

Weak Gravitational Lensing

Part II/II

Martin Kilbinger

CEA Saclay, Irfu/SAp - AIM, CosmoStat; IAP

Euclid Summer School, Roscoff
August 2018

`martin.kilbinger@cea.fr`

`www.cosmostat.org/kilbinger`

Slides: `http://www.cosmostat.org/events/ecole18`



`@energie_sombre`

`#EuclidRoscoff2018`



COSMOSTAT



université
PARIS-SACLAY



Overview

Part II day 1: E- and B-modes

- Very brief reminders from day I
- E-/B-mode decomposition recap
- E-/B-mode estimators
- Galaxy-galaxy lensing: motivation

Part II day 2: Shear estimation

- Galaxy-galaxy lensing in detail
- Back to the aperture mass: Filter function relation
- Spherical-sky lensing projections
- Shear calibration

Part II day 3: Cosmological parameter estimation

- Numerical simulations
- Covariance estimation
- Likelihood and parameter estimation

Reminder from last year ...

Books, Reviews and Lecture Notes

- Bartelmann & Schneider 2001, review **Weak gravitational lensing**, Phys. Rep., 340, 297 arXiv:9912508
- Kochanek, Schneider & Wambsganss 2004, book (Saas Fee) **Gravitational lensing: Strong, weak & micro**. Download Part I (Introduction) and Part III (Weak lensing) from my homepage <http://www.cosmostat.org/people/kilbinger>.
- Kilbinger 2015, review **Cosmology from cosmic shear observations** Reports on Progress in Physics, 78, 086901, arXiv:1411.0155
- Bartelmann & Maturi 2017, review **Weak gravitational lensing**, Scholarpedia 12(1):32440, arXiv:1612.06535
- Henk Hoekstra 2013, lecture notes (Varenna) arXiv:1312.5981
- Sarah Bridle 2014, lecture videos (Saas Fee) <http://archiveweb.epfl.ch/saasfee2014.epfl.ch/page-110036-en.html>
- Alan Heavens, 2015, lecture notes (Rio de Janeiro) www.on.br/cce/2015/br/arq/Heavens_Lecture_4.pdf

Science with gravitational lensing

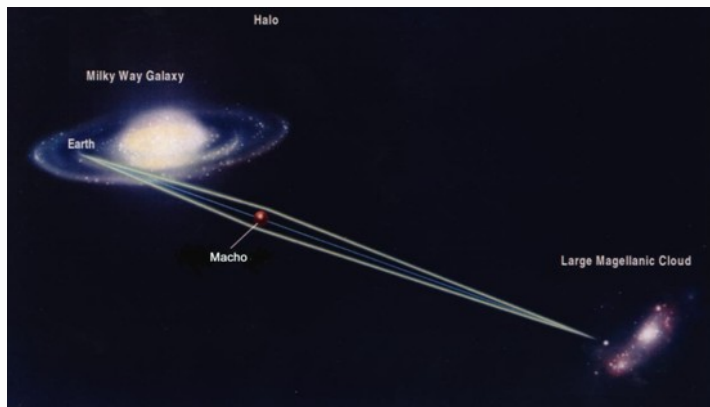
What has gravitational lensing ever done for us?



Science with gravitational lensing

Outstanding results

Dark matter is not in form of massive compact objects (MACHOs).
Microlensing rules out objects between 10^{-7} and few $10 M_{\odot}$.



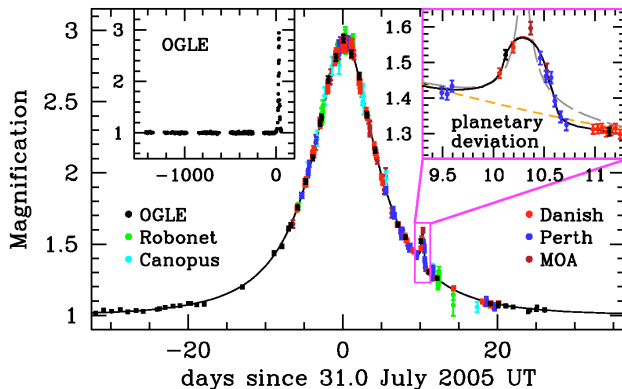
[Takahiro Sumi, Nagoya University]

Science with gravitational lensing

Outstanding results

Detection of Earth-like exoplanets with microlensing.

Masses and distances to host star similar to Earth.



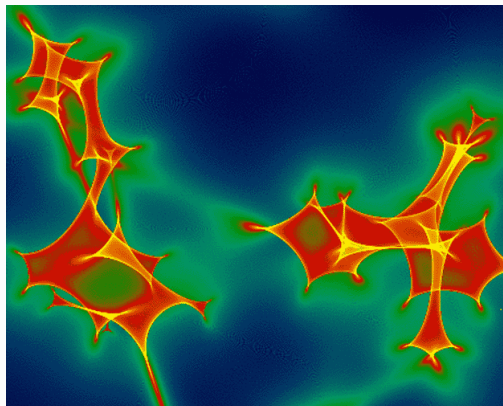
(Beaulieu et al. 2006)

Science with gravitational lensing

Outstanding results

Structure of QSO inner emission regions.

Microlensing by stars in lens galaxies.



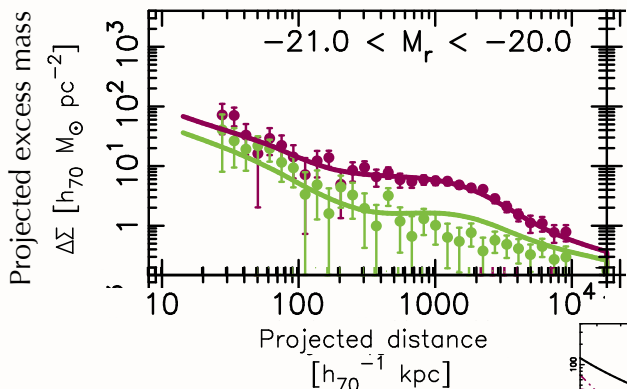
[J. Wambsganss]

Science with gravitational lensing

Outstanding results

Dark matter profiles in outskirts of galaxies.

Measuring halo mass to very large galactic scales.



(Velandier et al. 2014)

Science with gravitational lensing

Outstanding results

Galaxy clusters are dominated by dark matter.

Bullet cluster and others: bulk of mass is collisionless.



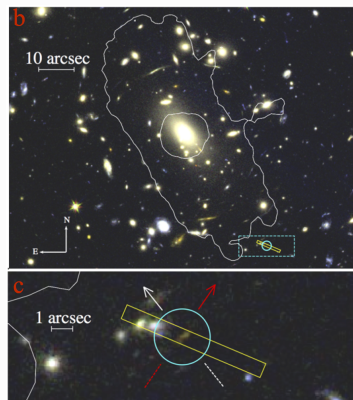
(Clowe et al. 2006)

Science with gravitational lensing

Outstanding results

Observation of very-high ($z \geq 7$) galaxies.

Galaxy clusters as “natural telescopes”.



(Hoag et al. 2017)

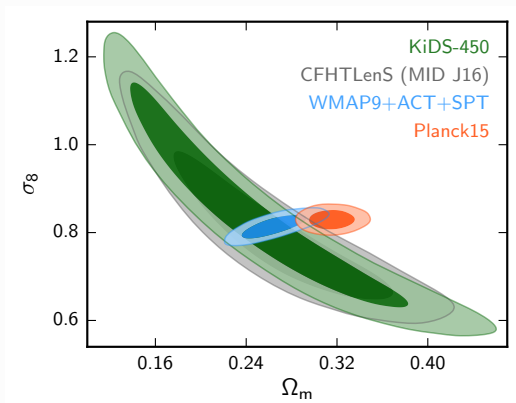
Science with gravitational lensing

Outstanding results

Hints of inconsistency of our cosmological model at low and high z ?

Planck and WL in tension? Also WL cluster masses for Planck SZ clusters;

H_0 from cepheids + SL.



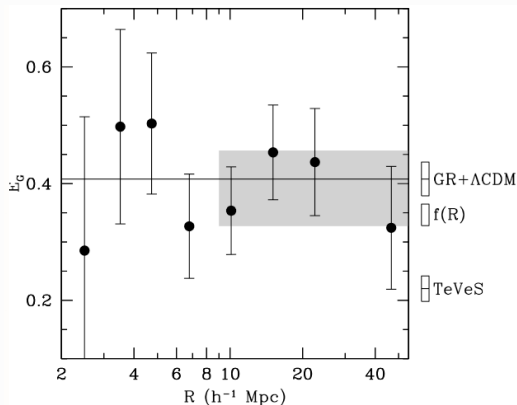
(Hildebrandt et al. 2017)

Science with gravitational lensing

Outstanding results

General relativity holds on cosmological scales.

Joint WL and galaxy clustering cosmology-independent GR test.



(Reyes et al. 2010)

Science with gravitational lensing

Outstanding results

Dark matter is not in form of massive compact objects (MACHOs).

Detection of Earth-mass exoplanets.

Structure of QSO inner emission regions.

Dark matter profiles in outskirts of galaxies.

Galaxy clusters are dominated by dark matter.

Observation of very-high ($z \geq 7$) galaxies.

Hints of inconsistency of our cosmological model at low and high z ?

General relativity holds on cosmological scales.

Most important properties of gravitational lensing

Lensing probes **total** matter, baryonic + dark.

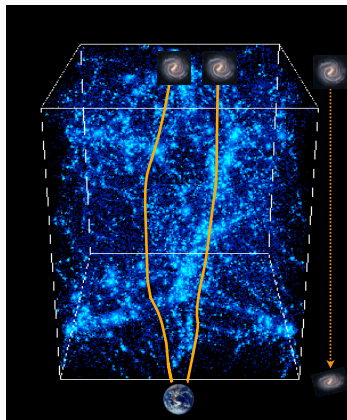
Independent of dynamical state of matter.

Independent of nature of matter.

Cosmic shear, or weak cosmological lensing

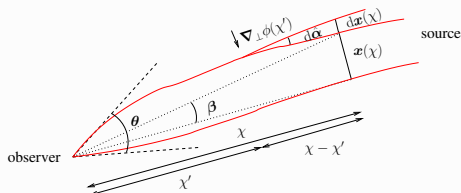
Light of distant galaxies is deflected while travelling through inhomogeneous Universe. Information about mass distribution is imprinted on observed galaxy images.

- Continuous deflection: sensitive to projected 2D mass distribution.
- Differential deflection: magnification, distortions of images.
- Small distortions, few percent change of images: need statistical measurement.
- Coherent distortions: measure correlations, scales few Mpc to few 100 Mpc.



Cosmic shear deflection angle

We derived the **deflection angle** as integral over the potential gradient (continuous deflection along the line of sight):



$$\alpha(\theta, \chi) = \frac{2}{c^2} \int_0^\chi d\chi' \frac{\chi - \chi'}{\chi} \left[\nabla_\perp \phi(\mathbf{x}(\chi'), \chi') - \nabla_\perp \phi^{(0)}(\chi') \right].$$

Geometrical relation: (Unobservable) unlensed source position β is observed lensed position (direction of incoming light ray) θ minus deflection angle α ,

$$\beta(\theta, \chi) = \theta - \alpha(\theta, \chi) = \theta - \nabla_\theta \psi(\theta);$$

with the **lensing potential**

$$\psi(\theta, \chi) = \frac{2}{c^2} \int_0^\chi d\chi' \frac{\chi - \chi'}{\chi \chi'} \phi(\chi' \theta, \chi').$$

Convergence and shear

The lens equation is the mapping from lens to source 2D coordinates. The linearized lens equation

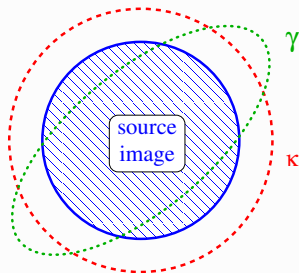
$$\frac{\partial \beta_i}{\partial \theta_j} \equiv A_{ij} = \delta_{ij} - \partial_i \partial_j \psi,$$

is described by the symmetrical 2×2 Jacobi matrix,

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

Which defines convergence κ and shear γ .

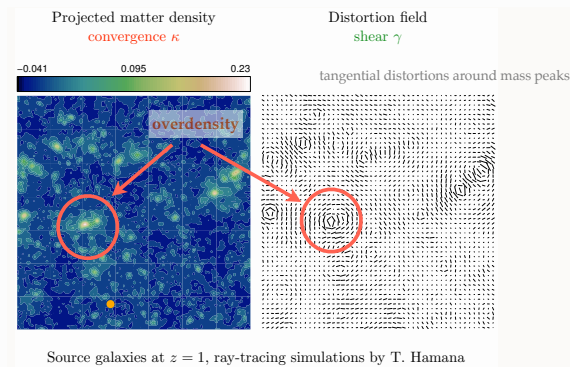
- **convergence** κ : isotropic magnification
- **shear** γ : anisotropic stretching



E- and B-modes: recap from part I

Shear patterns

We have seen tangential pattern in the shear field due to mass over-densities. Under-dense regions cause a similar pattern, but with opposite sign for γ . That results in radial pattern.

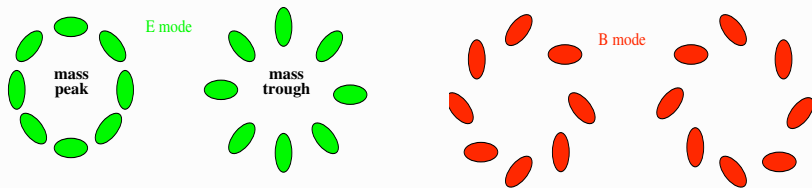


E- and B-modes: recap from part I

Shear patterns

We have seen tangential pattern in the shear field due to mass over-densities. Under-dense regions cause a similar pattern, but with opposite sign for γ . That results in radial pattern.

Under idealistic conditions, these are the only possible patterns for a shear field, the *E*-mode. A so-called *B*-mode is not generated.



E- and B-modes: recap I

Origins of a B-mode

Measuring a non-zero B-mode in observations is usually seen as indicator of residual systematics in the data processing (e.g. PSF correction, astrometry).

Other origins of a B-mode are small, of %-level:

- Higher-order terms beyond Born approximation (propagation along perturbed light ray, non-linear lens-lens coupling), and other (e.g. some ellipticity estimators)
- Lens galaxy selection biases (size, magnitude biases), and galaxy clustering
- Intrinsic alignment (although magnitude not well-known!)
- Varying seeing and other observational effects (table ronde topic!)
- Non-standard cosmologies (non-isotropic, TeVeS, ...)

E- and B-modes: recap II

Measuring E- and B-modes

Separating data into E- and B-mode is not trivial.

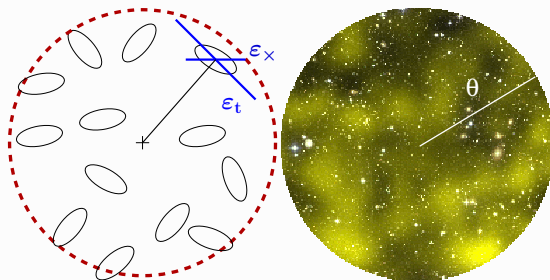
To directly obtain κ^E and κ^B from γ , there is leakage between modes due to the finite observed field (border and mask artefacts).

One can quantify the shear pattern, e.g. with respect to reference centre points, but the tangential shear γ_t is not defined at the center.

Solution: **filter** the shear map. (= convolve with a filter function Q). This also has the advantage that the spin-2 quantity shear is transformed into a scalar.

This is equivalent to filtering κ with a function U that is related to Q .

E- and B-modes: recap III



The resulting quantity is called **aperture mass** $M_{\text{ap}}(\theta)$, which is a function of the filter size, or smoothing scale, θ . It is only sensitive to the E-mode.

If one uses the cross-component shear γ_{\times} instead, the filtered quantity, M_{\times} captures the B-mode contribution only.

End of recap from part I.

Convergence as potential field

Again convergence κ and shear γ :

$$\frac{\partial \beta_i}{\partial \theta_j} \equiv A_{ij} = \delta_{ij} - \partial_i \partial_j \psi;$$

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

From this, write κ and γ as second derivatives of the potential.

Convergence as potential field

Again convergence κ and shear γ :

$$\frac{\partial \beta_i}{\partial \theta_j} \equiv A_{ij} = \delta_{ij} - \partial_i \partial_j \psi;$$

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

From this, write κ and γ as second derivatives of the potential.

$$\kappa = \frac{1}{2} (\partial_1 \partial_1 + \partial_2 \partial_2) \psi = \frac{1}{2} \nabla^2 \psi; \quad \gamma_1 = \frac{1}{2} (\partial_1 \partial_1 - \partial_2 \partial_2) \psi; \quad \gamma_2 = \partial_1 \partial_2 \psi.$$

Convergence as potential field

Again convergence κ and shear γ :

$$\frac{\partial \beta_i}{\partial \theta_j} \equiv A_{ij} = \delta_{ij} - \partial_i \partial_j \psi;$$

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

From this, write κ and γ as second derivatives of the potential.

$$\kappa = \frac{1}{2} (\partial_1 \partial_1 + \partial_2 \partial_2) \psi = \frac{1}{2} \nabla^2 \psi; \quad \gamma_1 = \frac{1}{2} (\partial_1 \partial_1 - \partial_2 \partial_2) \psi; \quad \gamma_2 = \partial_1 \partial_2 \psi.$$

We can now define a vector field \mathbf{u} for which the convergence is the “potential”, with

$$\mathbf{u} = \nabla \kappa.$$

Express \mathbf{u} in terms of the shear.

