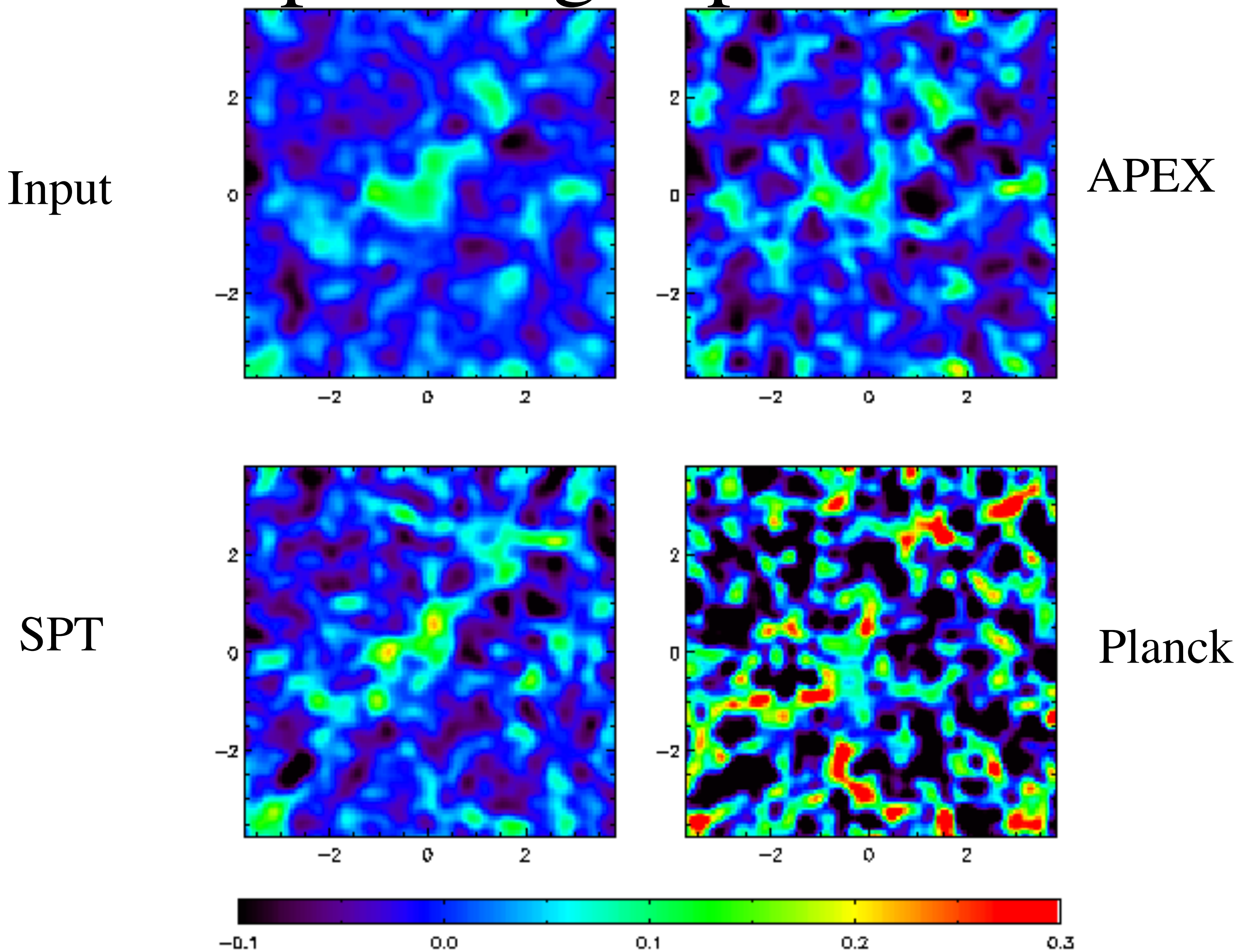
Maps have been smoothed from 0.8’ to 20’ to enhance the S/N.

