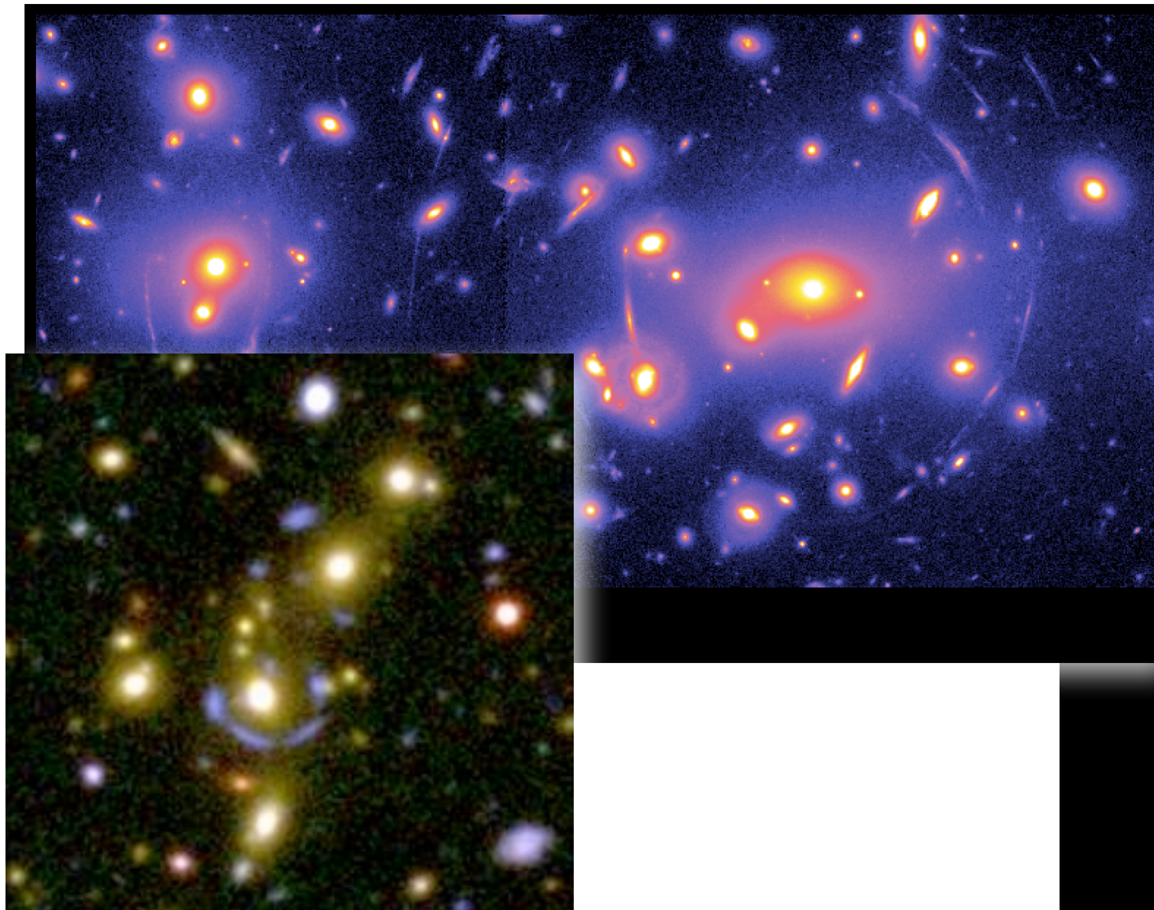
Weak gravitational lensing:
Higher-order statistics & new statistical approaches.

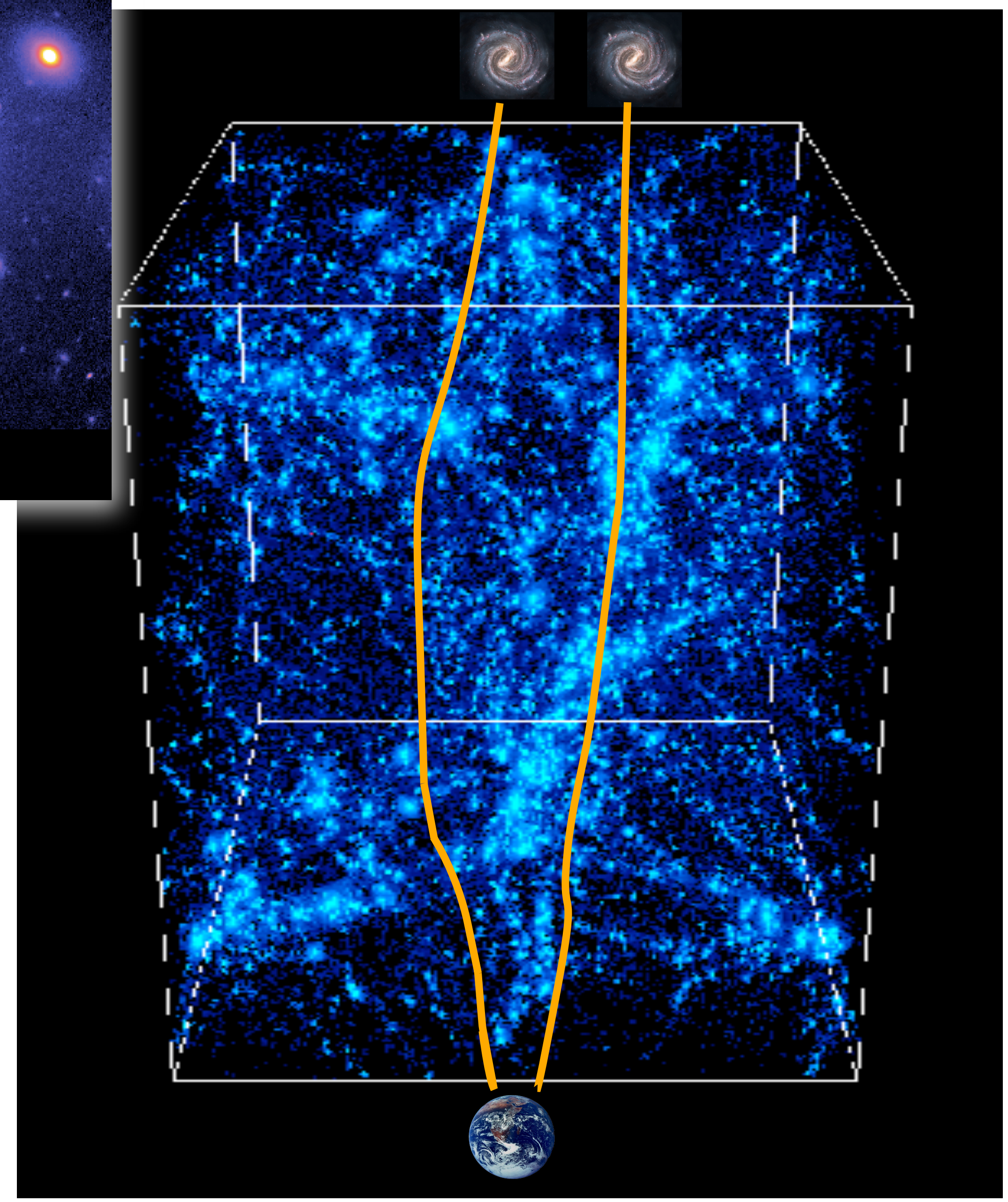
Martin Kilbinger
CEA Saclay-Irfu-SAp/AIM, IAP

Seminar, OCA, 22/02/2018

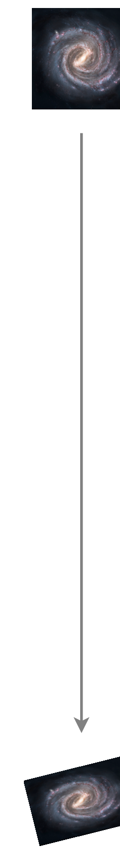
With Liping Fu (Shanghai NU) and the CFHTLenS collaboration;
Chieh-An Lin, Sandrine Pires, Austin Peel (CEA)

Weak cosmological lensing





$z_s \sim 1$



Lensing by LSS:
~ 3% distortion
$\kappa, \gamma \approx 0.03$

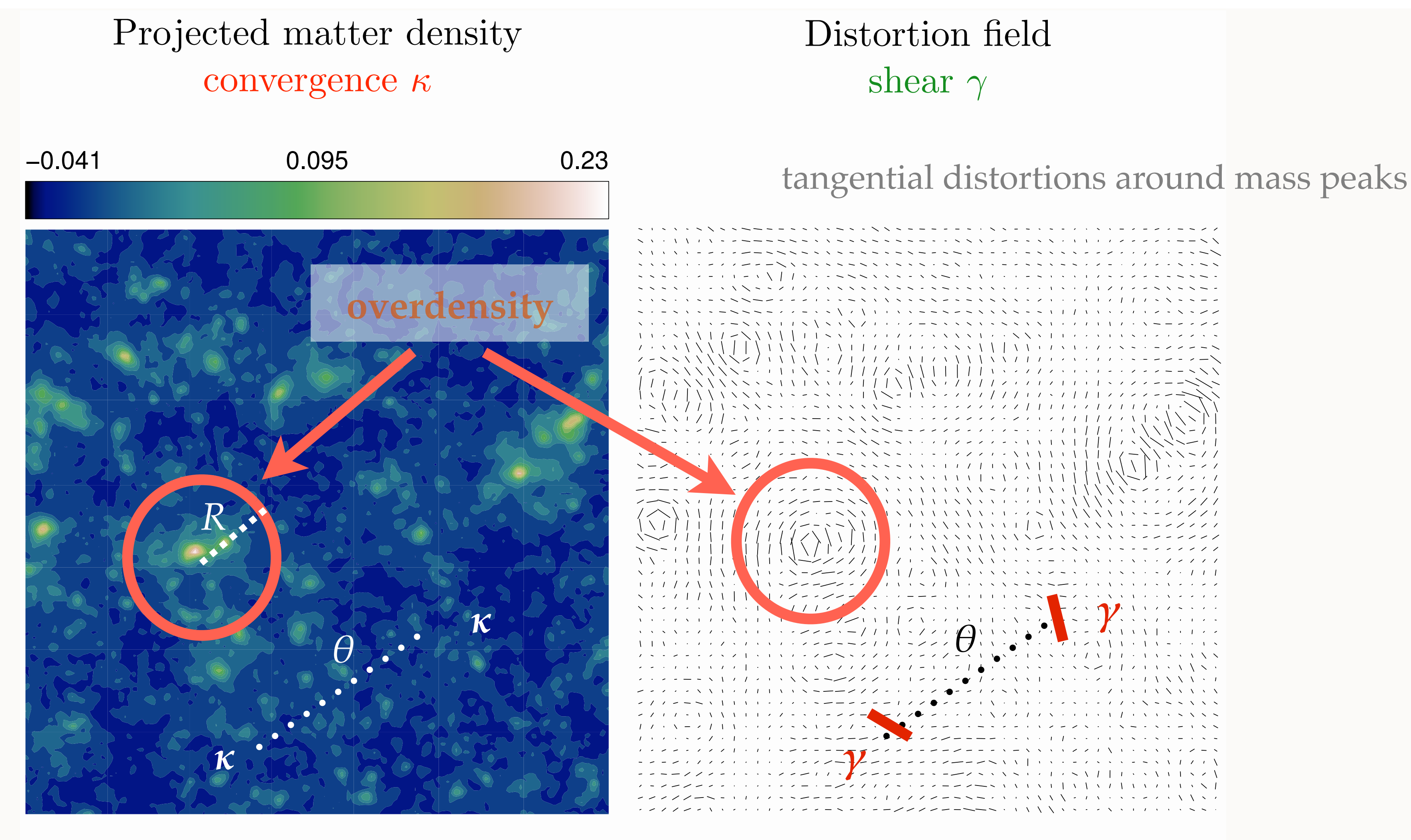
"Weak lensing"

- Sensitive to total (baryonic + dark) matter
  No need to assume relation (bias) between galaxies & DM
- Low (z~0.1 … 1) redshifts
  Epoch of acceleration
- Probes geometry & structure
  Modified gravity

WL: Gaussian statistics

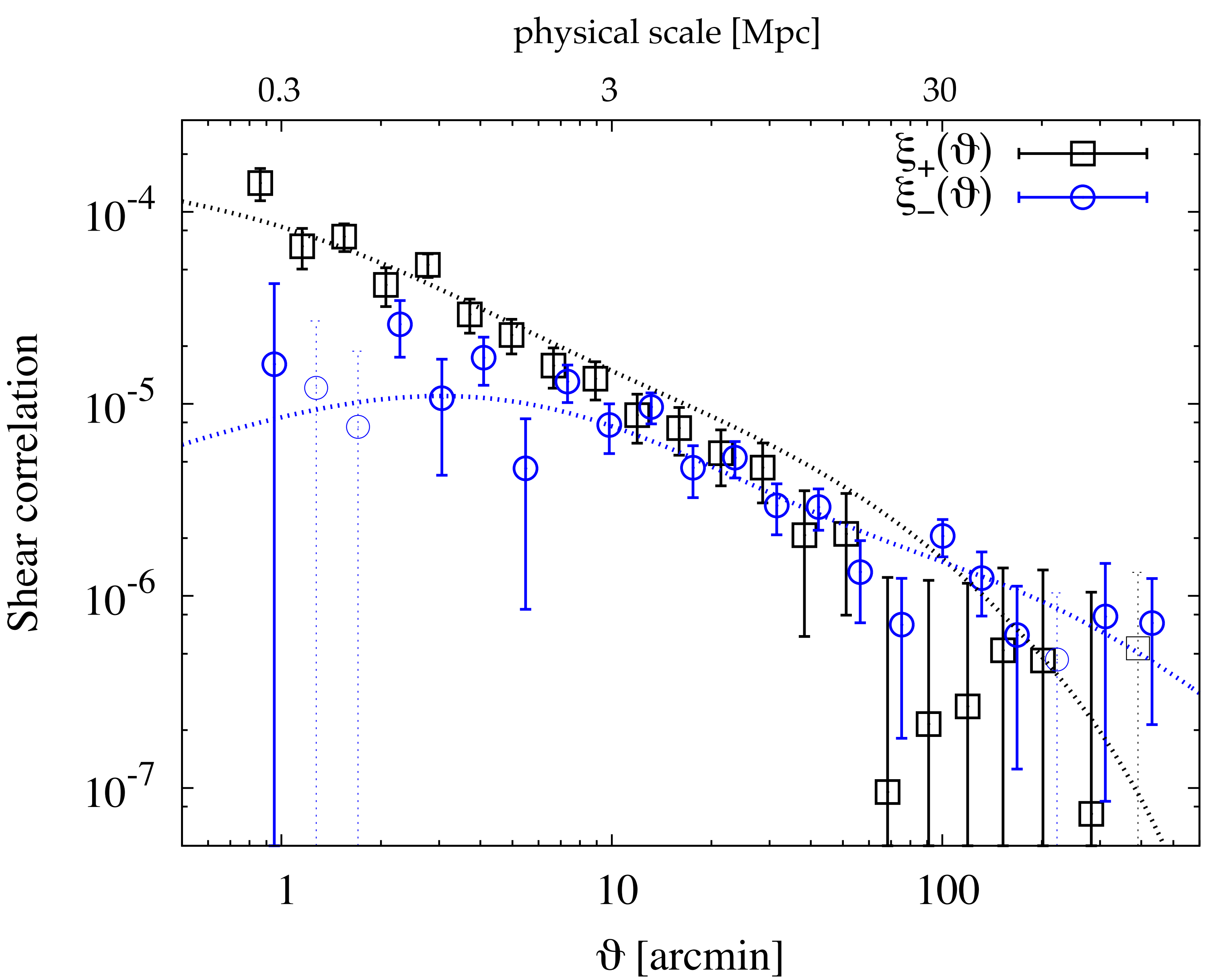
- Probes small scales, down to sub-Mpc at late time: **non-Gaussian structures**.
- However so far mainly used 2-point correlation function or other 2nd-order stats (functions of on power spectrum)

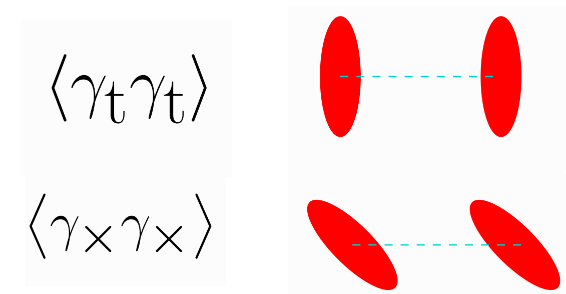
variance
smoothed at *R*:
$\sigma^2(\kappa) = \sigma^2(\gamma)$



Source galaxies at $z = 1$, ray-tracing simulations by T. Hamana

2-point shear correlation function





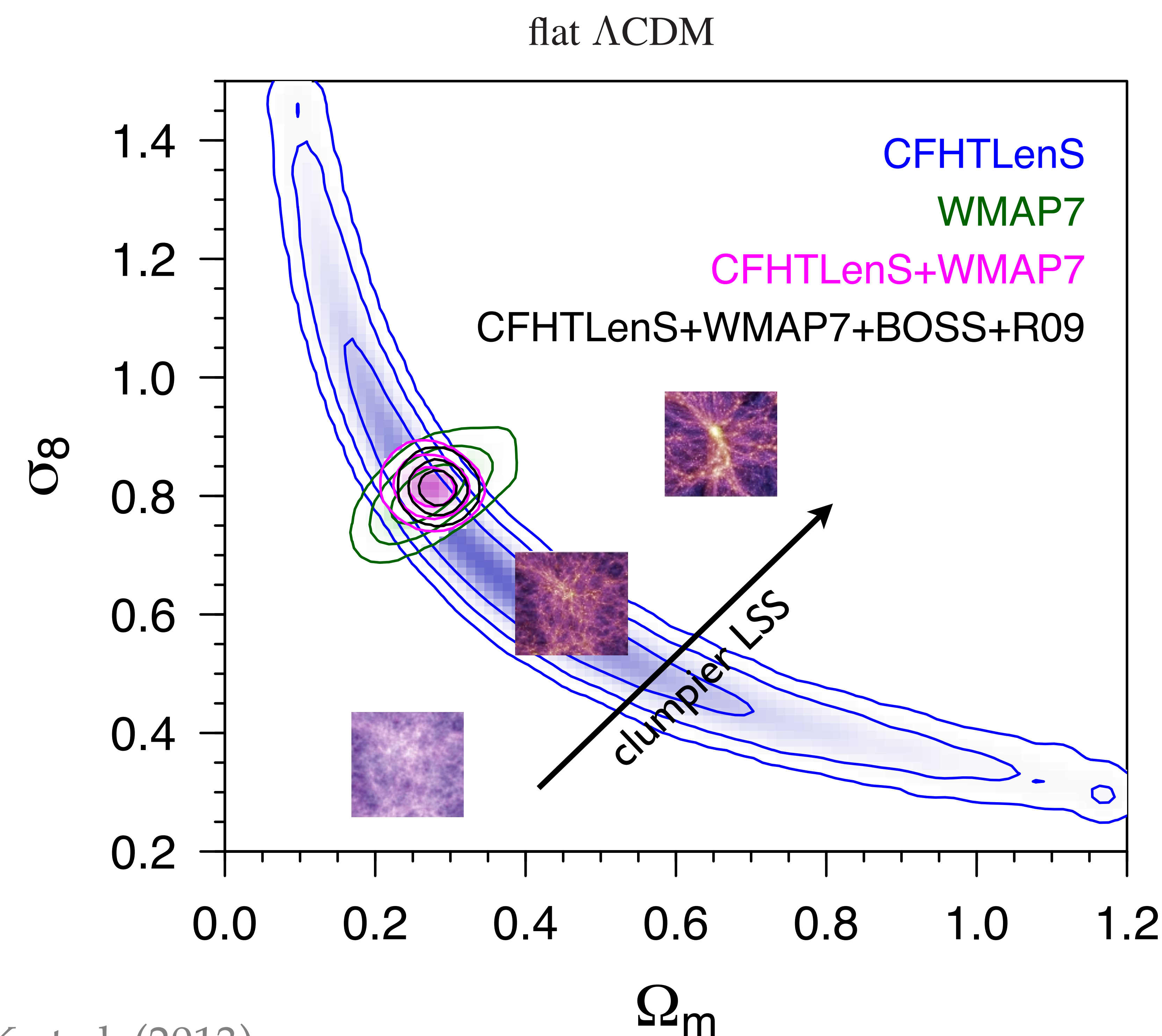
$$\xi_+(\vartheta) = \langle\gamma_t\gamma_t\rangle(\vartheta) + \langle\gamma_\times\gamma_\times\rangle(\vartheta)$$
$$\xi_-(\vartheta) = \langle\gamma_t\gamma_t\rangle(\vartheta) - \langle\gamma_\times\gamma_\times\rangle(\vartheta)$$

CFHTLenS, MK et al. (2013)

Prediction
(WMAP7 cosmology):
Flat Universe
$\Omega_m = 0.27$
$\sigma_8 = 0.8$
$w = -1$
Smith et al. (2003)
non-linear power spectrum

Cosmic shear: state of the art 2013

CFHTLenS



Normalisation consistent with WMAP, but lower than Planck!

Tension?