Convergence as potential field

Again convergence κ and shear γ :

$$\frac{\partial \beta_i}{\partial \theta_j} \equiv A_{ij} = \delta_{ij} - \partial_i \partial_j \psi;$$

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

From this, write κ and γ as second derivatives of the potential.

$$\kappa = \frac{1}{2} (\partial_1 \partial_1 + \partial_2 \partial_2) \psi = \frac{1}{2} \nabla^2 \psi; \quad \gamma_1 = \frac{1}{2} (\partial_1 \partial_1 - \partial_2 \partial_2) \psi; \quad \gamma_2 = \partial_1 \partial_2 \psi.$$

We can now define a vector field \mathbf{u} for which the convergence is the “potential”, with

$$\mathbf{u} = \nabla \kappa.$$

Express \mathbf{u} in terms of the shear.

$$\mathbf{u} = \begin{pmatrix} \partial_1 \kappa \\ \partial_2 \kappa \end{pmatrix} = \begin{pmatrix} \frac{1}{2} (\partial_1 \partial_1 \partial_1 + \partial_1 \partial_2 \partial_2) \kappa \\ \frac{1}{2} (\partial_1 \partial_1 \partial_2 + \partial_2 \partial_2 \partial_2) \kappa \end{pmatrix} = \begin{pmatrix} \partial_1 \gamma_1 + \partial_2 \gamma_2 \\ -\partial_2 \gamma_1 + \partial_1 \gamma_2 \end{pmatrix}.$$

E- and B-mode potential, convergence, and shear I

Thus, from a shear field γ , to linear order, the corresponding convergence is derived from a gradient field \mathbf{u} , and is curl-free, $\nabla \times \mathbf{u} = \partial_1 u_2 - \partial_2 u_1 = 0$, as can easily be seen.

This is the **E-mode**, in analogy to the **electric field**.

However, in reality, from an observed shear field, one might measure a non-zero curl component.

This is called the **B-mode**, in analogy to the **magnetic field**.

Definition:

$$\nabla^2 \kappa^{\text{E}} := \nabla \cdot \mathbf{u};$$

$$\nabla^2 \kappa^{\text{B}} := \nabla \times \mathbf{u},$$

and potentials

$$\nabla^2 \psi^{\text{E,B}} = 2\kappa^{\text{E,B}}.$$

Note that ψ^{B} and κ^{B} do not correspond to physical mass over-densities.

E- and B-mode potential, convergence, and shear II

These can be written in complex notation,

$$\psi = \psi^{\text{E}} + \text{i}\psi^{\text{B}}; \quad \kappa = \kappa^{\text{E}} + \text{i}\kappa^{\text{B}},$$

and the shear

$$\gamma_1 + \text{i}\gamma_2 = \frac{1}{2} (\partial_1 \partial_1 \psi^{\text{E}} - \partial_2 \partial_2 \psi^{\text{E}}) - \partial_1 \partial_2 \psi^{\text{B}} + \text{i} \left[\partial_1 \partial_2 \psi^{\text{E}} + \frac{1}{2} (\partial_1 \partial_1 \psi^{\text{B}} - \partial_2 \partial_2 \psi^{\text{B}}) \right].$$

Now, we can compute the E-, B-, and mixed EB-mode power spectrum.

$$\begin{aligned} \langle \hat{\kappa}^{\text{E}}(\ell) \hat{\kappa}^{\text{E}}(\ell') \rangle &= (2\pi)^2 \delta_{\text{D}}(\ell - \ell') P_{\kappa}^{\text{E}}(\ell), \\ \langle \hat{\kappa}^{\text{B}}(\ell) \hat{\kappa}^{\text{B}}(\ell') \rangle &= (2\pi)^2 \delta_{\text{D}}(\ell - \ell') P_{\kappa}^{\text{B}}(\ell), \\ \langle \hat{\kappa}^{\text{E}}(\ell) \hat{\kappa}^{\text{B}}(\ell') \rangle &= (2\pi)^2 \delta_{\text{D}}(\ell - \ell') P_{\kappa}^{\text{EB}}(\ell), \end{aligned}$$

and can derive (from $\hat{\gamma}(\ell) = \text{e}^{2\text{i}\beta} \hat{\kappa}(\ell)$, see last years' TD) for the correlators of γ in Fourier space

$$\begin{aligned} \langle \hat{\gamma}(\ell) \hat{\gamma}^*(\ell') \rangle &= (2\pi)^2 \delta_{\text{D}}(\ell - \ell') [P_{\kappa}^{\text{E}}(\ell) + P_{\kappa}^{\text{B}}(\ell)], \\ \langle \hat{\gamma}(\ell) \hat{\gamma}(\ell') \rangle &= (2\pi)^2 \delta_{\text{D}}(\ell + \ell') \text{e}^{4\text{i}\beta} [P_{\kappa}^{\text{E}}(\ell) - P_{\kappa}^{\text{B}}(\ell) + 2\text{i}P_{\kappa}^{\text{EB}}(\ell)]. \end{aligned}$$

Real-space correlation function (2PCF)

Fourier-transforming the last two expressions results in shear two-point correlators in real space,

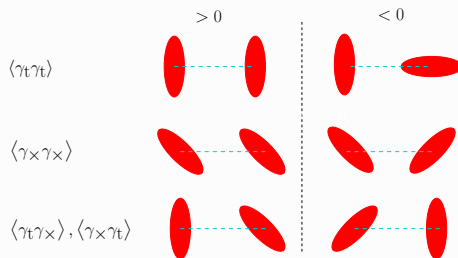
$$\begin{aligned}\langle \gamma(\boldsymbol{\theta}) \gamma^*(\boldsymbol{\theta} + \boldsymbol{\vartheta}) \rangle &= \langle \gamma \gamma^* \rangle(\boldsymbol{\vartheta}) = \mathcal{F} [\langle \hat{\gamma}(\boldsymbol{\ell}) \hat{\gamma}^*(\boldsymbol{\ell}') \rangle](\boldsymbol{\vartheta}); \\ \langle \gamma \gamma \rangle(\boldsymbol{\vartheta}) &= \mathcal{F} [\langle \hat{\gamma}(\boldsymbol{\ell}) \hat{\gamma}(\boldsymbol{\ell}') \rangle](\boldsymbol{\vartheta});\end{aligned}$$

But these correlators are very closely related to the shear two-point correlation functions ξ_+ and ξ_- , that we defined on day 1 (part I):

$$\begin{aligned}\xi_+(\boldsymbol{\vartheta}) &= \langle \gamma_t \gamma_t \rangle(\boldsymbol{\vartheta}) + \langle \gamma_{\times} \gamma_{\times} \rangle(\boldsymbol{\vartheta}) \\ \xi_-(\boldsymbol{\vartheta}) &= \langle \gamma_t \gamma_t \rangle(\boldsymbol{\vartheta}) - \langle \gamma_{\times} \gamma_{\times} \rangle(\boldsymbol{\vartheta})\end{aligned}$$

Recall: 2PCF

Correlation of the shear at two points yields four quantities



Parity conservation $\longrightarrow \langle \gamma_t \gamma_\times \rangle = \langle \gamma_\times \gamma_t \rangle = 0$

The two components of the shear **two-point correlation function** (2PCF) are defined as

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

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

Due to statistical isotropy & homogeneity, these correlators only depend on ϑ .

Real-space correlation function (2PCF)

Fourier-transforming the last two expressions results in shear two-point correlators in real space,

$$\begin{aligned}\langle \gamma(\boldsymbol{\theta}) \gamma^*(\boldsymbol{\theta} + \boldsymbol{\vartheta}) \rangle &= \langle \gamma \gamma^* \rangle(\boldsymbol{\vartheta}) = \mathcal{F} [\langle \hat{\gamma}(\boldsymbol{\ell}) \hat{\gamma}^*(\boldsymbol{\ell}') \rangle](\boldsymbol{\vartheta}); \\ \langle \gamma \gamma \rangle(\boldsymbol{\vartheta}) &= \mathcal{F} [\langle \hat{\gamma}(\boldsymbol{\ell}) \hat{\gamma}(\boldsymbol{\ell}') \rangle](\boldsymbol{\vartheta});\end{aligned}$$

But these correlators are very closely related to the shear two-point correlation functions ξ_+ and ξ_- , that we defined on day 1 (part I):

$$\begin{aligned}\xi_+(\vartheta) &= \langle \gamma_t \gamma_t \rangle(\vartheta) + \langle \gamma_x \gamma_x \rangle(\vartheta) \\ \xi_-(\vartheta) &= \langle \gamma_t \gamma_t \rangle(\vartheta) - \langle \gamma_x \gamma_x \rangle(\vartheta)\end{aligned}$$

Choose $\boldsymbol{\vartheta} = (\vartheta, 0)$. Then, $\gamma_t = -\gamma_1$ and $\gamma_x = -\gamma_2$.

$$\begin{aligned}\rightarrow \langle \gamma \gamma^* \rangle &= \langle \gamma_1 \gamma_1 \rangle + \langle \gamma_2 \gamma_2 \rangle = \xi_+; \\ \langle \gamma \gamma \rangle &= \langle \gamma_1 \gamma_1 \rangle - \langle \gamma_2 \gamma_2 \rangle + 2i \langle \gamma_1 \gamma_2 \rangle = \xi_- + 2i \xi_x.\end{aligned}$$

2PCF and E-/B-mode power spectra I

We generalize the relation between 2PCF and convergence power spectrum P_κ from day 1,

$$\begin{aligned}\xi_+(\vartheta) &= \frac{1}{2\pi} \int_0^\infty d\ell \ell J_0(\ell\vartheta) P_\kappa(\ell) \\ \xi_-(\vartheta) &= \frac{1}{2\pi} \int_0^\infty d\ell \ell J_4(\ell\vartheta) P_\kappa(\ell),\end{aligned}$$

to include E- and B-mode power spectra:

$$\begin{aligned}\xi_+(\vartheta) &= \frac{1}{2\pi} \int_0^\infty d\ell \ell J_0(\ell\vartheta) [P_\kappa^{\text{E}}(\ell) + P_\kappa^{\text{B}}(\ell)] \\ \xi_-(\vartheta) &= \frac{1}{2\pi} \int_0^\infty d\ell \ell J_4(\ell\vartheta) [P_\kappa^{\text{E}}(\ell) - P_\kappa^{\text{B}}(\ell)]\end{aligned}$$

(and we don't look any further at ξ_\times , which vanished for a parity-symmetric universe.)

2PCF and E-/B-mode power spectra II

We have thus two observables (ξ_+, ξ_-) and two unknowns (P_κ^E, P_κ^B) . Surely, these two power spectra can be deduced from the observations?

The above equations can be inverted using the orthogonality of the Bessel function:

$$\int_0^\infty d\vartheta \vartheta J_\nu(\ell\vartheta) J_\nu(\ell'\vartheta) = \frac{\delta_D(\ell - \ell')}{\ell},$$

(or, alternatively, go back to the 2D Fourier integrals and use the orthogonality of the plane wave basis functions $\exp(i\ell\vartheta)$) resulting in

$$P_\kappa^E(\ell) = \pi \int_0^\infty d\vartheta \vartheta [\xi_+(\vartheta) J_0(\ell\vartheta) + \xi_-(\vartheta) J_4(\ell\vartheta)],$$

$$P_\kappa^B(\ell) = \pi \int_0^\infty d\vartheta \vartheta [\xi_+(\vartheta) J_0(\ell\vartheta) - \xi_-(\vartheta) J_4(\ell\vartheta)].$$

So, **in principle**, the E-/ and B-mode power spectra can be computed separately, but **not in practice**, since this requires information about the shear correlation that is unobservable, towards 0 and ∞ separation.

→ We have to further filter the field for a better separation.

Aperture-mass

Yesterday we introduced the aperture-mass as convolution of the shear field with a filter Q ,

$$M_{\text{ap}}(\theta, \boldsymbol{\vartheta}) = \int d^2\vartheta' Q_{\theta}(|\boldsymbol{\vartheta} - \boldsymbol{\vartheta}'|) \gamma_{\text{t}}(\boldsymbol{\vartheta}')$$

and claimed that this was equivalent of convolving the convergence with another filter U ,

$$M_{\text{ap}}(\theta, \boldsymbol{\vartheta}) = \int d^2\vartheta' U_{\theta}(|\boldsymbol{\vartheta} - \boldsymbol{\vartheta}'|) \kappa^{\text{E}}(\boldsymbol{\vartheta}'), \quad (1)$$

(Kaiser et al. 1994, Schneider 1996).

Exercise for next session (where you'll need stuff from today's TD): What is the relation between U and Q ?

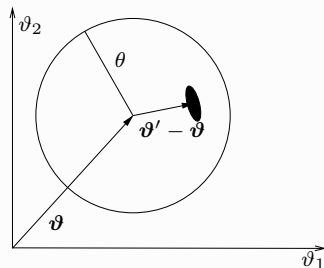
Convolution with shear

Parenthesis:

Eq. (3) involves the tangential shear γ_t with respect to the aperture centre ϑ ; it should be written $\gamma_t(\vartheta, \vartheta')$.

This “field” γ_t is thus defined locally, and cannot be represented globally.

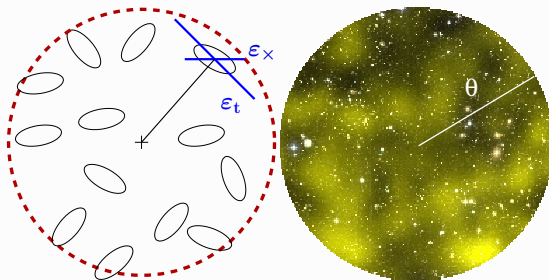
How can this expression be written as convolution with $\gamma = \gamma_1 + i\gamma_2$?



Solution:

$$\begin{aligned}\gamma_t(\vartheta, \vartheta') &= -\Re(\gamma e^{-2i\varphi}) = -\Re\left(\gamma e^{-2i \arctan(|\vartheta_2 - \vartheta'_2|/|\vartheta_1 - \vartheta'_1|)}\right) \\ \rightarrow M_{\text{ap}}(\theta, \vartheta) &= -\Re \int d^2\vartheta' \gamma(\vartheta') e^{-2i \arctan[|\vartheta_2 - \vartheta'_2|/|\vartheta_1 - \vartheta'_1|]} \\ &= \Re(Q'_\theta * \gamma)(\vartheta) \\ \text{with } Q'_\theta(\vartheta) &= -Q_\theta(\vartheta) e^{-2i \arctan[\vartheta_2/\vartheta_1]}.\end{aligned}$$

E-/B-mode separation with M_{ap} I



It is clear that M_{ap} (M_{\times}) is sensitive to the E-mode (B-mode) of the shear field γ .

When choosing Q such that its support is finite, with $Q(\theta) = 0$ for $\theta > \theta_{\text{max}}$, the E-/B-mode separation is achieved on a finite interval.

To get this separation at the second-order level, let's take the variance of the aperture-mass: Square $M_{\text{ap}}(\theta, \vartheta)$ and average over circle centres ϑ (Schneider et al. 1998).

E-/B-mode separation with M_{ap} II

Square $M_{\text{ap}}(\theta, \vartheta)$ and average over circle centres ϑ :

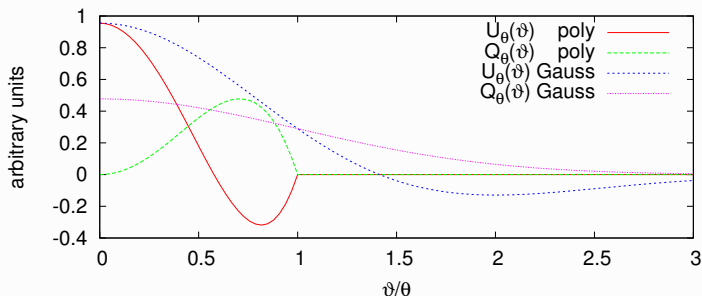
$$\begin{aligned}
 \langle M_{\text{ap}}^2 \rangle(\theta) &= \int d^2\vartheta' U_\theta(|\vartheta - \vartheta'|) \int d^2\vartheta'' U_\theta(|\vartheta - \vartheta''|) \langle \kappa^{\text{E}}(\vartheta') \kappa^{\text{E}}(\vartheta'') \rangle \\
 &= \int d^2\vartheta' U_\theta(\vartheta') \int d^2\vartheta'' U_\theta(\vartheta'') \langle \kappa^{\text{E}} \kappa^{\text{E}} \rangle(|\vartheta' - \vartheta''|) \\
 &= \int d^2\vartheta U_\theta(\vartheta) \int d^2\vartheta' U_\theta(\vartheta') \\
 &\quad \times \int \frac{d^2\ell}{(2\pi)^2} e^{-i\ell\vartheta} \int \frac{d^2\ell'}{(2\pi)^2} e^{-i\ell\vartheta'} (2\pi)^2 \delta_{\text{D}}(\ell - \ell') P_\kappa^{\text{E}}(\ell) \\
 &= \int \frac{d^2\ell}{(2\pi)^2} \left(\int d^2\vartheta e^{2i\ell\vartheta} U_\theta(\vartheta) \right)^2 P_\kappa^{\text{E}}(\ell) \\
 &= \frac{1}{2\pi} \int d\ell \ell \hat{U}^2(\theta\ell) P_\kappa^{\text{E}}(\ell).
 \end{aligned}$$

Note: Typically, the filter function U depends on the scale ϑ normalized to the radius θ , $U_\theta(\vartheta) = U(\vartheta/\theta)$. In Fourier space this then becomes $\hat{U}(\theta\ell)$.