kSZ and power spectrum

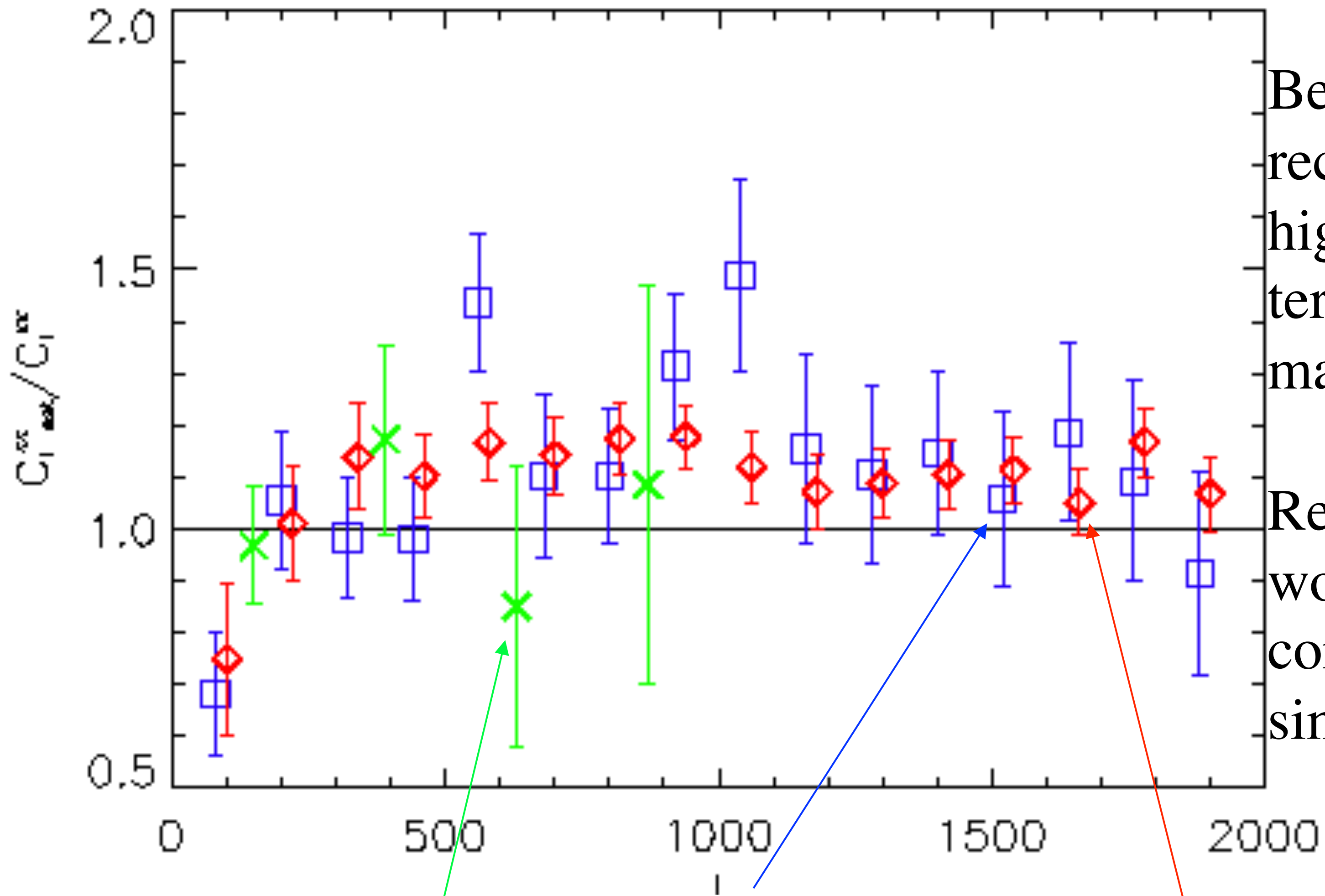
Foregrounds?!



Upcoming experiments



Power spectra



Best reconstruction: higher order terms and kSZ masked.

Remaining bias would need to be corrected by simulations.

Planck/10

APEX

SPT

Conclusions

- Cosmic shear has come of age!
- We can accurately simulate lensing fields on scales of arcminutes to degrees.
- *Non-Gaussianity offers rich opportunities and difficult challenges.*
- We may soon detect gravitational lensing of the cosmic microwave background.

The End