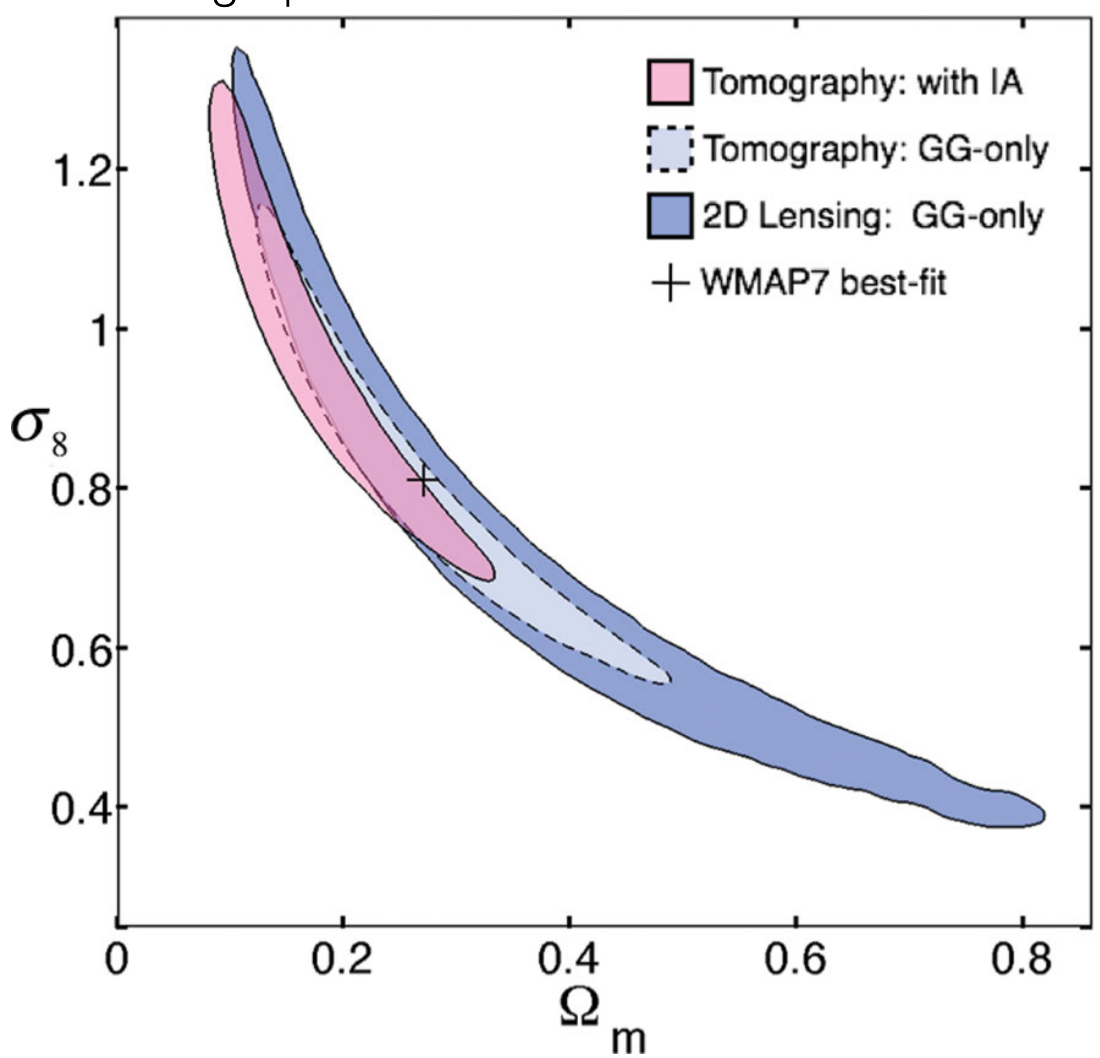
MK et al. (2013)

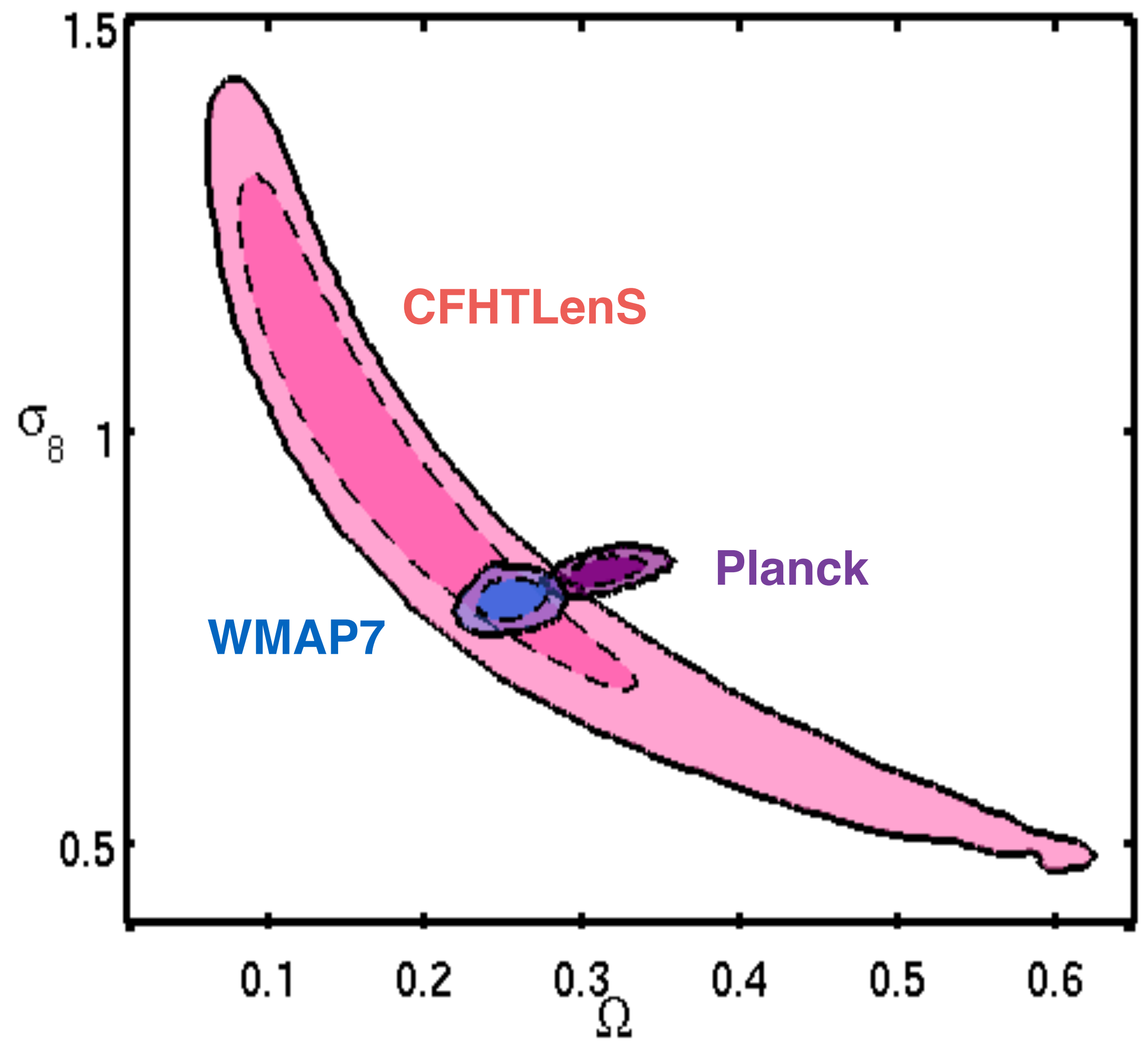
Cosmic shear: state of the art 2013

CFHTLenS

Adding individual galaxy redshifts → tomographic scanning of the LSS, information about z-evolution of clustering, dark energy.
But: account for intrinsic alignment (IA) of galaxies

Seven tomographic redshift bins

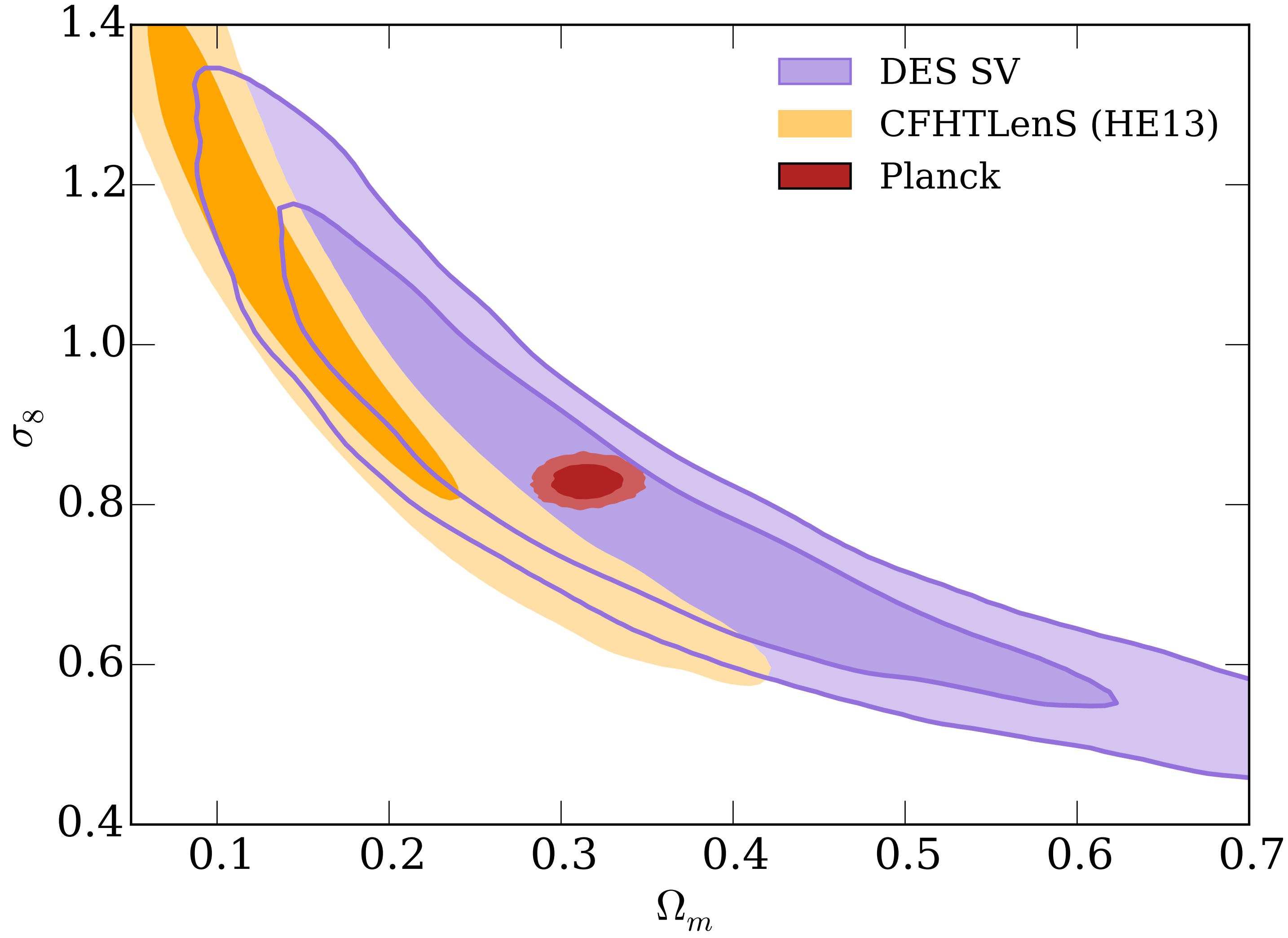




Heymans et al. (2013)

2015: DES

Three tomographic redshift bins



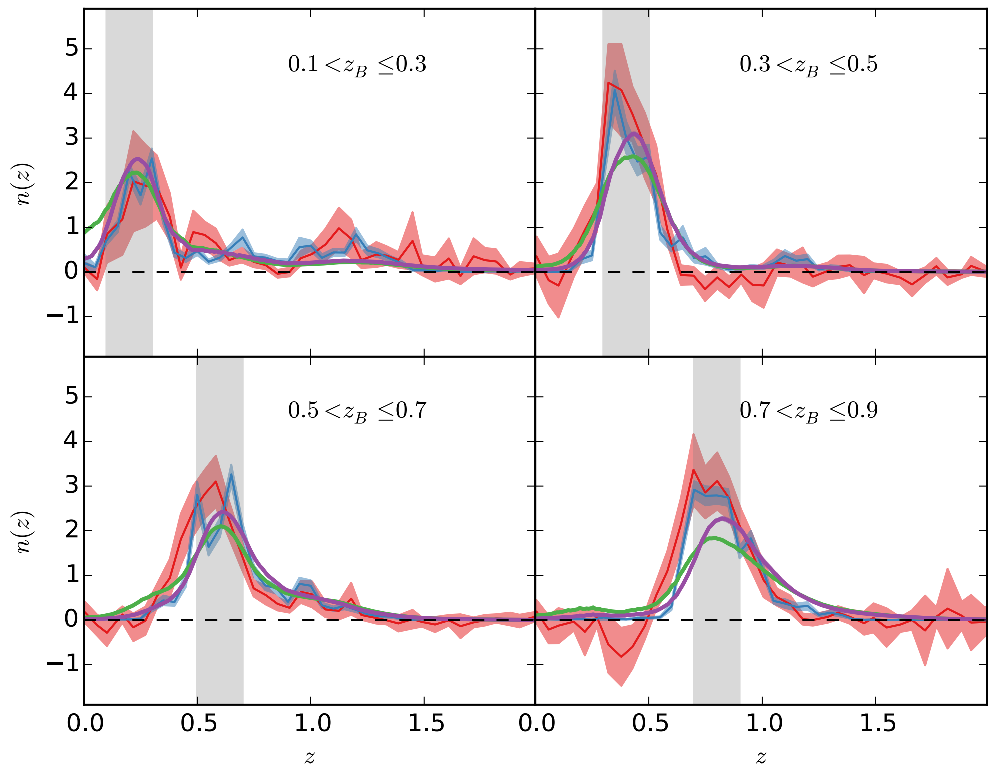
SVD, 139 deg^2

Dark Energy Collaboration (2015)

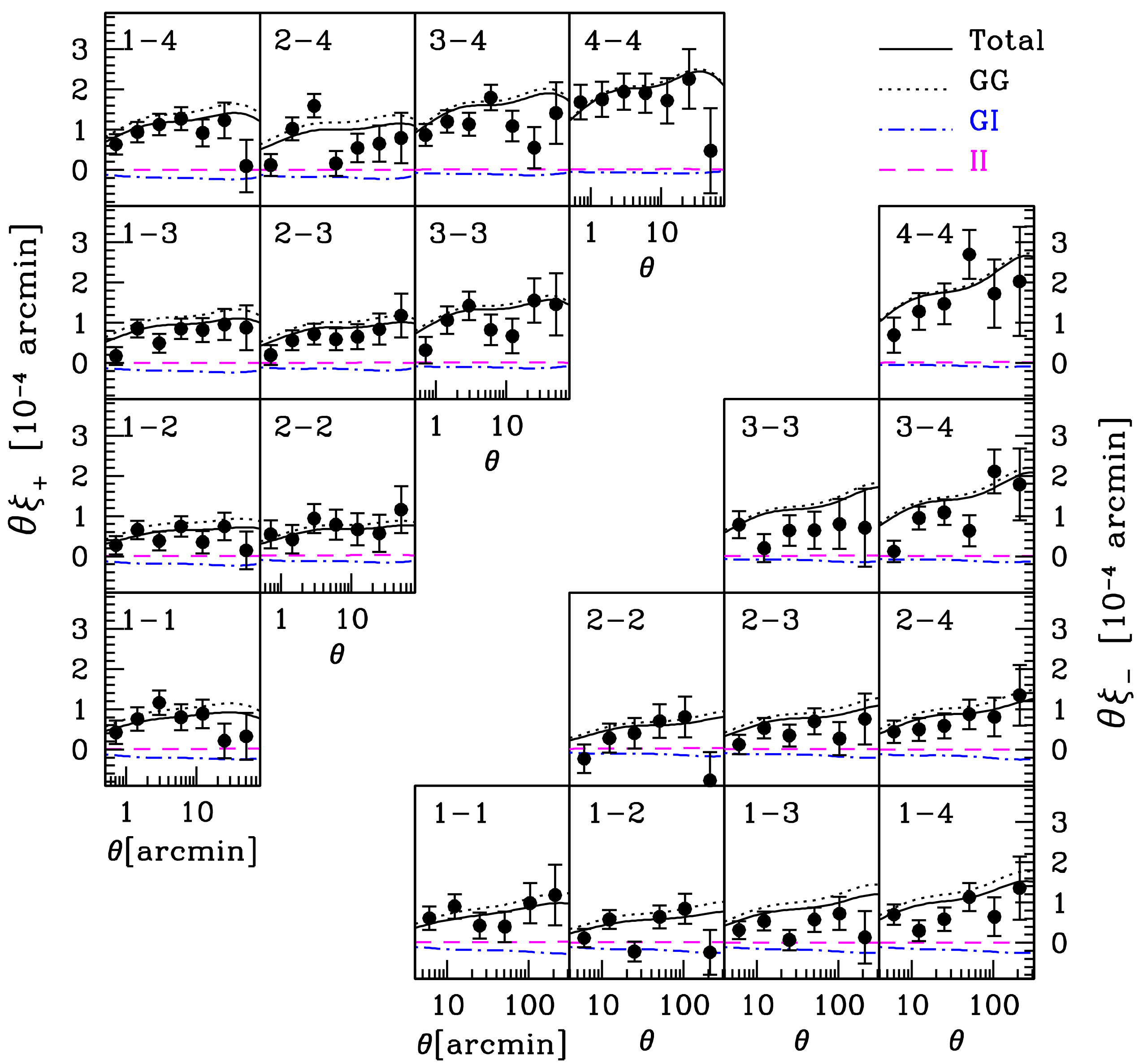
Cosmic shear: results since 2013

2016: KiDS

Four tomographic redshift bins



2-point correlation function,
all combinations between z-bins

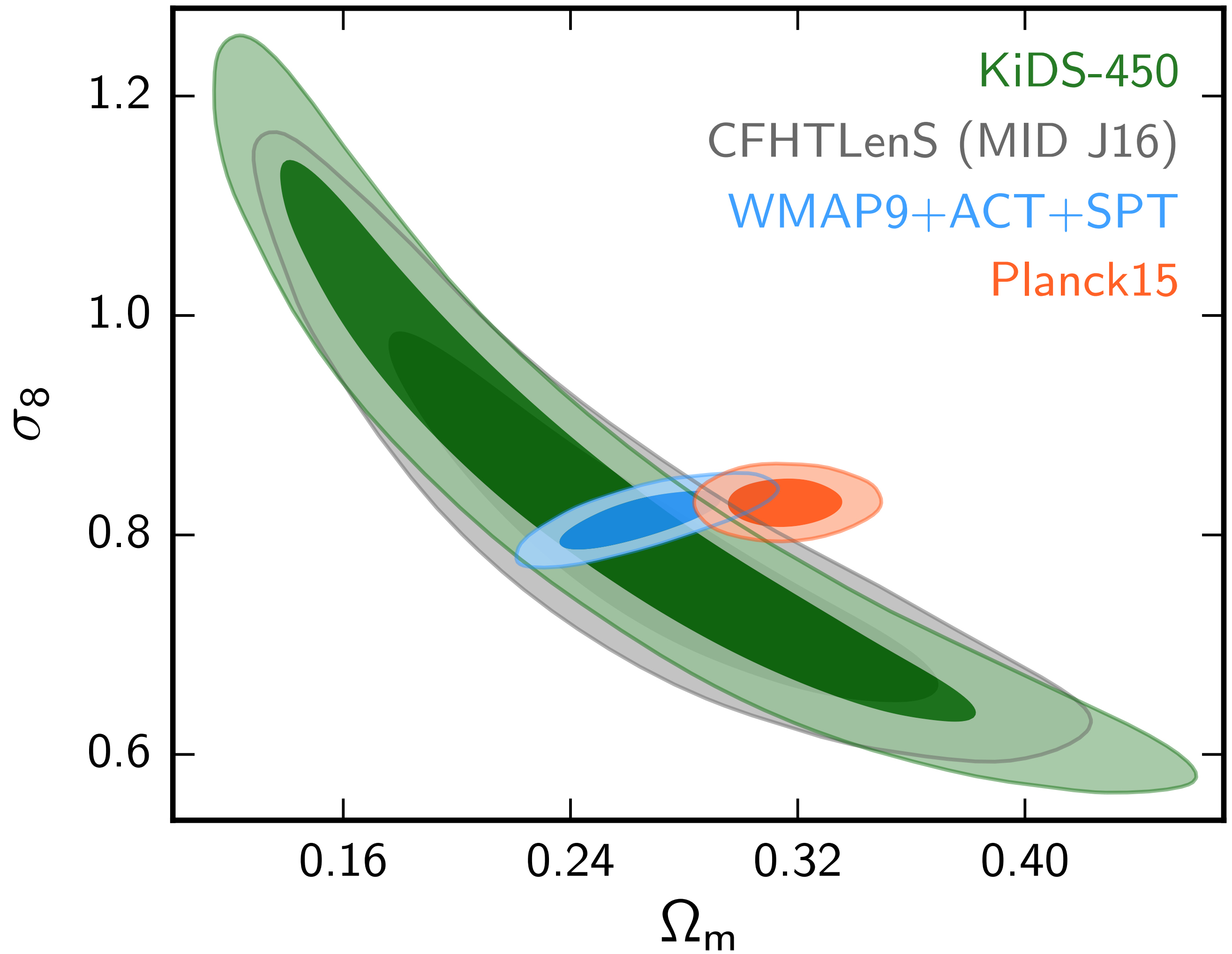


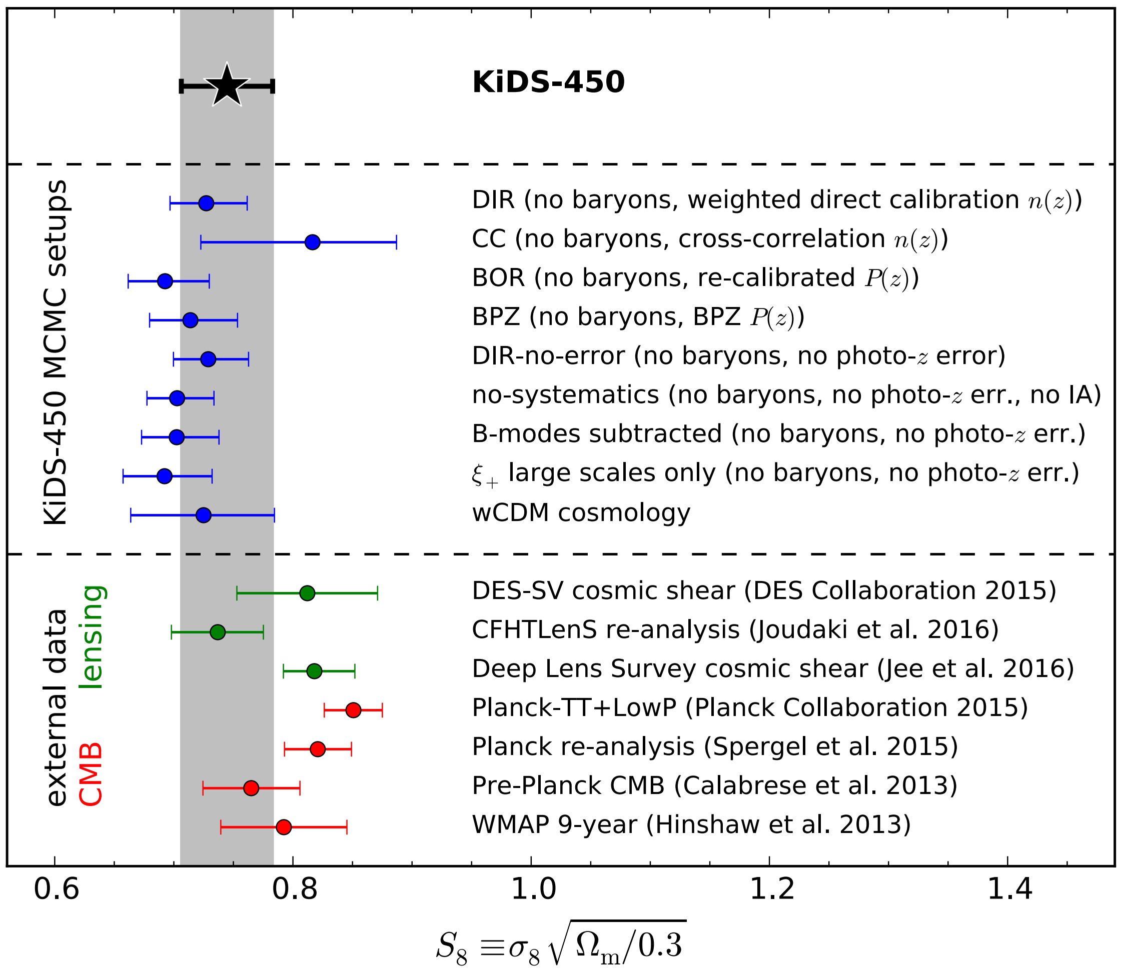
DR1/2, 450 deg^2

Hildebrandt et al. (2016)

Cosmic shear: results since 2013

2016: KiDS

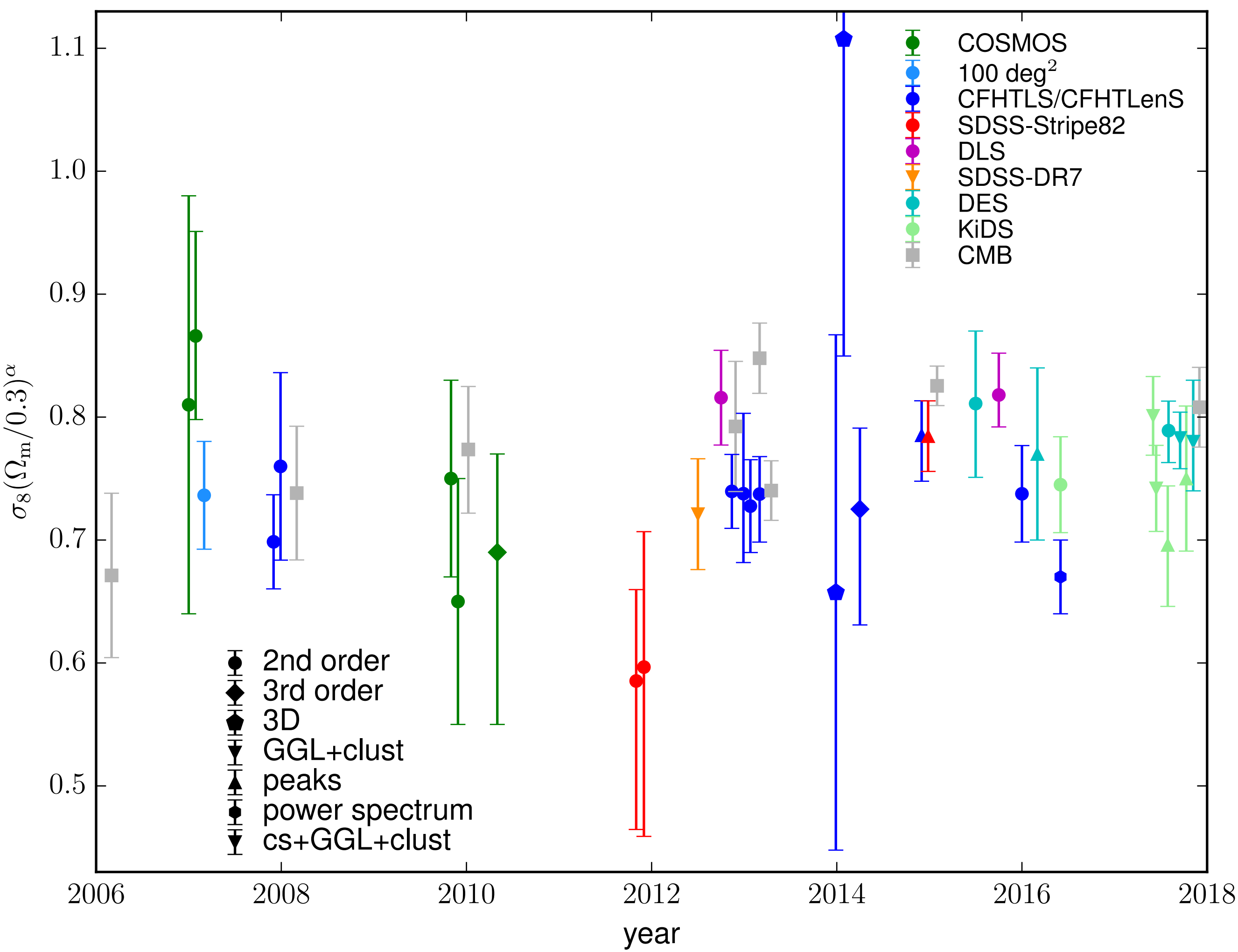




2.3σ tension with Planck (substantial discordance).