E-/B-mode separation with M_{ap} III

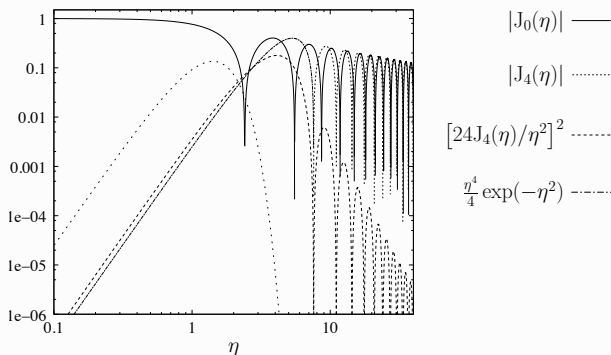
For popular choices of U , \hat{U}^2 is a narrow pass-band filter function.

	polynomial	Gaussian
$U_\theta(\vartheta)$	$\begin{cases} \frac{9}{\pi\theta^2} \left(1 - \frac{\vartheta^2}{\theta^2}\right) \left(\frac{1}{3} - \frac{\vartheta^2}{\theta^2}\right) & \vartheta < \theta \\ 0 & \text{else} \end{cases}$	$\frac{1}{2\pi\theta^2} \left(1 - \frac{\vartheta^2}{2\theta^2}\right) \exp\left(-\frac{\vartheta^2}{2\theta^2}\right)$
$Q_\theta(\vartheta)$	$\begin{cases} \frac{6}{\pi\theta^2} \frac{\vartheta^2}{\theta^2} \left(1 - \frac{\vartheta^2}{\theta^2}\right) & \vartheta < \theta \\ 0 & \text{else} \end{cases}$	$\frac{\vartheta^2}{4\pi\theta^4} \exp\left(-\frac{\vartheta^2}{2\theta^2}\right)$
$\hat{U}(\eta)$	$\frac{24J_4(\eta)}{\eta^2}$	$\frac{\eta^2}{2} \exp\left(\frac{-\eta^2}{2}\right)$



E-/B-mode separation with M_{ap} IV

Filter functions in Fourier space:



E-/B-mode separation with M_{ap} V

Thus, the aperture-mass dispersion filters out a small range of ℓ -modes around $\ell \sim \text{const } \theta^{-1}$.

For example, for the polynomial filter from (Schneider et al. 1998), the peak is $\theta\ell \approx 5$.

Analogous equations for B- and mixed modes are

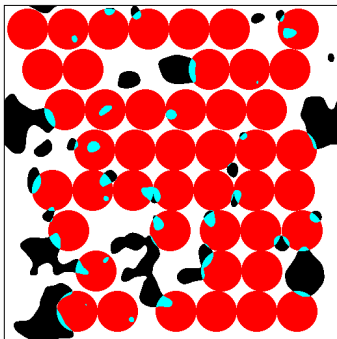
$$\begin{aligned}\langle M_{\times}^2 \rangle(\theta) &= \frac{1}{2\pi} \int d\ell \ell \hat{U}^2(\theta\ell) P_{\kappa}^{\text{B}}(\ell); \\ \langle M_{\text{ap}} M_{\times} \rangle(\theta) &= \frac{1}{2\pi} \int d\ell \ell \hat{U}^2(\theta\ell) P_{\kappa}^{\text{EB}}(\ell).\end{aligned}$$

In complex notation, the last three expressions can be written as

$$\langle M_{\text{ap}}^2 \rangle(\theta) \pm \langle M_{\times}^2 \rangle(\theta) + 2i\langle M_{\text{ap}} M_{\times} \rangle(\theta) = \frac{1}{2\pi} \int d\ell \ell \hat{U}^2(\theta\ell) [P_{\kappa}^{\text{E}} \pm P_{\kappa}^{\text{B}} + 2iP_{\kappa}^{\text{EB}}](\ell).$$

Aperture-mass dispersion and 2PCF I

The above recipe to get the aperture-mass variance can be implemented in an estimator as follows: For an aperture with center ϑ and radius θ , average the observed galaxy ellipticities weighted by the filter Q . Square, average over many centers ϑ :



This is however not very efficient due to masked regions and field boundaries.

Solutions:

- Inpainting of missing data (Starck et al. 2006), using fast algorithms for convolution (Leonard et al. 2012).
- Compute 2PCF first, integrate to get aperture-mass dispersion.

From [P. Simon, PhD thesis, 2005].

Aperture-mass dispersion and 2PCF II

Aperture-mass dispersion from 2PCF

M_{ap} depends on γ_t , thus we expect that $\langle M_{\text{ap}}^2 \rangle$ depends on $\langle \gamma_t \gamma_t \rangle \sim 2\text{PCF}$.

Simple calculation: Use

$$\langle M_{\text{ap}}^2 \rangle(\theta) = \frac{1}{2\pi} \int d\ell \ell \hat{U}^2(\theta\ell) P_{\kappa}^{\text{E}}(\ell)$$

and insert

$$P_{\kappa}^{\text{E}}(\ell) = \pi \int_0^{\infty} d\vartheta \vartheta [\xi_+(\vartheta) J_0(\ell\vartheta) + \xi_-(\vartheta) J_4(\ell\vartheta)] .$$

Result:

$$\langle M_{\text{ap}}^2 \rangle(\theta) = \int_0^{2\theta} d\vartheta \vartheta \left[T_+ \left(\frac{\vartheta}{\theta} \right) \xi_+(\vartheta) + T_- \left(\frac{\vartheta}{\theta} \right) \xi_-(\vartheta) \right] .$$

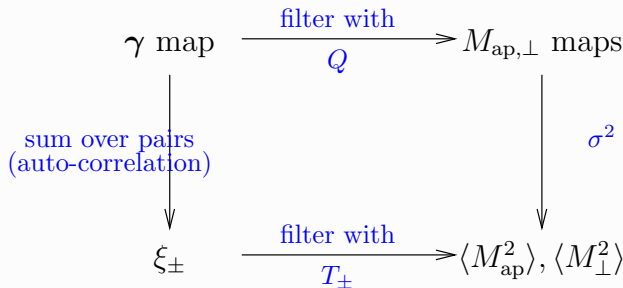
with

$$T_{\pm}(x) = \int_0^{\infty} dt t J_{0,4}(xt) \hat{U}^2(t) .$$

Aperture-mass dispersion and 2PCF III

The functions $T_{\pm}(x)$ have support $[0; 2]$, thus the above integral extends to 2θ . Therefore, the maximum distance to compute the shear correlation ξ_{\pm} is $\vartheta_{\max} = 2\theta$.

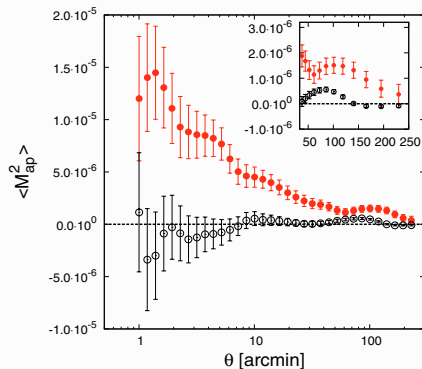
Remember the diagram from Part I?



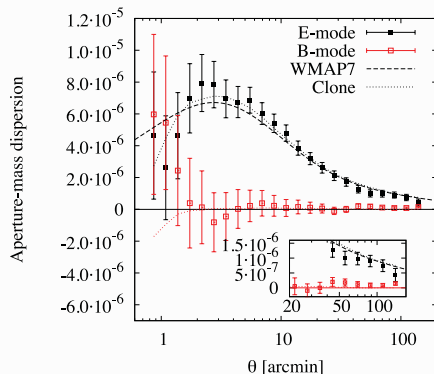
Maybe this makes a bit more sense now...

Aperture-mass dispersion measurements

CFHTLS 2007 versus CFHTlenS 2013.



From (Fu et al. 2008).



From (Kilbinger et al. 2013).

Ring statistic I

The problem of the inaccessible zero lag shear correlation for an E- and B-mode decomposition remains. How can we construct a E-/B-mode second-order correlation with a minimum galaxy separation $\vartheta_{\min} > 0$?

Solution: Correlate shear on two concentric rings (Schneider & Kilbinger 2007).

What are the minimum and maximum distances in this configuration?

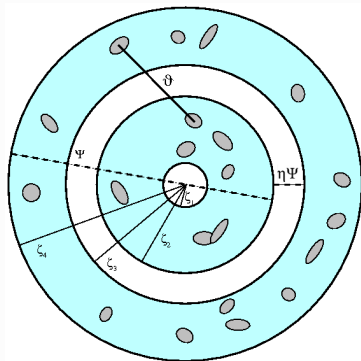


Figure from (Eifler et al. 2010).

Ring statistic II

Filter functions (in the original paper called Z_{\pm} instead of T_{\pm}) depend on geometry of circles, and free-to-choose weight functions over the rings.

$$\langle \mathcal{R}\mathcal{R} \rangle_{\text{E,B}} = \int_{\eta}^1 \frac{dx}{2x} [\xi_+(x\Psi) T_+(x, \eta) \pm \xi_-(x\Psi) T_-(x, \eta)] .$$

where $\eta = \vartheta_{\min}/\vartheta_{\max} < 1$ is ratio of minimum to maximum separation of the configuration.

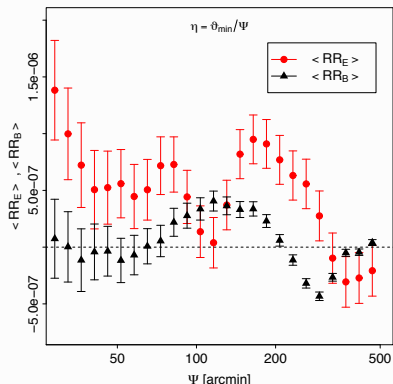
General E-/B-mode decomposition on a finite interval (in $\log \vartheta$).

(Schneider & Kilbinger 2007) worked out the conditions on T_{\pm} to have finite support, with $0 < \vartheta_{\min} < \vartheta_{\max} < \infty$:

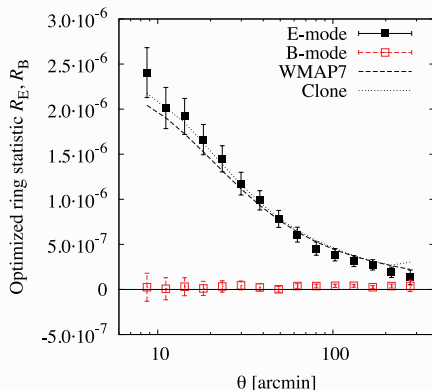
$$\begin{aligned} \int_{\vartheta_{\min}}^{\vartheta_{\max}} d\vartheta \, \vartheta \, T_+(\vartheta) &= 0 = \int_{\vartheta_{\min}}^{\vartheta_{\max}} d\vartheta \, \vartheta^3 \, T_+(\vartheta) ; \\ \int_{\vartheta_{\min}}^{\vartheta_{\max}} \frac{d\vartheta}{\vartheta} \, T_-(\vartheta) &= 0 = \int_{\vartheta_{\min}}^{\vartheta_{\max}} \frac{d\vartheta}{\vartheta^3} \, T_-(\vartheta) . \end{aligned}$$

Ring statistic measurements

CFHTLS 2007 versus CFHTLenS 2013.



From (Eifler et al. 2010).



From (Kilbinger et al. 2013), optimised ring statistic following (Fu & Kilbinger 2010).

COSEBIs I

$$\int_{\vartheta_{\min}}^{\vartheta_{\max}} d\vartheta \vartheta T_+(\vartheta) = 0 = \int_{\vartheta_{\min}}^{\vartheta_{\max}} d\vartheta \vartheta^3 T_+(\vartheta) ;$$

$$\int_{\vartheta_{\min}}^{\vartheta_{\max}} \frac{d\vartheta}{\vartheta} T_-(\vartheta) = 0 = \int_{\vartheta_{\min}}^{\vartheta_{\max}} \frac{d\vartheta}{\vartheta^3} T_-(\vartheta) .$$

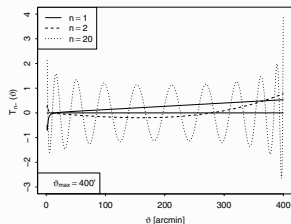
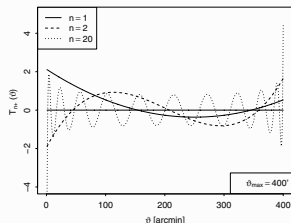
Under these conditions the functions T_{\pm} can be freely chosen. Idea of (Schneider et al. 2010): Define modes E_n, B_n using polynomials of order $n + 1$. Define family of orthogonal polynomials that provide all information about E-/B-modes on finite interval:

Complete Orthogonal Set of E-/B-mode Integrals.

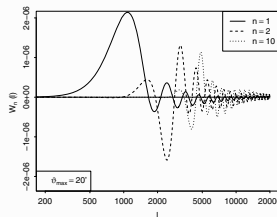
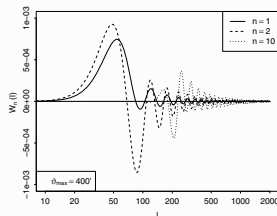
The COSEBIs contain nearly all information that is in ξ_+ and ξ_- , except the very large scales. These are outside the survey, and cannot be decomposed into E-/B-modes, but form an ambiguous mode. This mode is contained in $\xi_+(\theta)$, for which the filter $J_0(\theta\ell) \rightarrow \text{const}$ for arbitrarily large $\ell \rightarrow 0$.

COSEBIs II

Polynomials can be linear in θ (Lin-COSEBIs), or linear in $z = \log \theta$ (Log-COSEBIs).

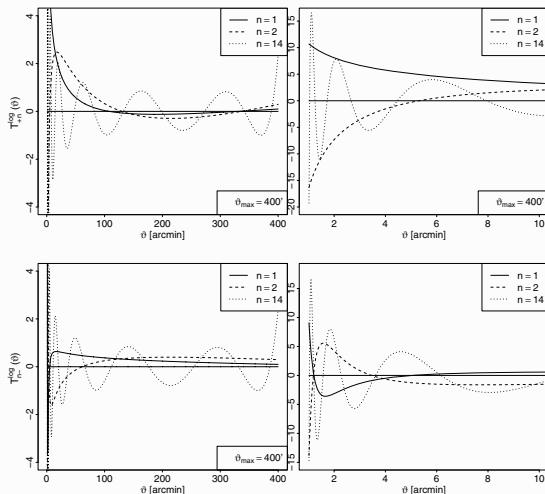


COSEBI linear filter functions $T_{\pm n}$ in real space.



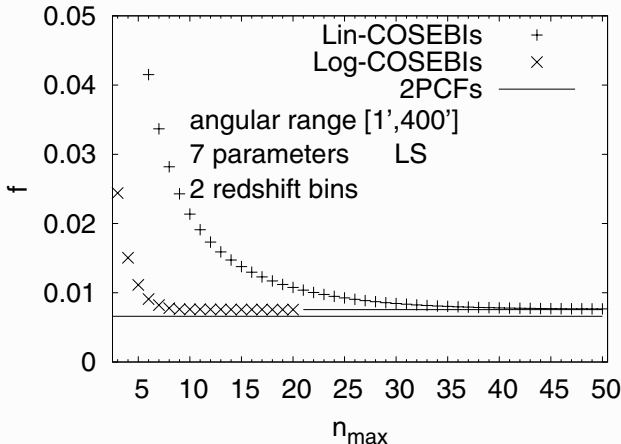
COSEBI linear filter functions $W_n (= \hat{U}^2)$ in Fourier space. From (Schneider et al. 2010).

COSEBIs III



COSEBI logarithmic filter functions $T_{\pm n}$ in real space.

COSEBIs IV



Inverse Fisher-matrix (allowed parameter) volume as function of COSEBIs maximum mode.
From (Asgari et al. 2012).