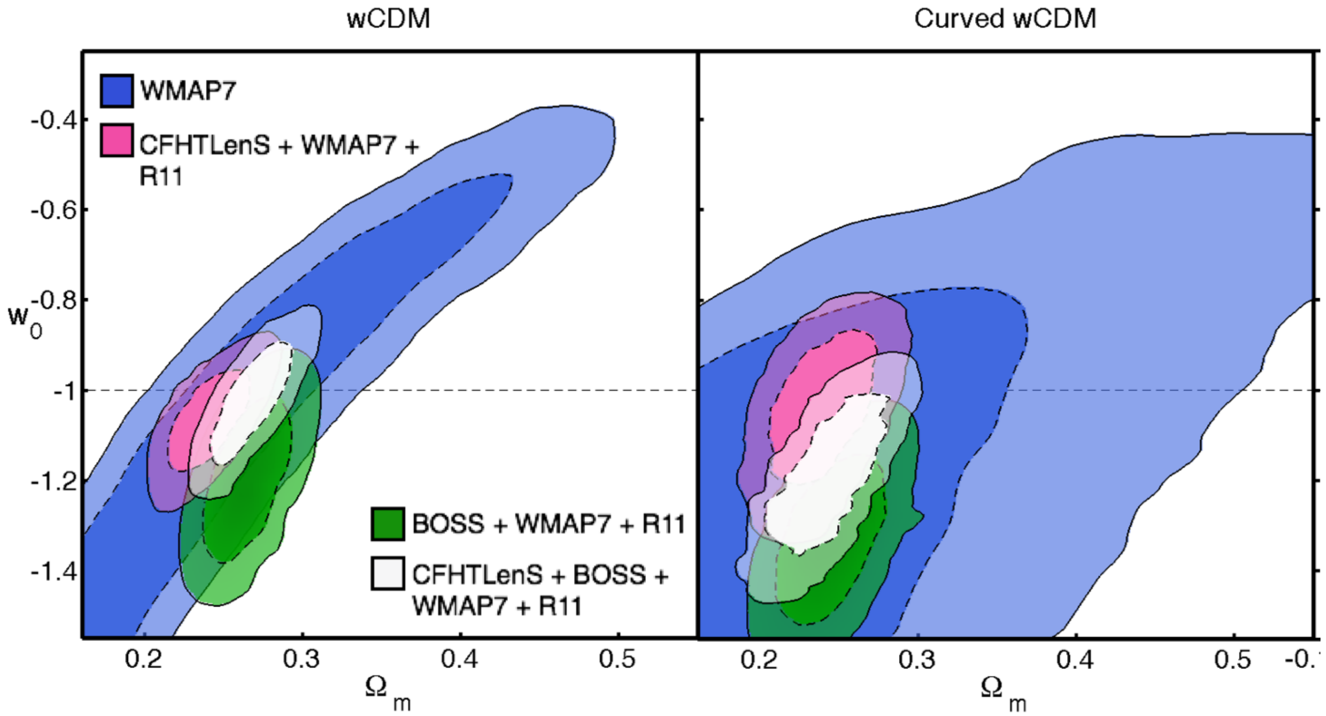
Thorough testing:

- Shape measurement bias (would need unaccounted Δm~0.16 to get to Planck)
- Photometric redshifts (compared 4 methods; would need Δz~0.14)
- Covariance matrix (super-survey modes; analytical vs. N-body)
- Intrinsic alignment & baryons (used large scales only)
- Blinded analysis

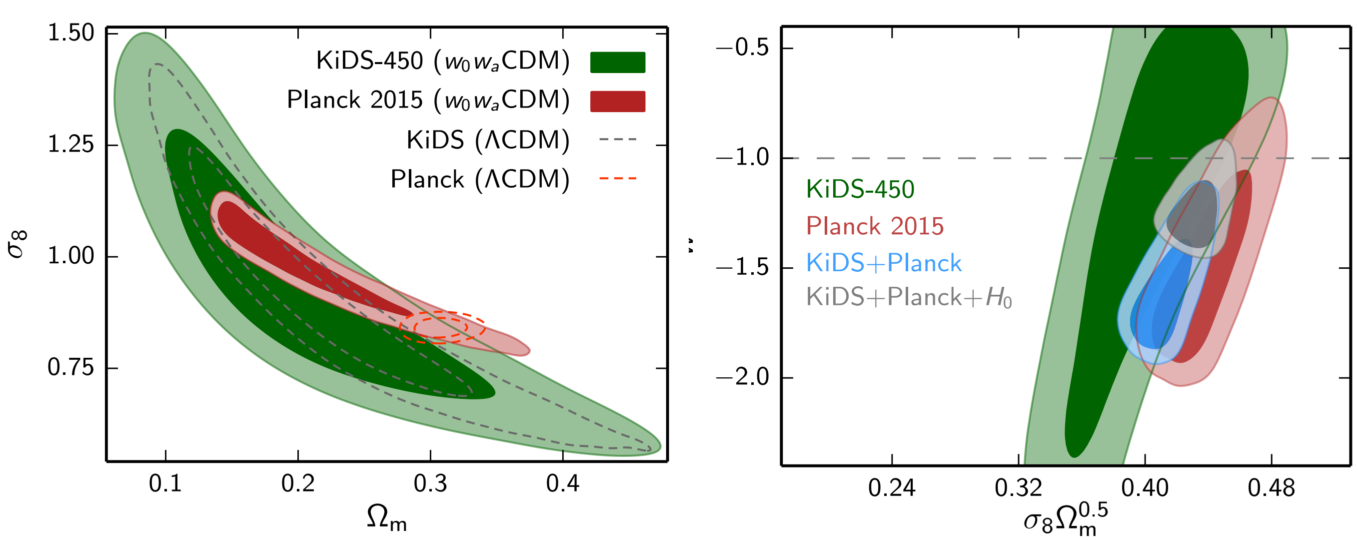

COSMOS
100 deg^2
CFHTLS/CFHTLenS
SDSS-Stripe82
DLS
SDSS-DR7
DES
KiDS
CMB
2nd order
3rd order
3D
GGL+clust
peaks
power spectrum
cs+GGL+clust
$\sigma_8(\Omega_\mathrm{m}/0.3)^\alpha$
year
2006
2008
2010
2012
2014
2016
2018
0.5
0.6
0.7
0.8
0.9
1.0
1.1

Dark energy/modified gravity

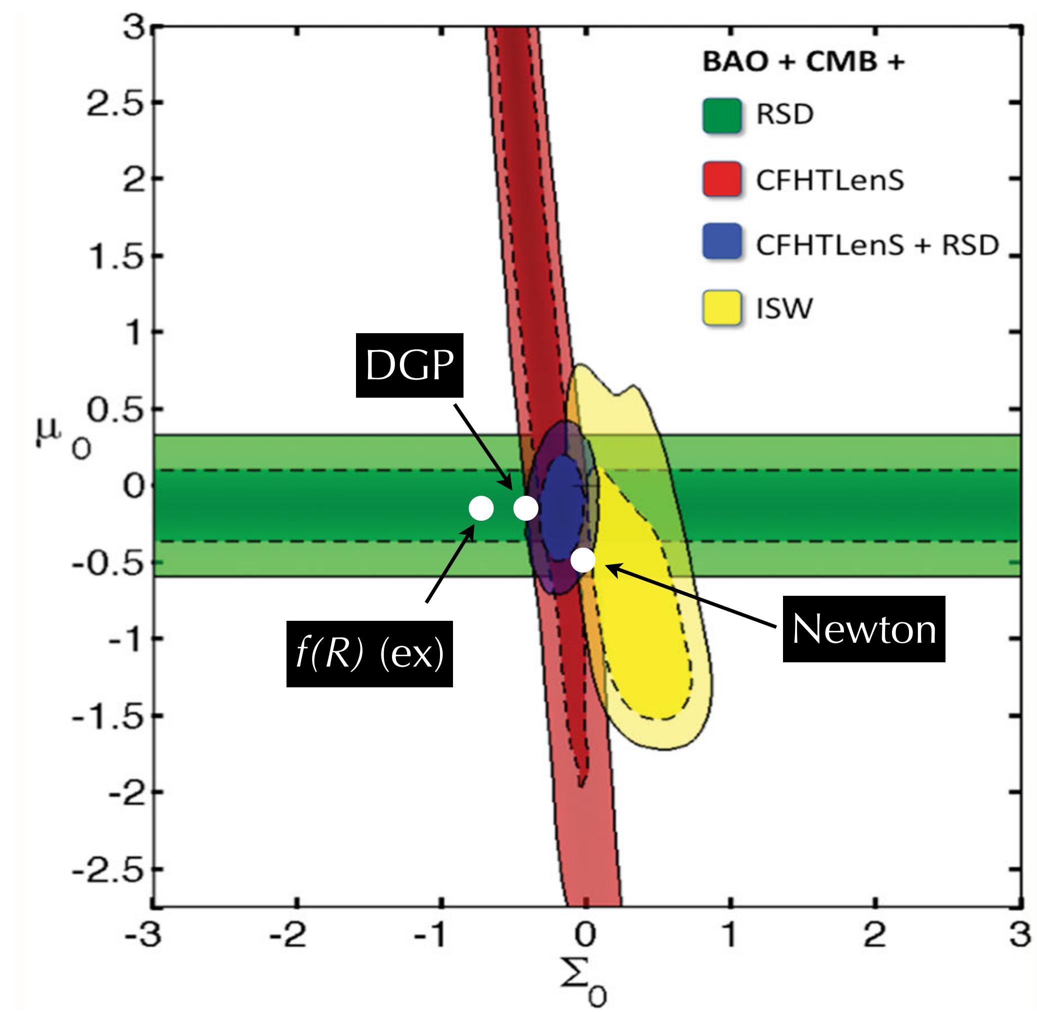
CFHTLenS (Heymans et al. 2013)



KiDS (Joudaki et al. 2016)



CFHTLenS (Simpson et al. 2013)



WL: Non-Gaussian statistics

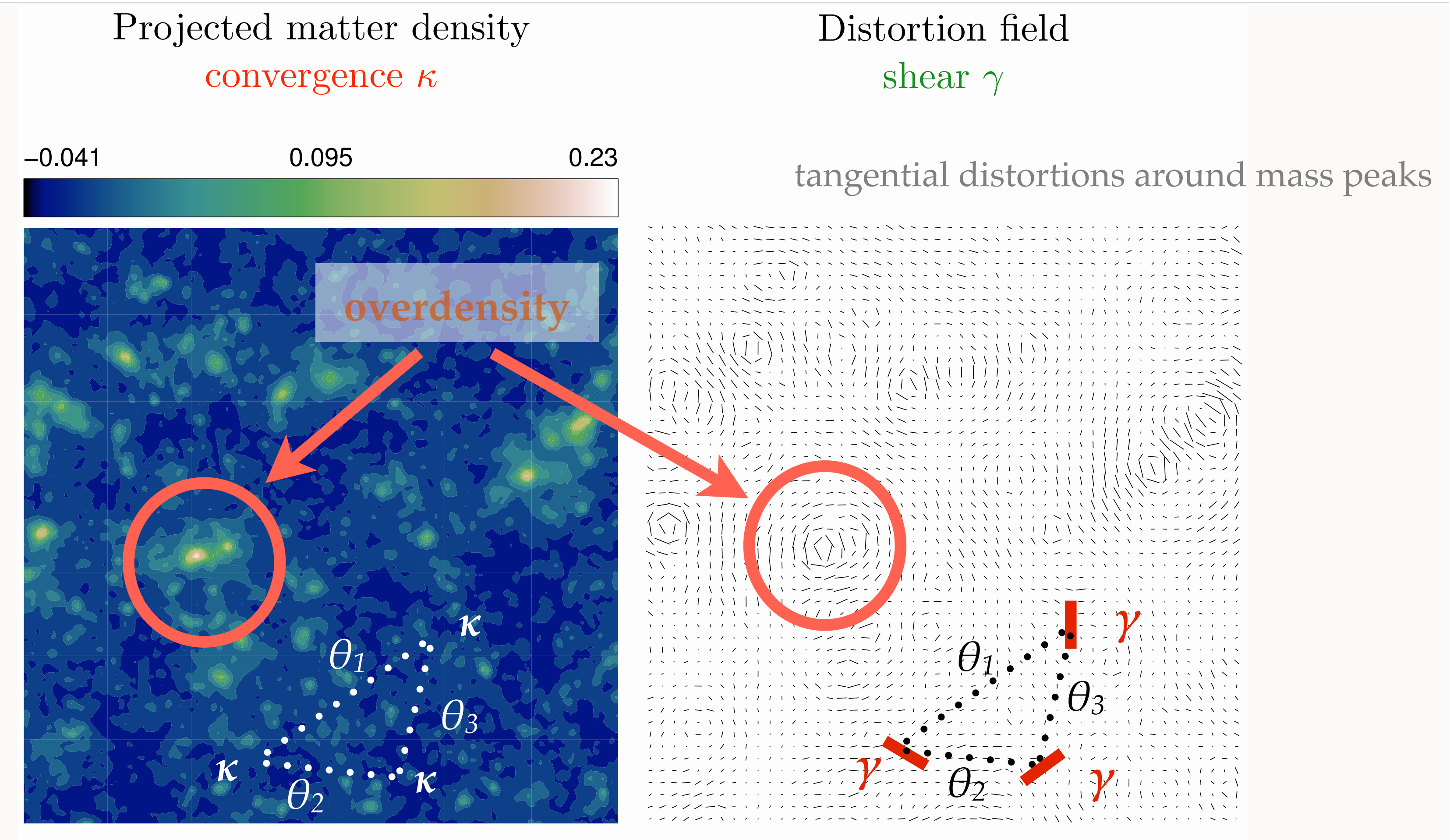
Non-Gaussian observables:

- Bispectrum, three-point correlation function, aperture-mass 3rd moment (Jarvis et al. 2004, Semboloni et al. 2011, Fu, MK et al. 2014)

- Minkowski functions, higher-order κ moments (Petri et al. 2015)

- Peak counts (Liu J. et al. 2014, Liu X. et al. 2014)

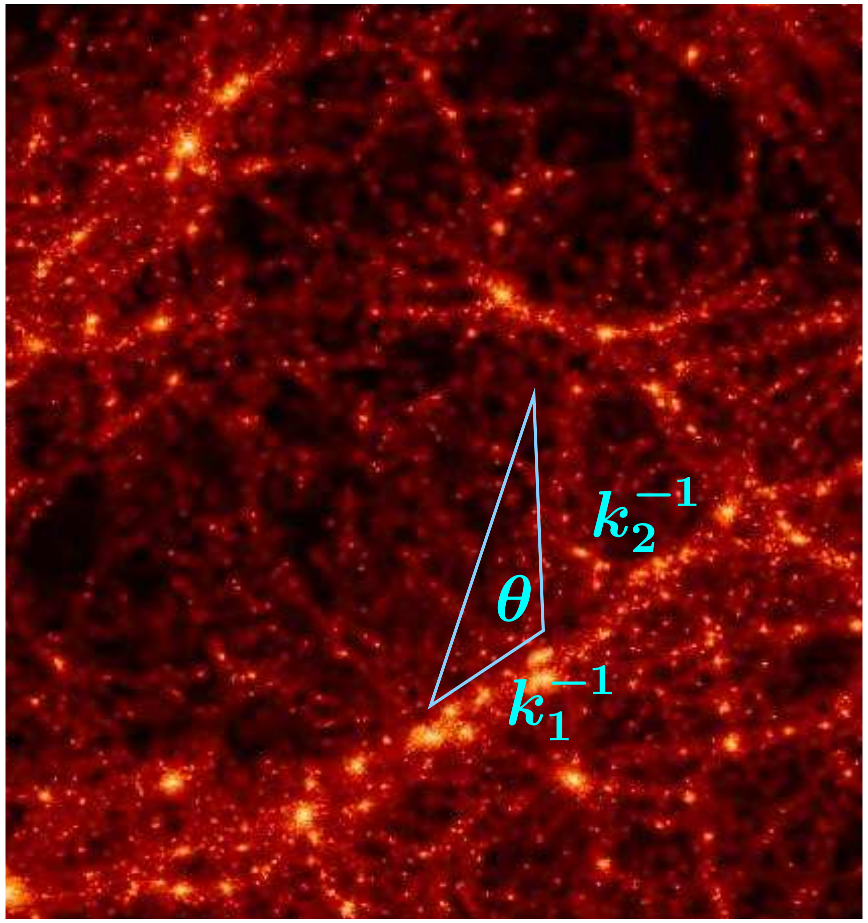
3rd moment smoothed at R: $\sigma^3(\kappa) \neq \sigma^3(\gamma)$



Source galaxies at $z = 1$, ray-tracing simulations by T. Hamana

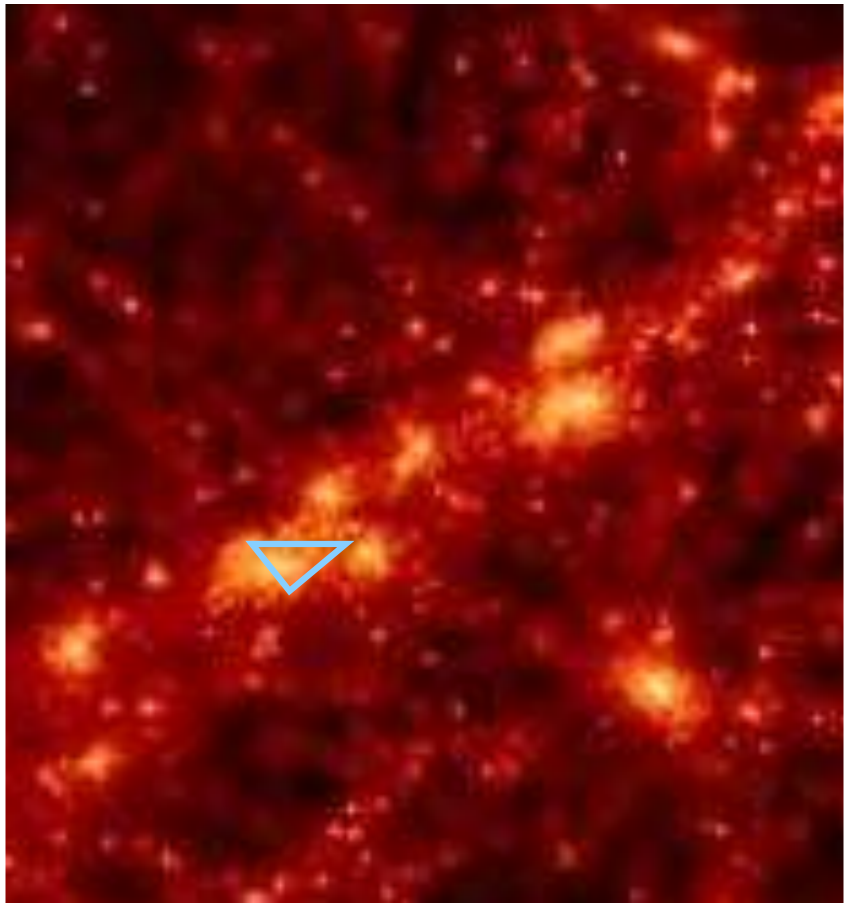
Bispectrum and LSS: cartoon

Large scales



Bispectrum probes filaments and voids
Modulation with opening angle

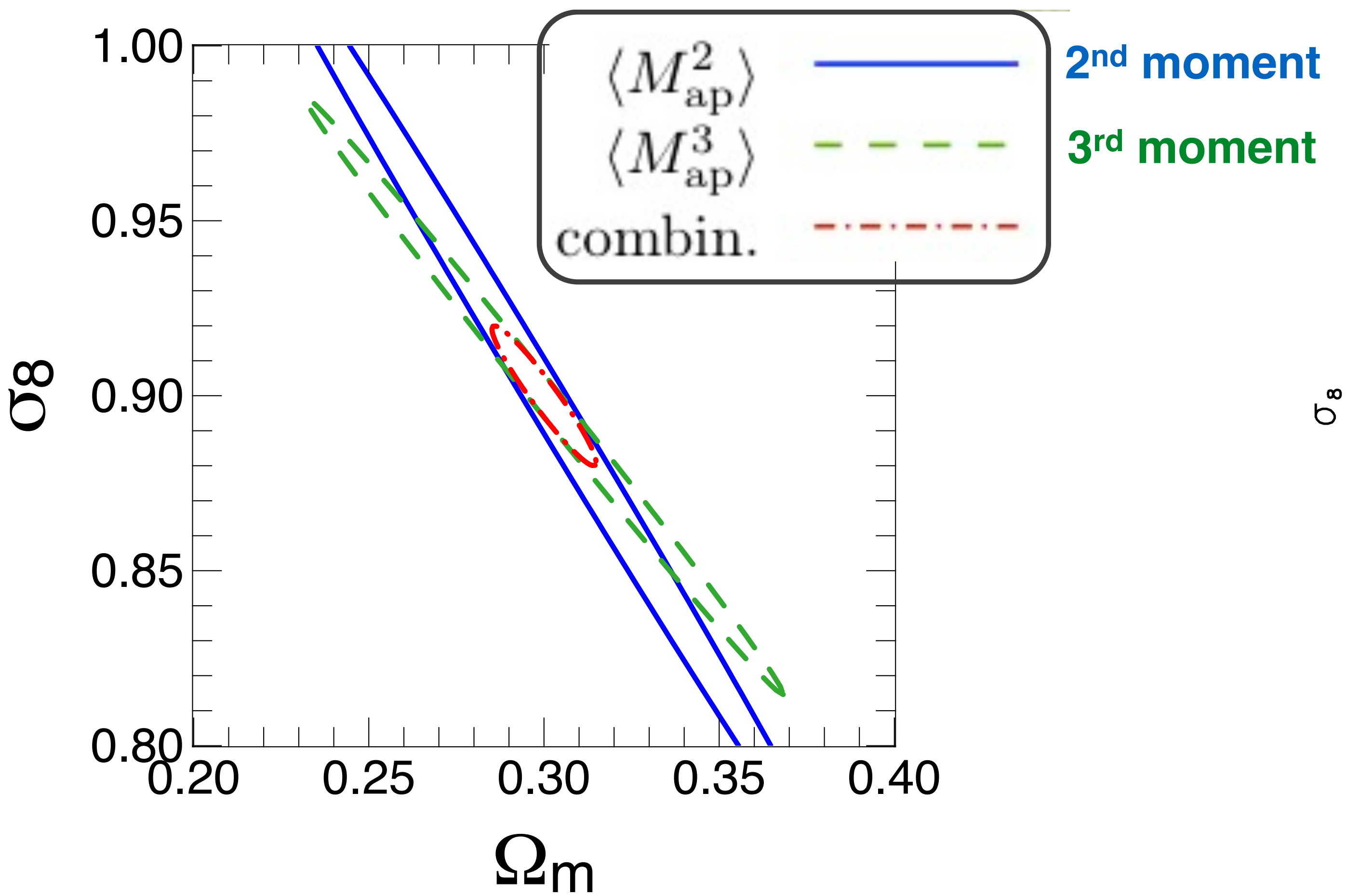
Small scales



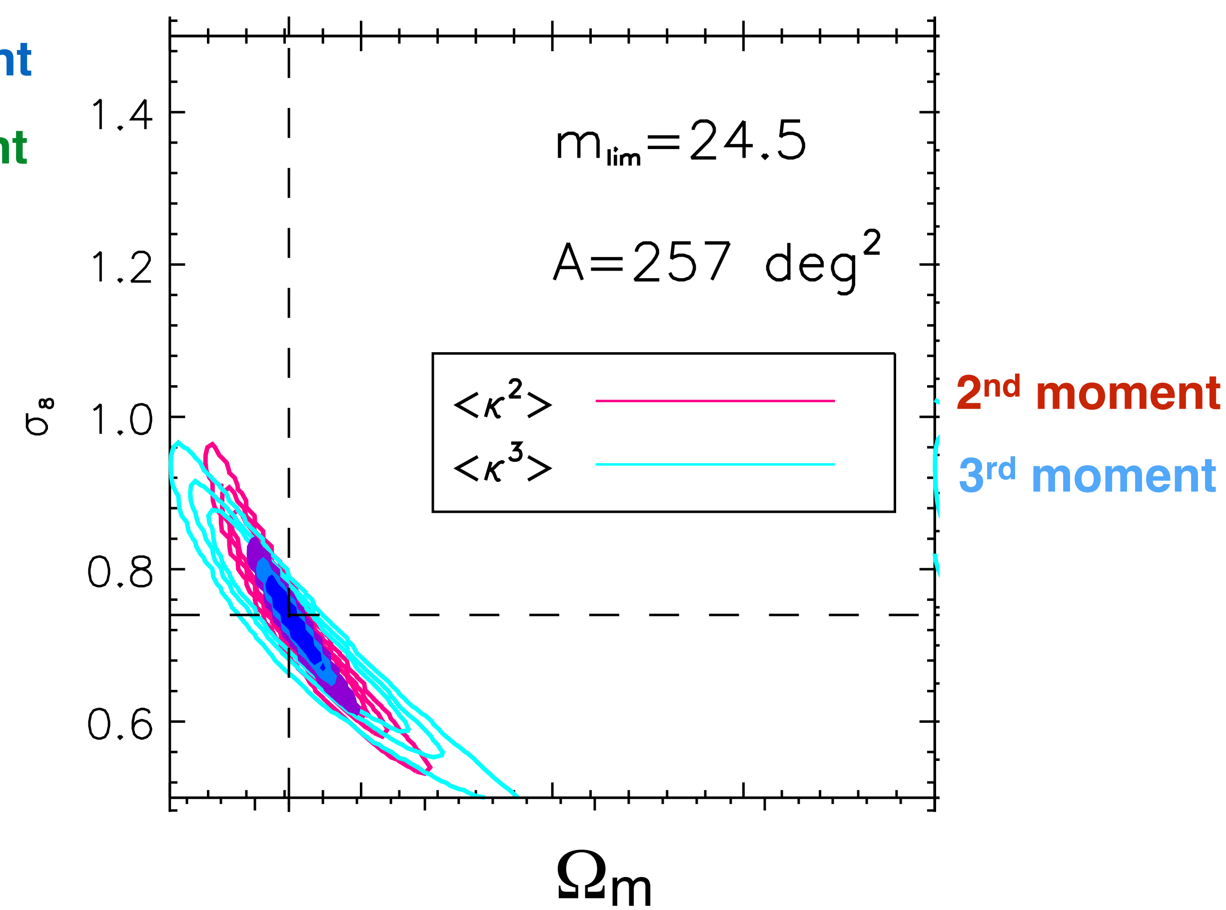
Bispectrum probes halos
No modulation

Higher-order statistics: lifting parameter degeneracies

following
MK & Schneider et al. 2005



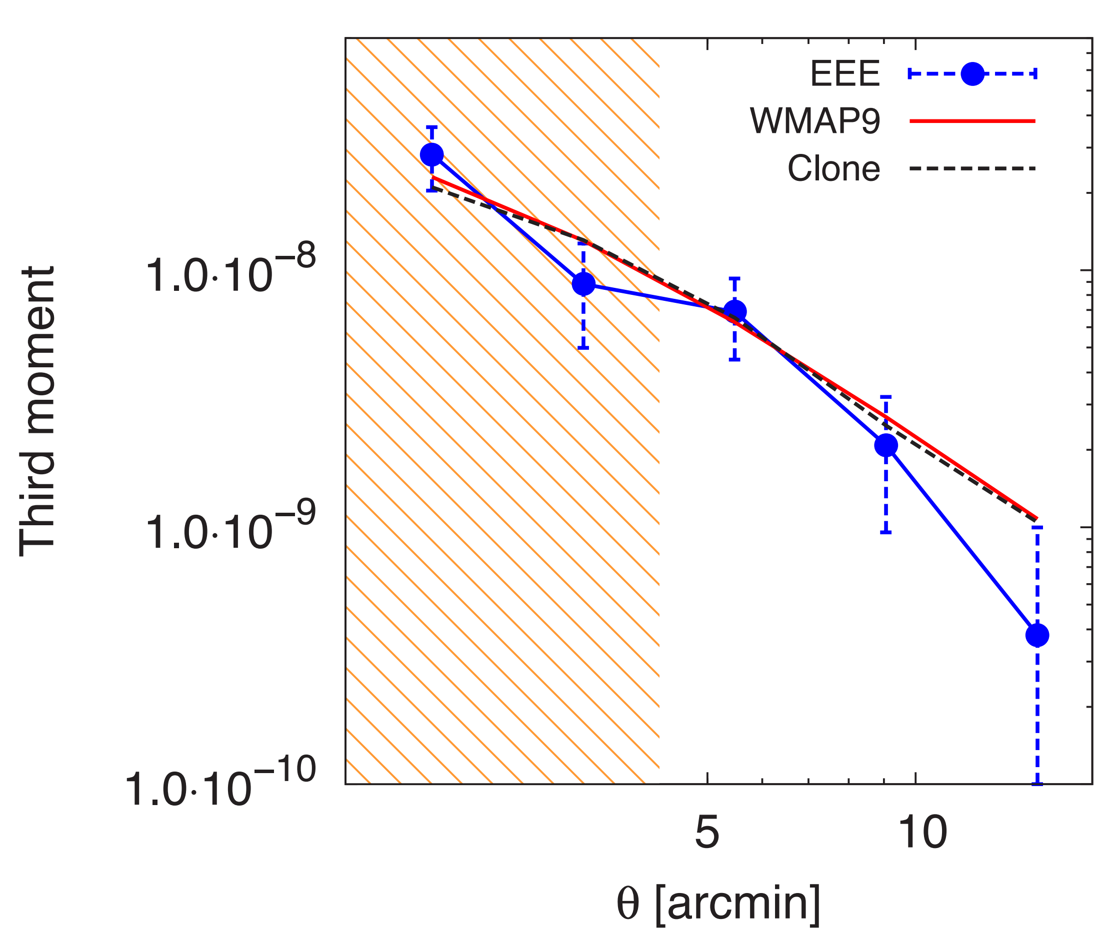
Vafaei et al. 2010



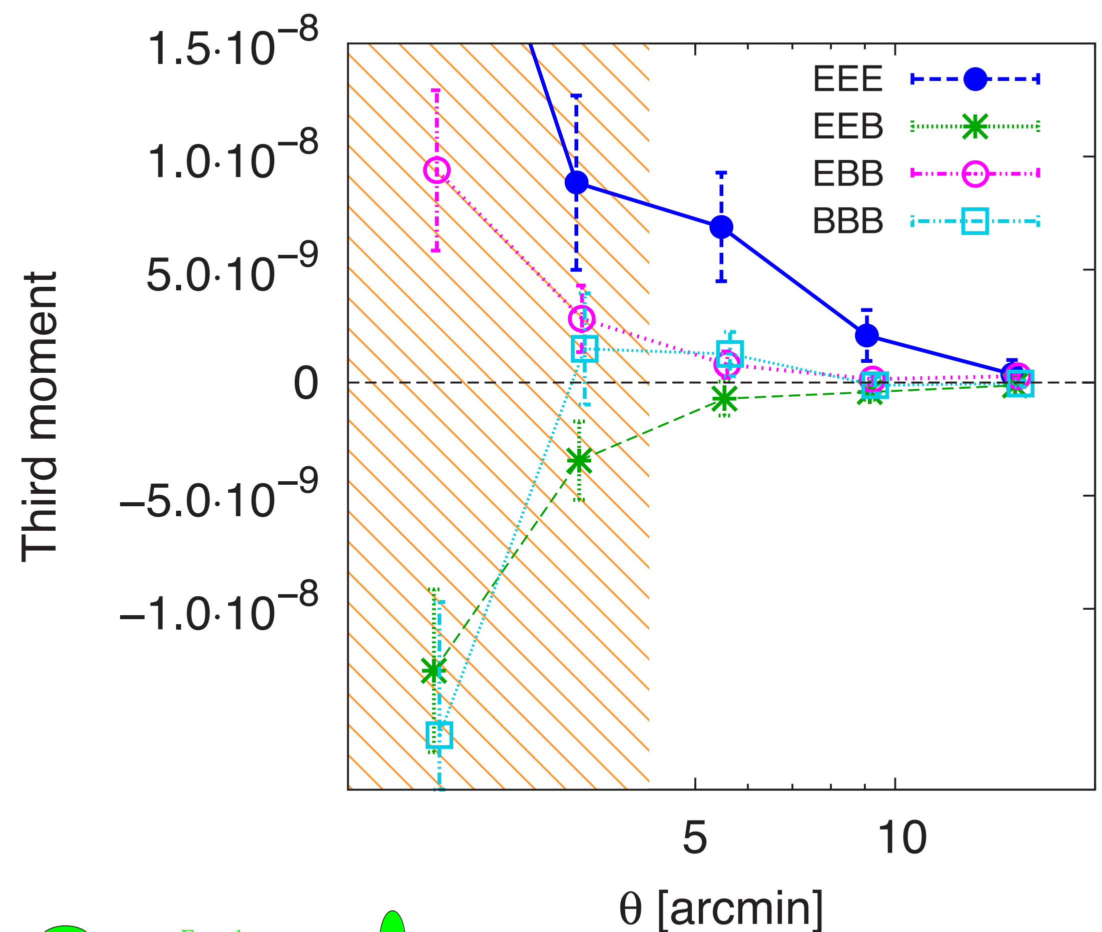
2^{nd} + 3^{rd}-order: Lifting parameter degeneracies (simulations)

CFHTlenS: skewness measurements

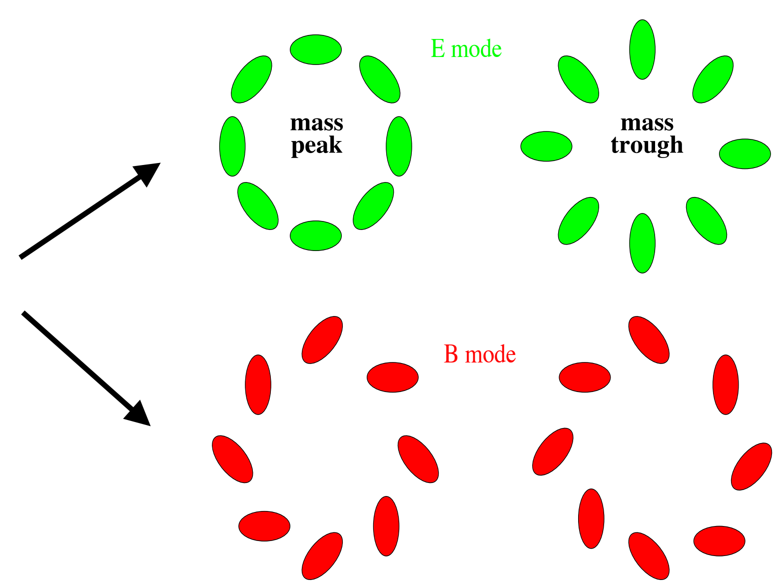
E-mode



B-modes



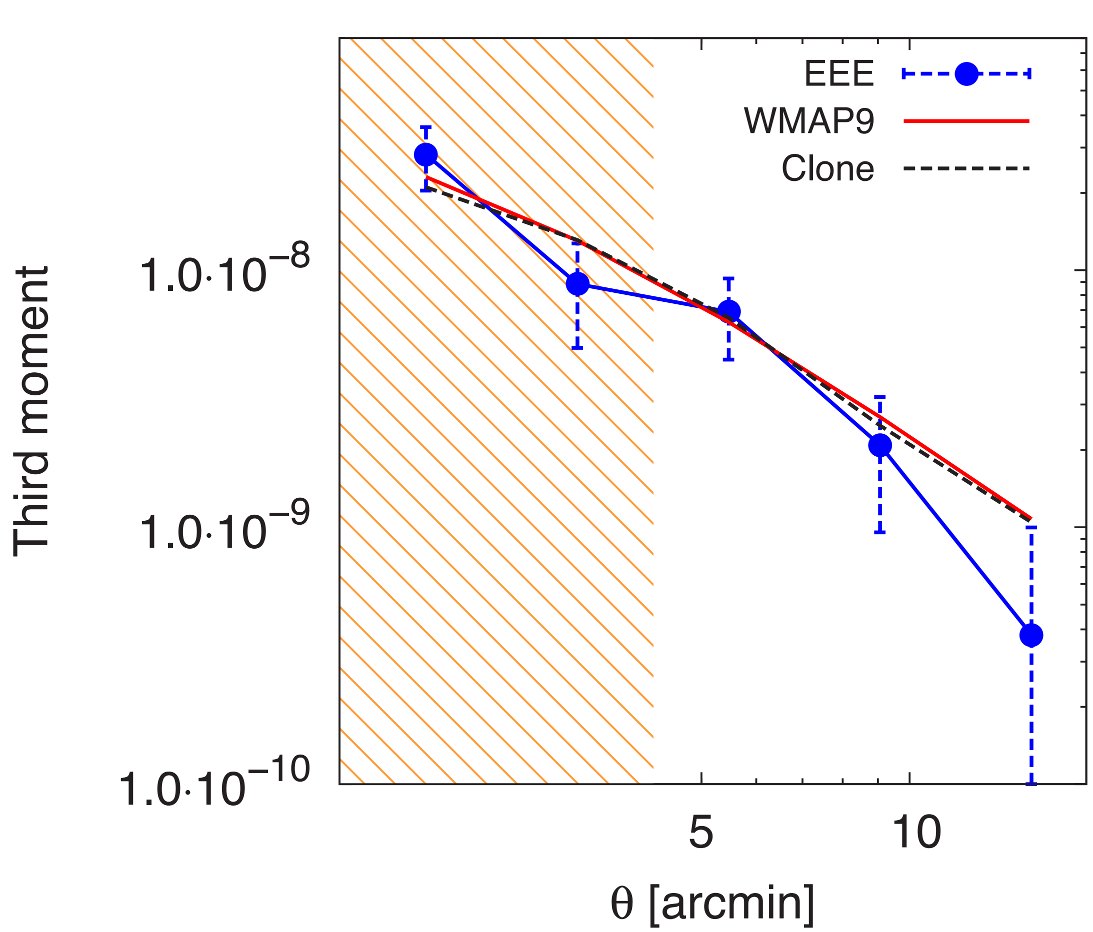
project out



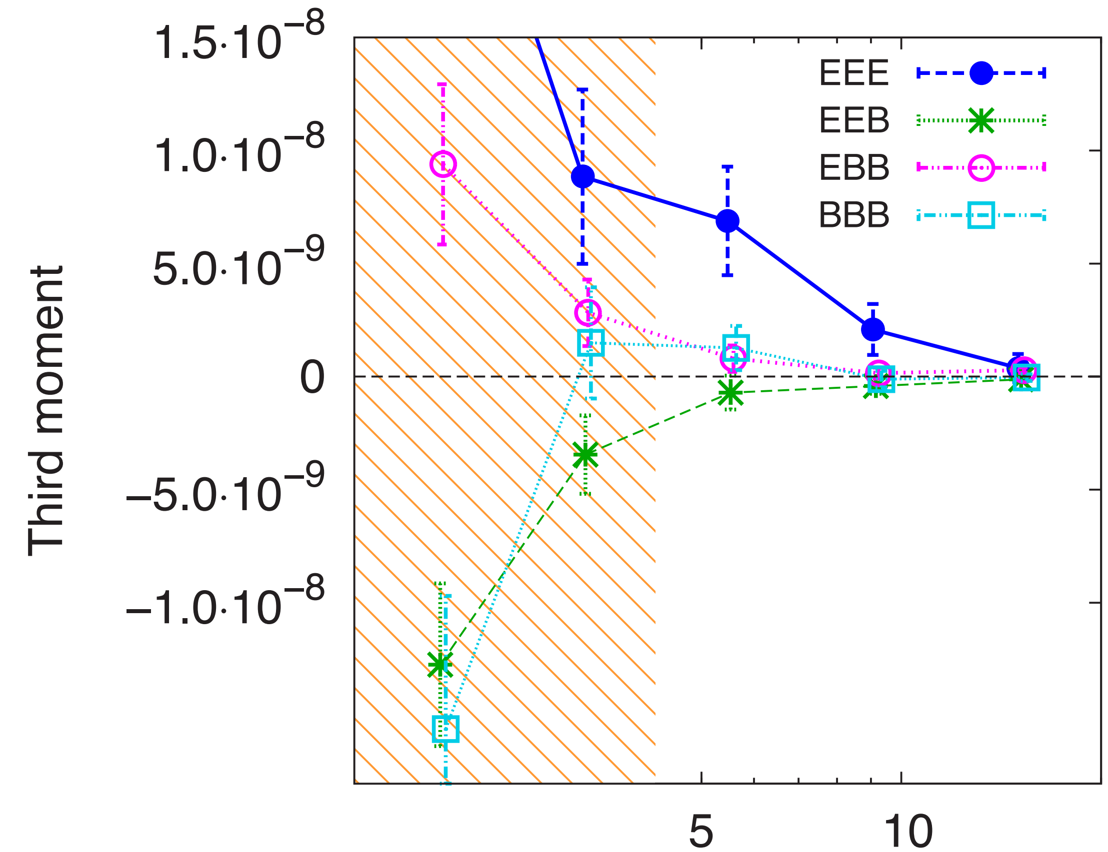
Fu, MK et al. (2014)

CFHTLenS: 3rd-order measurements

E-mode



B-modes



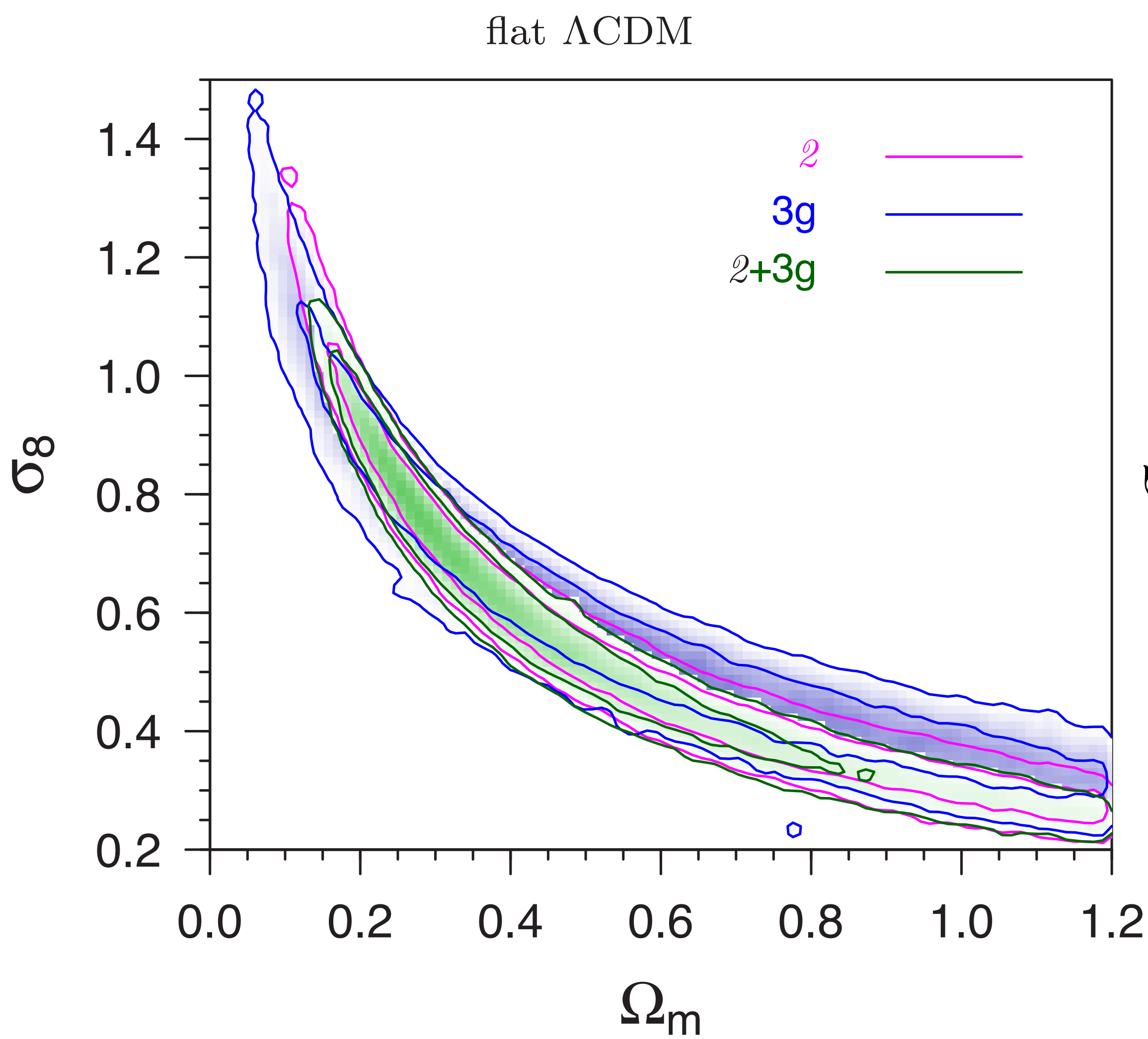
	only Poisson		Poisson + cosmic variance	
	3d	3g	3d	3g
EEE	8.33	10.72	2.02	1.66
EEB	0.48	0.46		
EBB	1.87	3.09		
BBB	0.70	2.63		

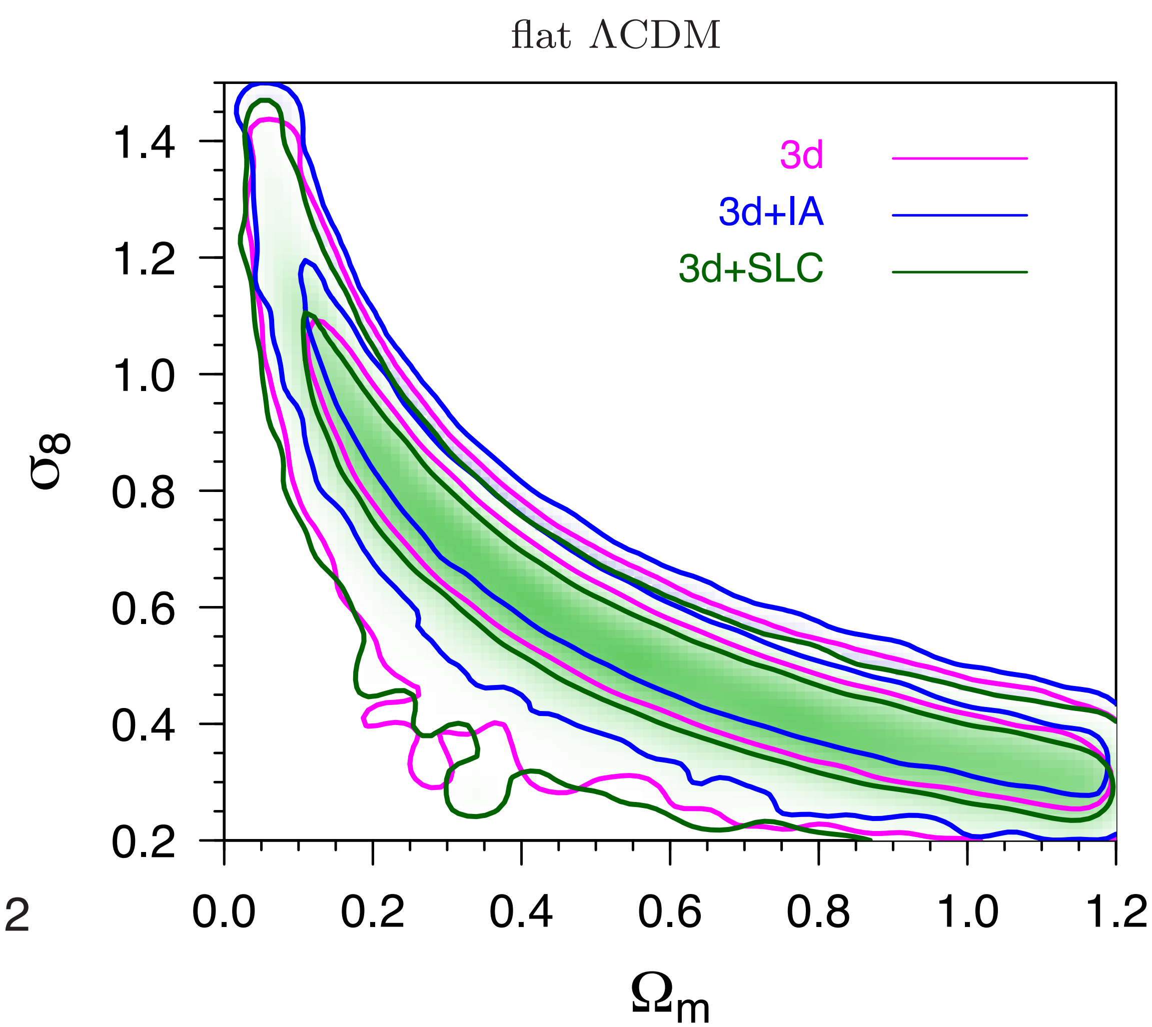
Detection significance [σ]