Log-COSEBIs show faster convergence of available information with n .

Band-power spectrum I

The power spectrum P_κ can be estimated from shear data using methods from the CBM,

(Pseudo- C_ℓ , Bayesian, ...)

from pixellised maps.

A much faster **but biased** method is a band-power estimate from the 2PCF.

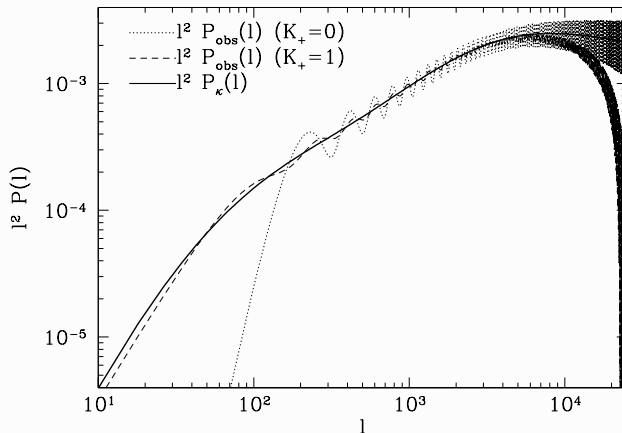
Recall the expressions

$$P_\kappa^{\text{E}}(\ell) = \pi \int_0^\infty d\vartheta \, \vartheta [\xi_+(\vartheta) J_0(\ell\vartheta) + \xi_-(\vartheta) J_4(\ell\vartheta)],$$

$$P_\kappa^{\text{B}}(\ell) = \pi \int_0^\infty d\vartheta \, \vartheta [\xi_+(\vartheta) J_0(\ell\vartheta) - \xi_-(\vartheta) J_4(\ell\vartheta)].$$

To estimate these improper integrals as direct sums over observed ξ_\pm between ϑ_{\min} and ϑ_{\max} would introduce large biases.

Band-power spectrum II



$$\hat{P}(\ell) = 2\pi \int_{\theta_{\min}}^{\theta_{\max}} d\theta \theta [K_+ \xi_+(\theta) J_0(\ell\theta) + (1 - K_+) \xi_-(\theta) J_4(\ell\theta)]$$

From (Schneider et al. 2002).

Band-power spectrum III

However, we can add another integration in bands of ℓ , between ℓ_{\min} and ℓ_{\max} ,

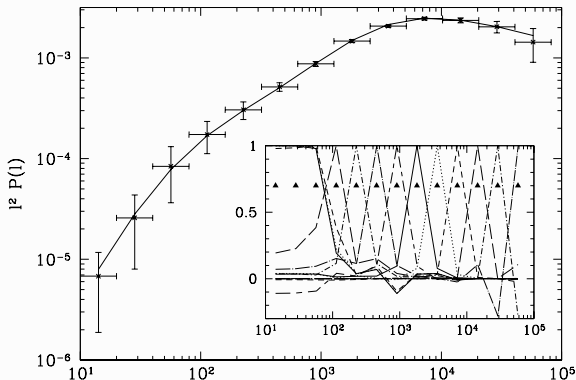
$$\mathcal{P}_i := \frac{1}{\Delta_i} \int_{\ell_{i1}}^{\ell_{iu}} d\ell \ell \hat{P}(\ell) = \frac{2\pi}{\Delta_i} \int_{\theta_{\min}}^{\theta_{\max}} \frac{d\theta}{\theta} \left\{ K_+ \xi_+(\theta) \left[g_+(\ell_{iu}\theta) - g_+(\ell_{i1}\theta) \right] + (1 - K_+) \xi_-(\theta) \left[g_-(\ell_{iu}\theta) - g_-(\ell_{i1}\theta) \right] \right\}$$

where $\Delta_i = \ln(\ell_{iu}/\ell_{i1})$ is the logarithmic width of the band, and

$$g_+(x) = x J_1(x) \quad ; \quad g_-(x) = \left(x - \frac{8}{x} \right) J_1(x) - 8 J_2(x) .$$

This strongly reduces
the bias.

You will use the
program `pallas.py` in
the TD this afternoon
that implements this
estimator.



Galaxy-galaxy lensing: Overview

Correlation between high- z galaxy shapes and low- z galaxy positions.

E.g. average tangential shear around massive galaxies.

Provides mass associated with galaxy sample.

- Galaxy halo profiles from kpc to Mpc
- Mass-to-light ratio

In combination with other tracers of matter (galaxy clustering, cosmic shear, velocity correlations, X-ray emission, ...):

- Galaxy bias. Properties such as linearity, scale-dependence, stochasticity
- Test of General Relativity

Can be done quasi model-independent since two or more observables trace same matter field, but with different biases.

Tangential shear and projected overdensity

Tangential shear at distance θ is related to total overdensity within this radius:

$$\langle \gamma_t \rangle (\theta) = \bar{\kappa}(\leq \theta) - \langle \kappa \rangle (\theta).$$

No assumption about mass distribution is made here!

End of day 1.

Reminder: Overview

Part II day 1: E- and B-modes

- Very brief reminders from day I
- E-/B-mode decomposition recap
- E-/B-mode estimators
- Galaxy-galaxy lensing: motivation

Part II day 2: Shear estimation

- Galaxy-galaxy lensing in detail
- Back to the aperture mass: Filter function relation
- Spherical-sky lensing projections
- Shear calibration

Part II day 3: Cosmological parameter estimation

- Numerical simulations
- Covariance estimation
- Likelihood and parameter estimation

Tangential shear and surface mass I

In an exercise you have derived the relation between tangential shear and encompassed projected surface mass,

$$\langle \gamma_t \rangle (\theta) = \bar{\kappa}(\leq \theta) - \langle \kappa \rangle (\theta).$$

We will re-write equation defining the *surface mass excess* $\Delta\Sigma$.

Surface mass excess

Assume a single lens at (angular diameter) distance D_l . Approximate for this case the expression of the convergence

$$\kappa(\boldsymbol{\theta}, \chi) = \frac{3}{2} \Omega_m \left(\frac{H_0}{c} \right)^2 \int_0^\chi d\chi' \frac{(\chi - \chi')\chi'}{\chi a(\chi')} \delta(\chi' \boldsymbol{\theta}, \chi').$$

and write D_s for the distance of the source, and D_{ls} for the distance between lens and source. Write all distances as proper, not comoving distances, express the density contrast in terms of the density, $\delta = \Delta\rho/\bar{\rho}$, and use the critical density ρ_{crit} .

Tangential shear and surface mass II

Assume that the lens mass distribution ρ extends over the interval $[D_1 - \Delta D/2; D_1 + \Delta D/2]$.

$$\kappa(\boldsymbol{\theta}) = \frac{4\pi G}{c^2} \frac{D_1 D_{\text{ls}}}{D_s} \int_{D_1 - \Delta D/2}^{D_1 + \Delta D/2} dD \Delta\rho(D\boldsymbol{\theta}, D).$$

Define the *critical surface mass density*

$$\Sigma_{\text{cr}}^{-1}(\theta) := \frac{4\pi G}{c^2} \frac{D_1 D_{\text{ls}}}{D_s}$$

to write convergence as

$$\kappa(\boldsymbol{\theta}) = \frac{\Sigma(\boldsymbol{\theta})}{\Sigma_{\text{cr}}}. \quad (2)$$

[Why is Σ_{cr} called *critical* surface mass?]

With that, we define the surface mass excess

$$\Delta\Sigma(\leq \theta) := \langle \gamma_t \rangle(\theta) \Sigma_{\text{cr}} = \bar{\Sigma}(\theta) - \langle \Sigma \rangle(\theta).$$

Statistical galaxy-galaxy lensing (GGL) I

The convergence or tangential shear defined in the last slides depend linearly on the mass distribution ρ , or Σ . So it seems to be a first-order statistic.

However, when measured statistically using a population of foreground galaxies, it can be written as two-point correlation function. The convergence is then the correlation of background lensing convergence and foreground galaxy position.

If we write the latter as galaxy over-density δ_g , we get

$$\begin{aligned}
 \langle \kappa \rangle(\theta) &= \langle \kappa(\boldsymbol{\vartheta}) \delta_g(\boldsymbol{\vartheta} + \boldsymbol{\theta}) \rangle_{\boldsymbol{\vartheta}} \\
 &= \Sigma_{\text{cr}}^{-1} \bar{\rho} \int dD \langle \delta(D\boldsymbol{\theta}, D) \delta_g(D_1\boldsymbol{\theta}, D_1) \rangle \\
 &= \Sigma_{\text{cr}}^{-1} \bar{\rho} \int dD \xi_{\delta_g}(\sqrt{(D\theta)^2 + (D - D_1)^2}).
 \end{aligned}$$

Statistical galaxy-galaxy lensing (GGL) II

Properties of statistical GGL

- Circular averages of tangential shear: robust against (some) systematic, e.g. large-scale modes of PSF residuals cancel out.
CFHTLenS: 25% fields had to be discarded for cosmic shear, none for GGL.
- Simple null tests:
 $\langle \gamma_{\times} \rangle$ around foreground objects (parity mode, should vanish).
 $\langle \gamma_t \rangle$ around random points, or special points that should not be correlated with foreground sample such as chip corners, field centres, stars.
- Higher SNR compared to cosmic shear:
correlation with tracers of dense matter regions;
one shape instead of two;
- Can use spectroscopic galaxies for foreground sample.

Parenthesis: galaxy bias I

Simple bias

GGL measures the cross-correlation between galaxies and dark (more precisely: total) matter, $\langle \delta_g \delta \rangle$. This correlation is non-zero since galaxies trace the underlying matter.

Simplest model: linear, constant, deterministic bias:

$$\delta_g = b\delta.$$

From that it follows that

$$\langle \delta_g \delta_g \rangle(\theta) = b^2 \langle \delta \delta \rangle(\theta); \quad \langle \delta_g \delta \rangle(\theta) = b \langle \delta \delta \rangle(\theta),$$

or in Fourier space

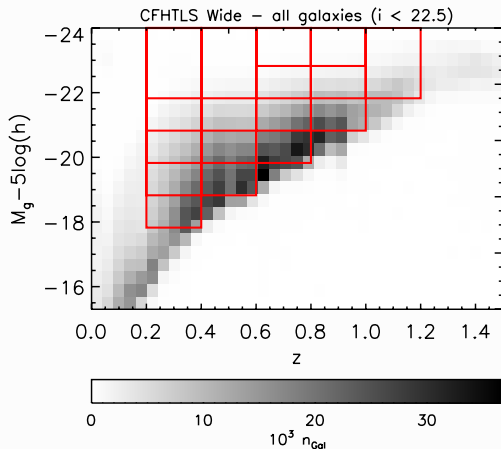
$$P_{gg}(k) = b^2 P_{mm}(k); \quad P_{gm}(k) = b P_{mm}(k).$$

Parenthesis: galaxy bias II

Properties

- The bias depends on the galaxy properties (type, color, luminosity, ..., and can be measured for different populations (e.g. early/late-type).
- Bias is redshift-dependent. Difficult to measure since degenerate with z -dependent selection effects. Volume-limited samples: Bias tends to increase with z : galaxies are more rare objects at higher z , situated in more extreme environments (halo centres).

Sample selection for galaxy bias measurement



Sample selection in absolute magnitude and redshift, from (Coupon et al. 2012).
Samples in horizontal boxes have same absolute magnitudes and are volume-limited.

Galaxy bias extended I

More complex bias models

- Scale-dependence, $b(\theta)$, or $\hat{b}(\hat{k})$. In particular on small scales, bias is not constant.
- Non-linear bias

$$\delta_g = b_1\delta + b_2\delta^2 + b_3\delta^3 + \dots$$

- Stochastic bias

Relation between δ_g is not deterministic ($\delta_g = b\delta$) but stochastic. In a statistical picture, the two fields δ_g and δ can be interpreted as realizations of random fields with joint pdf $p(\delta_g, \delta)$. The study of stochastic biasing is trying to quantify this joint pdf.

Galaxy bias extended II

At second-order level, one can measure the variances of both fields, and their cross-correlation. If the fields are correlated, one can write down the following two relations:

$$b = \frac{\sigma_g}{\sigma} = \sqrt{\frac{\langle \delta_g^2 \rangle}{\langle \delta^2 \rangle}}; \quad r = \frac{\sigma_{gm}^2}{\sigma_g \sigma} = \frac{\langle \delta_g \delta \rangle}{\sqrt{\langle \delta_g^2 \delta \rangle \langle \delta^2 \delta \rangle}}$$

introducing a correlation coefficient $r = -1 \dots 1$ between both fields.

In the above ratio cosmology dependence (dm correlation function or power spectrum) mainly drops out!

Allows for model-independent measurement.

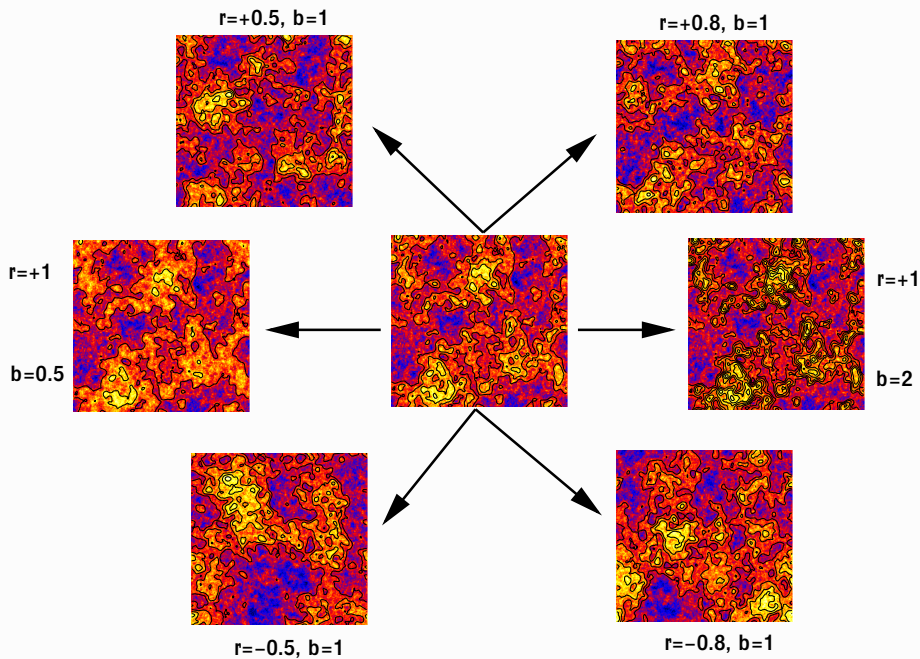


Illustration of correlated fields, from [P. Simon, PhD thesis, 2005].

Galaxy bias II

Question: How would the correlation between δ_g and δ look like for negative bias $b < 0$? For example $b = -1, r = 1$.

Non-linear and stochastic bias

A non-linear bias can mimic stochasticity.

Consider the (made-up) example of deterministic bias with $\delta_g = \delta^3$.

Exercise:

Calculate r in the case where both fields follow Gaussian pdf's.

Galaxy bias II

Question: How would the correlation between δ_g and δ look like for negative bias $b < 0$? For example $b = -1, r = 1$.

Non-linear and stochastic bias

A non-linear bias can mimic stochasticity.

Consider the (made-up) example of deterministic bias with $\delta_g = \delta^3$.

Exercise:

Calculate r in the case where both fields follow Gaussian pdf's.

$$r = \frac{\langle \delta_g \delta \rangle}{\sqrt{\langle \delta_g \delta_g \rangle \langle \delta \delta \rangle}} = \frac{\langle \delta^4 \rangle}{\sqrt{\langle \delta^6 \rangle \langle \delta^2 \rangle}} = \frac{3\sigma^4}{\sqrt{15\sigma^6 \sigma^2}} = \frac{3}{\sqrt{3 \cdot 5}} = \sqrt{\frac{3}{5}} \approx 0.77 \leq 1!$$

Final note: The density field *cannot* be a Gaussian, since $\delta \leq -1$.

GGL: model-independent measurement of b/r

Idea:

Combine weak lensing and galaxy clustering to determine b and r .

- Galaxy clustering $\langle \delta_g^2 \rangle$
- Galaxy-galaxy lensing $\langle \delta_g \delta \rangle$
- Cosmic shear $\langle \delta^2 \rangle$

Cosmic shear is the most difficult to measure, so first measurements only used GC and GGL.

Form ratio:

$$\frac{\langle \delta_g \delta \rangle(\theta)}{\langle \delta_g \delta_g \rangle(\theta)} = \frac{br}{b^2} = \frac{b}{r}.$$

Any cosmology-dependence, e.g. of clustering, drops out in the ratio. These density correlations are projected to weak-lensing observables, and b and r (if constant) can directly be measured.

GGL: Aperture measures I

How can we trace the galaxy and dark-matter over-densities with weak lensing?