Dectection of B-modes and parity-modes on small scales

Fu, MK et al. (2014)

CFHTLenS: 2nd- and 3rd-order combined





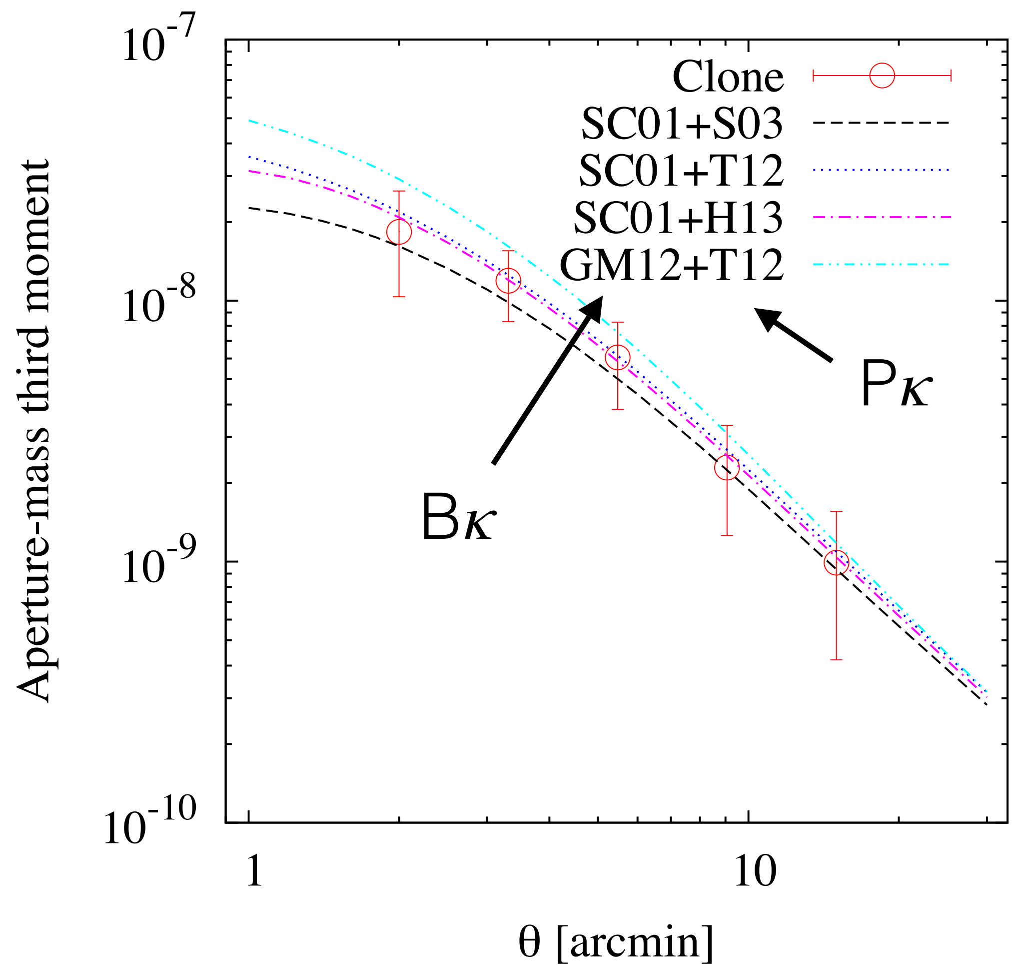
Statistics:
2nd-order statistics
3rd-order statistics
combination (with xcorrelation)

Systematics:
+ Intrinsic alignment (IA)
+ Source-lens clustering (SLC)
- Baryonic physics (3%-5% bias on σ_8)

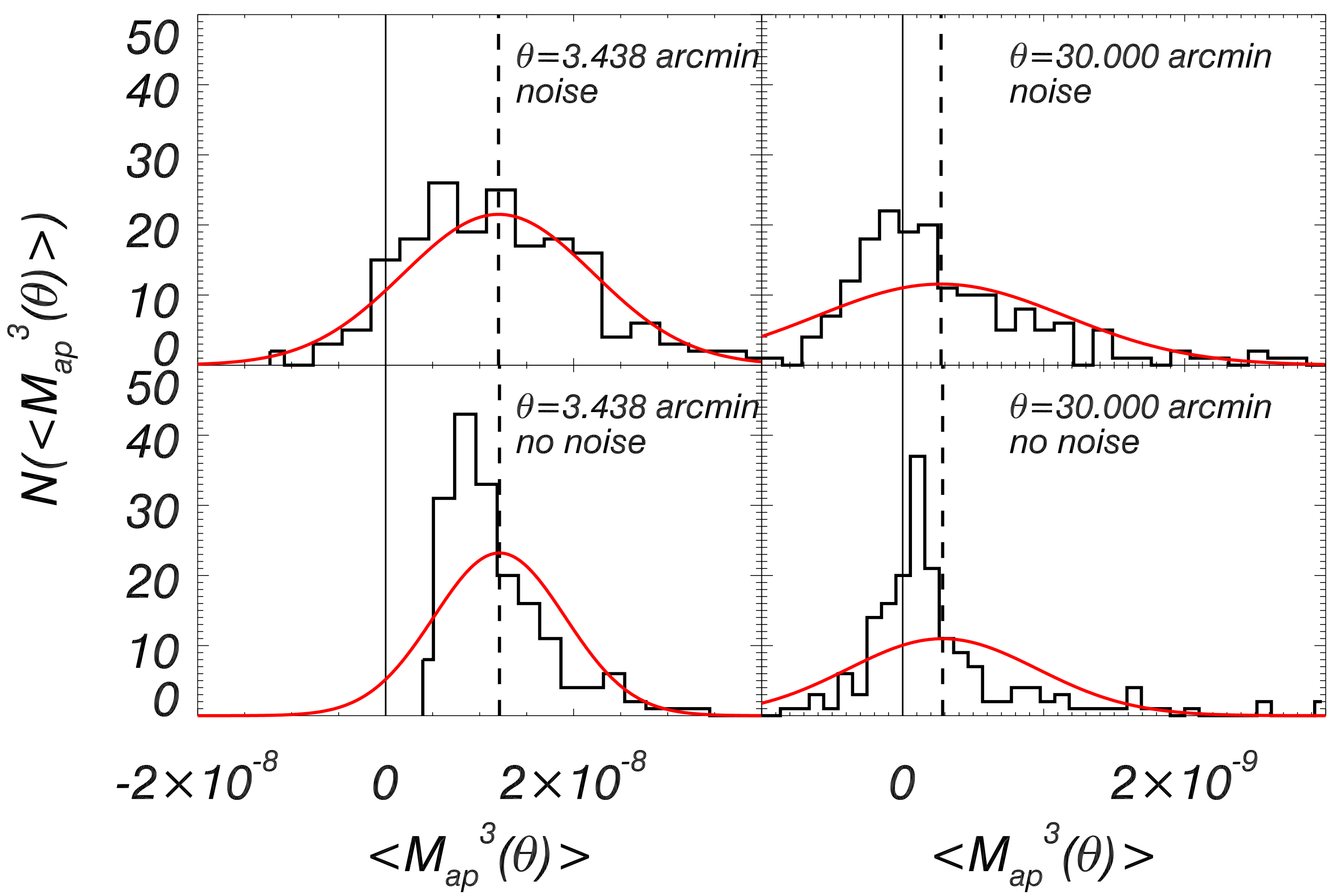
Fu, MK et al. (2014)

3rd-order statistics: further uncertainties

Non-linear model predictions



pdf(3rd-order moments) from *N*-body



Bispectrum $B\kappa = F[(P\kappa)^2]$

Fits to *N*-body simulations accurate to ~ 30%.

[Simon et al. 2015]

3^{rd}-order WL statistics: summary

- Additional (non-Gaussian) information about LSS ✓
- Lift parameter degeneracies ✓

- Prone to residual systematics in data ✗
- Astrophysical systematics (IA, SLC, baryons) ✗
- Model uncertainties ✗
- non-Gaussian pdf, likelihood? ✗

WL peak counts interesting alternative?

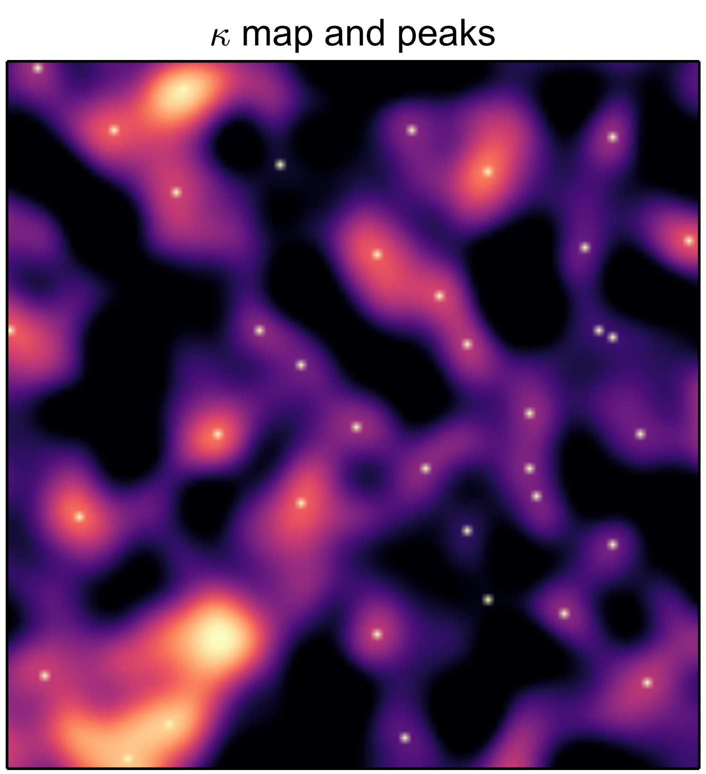
3rd-order WL statistics: summary

	3rd	peaks
• Additional (non-Gaussian) information about LSS	✓	✓
• Lift parameter degeneracies	✓	?
• Prone to residual systematics in data	✗	✓
• Astrophysical systematics (IA, SLC, baryons)	✗	?
• Model uncertainties	✗	✓
• non-Gaussian pdf, likelihood?	✗	✓

WL peak counts interesting alternative?

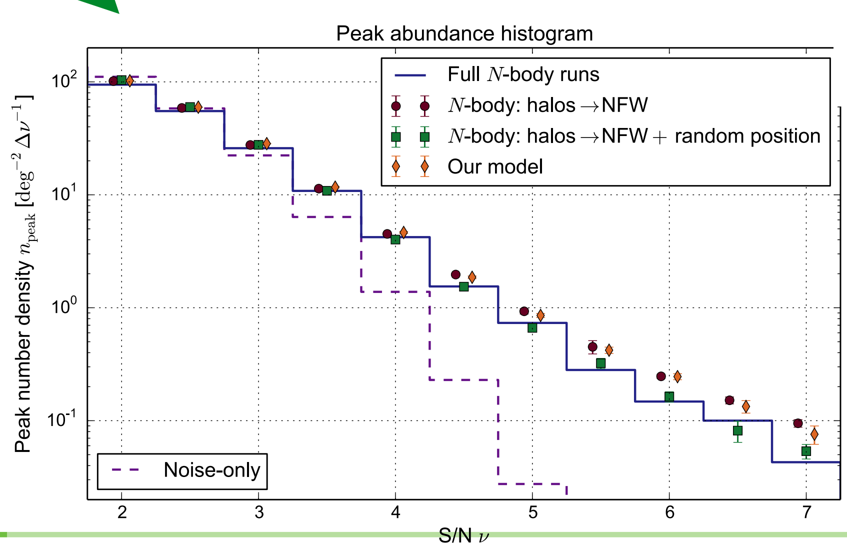
Weak-lensing peak counts

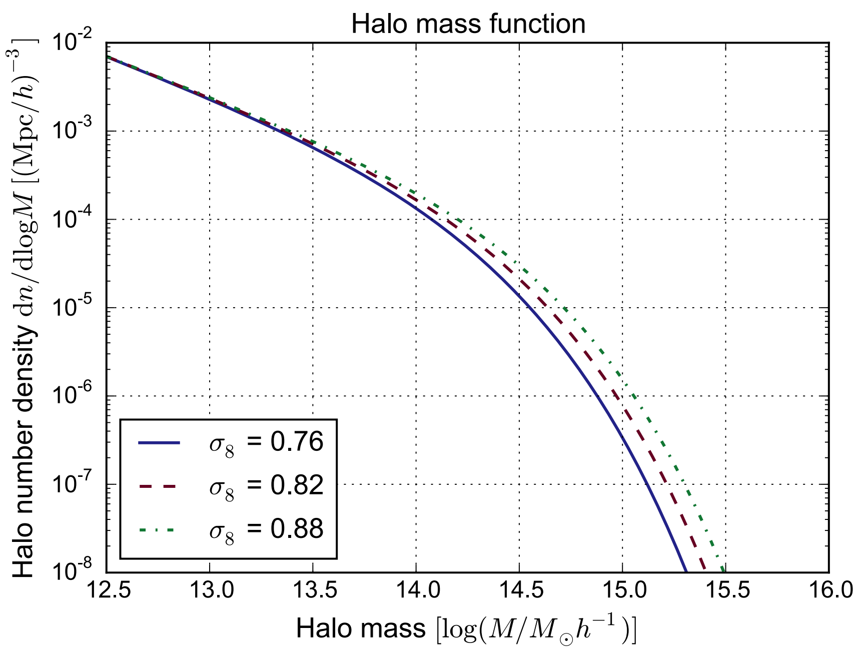
- WL peaks probe high-density regions ↔ **non-Gaussian** tail of LSS
- **First-order** in observed shear: less sensitive to systematics, circular average!
- High-density regions ↔ **halo mass function**, but **indirect** probe:
  - Intrinsic ellipticity **shape noise**, creating false positives, up-scatter in S/N
  - **Projections** along line of sight



modelling

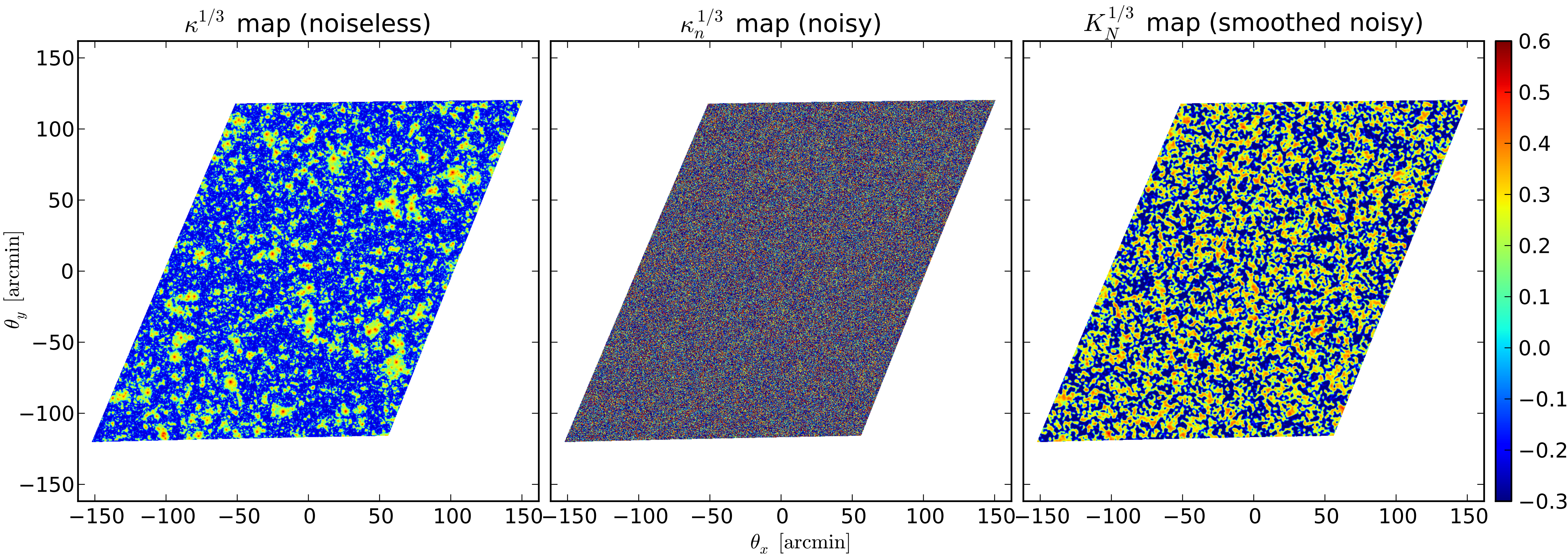
counting





interpretation ?

WL peaks: measurement



WL basic observable is shear γ.

To convert to convergence κ: inverse problem; need to regularise/filter/pixelize.

Both γ and κ are very noisy.

Clusters have characteristic scale(s).

Need to filter WL data!

E.g. Gaussian, compensated filter (aperture-mass).

Non-Gaussian filters: combine scales, denoise.

WL peak modelling

[Early studies: cluster detection, can only use at very high S/N (>5)]

Analytical modelling

Gaussian random field theory (Fan, Shan & Liu J. 2011; Maturi et al. 2010)

Limitations: Additions for high-end tail; Gaussian filtering only; difficult to model systematics (e.g. photo-z errors, masks; intrinsic alignment, baryons, halo substructure, triaxiality)

Forward modelling

N-body simulations

(Haiman group; Dietrich & Hartlap (2010).

Limitations: Large computation time; limited to few std parameters. Need large boxes (High S/N peaks are rare events) and good resolution (resolve group-scale halos)

Fast simulations

Lin & MK (2015a, b) Lin, MK & Pires (2016), Peel et al. (2017)

WL peak modelling

[Early studies: cluster detection, can only use at very high S/N (>5)]

Analytical modelling

Gaussian random field theory (Fan, Shan & Liu J. 2011; Maturi et al. 2010)

Limitations: Additions for high-end tail; Gaussian filtering only; difficult to model systematics (e.g. photo-z errors, masks; intrinsic alignment, baryons, halo substructure, triaxiality)

Forward modelling

Straight-forward to include systematics & complex distributions

N-body simulations

(Haiman group; Dietrich & Hartlap (2010).

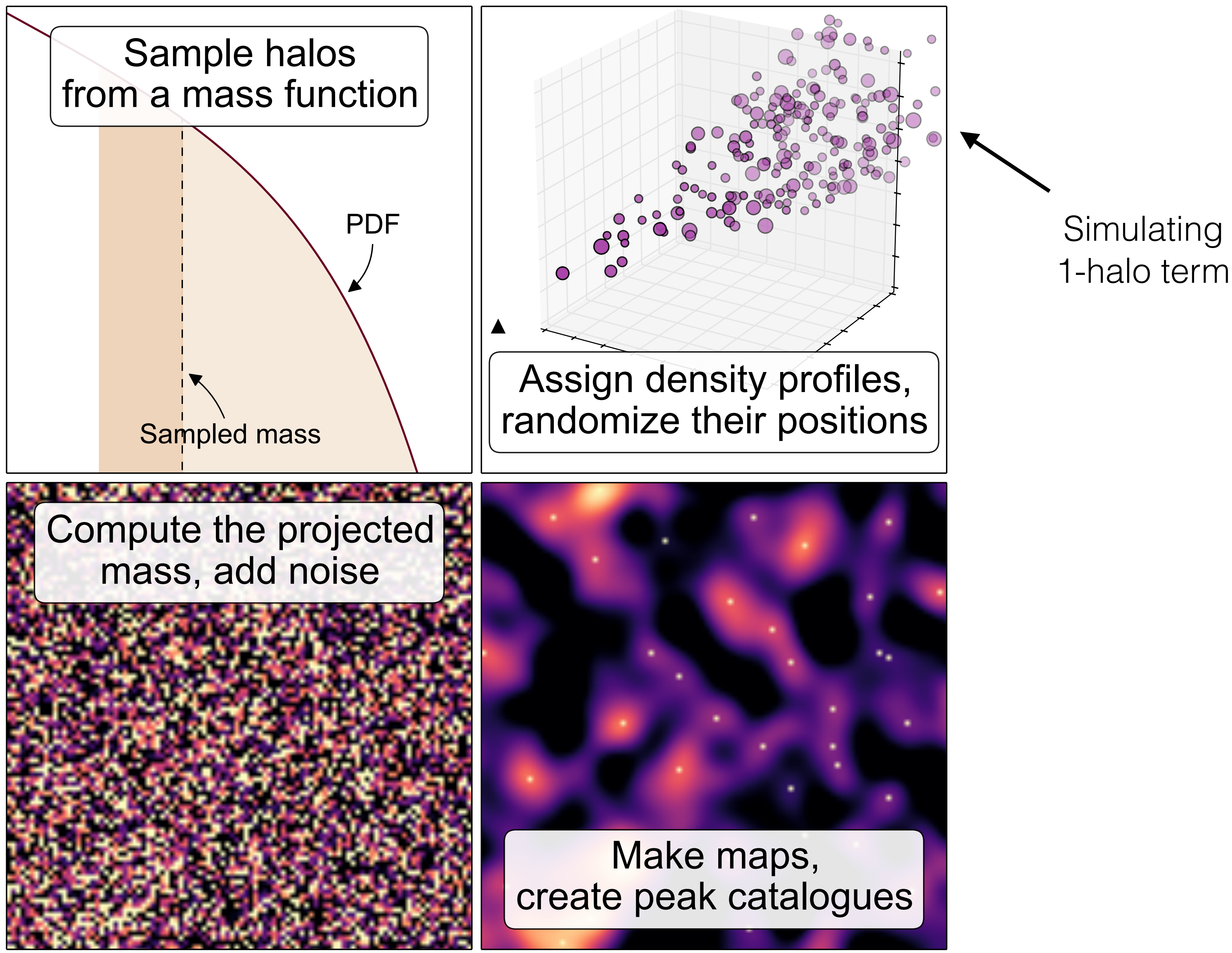
Limitations: Large computation time; limited to few std parameters. Need large boxes (High S/N peaks are rare events) and good resolution (resolve group-scale halos)

Fast simulations

Lin & MK (2015a, b), Lin, MK & Pires (2016), Peel et al. (2017)