Use aperture measures

$$\langle N^2 \rangle(\theta), \langle NM_{\text{ap}} \rangle(\theta), \langle M_{\text{ap}}^2 \rangle(\theta)$$

to trace

$$\langle \delta_g^2 \rangle, \langle \delta_g \delta \rangle, \langle \delta^2 \rangle.$$

Difficulty: Structure along all redshifts contribute to cosmic shear $\langle M_{\text{ap}}^2 \rangle$, not only mass associated with foreground galaxy sample δ_g .

Solutions:

- Choose background sample such that maximum lensing efficiency coincides with foreground redshift.
- Add correction functions with minor dependency on cosmology (geometry).

Redshift calibration factors

Aperture measure ratios

$$b(\theta) = \sqrt{\frac{\langle N^2 \rangle(\theta)}{\langle M_{\text{ap}}^2 \rangle(\theta)}} \quad ?$$
$$r(\theta) = \frac{\langle NM_{\text{ap}} \rangle(\theta)}{\sqrt{\langle N^2 \rangle(\theta) \langle M_{\text{ap}}^2 \rangle(\theta)}} \quad ?$$

Redshift calibration factors

Aperture measure ratios

$$b(\theta) = f_1(\theta) \sqrt{\frac{\langle N^2 \rangle(\theta)}{\langle M_{\text{ap}}^2 \rangle(\theta)}}$$

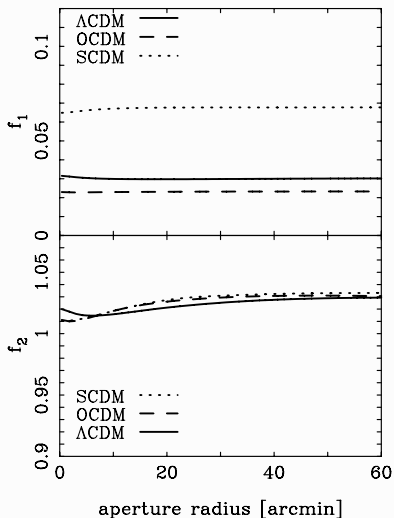
$$r(\theta) = f_2(\theta) \frac{\langle N M_{\text{ap}} \rangle(\theta)}{\sqrt{\langle N^2 \rangle(\theta) \langle M_{\text{ap}}^2 \rangle(\theta)}}$$

Calibration factors f_1, f_2 to account for different redshifts/lensing efficiency (Hoekstra et al. 2001). Calculate those using theoretical model for fiducial cosmology (fixing power spectrum, geometry), setting $b = r = 1$:

$$f_1(\theta) = \sqrt{\frac{\langle M_{\text{ap}}^2 \rangle(\theta)}{\langle N^2 \rangle(\theta)}} \Big|_{\text{fid}, b=1}$$

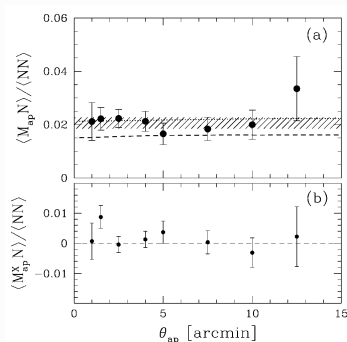
$$f_2(\theta) = \frac{\sqrt{\langle N^2 \rangle(\theta) \langle M_{\text{ap}}^2 \rangle(\theta)}}{\langle N M_{\text{ap}} \rangle(\theta)} \Big|_{\text{fid}, b=r=1}$$

Redshift calibration factors



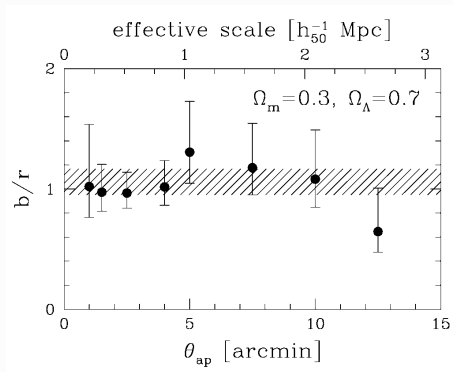
Scale-and cosmology-dependence of calibration factors. From (Simon et al. 2007), GaBoDS (Garching-Bonn Deep Survey).

GGL results: model-independent measurement of b/r



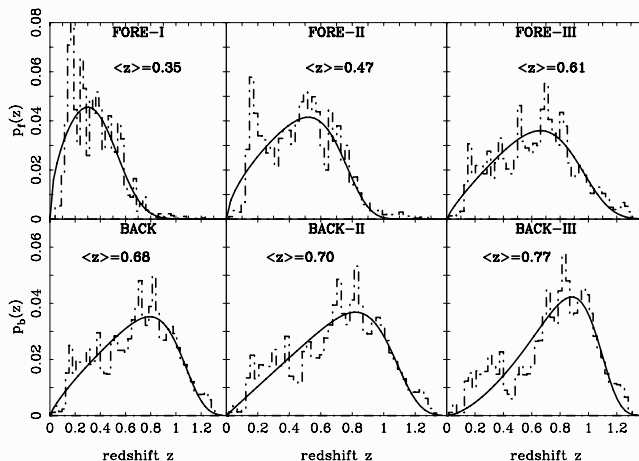
$$\mathcal{R} = \frac{r\Omega_m}{100b} [(5.8 - 1.6\Omega_m^{0.63}) + (4.6 - 2.6\Omega_m^{0.63})\Omega_\Lambda^{1.23}].$$

Observed ratio \mathcal{R} (a), and B-mode (b); b/r (right) from (Hoekstra et al. 2001).



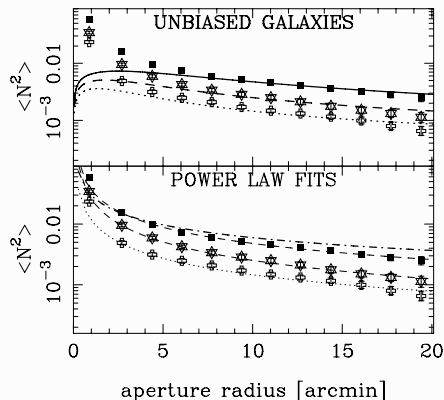
Main result: no scale-dependence found (on observed scales).

GGL results: model-indep. measurement of b and r I



Redshift distributions for GaBoDS samples, estimated from COMBO-17. From (Simon et al. 2007), GaBoDS (Garching-Bonn Deep Survey).

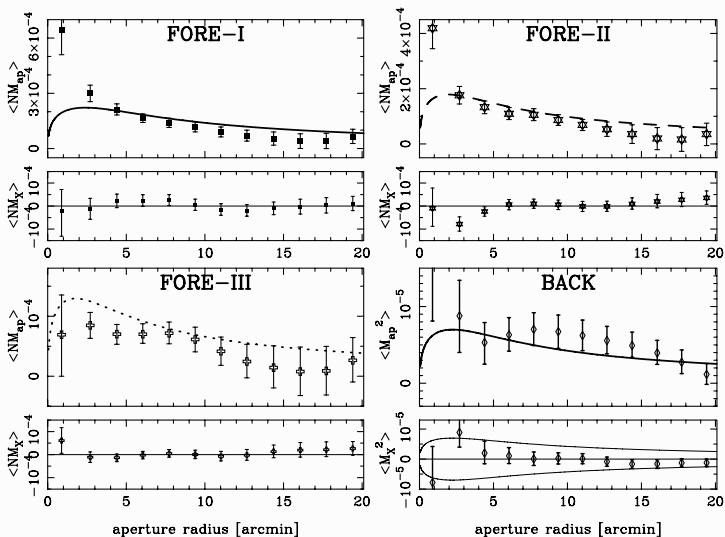
GGL results: model-indep. measurement of b and r II



Filled boxes, open stars, open crosses = FORE-I, FORE-II, FORE-III.

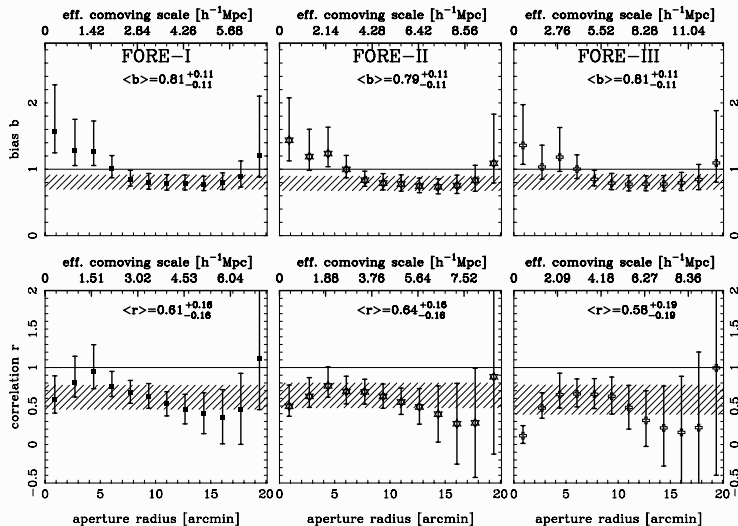
Galaxy clustering: Bias on small scales is not constant, but scale-dependent. Stronger galaxy clustering than from constant bias. (Simon et al. 2007), GaBoDS (Garching-Bonn Deep Survey).

GGL results: model-indep. measurement of b and r III



GGL and cosmic shear. (Simon et al. 2007), GaBoDS (Garching-Bonn Deep Survey).

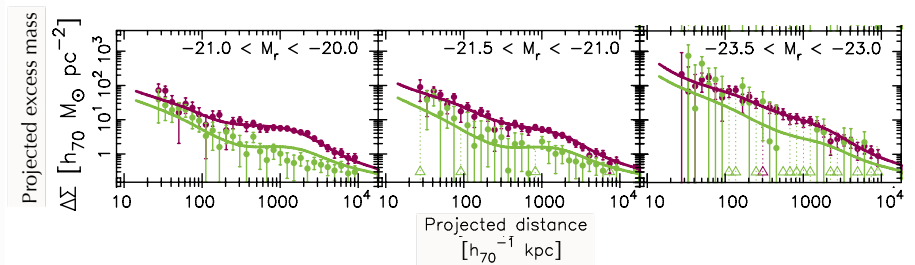
GGL results: model-indep. measurement of b and r IV



Bias and correlation coefficient. (Simon et al. 2007), GaBoDS (Garching-Bonn Deep Survey).

GGL: HOD model measurements

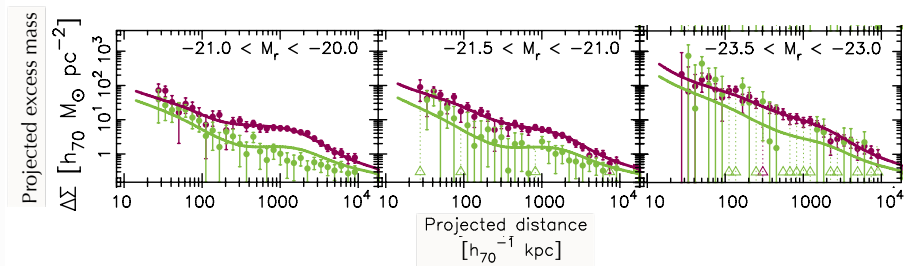
increasing luminosity \rightarrow



Purple=red early-type galaxies; Green=blue late-type galaxies. From (Velandier et al. 2014).

GGL: HOD model measurements

increasing luminosity \rightarrow

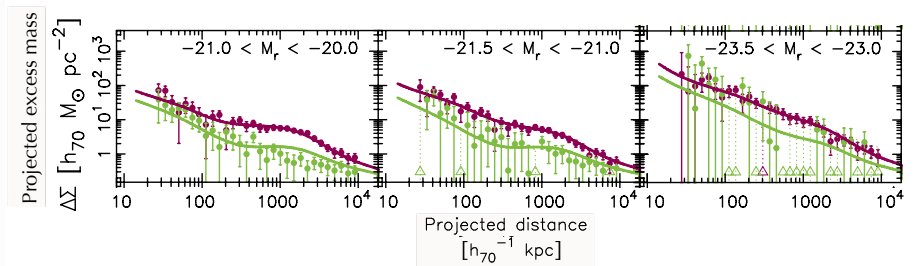


Purple=red early-type galaxies; Green=blue late-type galaxies. From (Vander et al. 2014).

- Red galaxies have larger associated mass than blue galaxies.
- Excess mass increases with luminosity. **Light traces mass.**

GGL: HOD model measurements

increasing luminosity \rightarrow

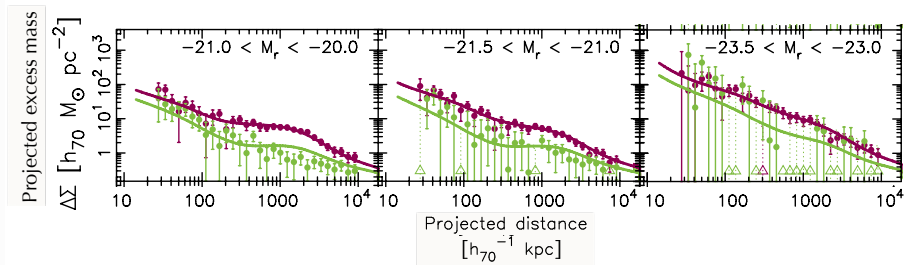


Purple=red early-type galaxies; Green=blue late-type galaxies. From (Vander et al. 2014).

- Red galaxies have larger associated mass than blue galaxies.
- Excess mass increases with luminosity. **Light traces mass.**
- Bump at 1 Mpc for low-luminosity red galaxies, disappears at higher L .
Red satellite galaxies.

GGL: HOD model measurements

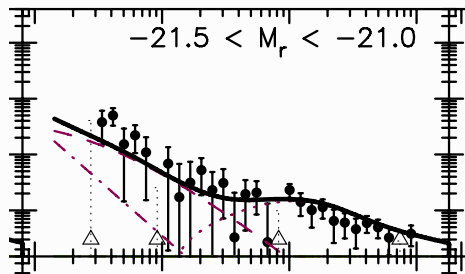
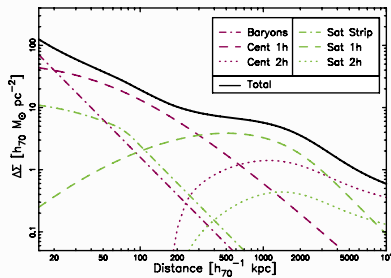
increasing luminosity \rightarrow



Purple=red early-type galaxies; Green=blue late-type galaxies. From (Velandier et al. 2014).

- Red galaxies have larger associated mass than blue galaxies.
- Excess mass increases with luminosity. **Light traces mass.**
- Bump at 1 Mpc for low-luminosity red galaxies, disappears at higher L . **Red satellite galaxies.**
- Bump at slightly larger scale for blue galaxies. **2-halo term, from clustered nearby galaxies.**

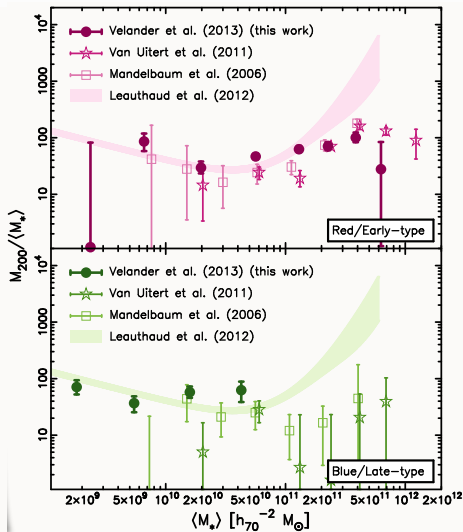
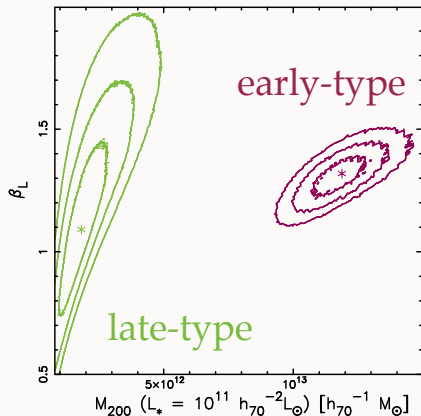
GGL: HOD model



HOD model, (Vander et al. 2014).

GGL: M/L parameters

$$M_{200} = M_{0,L} \left(\frac{L}{L_{\text{fid}}} \right)^{\beta_L}$$



(Velander et al. 2014).

Modified gravity

General, perturbed Friedmann-Lemaître Robertson Walker (FLRW) metric:

$$ds^2 = \left(1 + \frac{2\Psi}{c^2}\right) c^2 dt^2 - a^2(t) \left(1 - \frac{2\Phi}{c^2}\right) dl^2,$$

Valid for weak fields, (Bardeen) potentials $\Psi, \Phi \ll c^2$.

- In GR, and absence of anisotropic stress: $\Psi = \Phi$.
- In most modified gravity models: $\Psi \neq \Phi$! Very generic signature for MoG.

Some characteristics

- Ψ is Newtonian potential. Time-like. Quantifies time dilation.
- Ψ is gravitational action on non-relativistic objects (e.g. galaxies).
- Φ is space-like. Describes spatial curvature.
- $\Psi + \Phi$ is gravitational action on relativistic objects (e.g. photons; lensing!). [Photons travel equal parts of space and time. This is the origin for the factor two in GR equations compared to Newtonian mechanics!]

Testing GR I

Idea of a null test

Measure difference in potentials to test GR: Galaxy clustering for Ψ , weak lensing for $\Psi + \Phi$.

Modified Poisson equation

Potentials are related to density contrast δ via Poisson equation. Generalise to account for MoG, and write in Fourier space:

$$\begin{aligned}k^2 \tilde{\Psi}(k, a) &= 4\pi G a^2 [1 + \mu(k, a)] \rho \tilde{\delta}(k, a); \\k^2 [\tilde{\Psi}(k, a) + \tilde{\Phi}(k, a)] &= 8\pi G a^2 [1 + \Sigma(k, a)] \rho \tilde{\delta}(k, a).\end{aligned}$$

With free parameters/functions μ, Σ . GR: $\mu = \Sigma = 0$.

Testing GR II

Probes of Bardeen potentials

Assuming linear, deterministic bias ($b = \text{const}$, $r = 1$).

- Galaxy clustering measures Ψ and b ; $\langle \delta_g^2 \rangle \propto b^2 P_\Psi$.
- GGL measures $\Psi + \Phi$ and b ; $\langle \delta_g \delta \rangle \propto b P_{\Psi+\Phi}$.

→ form ratio to get rid of cosmology dependence!

However, bias still remains, need another observable.

- RSD anisotropy parameter; $\beta = \frac{1}{b} \frac{d \ln D_+(a)}{d \ln a}$.

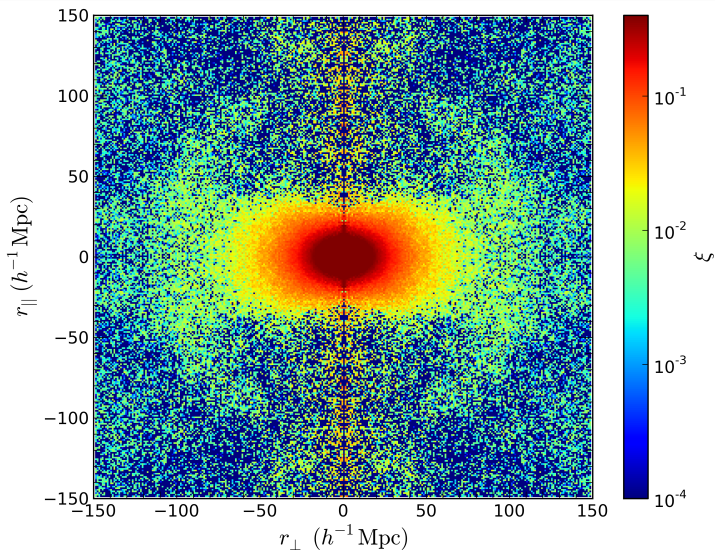
Can be measured from redshift space galaxy clustering along ($\mu = \cos \theta = 1$) and perpendicular ($\mu = 0$) to line of sight. Linear power spectrum:

$$P(k, \mu) = P(k) (1 + \beta \mu^2)^2.$$

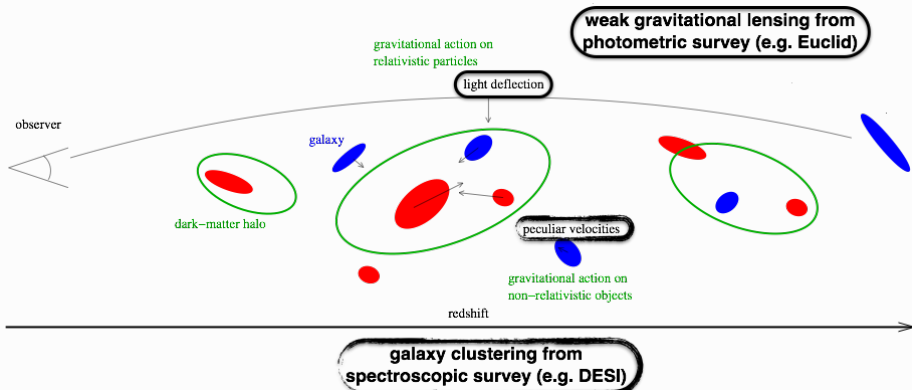
E_G parameter

$$E_G \simeq \frac{1}{\beta} \frac{\langle \delta_g \delta \rangle}{\langle \delta_g^2 \rangle}$$

Parenthesis: Anisotropic clustering



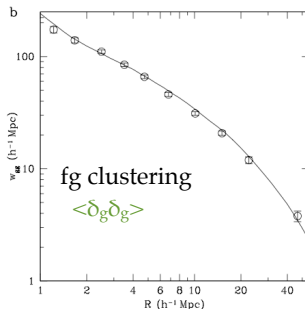
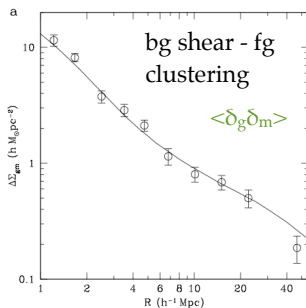
BOSS, from (?).



Testing GR: results I

SDSS

(Reyes et al. 2010)



$$\beta = 0.309 \pm 0.035$$

from SDSS galaxy clustering
(redshift-space distortions)
Tegmark et al. (2006)

Testing GR: results II

Introducing new observable to exclude small scales:

$$\begin{aligned}\Upsilon_{\text{gm}}(R) &= \Delta\Sigma_{\text{gm}}(R) - \left(\frac{R_0}{R}\right)^2 \Delta\Sigma_{\text{gm}}(R_0) \\ &= \frac{2}{R^2} \int_{R_0}^R dR' R' \Sigma_{|rmgm}(R') - \Sigma_{\text{gm}}(R') + \left(\frac{R_0}{R}\right)^2 \Sigma_{\text{gm}}(R_0),\end{aligned}$$

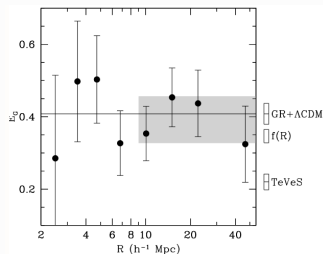
(Baldauf et al. 2010).

Define in analogy Σ_{gg} .

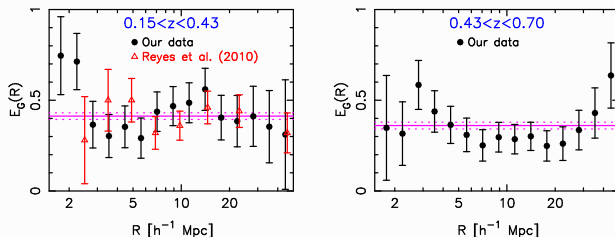
Then modified E_{G} probe of gravity:

$$E_{\text{G}}(R) = \frac{1}{\beta} \frac{\Sigma_{\text{gm}}(R)}{\Sigma_{\text{gg}}(R)}.$$

Testing GR: results III



From (Reyes et al. 2010).



From (Blake et al. 2016).

Recall: Aperture-mass definition

Yesterday we introduced the aperture-mass as convolution of the shear field with a filter Q ,

$$M_{\text{ap}}(\theta, \boldsymbol{\vartheta}) = \int d^2\vartheta' Q_{\theta}(|\boldsymbol{\vartheta} - \boldsymbol{\vartheta}'|) \gamma_t(\boldsymbol{\vartheta}')$$

and claimed that this was equivalent of convolving the convergence with another filter U ,

$$M_{\text{ap}}(\theta, \boldsymbol{\vartheta}) = \int d^2\vartheta' U_{\theta}(|\boldsymbol{\vartheta} - \boldsymbol{\vartheta}'|) \kappa^{\text{E}}(\boldsymbol{\vartheta}'), \quad (3)$$

(Kaiser et al. 1994, Schneider 1996).

Relation between U and Q I

First, place aperture at center, $\vartheta = 0$. Assume that the filter function $U_\theta(\vartheta)$ has support $[0; \theta]$, (note that θ can be ∞ .)

Introduce angle-averaged convergence,

$$\langle \kappa \rangle(\vartheta) := \frac{1}{2\pi} \int_0^{2\pi} d\varphi \kappa(\vartheta, \varphi),$$

and write aperture-mass

$$M_{\text{ap}}(\theta) = \int_0^\theta d\vartheta \vartheta U_\theta(\vartheta) \langle \kappa \rangle(\vartheta).$$

Integrate in parts, defining

$$\vartheta U_\theta(\vartheta) =: \frac{dX_\theta(\vartheta)}{d\vartheta} \quad \leftarrow \quad X_\theta(\vartheta) = \int_0^\vartheta d\vartheta' \vartheta' U_\theta(\vartheta')$$

to get

$$M_{\text{ap}}(\theta) = [X_\theta(\vartheta) \langle \kappa \rangle(\vartheta)]_0^\theta - \int_0^\theta d\vartheta U_\theta(\vartheta) \frac{d\langle \kappa \rangle(\vartheta)}{d\vartheta}.$$

Relation between U and Q II

To get rid of the boundary term, we demand that U be a *compensated filter* function, i.e.

$$X_{\theta}(\theta) = \int_0^{\theta} d\vartheta \vartheta U_{\theta}(\vartheta) = 0.$$

This means, that M_{ap} is not sensitive to a constant convergence κ_0 .

Why?

This makes it independent of the *mass-sheat degeneracy*.

We insert the expression for the derivative of the circularly averaged convergence from the TD,

$$\frac{d\langle\kappa\rangle(\vartheta)}{d\vartheta} = \frac{d\bar{\kappa}(\leq \vartheta)}{d\vartheta} - \frac{d\bar{\gamma}(\leq \vartheta)}{d\vartheta} = \frac{2}{\vartheta} \langle\gamma_t\rangle(\vartheta) - \frac{d\langle\gamma\rangle(\vartheta)}{d\vartheta}$$

to get

$$M_{\text{ap}}(\theta) = \int_0^{\theta} d\vartheta U_{\theta}(\vartheta) \left[\frac{2}{\vartheta} X_{\theta}(\vartheta) \langle\gamma_t\rangle(\vartheta) - \frac{d\langle\gamma\rangle(\vartheta)}{d\vartheta} \right].$$

Relation between U and Q III

The second term is again integrated by parts. The boundary term vanishes as before and we are left with

$$M_{\text{ap}}(\theta) = \int_0^\theta d\vartheta \left[\vartheta \frac{2}{\vartheta^2} X_\theta(\vartheta) - \frac{dX(\vartheta)}{d\vartheta} \right] \langle \gamma \rangle(\vartheta).$$

This can be transformed back to the form

$$M_{\text{ap}}(\theta) = \int d^2\vartheta Q_\theta(\vartheta) \langle \gamma_t \rangle(\vartheta)$$

and we get the relation between U and Q :

$$Q_\theta(\vartheta) = \frac{2}{\vartheta^2} \int_0^\vartheta d\vartheta' \vartheta' U_\theta(\vartheta') - U_\theta(\vartheta).$$

Relation between U and Q IV

Some properties

- If U has finite support, so has Q . [This follows from U being compensated]. That means that aperture-mass can be obtained from shear on finite region.
[This is not true when computing κ from γ without filters. Formally, this relation requires all of \mathbb{R}^2 .]

Reminders: lensing potential and convergence I

On day I we defined the lensing potential ψ at sky 2D coordinate $\boldsymbol{\theta}$ for a source galaxy at comoving distance χ , in a flat universe

$$\psi(\boldsymbol{\theta}, \chi) = \frac{2}{c^2} \int_0^\chi d\chi' \frac{\chi - \chi'}{\chi\chi'} \phi(\chi' \boldsymbol{\theta}, \chi').$$

The lensing convergence κ is given by a 2D Poisson equation,

$$\kappa = \frac{1}{2} \Delta \psi.$$

We pulled the 2D Laplacian through the integral, and add the 3-component $\Delta_{\chi'\chi'}$ to yield the 3D Laplacian.

We then used the 3D Poisson equation to transform the 3D potential ϕ to the density contrast δ ,

$$\Delta \Phi = \frac{3H_0^2 \Omega_m}{2a} \delta,$$

Reminders: lensing potential and convergence II

and obtained

$$\kappa(\boldsymbol{\theta}, \chi) = \frac{3}{2} \Omega_{\text{m}} \left(\frac{H_0}{c} \right)^2 \int_0^\chi d\chi' \frac{(\chi - \chi')\chi'}{\chi a(\chi')} \delta(\chi' \boldsymbol{\theta}, \chi').$$

Finally, we introduced a source galaxy distribution $p(\chi)d\chi = p(z)dz$ to get the convergence for a population of galaxies

$$\kappa(\boldsymbol{\theta}) = \int_0^{\chi_{\text{lim}}} d\chi p(\chi) \kappa(\boldsymbol{\theta}, \chi) = \int_0^{\chi_{\text{lim}}} d\chi G(\chi) \chi \delta(\chi \boldsymbol{\theta}, \chi)$$

with the lensing efficiency

$$G(\chi) = \frac{3}{2} \left(\frac{H_0}{c} \right)^2 \frac{\Omega_{\text{m}}}{a(\chi)} \int_\chi^{\chi_{\text{lim}}} d\chi' p(\chi') \frac{\chi' - \chi}{\chi'}.$$

We then introduced the variance of the convergence,

$$\langle \kappa(\boldsymbol{\vartheta} + \boldsymbol{\theta}) \kappa(\boldsymbol{\vartheta}) \rangle = \langle \kappa \kappa \rangle(\boldsymbol{\theta}),$$

Reminders: lensing potential and convergence III

and wrote it in Fourier space to define the convergence power spectrum

$$\langle \hat{\kappa}(\ell) \hat{\kappa}^*(\ell') \rangle = (2\pi)^2 \delta_D(\ell - \ell') P_\kappa(\ell).$$

End reminder.

Spherical transformations

The Fourier transformation is only defined on a flat space. To perform Fourier transforms on fields defined on the spherical sky is fine on small scales, but breaks down on very large angles. The Fourier transform should be replaced by a *spherical harmonic transformation*.

However, we have to go back one step further: Convergence and shear are defined as second derivatives of the lensing potential,

$$\kappa = \frac{1}{2} (\partial_1 \partial_1 + \partial_2 \partial_2) \psi = \frac{1}{2} \nabla^2 \psi; \quad \gamma_1 = \frac{1}{2} (\partial_1 \partial_1 - \partial_2 \partial_2) \psi; \quad \gamma_2 = \partial_1 \partial_2 \psi.$$

These derivatives are defined in flat space and should also be replaced on the sphere.

So, we have to start again with the lensing potential.

Lensing potential on the sphere I

The lensing potential ψ from a population of source galaxies with redshift distribution $p_i(z)$ is given, in analogy to $\kappa(\boldsymbol{\theta})$ defined above, as

$$\psi(\boldsymbol{\theta}) = \frac{2}{c^2} \int_0^\infty \frac{d\chi}{\chi} \Phi[\chi, \chi\boldsymbol{\theta}; \chi] q(\chi),,$$

where the lensing efficiency q_i is given as

$$q(\chi) = \int_\chi^{\chi_{\text{lim}}} d\chi' p(\chi') \frac{\chi' - \chi}{\chi'}.$$

[Note: On day I we defined the lensing efficiency G for the convergence, which is different from q by just the “Poisson” prefactor,

$$G(\chi) = \frac{3}{2} \left(\frac{H_0}{c} \right)^2 \frac{\Omega_m}{a(\chi)} \int_\chi^{\chi_{\text{lim}}} d\chi' p(\chi') \frac{\chi' - \chi}{\chi'} = \frac{3}{2} \left(\frac{H_0}{c} \right)^2 \frac{\Omega_m}{a(\chi)} q(\chi).]$$

Lensing potential on the sphere II

Let us now derive the angular harmonics spectrum of ψ (spherical analogue of power spectrum)

Decompose potential into spherical harmonics,

$$\psi(\boldsymbol{\theta}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \psi_{\ell m} Y_{\ell m}(\boldsymbol{\theta}); \quad \psi_{\ell m} = \int_{\mathbb{S}^2} d\Omega \psi(\boldsymbol{\theta}) Y_{\ell m}^*(\boldsymbol{\theta}).$$

Completely analogous to CMB temperature — both ψ and T are scalar fields.