Fast simulations for WL peak counts

Replace N-body simulations by Poisson distribution of halos



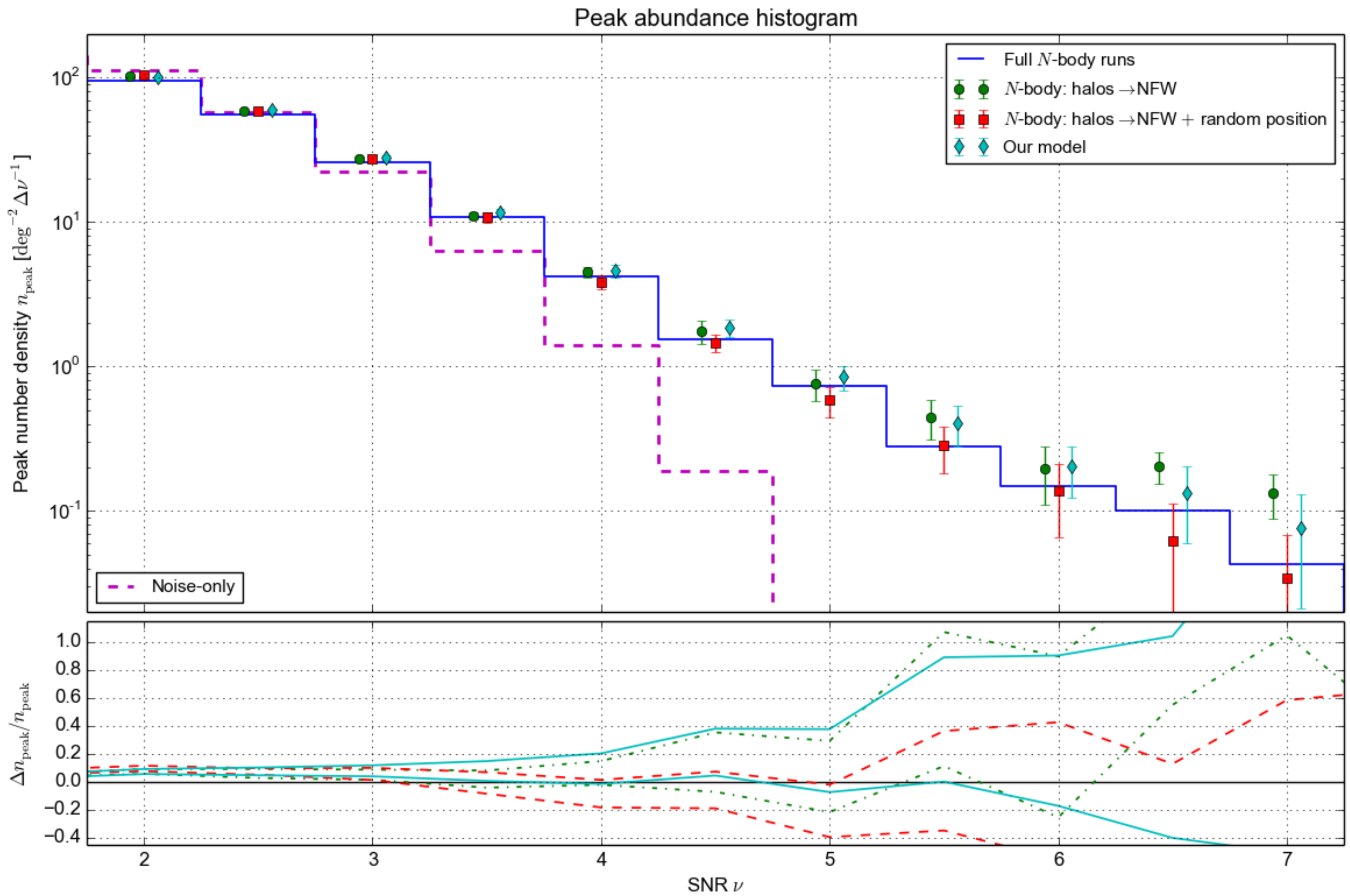
Lin, MK & Pires 2016

Fast simulations for WL peak counts

Hypotheses:

1. Clustering of halos not important for counting peaks (along los: Marian et al. 2013)
2. Unbound LSS does not contribute to WL peaks

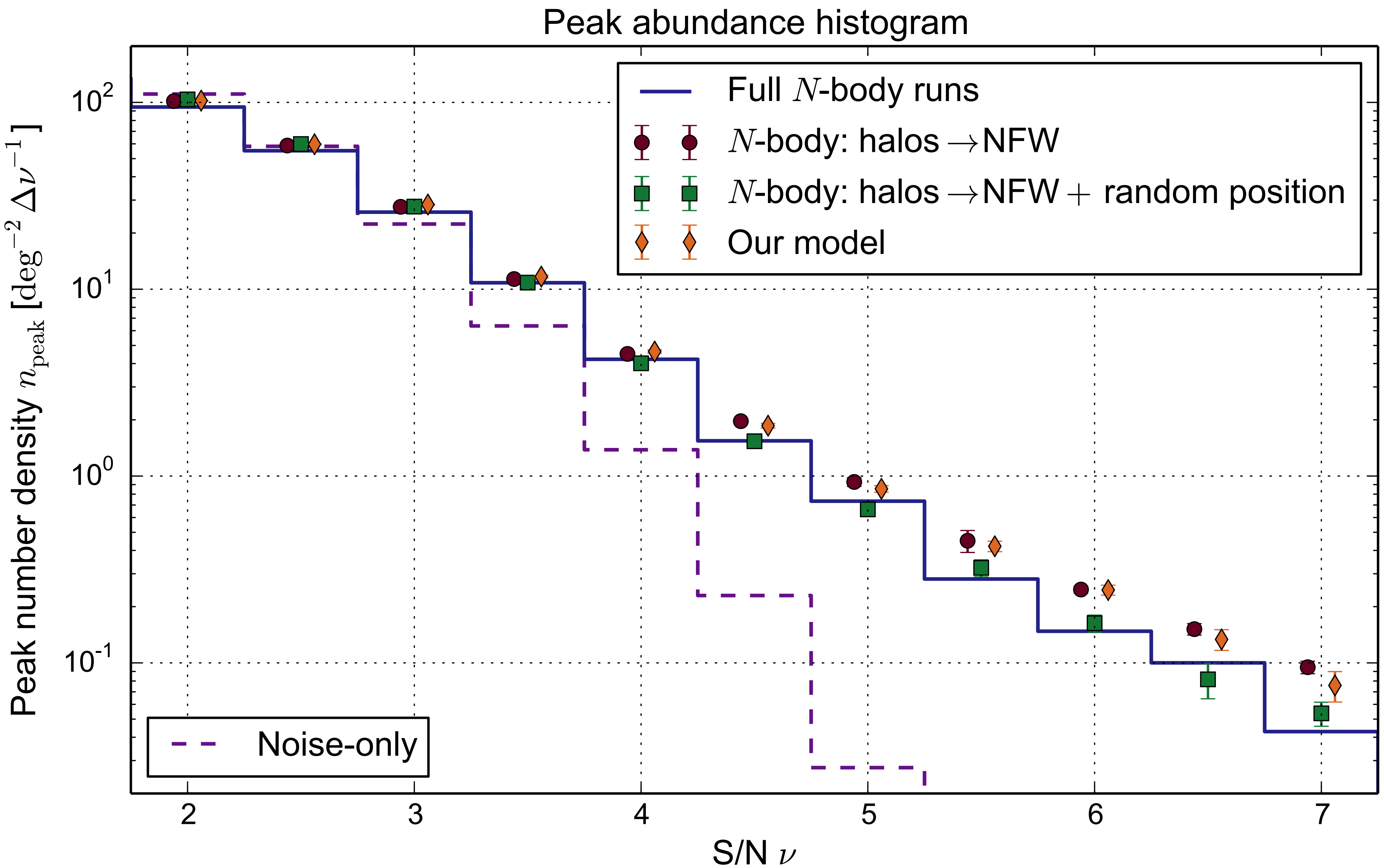
Test:



Field of view = 54 deg^2; 10 halo redshift bins from z = 0 to 1; galaxies on regular grid, $z_{\rm s}$ = 1.0

Fast simulations for WL peak counts

Test with larger field of view (but same for full N-body):



Lin & Kilbinger (2015a)

WL peak model properties

Our model is:

Fast.

Few sec for 25 deg^2 field on single CPU

Flexible.

Easy to include astrophysical & observational effects, non-Gaussian filters.

Full pdf:

Model is stochastic, drawing samples from distribution of the observables.

Constraining parameters

Bayes' theorem

Likelihood:
probability of data given
parameters and model

Prior

$$p(\boldsymbol{\pi}|\boldsymbol{x}, m) = \frac{L(\boldsymbol{x}|\boldsymbol{\pi}, m)P(\boldsymbol{\pi}|m)}{E(\boldsymbol{x}|m)}$$

Posterior:
probability of parameters
given data and model

Evidence

$\boldsymbol{\pi}$: parameters
$\boldsymbol{x}$: data
m : *model*

Parameter constraints = integrals over the posterior

$$\int \mathrm{d}^n\pi \, h(\boldsymbol{\pi}) p(\boldsymbol{\pi}|\boldsymbol{x}, m)$$

For example:

$$h(\boldsymbol{\pi}) = \boldsymbol{\pi} : \quad \text{mean}$$
$$h(\boldsymbol{\pi}) = 1_{68\%} : \quad 68\% \text{ credible region}$$

Approaches: Sampling (Monte-Carlo integration), Fisher-matrix approximation, frequentist evaluation, ABC, …

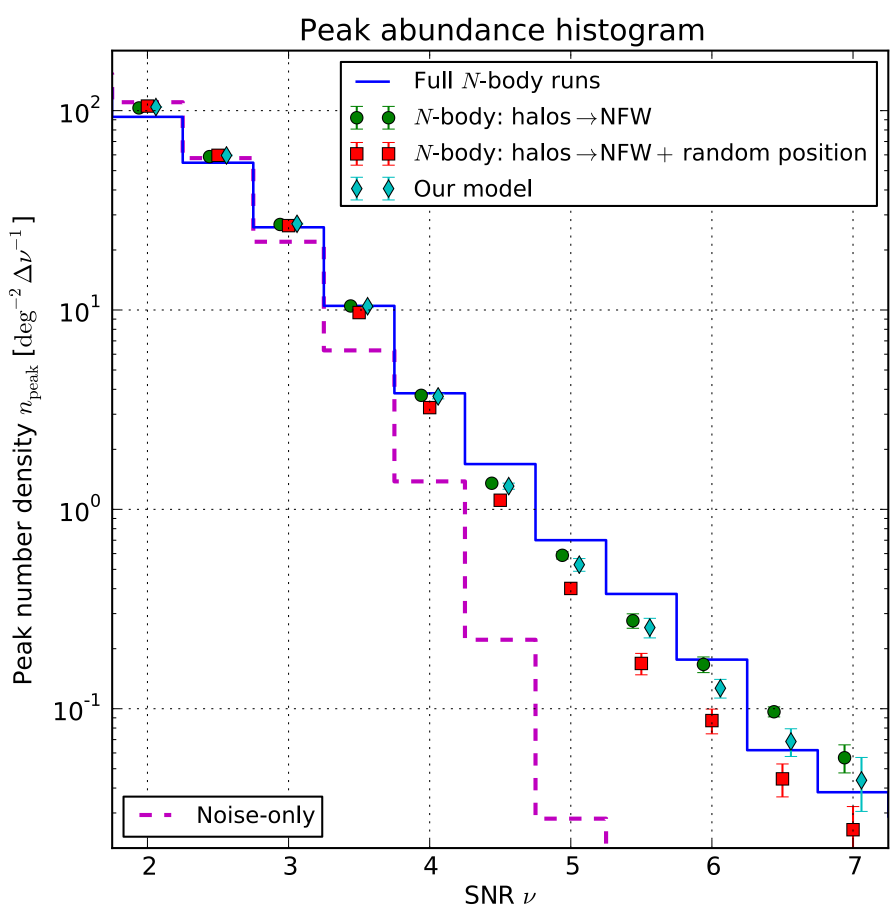
WL peak counts:
parameter constraint strategies

Data vector $x = x(t_i)$. Different cases:

- **Abundance** of peaks n_i as fct. of SNR ν (PDF; binned histogram) **or**
- **SNR values** ν_i at some percentile values of peak CDF)
  - with or without lower cut $\nu_{\min}$.

Lin & MK 2015a

x_i = abundance values …

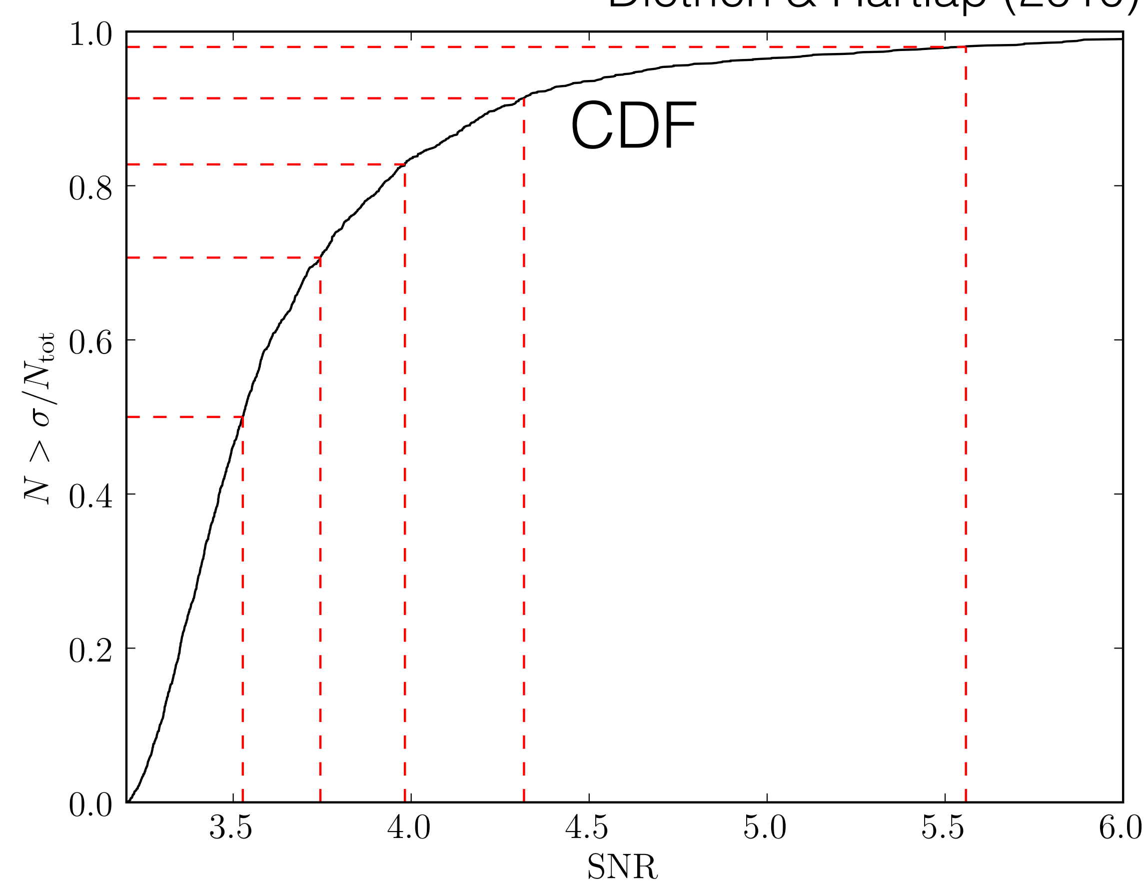


… for given SNR t_i

or

Dietrich & Hartlap (2010)

… for given percentiles t_i



x_i = SNR values …

WL peak counts: parameter constraint strategies

Data vector $\boldsymbol{x}$, model $\boldsymbol{\pi}$, covariance $\boldsymbol{C}$.

$$\Delta \boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x}^{\mathrm{mod}}(\boldsymbol{\pi}|m)$$
$$= \mathrm{data} - \mathrm{model}$$

Parameter constraints:

log-likelihood

Analytical likelihood (approximation)

- Gaussian

$$L_{\mathrm{cg}} \equiv \Delta \boldsymbol{x}^T(\boldsymbol{\pi})\, \widehat{\boldsymbol{C}^{-1}}(\boldsymbol{\pi}^{\mathrm{obs}})\, \Delta \boldsymbol{x}(\boldsymbol{\pi}),$$
$$L_{\mathrm{svg}} \equiv \Delta \boldsymbol{x}^T(\boldsymbol{\pi})\, \widehat{\boldsymbol{C}^{-1}}(\boldsymbol{\pi})\, \Delta \boldsymbol{x}(\boldsymbol{\pi}), \quad \text{and}$$
$$L_{\mathrm{vg}} \equiv \ln\left[\det \widehat{\boldsymbol{C}}(\boldsymbol{\pi})\right] + \Delta \boldsymbol{x}^T(\boldsymbol{\pi})\, \widehat{\boldsymbol{C}^{-1}}(\boldsymbol{\pi})\, \Delta \boldsymbol{x}(\boldsymbol{\pi}).$$

- Copula

$$P(\boldsymbol{x}) = \phi(\boldsymbol{q}) \frac{P_1(x_1) \cdots P_d(x_d)}{\phi_1(q_1) \cdots \phi_d(q_d)}.$$

Non-analytical likelihood (exact)

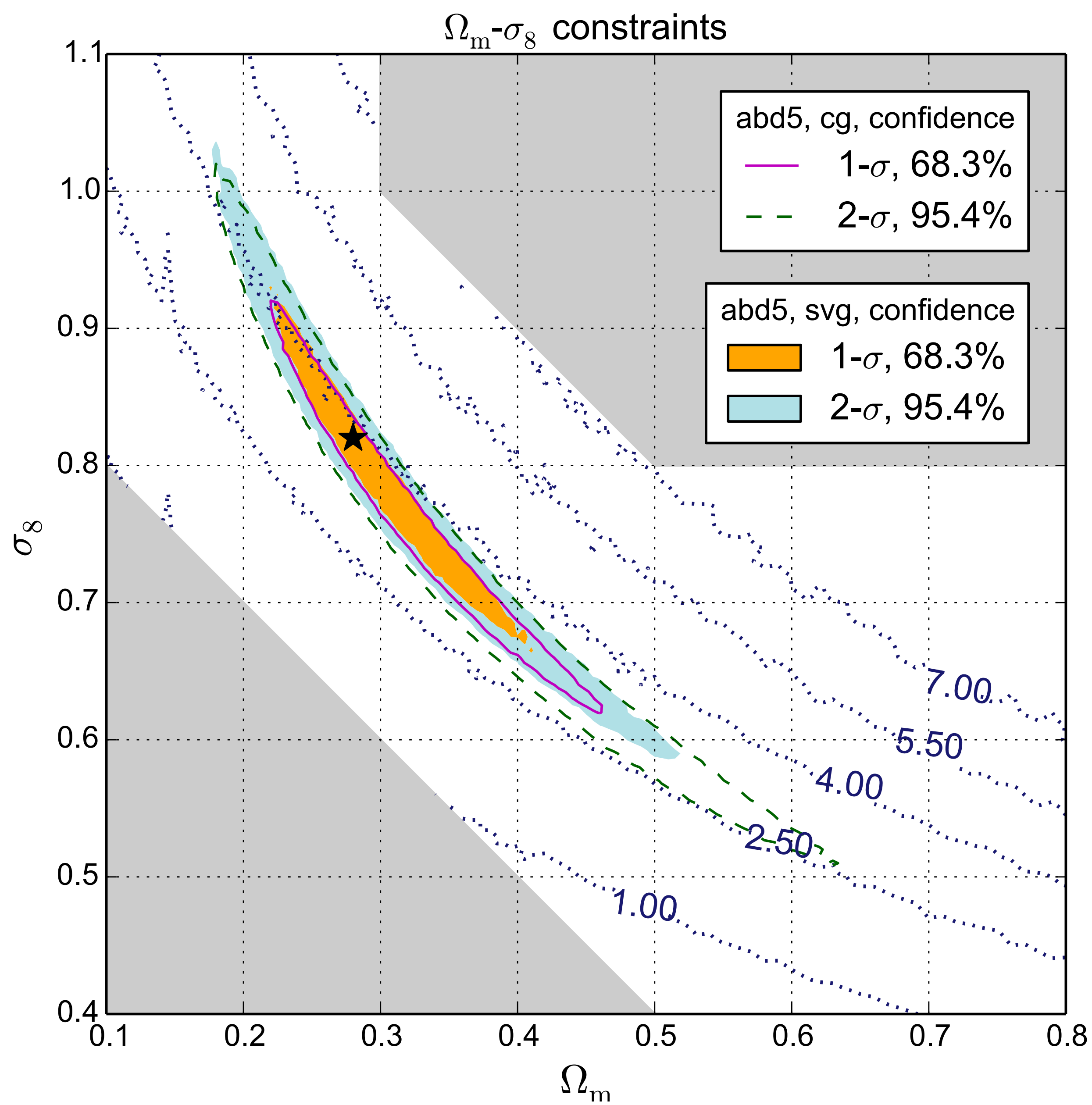
- \`True' likelihood (+ KDE)

$$L(\boldsymbol{x}|\boldsymbol{\pi}, m) \approx \frac{1}{N} \sum_{k=1}^{N} K\left(\boldsymbol{x} - x^{(k)}\right)$$

Likelihood-free

- ABC

Parameter constraints: Gaussian



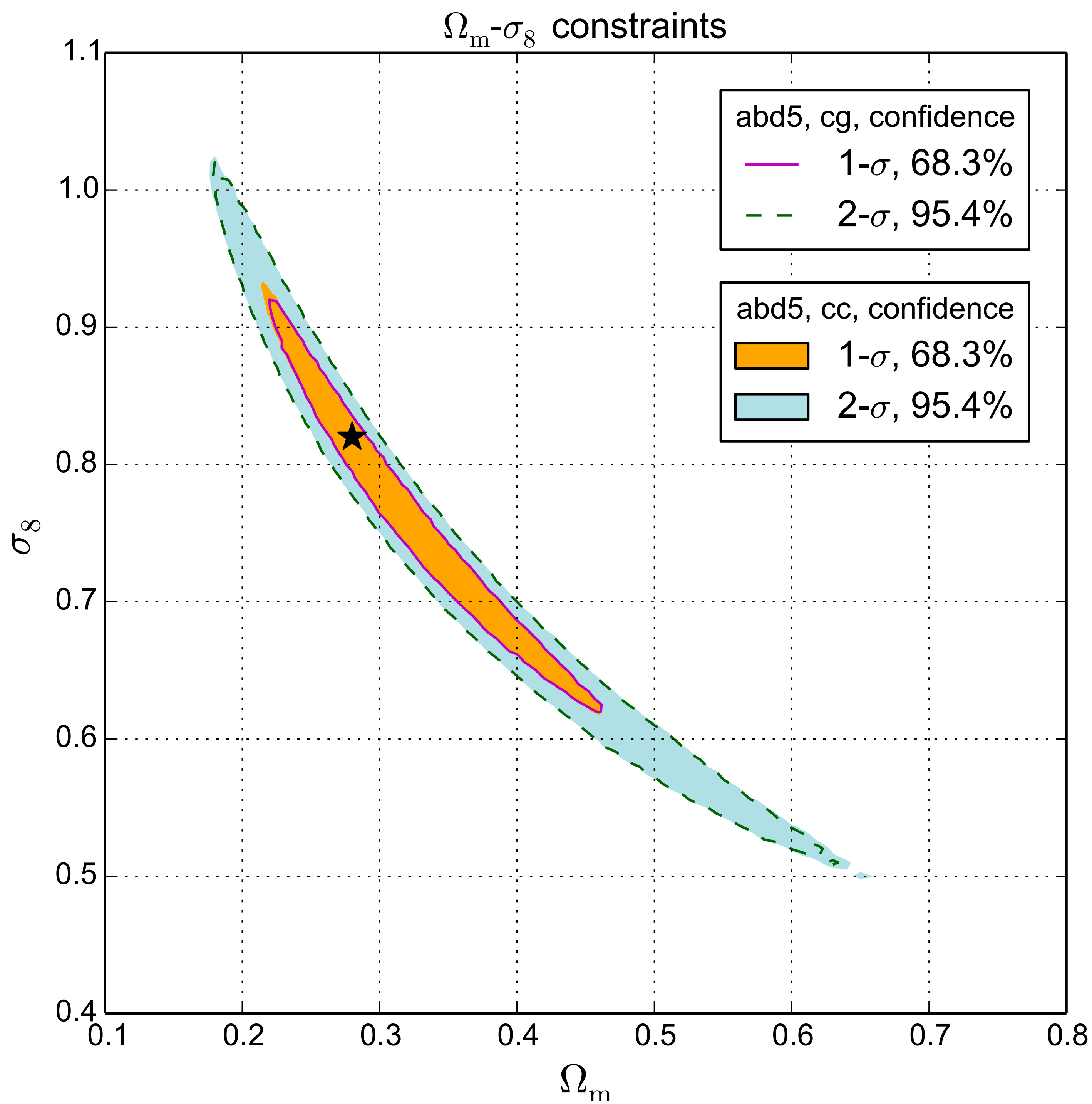
$$L_{\rm cg} \equiv \Delta \boldsymbol{x}^T(\boldsymbol{\pi})\, \widehat{\boldsymbol{C}^{-1}}(\boldsymbol{\pi}^{\rm obs})\, \Delta \boldsymbol{x}(\boldsymbol{\pi}),$$

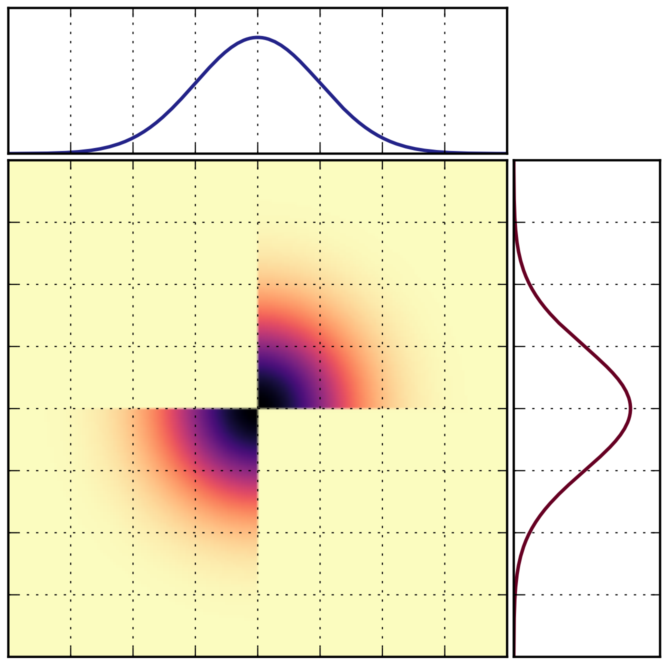
$$L_{\rm svg} \equiv \Delta \boldsymbol{x}^T(\boldsymbol{\pi})\, \widehat{\boldsymbol{C}^{-1}}(\boldsymbol{\pi})\, \Delta \boldsymbol{x}(\boldsymbol{\pi}), \quad \text{and}$$

$$L_{\rm vg} \equiv \ln\left[\det \widehat{\boldsymbol{C}}(\boldsymbol{\pi})\right] + \Delta \boldsymbol{x}^T(\boldsymbol{\pi})\, \widehat{\boldsymbol{C}^{-1}}(\boldsymbol{\pi})\, \Delta \boldsymbol{x}(\boldsymbol{\pi}).$$

Cosmology-dependent covariance [(s)vg] reduces error area by 20%.