The harmonics expansion coefficient is, after insertion of the expression for ψ , and Fourier-transforming the 3D potential (note: in \mathbb{R}^3 , not on the sphere),

$$\psi_{\ell m} = \frac{2}{c^2} \int d\Omega Y_{\ell m}^*(\theta, \varphi) \int_0^\infty \frac{d\chi}{\chi} q(\chi) \int \frac{d^3k}{(2\pi)^3} \hat{\Phi}(\mathbf{k}; \chi) e^{-i\mathbf{k} \cdot \mathbf{r}}.$$

Insert spherical harmonics expansion of the plane wave basis function,

$$e^{i\mathbf{k} \cdot \mathbf{r}} = 4\pi \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} i^\ell j_\ell(k\chi) Y_{\ell m}(\theta, \varphi) Y_{\ell m}(\theta_k, \varphi_k),$$

Lensing potential on the sphere III

and make use of orthogonality relation of the spherical harmonics

$$\int d\Omega Y_{\ell m}(\theta, \varphi) Y_{\ell' m'}^*(\theta, \varphi) = \delta_{\ell\ell'} \delta_{mm'},$$

to yield

$$\psi_{\ell m} = \frac{i^\ell}{c^2 \pi^2} \int_0^\infty \frac{d\chi}{\chi} q(\chi) \int d^3k \hat{\Phi}(\mathbf{k}; \chi) j_\ell(k\chi) Y_{\ell m}(\theta_k, \varphi_k).$$

Angular harmonics (cross-)spectrum (between redshift bins i and j) of the lensing potential is defined as

$$\langle \psi_{\ell m, i} \psi_{\ell' m', j}^* \rangle = \delta_{\ell\ell'} \delta_{mm'} C_{ij}^\psi(\ell).$$

Using once more the orthogonality of the $Y_{\ell m}$'s, we get finally

$$\begin{aligned} C_{ij}^\psi(\ell) &= \frac{8}{c^4 \pi} \int_0^\infty \frac{d\chi}{\chi} q_i(\chi) \int_0^\infty \frac{d\chi'}{\chi'} q_j(\chi') \int dk k^2 j_\ell(k\chi) j_\ell(k\chi') P_\Phi(k; \chi, \chi') \\ &= \frac{8}{\pi} \left[\frac{3}{2} \Omega_m \left(\frac{H_0}{c} \right)^2 \right]^2 \int_0^\infty \frac{d\chi}{\chi} \frac{q_i(\chi)}{a(\chi)} \int_0^\infty \frac{d\chi'}{\chi'} \frac{q_j(\chi')}{a(\chi')} \int \frac{dk}{k^2} j_\ell(k\chi) j_\ell(k\chi') P_m(k; \chi, \chi'); \end{aligned}$$

Shear on the sphere I

Preparation

Define complex derivative operator

$$\partial := \partial_1 + i\partial_2.$$

From that we get

$$\partial\partial = \partial_1\partial_1 - \partial_2\partial_2 + 2i\partial_1\partial_2.$$

Thus, we can rewrite the shear

$$\gamma_1 = \frac{1}{2}(\partial_1\partial_1 - \partial_2\partial_2)\psi; \quad \gamma_2 = \partial_1\partial_2\psi.$$

in complex form

$$\gamma = \frac{1}{2}\partial\partial\psi; \quad \gamma^* = \frac{1}{2}\partial\partial^*\psi.$$

The corresponding derivative on the sphere is called *edth* derivative $\bar{\partial}$ (Castro et al. 2005).

Shear on the sphere II

We write

$$\gamma(\boldsymbol{\theta}) = \frac{1}{2} \bar{\partial} \partial \psi(\boldsymbol{\theta}); \quad \gamma^*(\boldsymbol{\theta}) = \frac{1}{2} \bar{\partial}^* \partial^* \psi(\boldsymbol{\theta}).$$

Inserting the spherical harmonics expansion of $\psi \rightarrow 2^{\text{nd}}$ edth derivatives of $Y_{\ell m}$.

This defines a new object, the *spin-weighted spherical harmonics* ${}_2Y_{\ell m}$.

[Note: Spin $s = 2$ because second-derivatives; each derivative $\bar{\partial}$ ($\bar{\partial}^*$) raises (lowers) spin by one.]

Therefore,

$$\begin{aligned} (\gamma_1 \pm i\gamma_2)(\boldsymbol{\theta}) &= \sum_{\ell m} \pm {}_2\gamma_{\ell m} \pm {}_2Y_{\ell m}(\boldsymbol{\theta}); \\ {}_2\gamma_{\ell m} &= \int_{\mathbb{S}^2} d\Omega \gamma(\boldsymbol{\theta}) {}_2Y_{\ell m}^*(\boldsymbol{\theta}); \\ -{}_2\gamma_{\ell m} &= \int_{\mathbb{S}^2} d\Omega \gamma^*(\boldsymbol{\theta}) {}_2Y_{\ell m}^*(\boldsymbol{\theta}). \end{aligned}$$

Shear on the sphere III

These objects ${}_{\pm 2}Y_{\ell m}$ are eigen functions of $\tilde{\partial}$:

$$t(\ell, 2) {}_2Y_{\ell m}(\boldsymbol{\theta}) = \tilde{\partial}^2 Y_{\ell m}(\boldsymbol{\theta}); \quad t(\ell, 2) {}_{-2}Y_{\ell m}(\boldsymbol{\theta}) = (\tilde{\partial}^*)^2 Y_{\ell m}(\boldsymbol{\theta}).$$

with the spin pre-factor (Bernardeau et al. 2012)

$$t(\ell, 2) = \sqrt{\frac{(\ell+2)!}{(\ell-2)!}} = \sqrt{(\ell-1)\ell(\ell+1)(\ell+2)},$$

And we get the relation between shear and potential coefficients,

$${}_{\pm 2}\gamma_{\ell m} = \frac{1}{2}t(\ell, 2)\psi_{\ell m}.$$

Shear on the sphere IV

Now it is easy to write down the shear angular harmonics spectrum, again for bins i and j to be general:

$$C_{ij}^{\gamma}(\ell) = \frac{2}{\pi} t^2(\ell, 2) \left[\frac{3}{2} \Omega_{\text{m}} \left(\frac{H_0}{c} \right)^2 \right]^2 \int_0^{\infty} \frac{d\chi}{\chi} \frac{q_i(\chi)}{a(\chi)} \int_0^{\infty} \frac{d\chi'}{\chi'} \frac{q_j(\chi')}{a(\chi')} \\ \times \int_0^{\infty} \frac{dk}{k^2} P_{\text{m}}(k, \chi, \chi') j_{\ell}(k\chi) j_{\ell}(k\chi')$$

Flat-sky approximation

Going back to flat sky from the full spherical expression, we replace again the edth by the ordinary flat-space derivatives.

(Hu 2000) calculates the derivatives of the spherical harmonics as

$$\ell^2 {}_{\pm 2}Y_{\ell m}(\theta, \varphi) \approx e^{\mp 2i\phi_\ell} (\partial_1 \pm i\partial_2)^2 Y_{\ell m}(\theta, \varphi)$$

and we get a slightly different expression for the shear power spectrum, with the replacement

$$t^2(\ell, 2) = (\ell - 1)\ell(\ell + 1)(\ell + 2) \rightarrow \ell^4.$$

[Note: this is a slightly strange approach, since we first expand the field into spherical harmonics, and then perform the flat-sky approximation of the derivatives. More consistent would be to start with the Fourier transform. But I don't know how to derive the ℓ^4 factor in this case without making additional assumptions, see later.]

Limber approximation I

In short:

We use the identity of the Bessel functions

$$j_\ell(x) = \sqrt{\frac{\pi}{2x}} J_{\ell+1/2}(x)$$

and replace Bessel function $J_{\ell+1/2}(k\chi)$ by a Dirac delta $\delta_D(\ell + 1/2 - k\chi)$ (maximum of Bessel function).

Thus:

- Only modes $\ell + 1/2 \approx k\chi$ contribute.
- Only modes at $\chi \approx \chi'$ contribute.

Limber approximation II

Note: From linear perturbation theory, which holds on large scales:

$$\hat{\delta}(\mathbf{k}, \chi) = D_+(\chi) \hat{\delta}_0(\mathbf{k}) \rightarrow P_\delta(k, \chi, \chi') = D_+(\chi) D_+(\chi') P_{\delta,0}(k),$$

modes at arbitrary distances $\chi \neq \chi'$ are correlated.

The shear harmonic spectrum then simplifies to:

$$C_{ij}^\gamma(\ell) = t^2(\ell, 2) \left[\frac{3}{2} \Omega_m \left(\frac{H_0}{c} \right)^2 \right]^2 \int_0^\infty \frac{dk}{k^3} \\ \times \int_0^\infty \frac{d\chi}{\chi^{3/2}} \frac{q_i(\chi)}{a(\chi)} \delta_D(\ell + 1/2 - k\chi) \int_0^\infty \frac{d\chi'}{\chi'^{3/2}} \frac{q_j(\chi')}{a(\chi')} \delta_D(\ell + 1/2 - k\chi') P_m(k, \chi)$$

The comoving integrals are solved trivially with the Dirac delta, yielding a further k^{-2} due to a variable transformation $d\chi = d(k\chi)/k$.

Limber approximation III

We then substitute $(\ell + 1/2)^3 = (k\chi)^3$, and perform another variable transformation

$dk k^{-2} = d[(\ell + 1/2)\chi^{-1}]k^{-2} = (\ell + 1/2)d\chi\chi^{-2}k^{-2} = d\chi(\ell + 1/2)^{-1}$, and get

$$C_{ij}^{\gamma}(\ell) = \frac{t^2(\ell, 2)}{(\ell + 1/2)^4} \left[\frac{3}{2}\Omega_{\text{m}} \left(\frac{H_0}{c} \right)^2 \right]^2 \int d\chi \frac{q_i(\chi)q_j(\chi)}{a^2(\chi)} P_{\text{m}} \left(\frac{\ell + 1/2}{\chi}; \chi \right).$$

Approximations accuracy I

Most pre-2014 used the flat-sky approximation ($l^2(\ell, 2) \approx \ell^4$) and further $\ell + 1/2 \approx \ell$. Then the prefactor cancels.

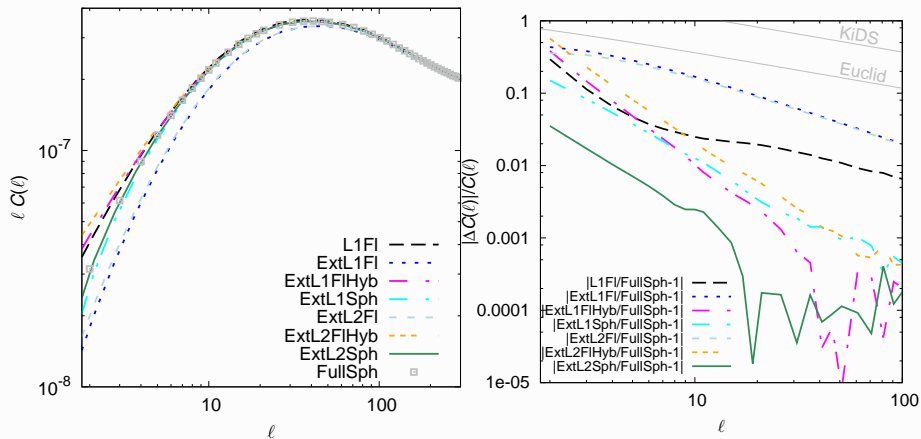
This *standard Limber* approximation is accurate to 1% (10%) for $\ell < 60(4)$.

The next logical approximation is *extended Limber*, with $\ell + 1/2$ kept in prefactor and power-spectrum argument. This is actually a worse approximation than standard Limber, since the approximated prefactor converges only with $\mathcal{O}(\ell^1)$.

Better is *hybrid*, with ℓ in prefactor denominator, but $\ell + 1/2$ in integral.

Even better is *second-order Limber*.

Approximations accuracy II



From (Kilbinger et al. 2017).

Spherical correlation function I

For the correlation function on the sphere, the Bessel functions $J_{0,4}$ are replaced by the *reduced Wigner D-matrices*,

$$\xi_+(\theta) = \frac{1}{4\pi} \sum_{\ell=2}^{\infty} (2\ell+1) C^\gamma(\ell) d_{22}^\ell(\theta); \quad \xi_0(\theta) = \frac{1}{4\pi} \sum_{\ell=2}^{\infty} (2\ell+1) C^\gamma(\ell) d_{2-2}^\ell(\theta).$$

These are defined as follows:

$$\begin{aligned} D_{ss'}^\ell(\alpha, \beta, -\gamma) &= \sum_m \frac{4\pi}{2\ell+1} Y_{\ell m}^*(\theta, \varphi) {}_s' Y_{\ell m}^*(\theta', \varphi') \\ &= \exp^{-is'\alpha} d_{ss'}^\ell(\beta) \exp^{is\gamma}. \end{aligned}$$

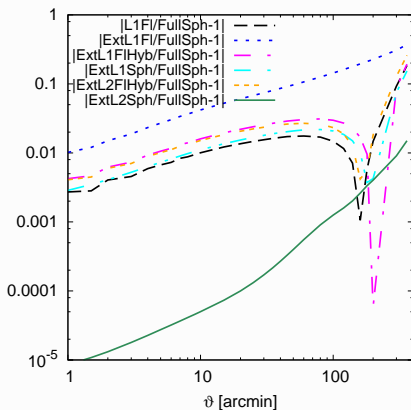
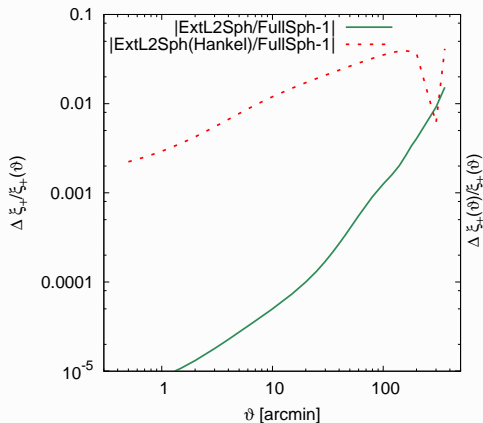
Angles:

β = angle between (θ, φ) and (θ', φ') .

$\alpha[\gamma]$: angle to rotate \hat{e}_θ about (θ, φ) $[(\theta', \varphi')]$ perpendicular to connecting line between (θ, φ) and (θ', φ') .

(Chon et al. 2004).

Spherical correlation function II



From (Kilbinger et al. 2017).

Shear bias

For ~~basically~~ all shape measurement methods: observed shear \neq true shear.
This is called **shear bias**.

Reminder: Write as multiplicative and additive bias:

$$\langle \varepsilon_{\alpha}^{\text{obs}} \rangle = g_{\alpha}^{\text{obs}} = (1 + m_{\alpha}) g_{\alpha}^{\text{true}} + c_{\alpha}; \quad \alpha = 1, 2.$$

There is also ellipticity bias, which is different:

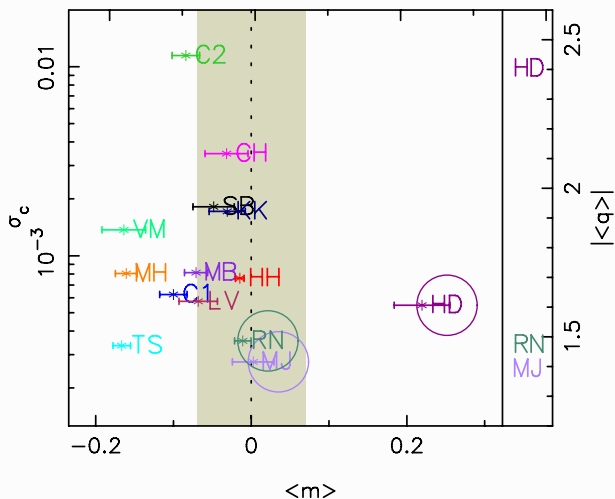
$$\varepsilon_i^{\text{obs}} = (1 + m'_i) \varepsilon_i^{\text{true}} + c'_i; \quad i = 1, 2.$$

Typical values:

year	program	m	c	$\sigma(c)$
2006	STEP I	0.1		10^{-3}
2012	CFHTLenS	0.06	0.002	
2013	great3	0.01	10^{-3}	
2014	DES	0.03–0.04	10^{-3}	
2016	KiDS	0.01–0.02	$8 \cdot 10^{-4}$	
2021	Euclid required	$2 \cdot 10^{-3}$	$5 \cdot 10^{-4}$	

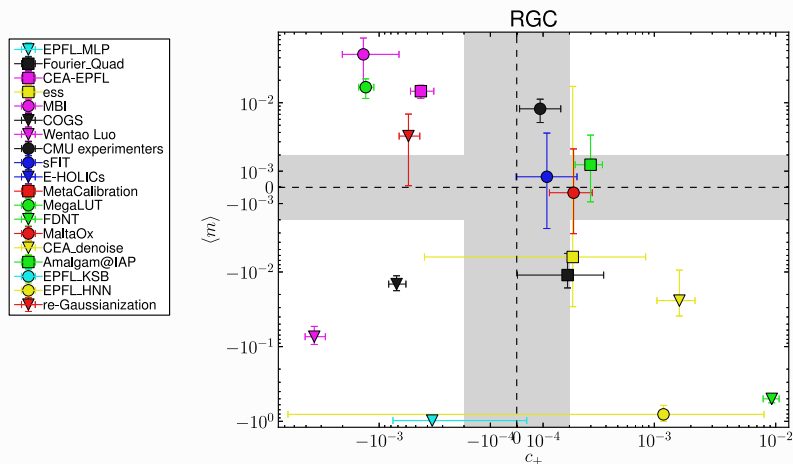
Shear bias and simulations I

From the STEP I shear measurement challenge (Heymans et al. 2006).