Parameter constraints: Copula

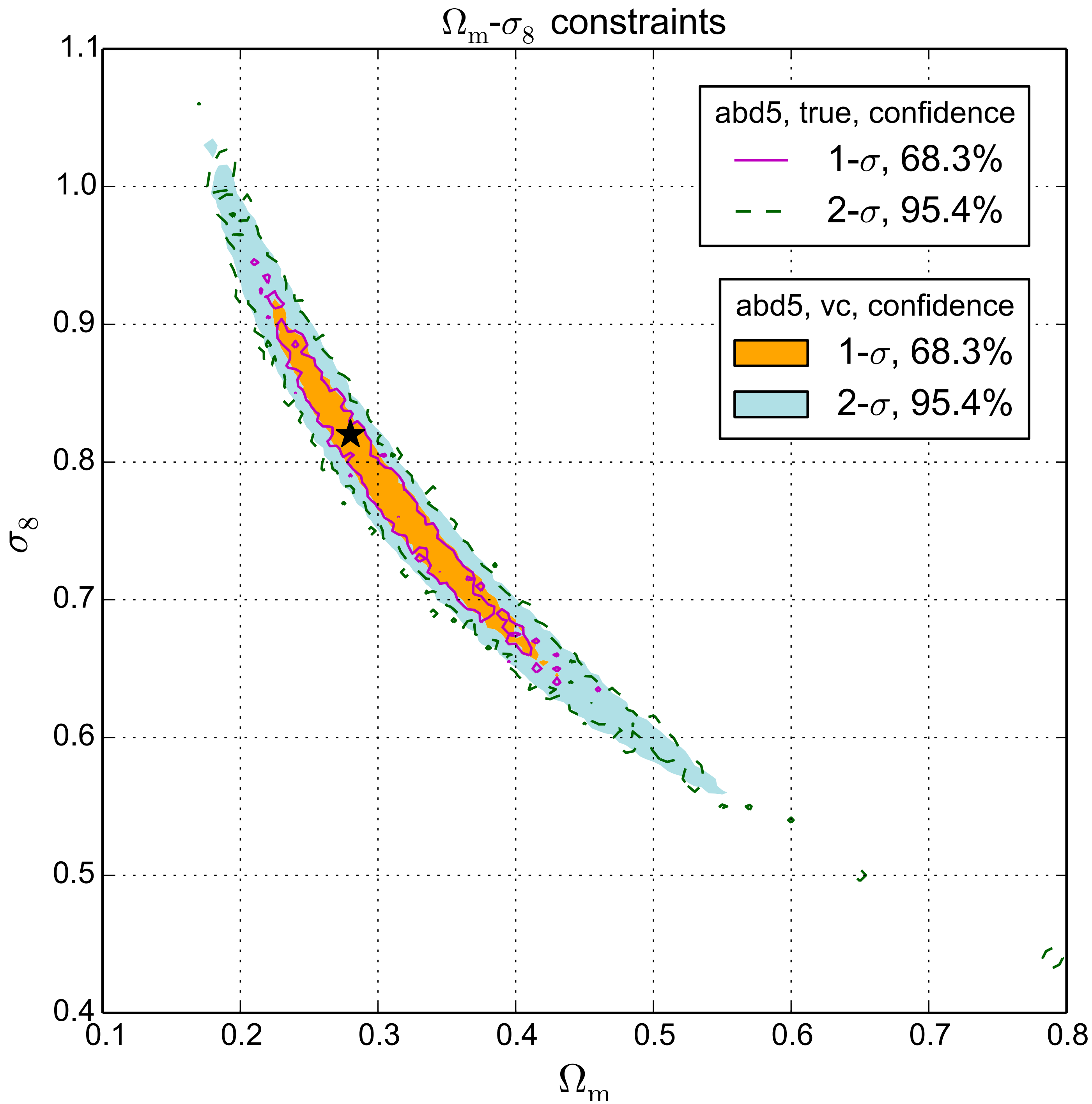




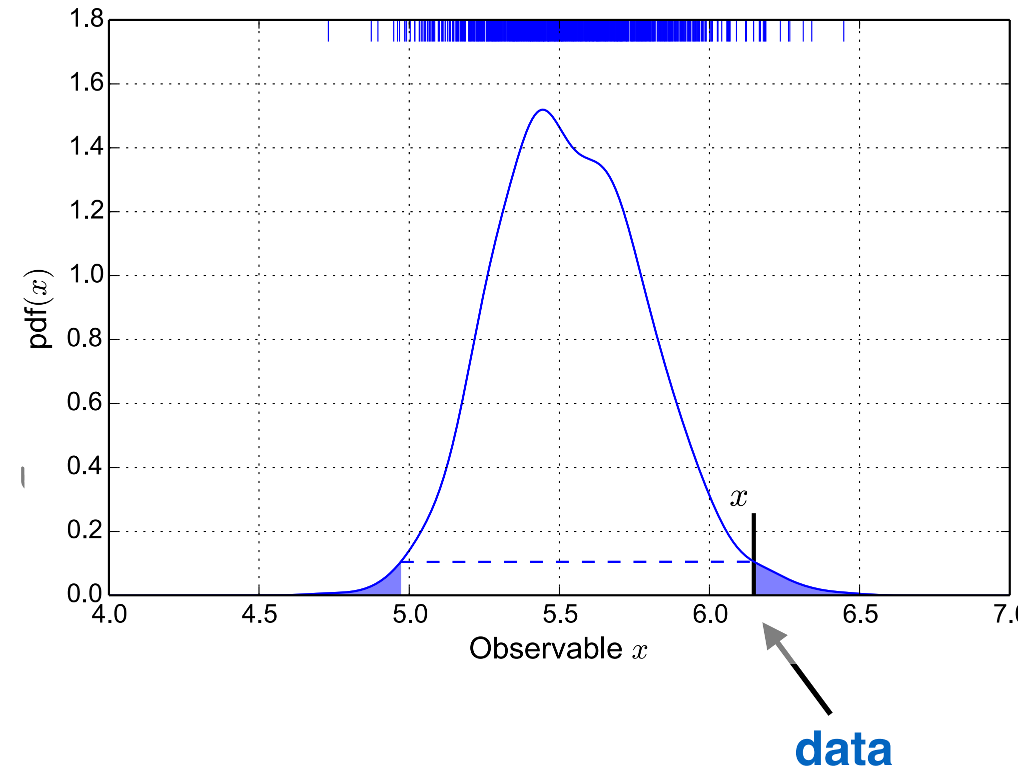
$$P(\boldsymbol{x}) = \phi(\boldsymbol{q}) \frac{P_1(x_1) \cdots P_d(x_d)}{\phi_1(q_1) \cdots \phi_d(q_d)}.$$

Copula provides similar constraints than Gaussian.

Parameter constraints: true likelihood (+ KDE)



models for given parameter π
= samples under distribution, from stochastic process



sample density = likelihood probability

$$L(\boldsymbol{x}|\boldsymbol{\pi}, m) \approx \frac{1}{N} \sum_{k=1}^{N} K(\boldsymbol{x} - \boldsymbol{x}_k^{\rm mod}(\boldsymbol{\pi}|m))$$

(Kernel density estimate)

True likelihood similar to Copula with varying covariance

Approximate Bayesian Computation (ABC)

Likelihood:
probability of data given parameters and model

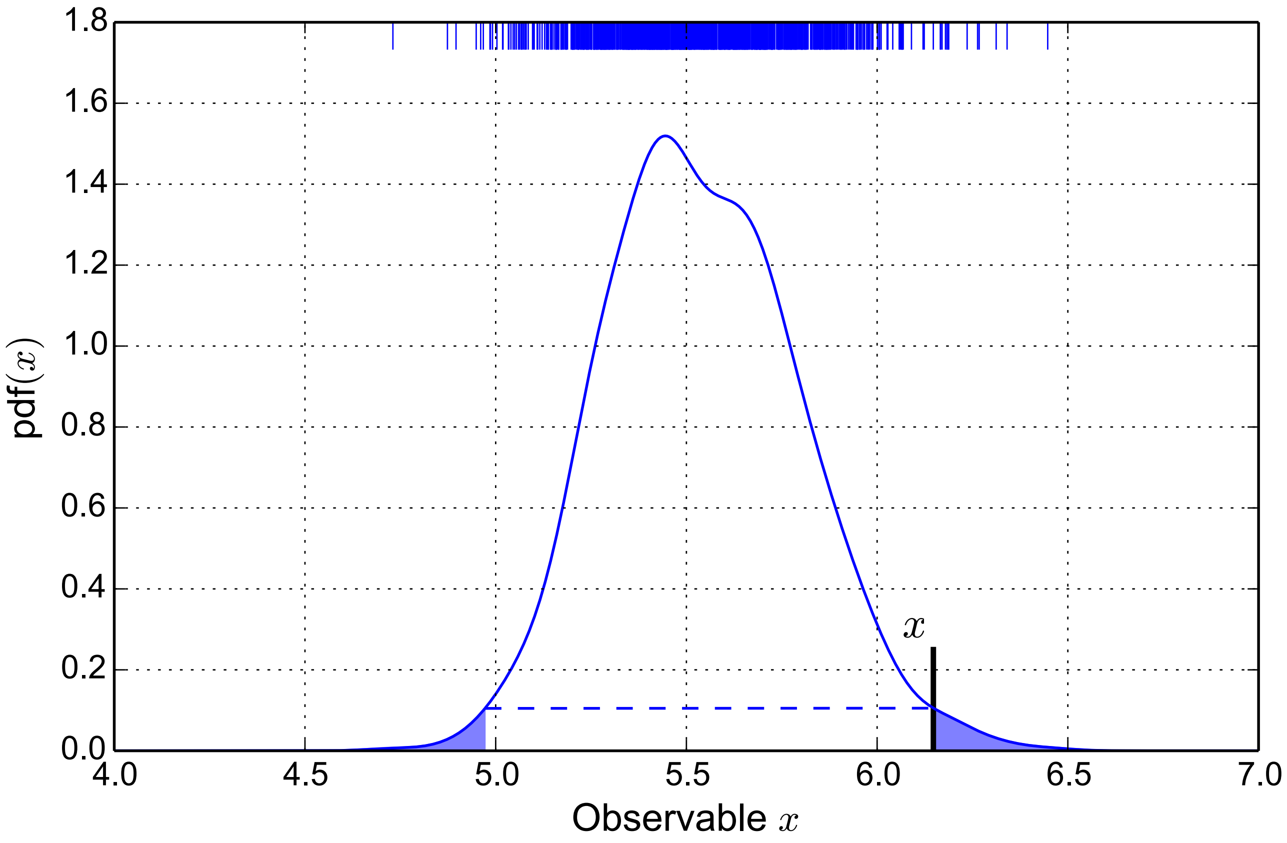
$$p(\boldsymbol{\pi}|\boldsymbol{x}, m) = \frac{L(\boldsymbol{x}|\boldsymbol{\pi}, m) P(\boldsymbol{\pi}|m)}{E(\boldsymbol{x}|m)}$$

$\boldsymbol{\pi}$: parameters
$\boldsymbol{x}$: data
m : *model*

Likelihood: how likely is it that model prediction $\boldsymbol{x}^{\rm mod}(\boldsymbol{\pi})$ reproduces data $\boldsymbol{x}$?

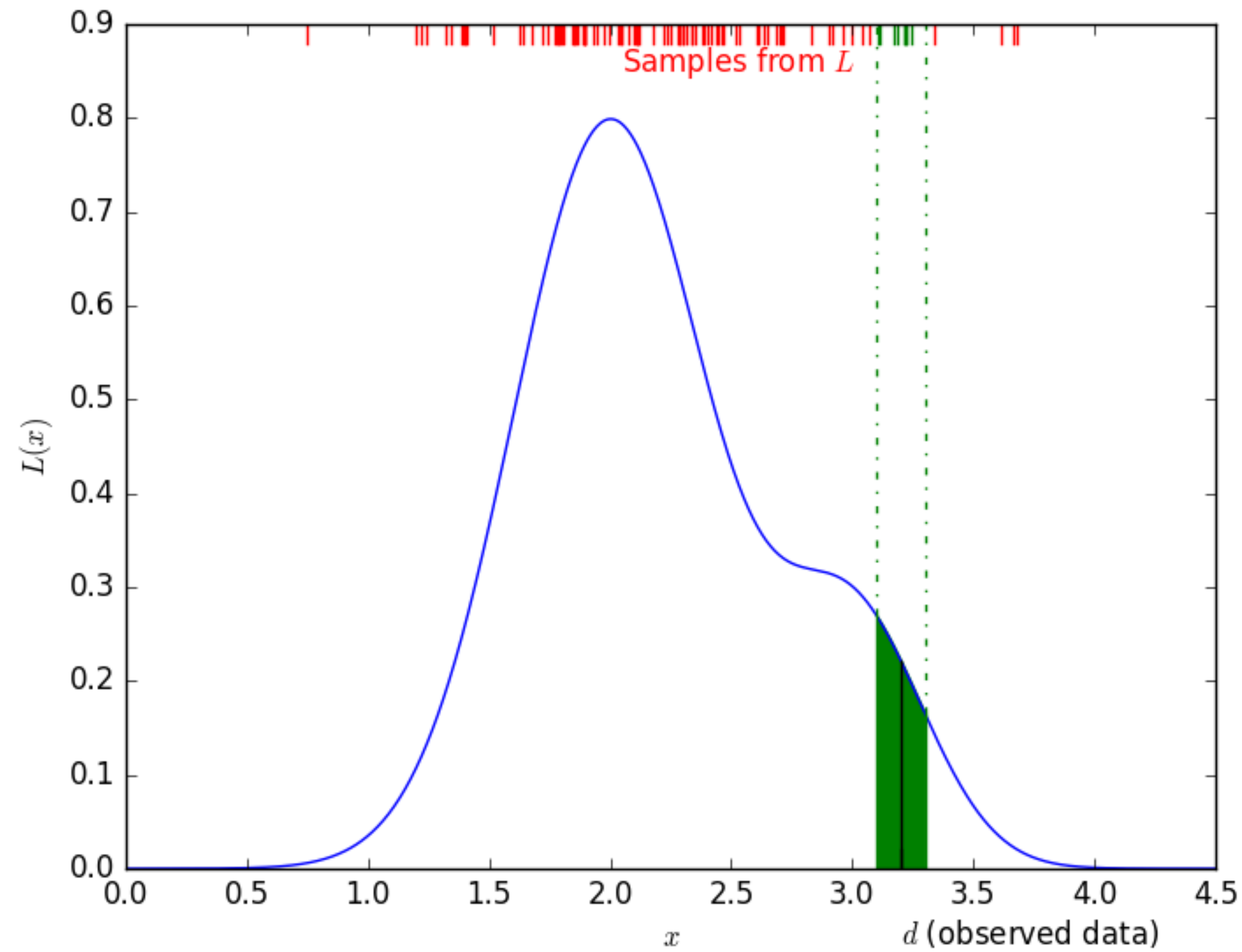
Classical answer: evaluate function L at $\boldsymbol{x}$.

Alternative: compute fraction of models that are equal to the data $\boldsymbol{x}$.



Probability = *p*/*N* in frequentist sense.

Magic: Don't need to sample *N* models. **One** per parameter π is sufficient with accept-reject algorithm.



ABC can be performed if:

- it is possible and easy to sample from *L*

ABC is useful when:

- functional form of *L* is unknown
- evaluation of *L* is expensive
- model is intrinsically stochastic

Example: let's make soup.

Goal: Determine ingredients from final result.
Model physical processes? Complicated.

Example: let's make soup.

Goal: Determine ingredients from final result.
Model physical processes? Complicated.
Easier: Make lots of soups with different ingredients, compare.

Example: let's make soup.

Questions:

- What aspect of data and simulations do we compare? (**summary statistic**)
- How do we compare? (**metric, distance**)
- When do we accept? (**tolerance**)

Parameter constraints: ABC

- Summary statistic

  $\boldsymbol{s} = \boldsymbol{x}$ (data vector for 2 cases)

- Metric D: two cases

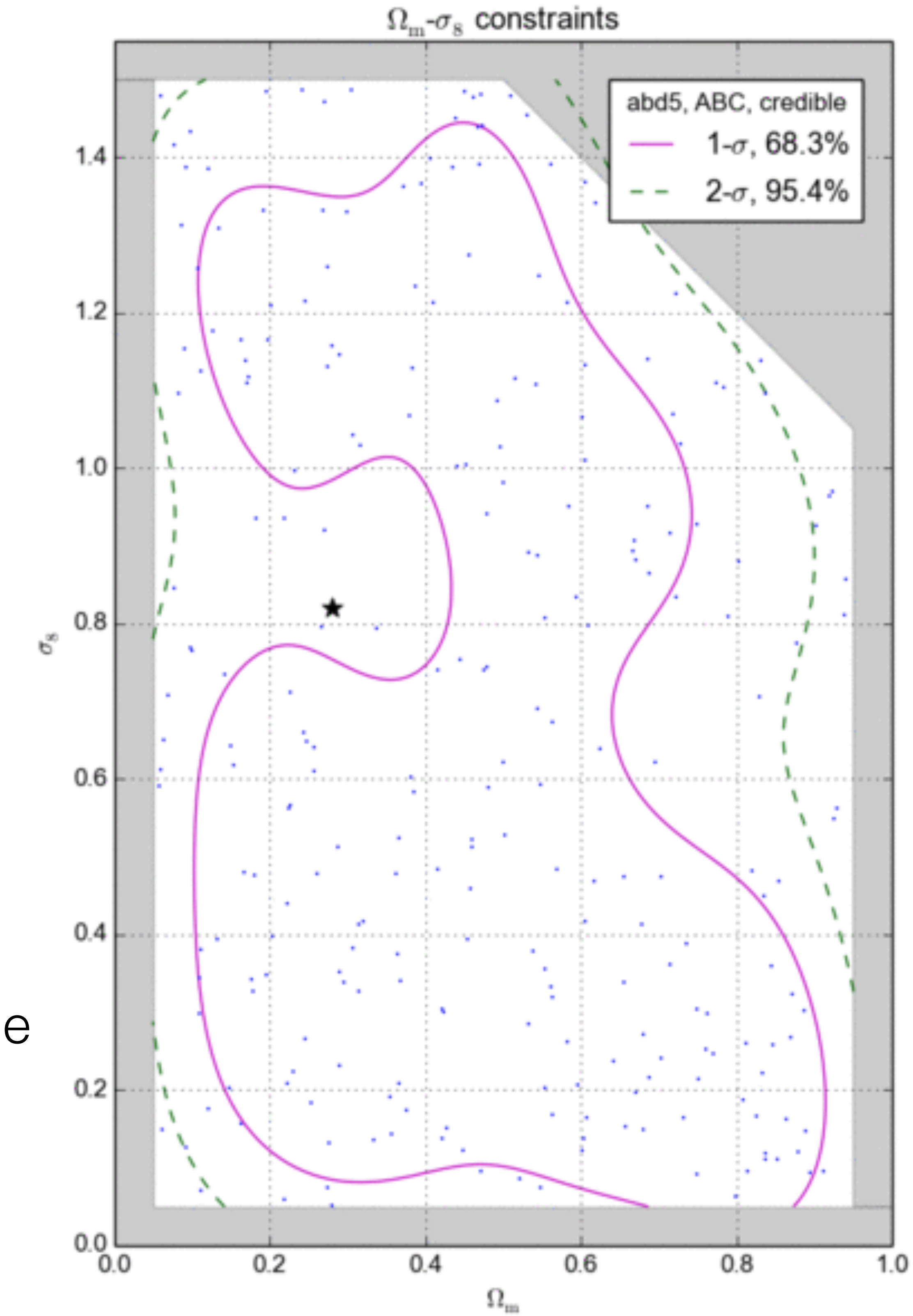
$$D_1\left(\boldsymbol{x}, \boldsymbol{x}^{\rm obs}\right) \equiv \sqrt{\sum_i \frac{\left(x_i - x_i^{\rm obs}\right)^2}{C_{ii}}},$$

$$D_2\left(\boldsymbol{x}, \boldsymbol{x}^{\rm obs}\right) \equiv \sqrt{\left(\boldsymbol{x} - \boldsymbol{x}^{\rm obs}\right)^T \boldsymbol{C}^{-1} \left(\boldsymbol{x} - \boldsymbol{x}^{\rm obs}\right)},$$

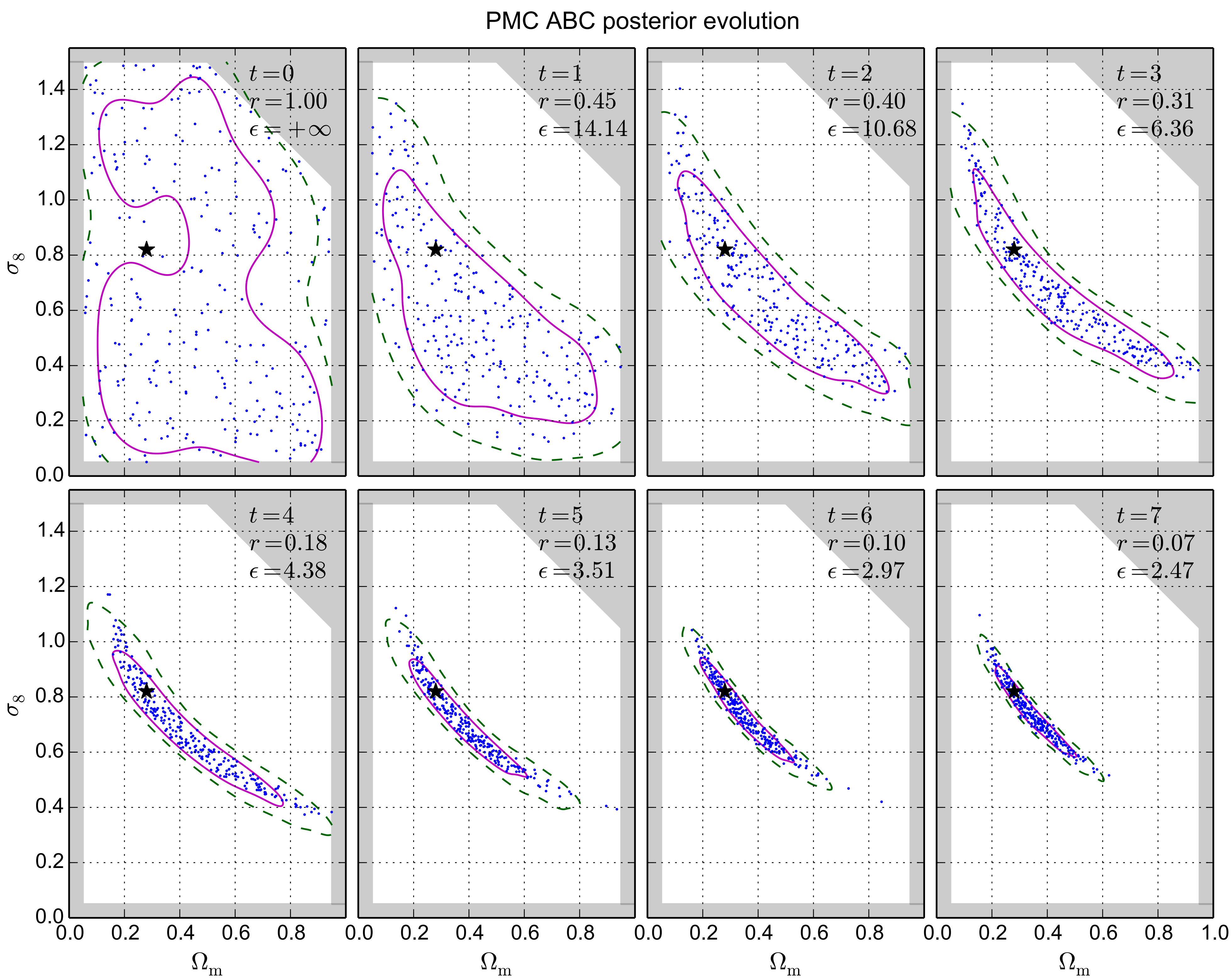
D_1 in Lin & MK 2015b

D_1 + D_2 in Lin, MK & Pires 2016

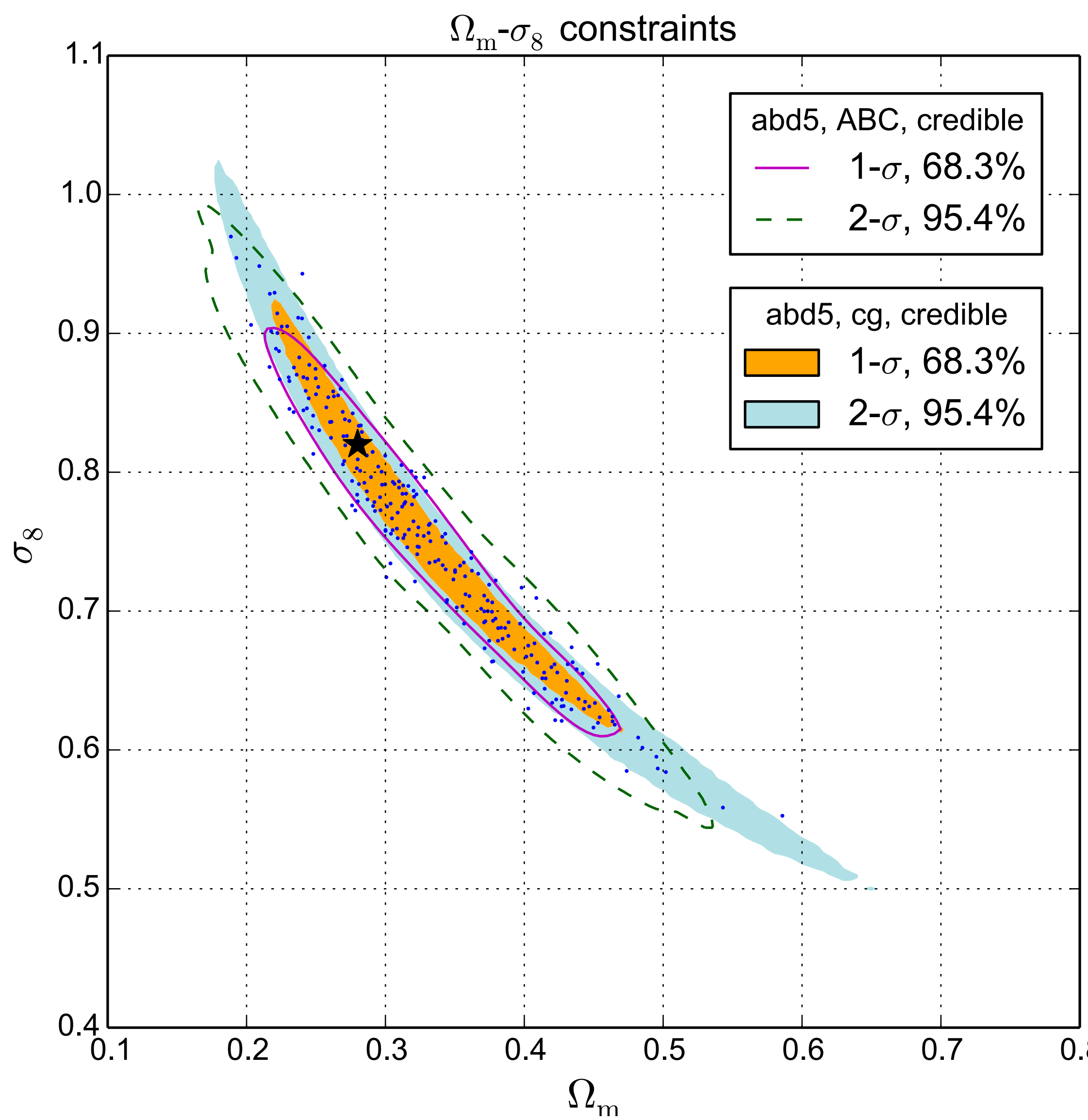
- ABC algorithm: iterative importance sampling (PMC) with decreasing tolerance

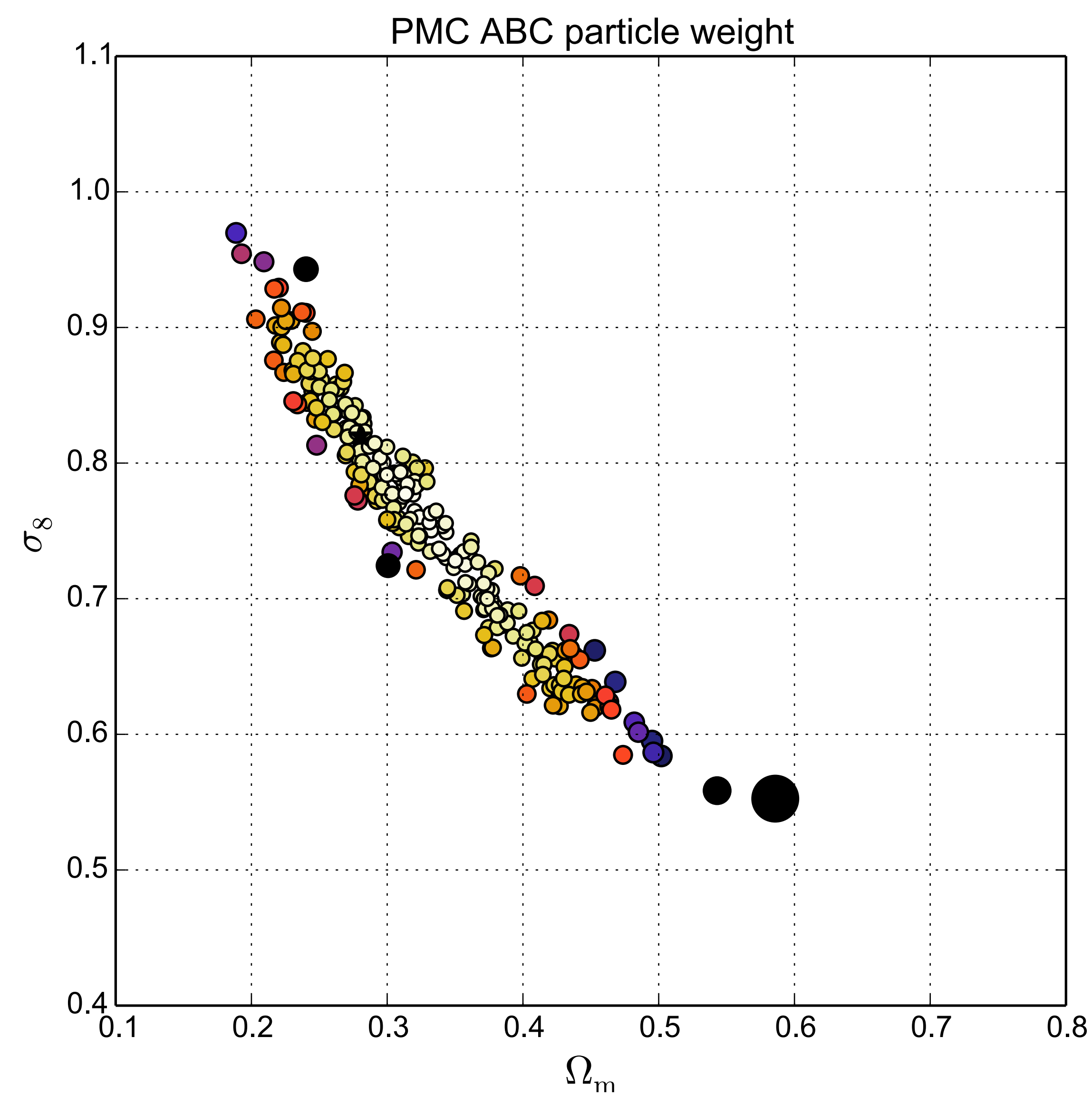


Parameter constraints: ABC



Parameter constraints: comparison





ABC wider but less elongated and less bent contours than Gaussian with const cov.
KDE smoothing effect?

Parameter constraints: comparison

correlated data points

difference between metrics

Filter	Constraints	$\Delta\Sigma_8$ (contour width)	FoM (contour area)
Gaussian	Likelihood	0.046	20.7
Gaussian	ABC, D_1	0.043	16.3
Gaussian	ABC, D_2	0.059	11.7
Starlet	Likelihood	0.054	23.4
Starlet	ABC, D_1	0.050	15.5
Starlet	ABC, D_2	0.054	15.7
$M_{\rm ap}$ tanh	ABC, D_1	0.037	19.4
$M_{\rm ap}$ tanh	ABC, D_2	0.043	15.5
MRLENS	Likelihood	0.046	16.2
MRLENS	ABC, D_1	0.045	11.5
MRLENS	ABC, D_2	0.045	12.5

smaller larger
is better

$$\Sigma_8 \equiv \left(\frac{\Omega_{\rm m} + \beta}{1 - \alpha}\right)^{1-\alpha} \left(\frac{\sigma_8}{\alpha}\right)^{\alpha}$$

Code(s)

Fast weak-lensing peak counts modelling in C
with PMC ABC

CAMELUS

Camelus@GitHub

https://github.com/Linc-tw/camelus

uses

nicaea:
NumerIcal Cosmology And
lEnsing cAlculations

cosmostat.org/software/nicaea

implements?

part of

CosmoPMC

www2.iap.fr/users/kilbinge/CosmoPMC

Summary

- Higher-order statistics in WL tighten constraints in CFHTLenS
- WL peak counts new, promising non-Gaussian probe for cosmology
- ABC more and more popular sampling method in astrophysics and cosmology
    - Cameron & Pettitt (2012): galaxy morphology
    - Weyant, Schafer, & Wood-Vasey (2013): SNIa
    - Ishida et al. (2015; cosmoABC): galaxy cluster counts
    - Akeret et al. (2015): image simulations
    - Lin & MK (2015b; camelus): WL peak counts
    - Killedar et al. (2015): strong lensing
    - Hahn et al. (2016): HOD
    - Jennings & Madigan (2017; astroABC): SNIa

Backup slides

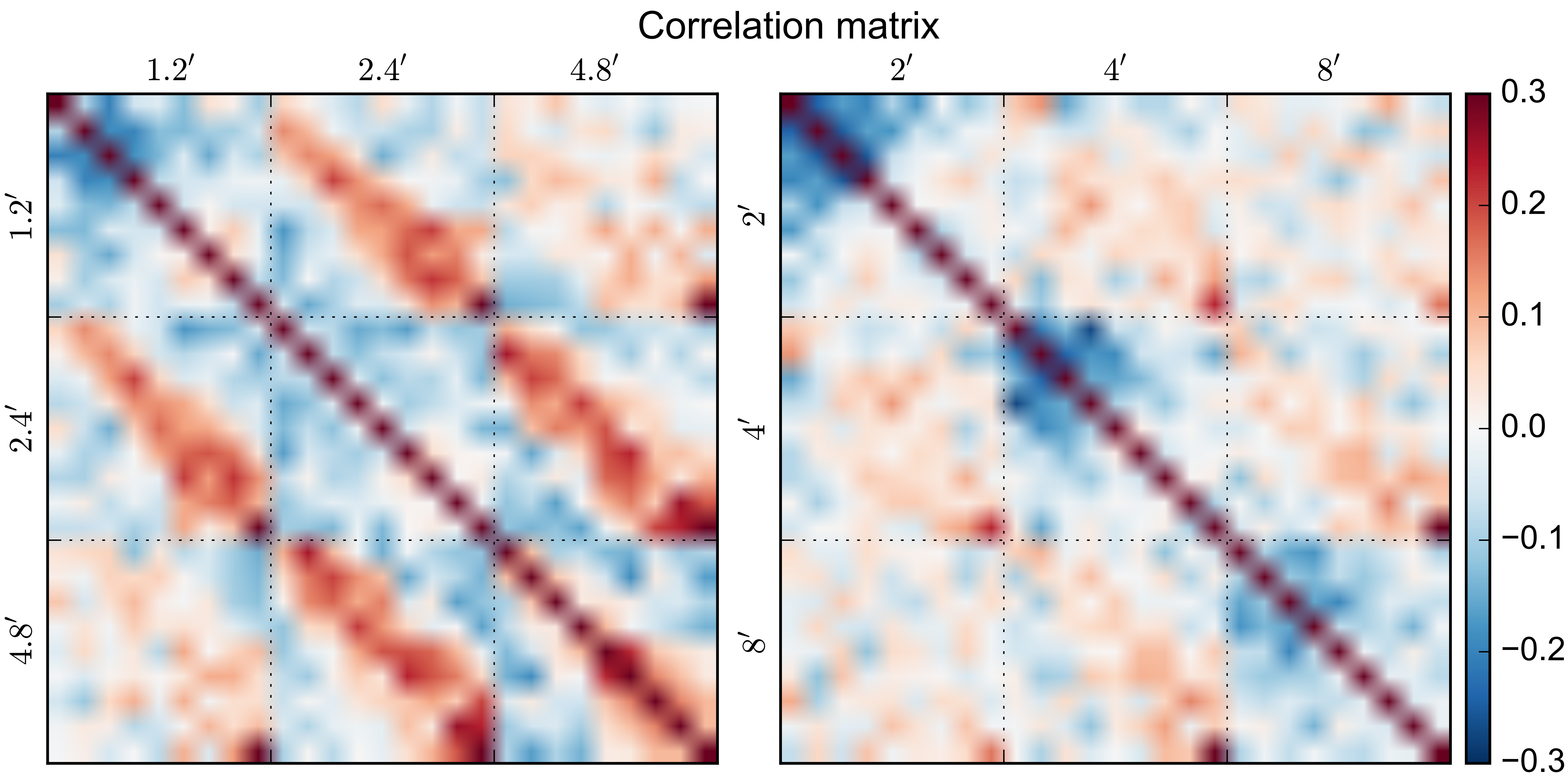
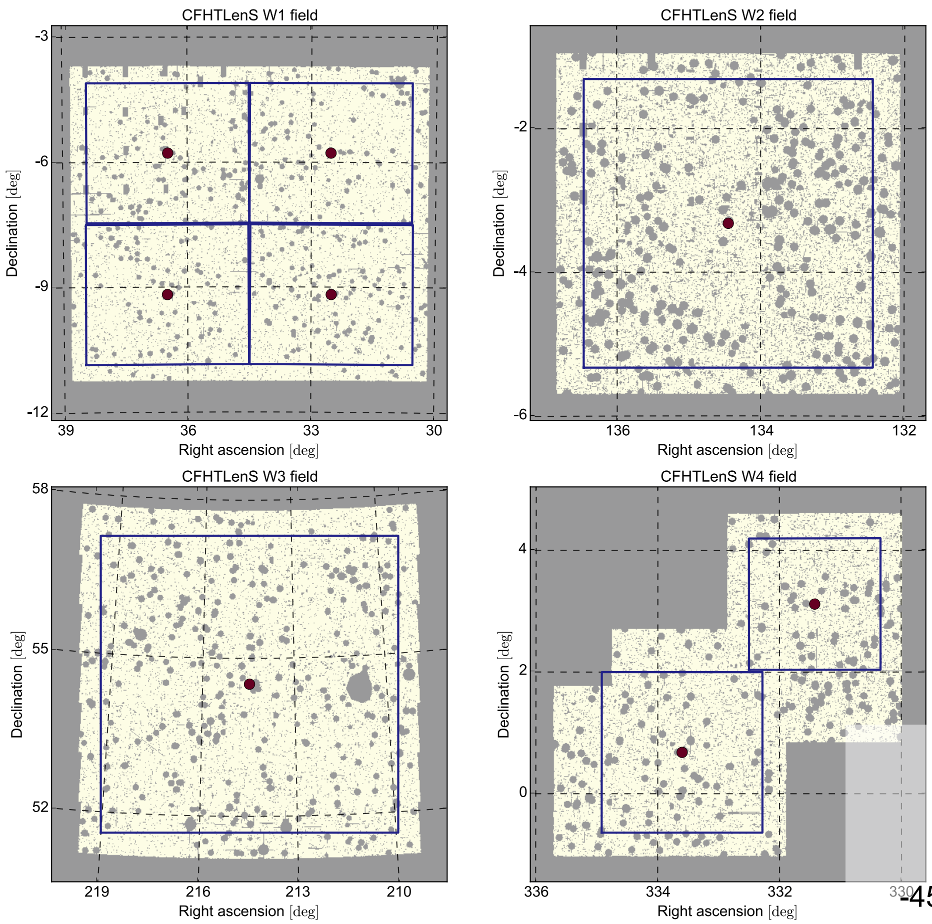
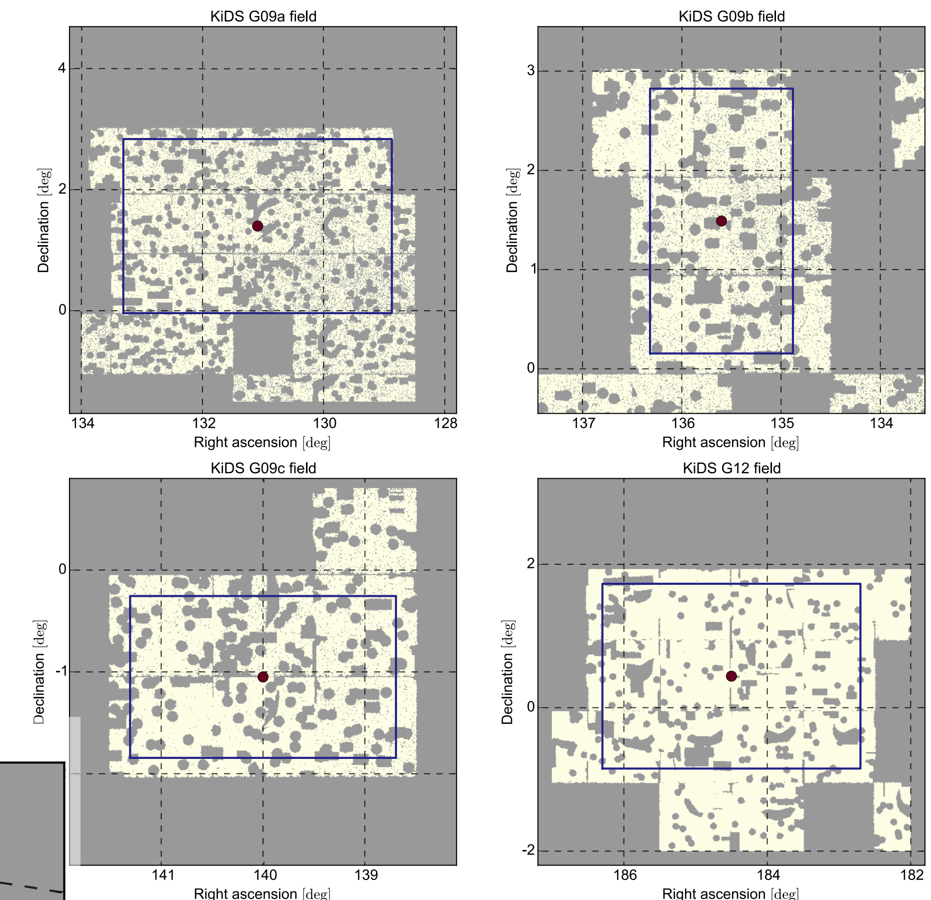


FIGURE 8.9: Correlation coefficient matrices under the input cosmology. Left panel: the Gaussian case with $\theta_{\rm ker} = 1.2$, 2.4, and 4.8 arcmin. Right panel: the starlet case with $\theta_{\rm ker} = 2$, 4, and 8 arcmin. Each of the 3×3 blocks corresponds to the correlations between two filter scales. With each block, the S/N bins are [1, 1.5, 2, . . ., 5, $+\infty$[. The data vector by starlet is less correlated.

Data

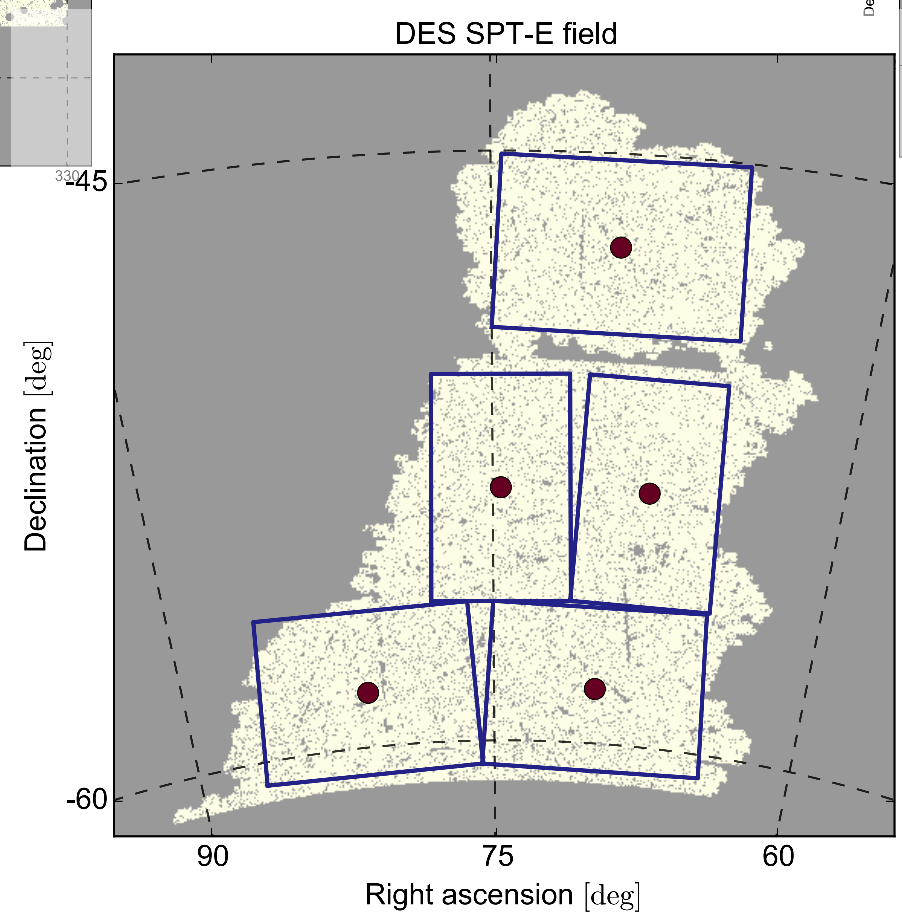


CFHTLenS

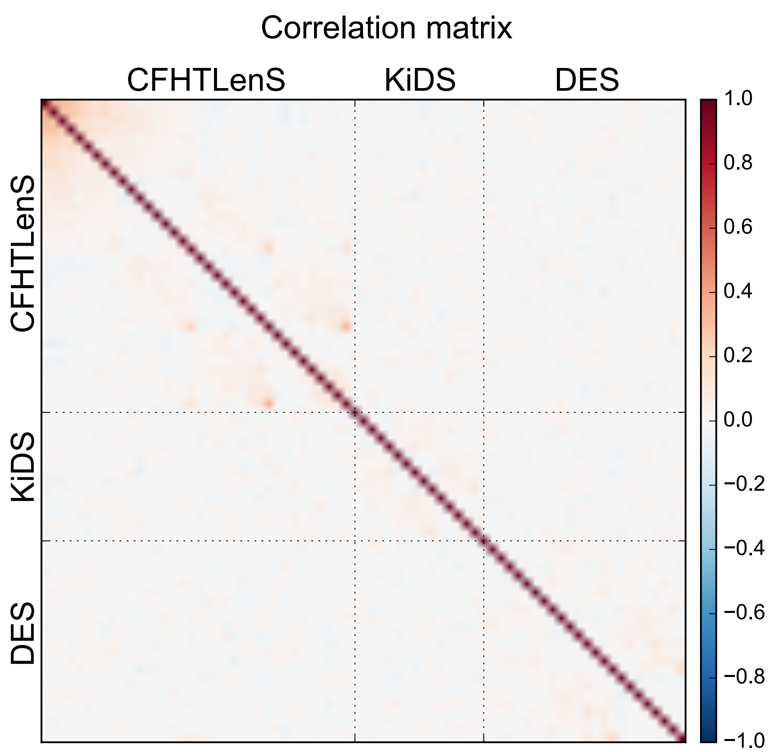


KiDS DR1/2

DES SV



Data


Correlation matrix
CFHTLenS
KiDS
DES
CFHTLenS
KiDS
DES
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
−1.0

Haiman group paper on covariance