Shear bias and simulations II

From the great3 shear measurement challenge ([Mandelbaum et al. 2015](#)).



Shear bias and simulations III

Interpret with caution!

- Small biases because simulations are not realistic enough? E.g. constant PSF, analytical galaxy light distributions, simplistic noise, (constant shear)
- Simulation (challenges) only address part of the problem. Usually no blended galaxy images, star-galaxy separation, color effects, ...
- Calibrated or un-calibrated?

Amplitude of m, c not that important, since they can be calibrated empirically.
What counts are $\Delta m, \Delta c$ after calibration!

More on this in a few slides.

Shear bias and simulations IV

A very general statement (see Part I day 2):

Most ellipticity estimators are non-linear pixel light distribution. Noise then creates biases in the estimator. This is called **noise bias**.

Thus, observed shear needs to be de-biased (calibrated) using simulations.

There are a few unbiased estimators:

- Not normalised to total flux: maybe unbiased, but very large variance
- Bayesian estimators, sample posterior distribution, unbiased if correct model, likelihood and prior.

Prior needs to be estimated from simulations or deep survey!

Sources of bias

Reminder:

- Noise bias
- Model bias
 - Model-fitting method: incorrect model, complex galaxy morphology
 - Direct estimation: inappropriate filter function for weighted moments; truncated eigenfunction decomposition
 - Ellipticity gradients
 - Color gradients
- PSF residuals
- CTI (charge transfer inefficiency)
- Selection effects (population biases). Detection probability depends on ellipticity, orientation with PSF, pixel scale
- **New:** Environmental effects
 - Unresolved faint galaxies

Shear calibration

The bias should be *robust* for method to be *calibratable*.

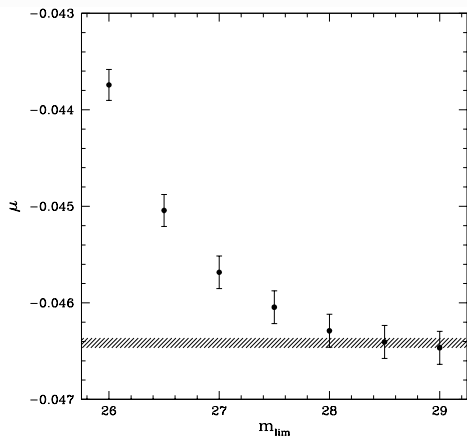
Define *sensitivity* as dependence of bias with respect to parameters, or

$$|\partial m / \partial p_i|, \quad \text{for } \mathbf{p} = \text{set of parameters.}$$

A method is calibratable, see (Hoekstra et al. 2017), if

- the sensitivity is small (otherwise simulation sampling in p too costly)
- does not depend on too many parameters
- those parameters can be measured accurately (e.g. intrinsic ellipticity dispersion σ_ε from Euclid Deep Survey \rightarrow requirement on accuracy of measured σ_ε sets area of calibration fields)
- those parameters can be reasonably simulated to estimate sensitivity
- difficult if parameter is correlated with shear signal (e.g. local galaxy density with large-scale structure, correlated with shear signal, magnification)

Shear calibration: Unresolved faint galaxies I



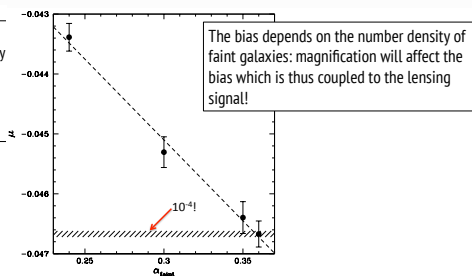
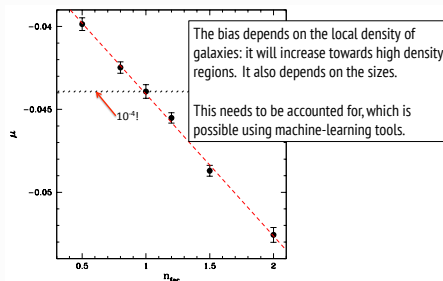
Overall values on y -axis (amplitude of m) not really important, will be corrected for.

Need simulation up to very high depth, until plateau in m is reached ($\partial m / \partial m_{\text{lim}} = 0$).

Error bars need to decrease to match hashed region.

Multiplicative bias m (here μ) for galaxies $20 < m < 24.5$ as function of limiting magnitude of simulated galaxies. From (Hoekstra et al. 2017).

Shear calibration: Unresolved faint galaxies II



Shear calibration using image simulations: tricks of the trade I

Again: multiplicative and additive bias,

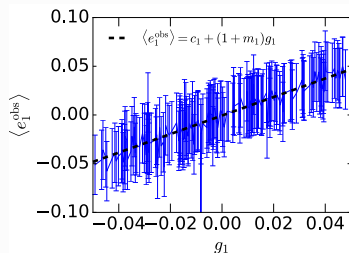
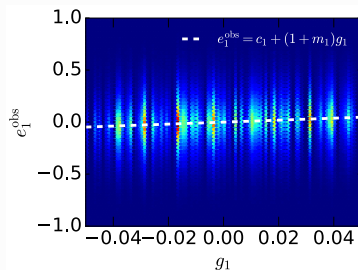
$$\langle \varepsilon_{\alpha}^{\text{obs}} \rangle = g_{\alpha}^{\text{obs}} = (1 + m_{\alpha}) g_{\alpha}^{\text{true}} + c_{\alpha}; \quad \alpha = 1, 2.$$

for sample of galaxies with vanishing intrinsic ellipticity $\langle \varepsilon_{\alpha}^{\text{I}} \rangle = 0$.

How can we determine the multiplicative bias?

Simple method

From linear fit of many simulated pairs $(\varepsilon_{\alpha}^{\text{obs}}, g_{\alpha}^{\text{true}})$.

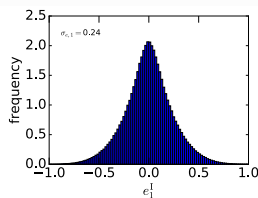


Shear calibration using image simulations: tricks of the trade II

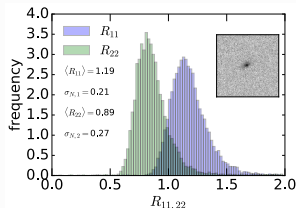
Error on best-fit m_α given by width in ε^{obs} (including measurement errors), g^{true} , and stochasticity of galaxy images (from pixel noise),

$$\sigma_{m,\alpha} = \frac{1}{\sqrt{N}} \sqrt{\sigma_{R,\alpha}^2 + \frac{\sigma_{S,\alpha}^2}{\sigma_{g,\alpha}^2}}$$

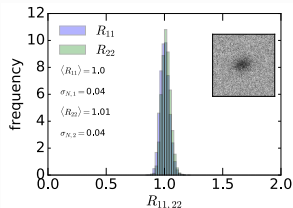
Second terms is dominant in most cases.



Ell. distribution.



Stochasticity for low SNR.



Stochasticity for high SNR.

Shear calibration using image simulations: tricks of the trade III

Noise suppression

Simulate pairs of galaxies with same shear and **orthogonal** intrinsic ellipticity (rotated by 90 degrees),

$$\varepsilon_A^I + \varepsilon_B^I = 0.$$

This however does not mean that the *observed* ellipticity vanishes, due to:

- Measurement stochasticity
- Ellipticity bias, if depends on galaxy orientation wrt PSF, shear, (pixelization)
- Selection effects, one pair member might drop out of sample

Shear calibration using image simulations: tricks of the trade IV

More advanced noise suppression: ring test. Simulate n galaxies with equidistant intrinsic ellipticity on ring around 0.

Derivative method

Write shear bias for individual galaxies, and as matrix equation (?):

$$\epsilon_{\alpha}^{\text{obs}} = \mathbf{R} \mathbf{g}^{\text{true}} + \mathbf{c}$$

The **shear response** tensor \mathbf{R} generalizes m : $1 + m_{\alpha} = R_{\alpha\alpha}$.

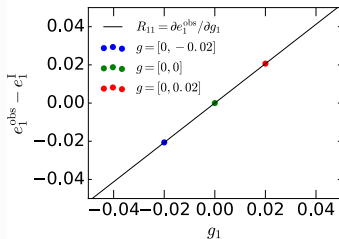
To get population bias, average over measured shear responses $\langle R \rangle$, and correct measured ellipticities by $\langle R \rangle^{-1}$.

Measure individual \mathbf{R} as numerical derivatives

$$R_{\alpha\beta} = \frac{\partial \epsilon_{\alpha}^{\text{obs}}}{\partial g_{\beta}}$$

by simulating the same galaxy several times with small added shear $\pm \Delta g_{\alpha} \sim 0.02$. With **same noise realisation** this measurement is extremely precise!

Shear calibration using image simulations: tricks of the trade V



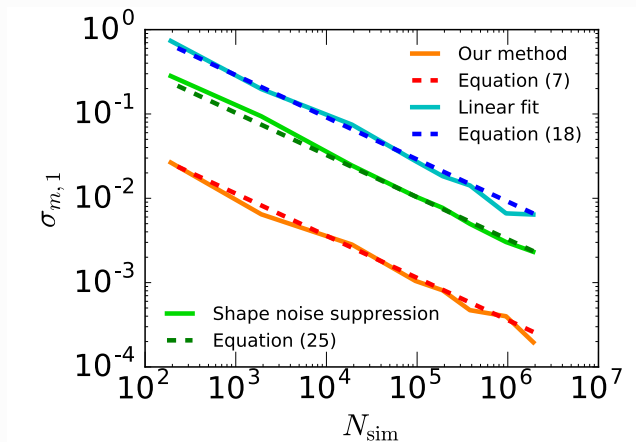
This measurement is independent of ellipticity (observed and intrinsic) and thus removes the main uncertainty of error!

Note: For a different noise realisation, the obtained \mathbf{R} can be quite different. But the use of many simulated galaxy images assures the sampling of the distribution of R , no additional error is introduced on the population bias. Error on bias estimate:

$$\sigma_{m,\alpha} = \frac{\sigma_{R,\alpha}}{\sqrt{N}}$$

Shear calibration using image simulations: tricks of the trade VI

This method requires a factor of several hundred fewer image simulations.



Shear calibration using image simulations: tricks of the trade VII

From (?).

Why do we need N -body simulations for WL I

- WL probes LSS on small, non-linear scales.
Cosmic shear: down to sub-Mpc. Surveys sensitive to $k \sim 50h/\text{Mpc}$.
Need theoretical prediction of non-linear power spectrum.
(Semi-)analytical approaches go to $k \sim 0.5h/\text{Mpc}$.
- Shear field follows non-Gaussian distribution.
Follows from the fact that δ in non-linear regime is non-Gaussian.
Complex survey geometry modify distribution.
At the least, need non-Gaussian covariance for likelihood.
Difficult from (semi-)analytical models (see previous point).
- Baryonic physics modifies dark-matter halo properties (profile, concentration, ...).
Model with hydro-dynamical simulations.
- Systematic effects that correlate to astrophysics or the LSS.
Can use forward modelling for complex physical processes in N -body simulation.

Why do we need N -body simulations for WL II

Examples

- Blended galaxy images lead to deselection of galaxies in crowded fields, which are correlated to high-density regions, that are then under-represented. This leads to biases in inferred $n(z)$, cosmological parameters.
- ξ_{sys} = correlation between stars and PSF-corrected galaxies = measure of PSF residuals in galaxy shapes. **But:** Need to account for chance alignment between PSF and LSS.
- Test mathematical useful approximations: Born, neglecting lens-lens coupling, reduced shear $g = \gamma/(1 - \kappa)$ versus shear γ . Most of these effects introduce higher-order correlations, again difficult to solve unless Gaussian limits.

Ray-tracing I

Principle of Ray-tracing

- Numerical evaluation of projection integral from particle distribution in N -body simulation.
- Most algorithms first project particles on *multiple lens planes* (Blandford & Narayan 1986), with $\Delta z/z$ of order 0.03 - 0.05.

Corresponds to a finite-sum discretization of the projection integral.

$$\alpha = \nabla_{\theta} \psi = \frac{2}{c^2} \int_0^{\chi} d\chi' \frac{\chi - \chi'}{\chi} \nabla_{\perp} \Phi(\mathbf{x}(\chi'), \chi').$$

On each lens plane, compute Jacobi matrix $A_{ij} = \partial \beta_i / \partial \theta_j = \delta_{ij} - \partial_i \partial_j \psi$.

Algorithm (Hilbert et al. 2009):

$$A_{ij}^{(k)}(\theta) = \delta_{ij} - \sum_{n=1}^{k-1} \frac{f_K^{(n,k)}}{f_K^{(k)}} U_{ij}^{(n)}(\theta). \quad U_{ij}^{(k)} = \frac{\partial^2 \psi^{(k)}(\boldsymbol{\beta}^{(k)})}{\partial \beta_i^{(k)} \partial \beta_j^{(k)}} = \frac{\partial \alpha_i^{(k)}(\boldsymbol{\beta}^{(k)})}{\partial \beta_j^{(k)}}.$$

Ray-shooting versus ray-tracing

Two methods are common to propagate photons for the projection:

1. **Ray shooting:**

Compute cumulative lensing potential ϕ on a grid. Light rays travel on (unperturbed) straight lines, corresponds to Born approximation.

2. **Ray tracing:**

Additionally compute deflection angle α , change direction of light ray accordingly.

Light rays travel on straight lines between lens planes, where they change direction.

Start at observer and shoot backwards. **Why?**

Ray-shooting versus ray-tracing

Two methods are common to propagate photons for the projection:

1. Ray shooting:

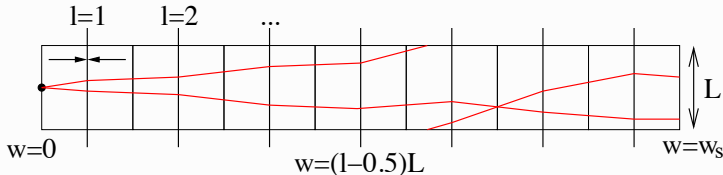
Compute cumulative lensing potential ϕ on a grid. Light rays travel on (unperturbed) straight lines, corresponds to Born approximation.

2. Ray tracing:

Additionally compute deflection angle α , change direction of light ray accordingly.

Light rays travel on straight lines between lens planes, where they change direction.

Start at observer and shoot backwards. **Why?**

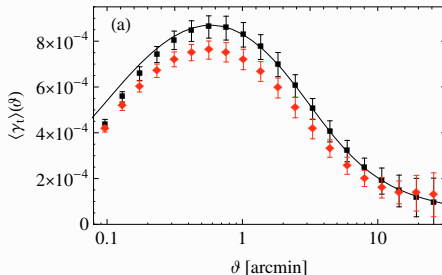


From Hartlap, PhD thesis 2005

Ray shooting

For cosmic shear ray-shooting is a very good (percent-level) approximation.

However, for **galaxy-galaxy lensing** this is not the case.



From (Hilbert et al. 2009).

This is because relative distance between light rays from two bg galaxies for cosmic shear not much affected by coherent deflection.

But distance between light ray from bg galaxy and fg galaxy position (impact parameter) is affected.

Ray-tracing approximations

- Ray-tracing through N -body output snapshot boxes: **Fixed cosmic time**, neglecting LSS evolution during photon travel time through box. Limit box size to $L \lesssim 300$ Mpc.

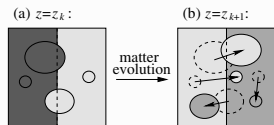
Larger boxes can be **split** and projected to more than one lens plane, but:

- Avoid cutting through halos
- Leads to loss of power on large scales

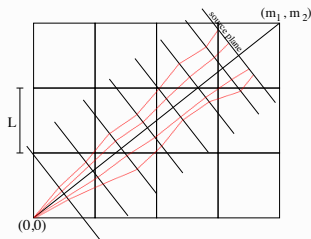
Use snapshots at different output times to account for time evolution.

If box size is small, boxes have to be concatenated. To avoid photons to encounter **repeated structures** at different epochs:

- Rotate and translate randomly.
- Shoot light rays under an skewed angle.



From (Hilbert et al. 2009).

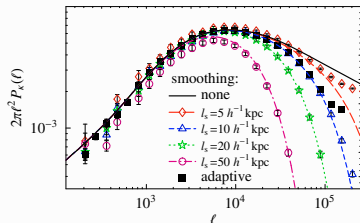


From Hartlap, PhD thesis 2005

Ray-tracing approximations

- Get shear and convergence by FFT, or finite differences in real space:
Smoothing is necessary to reduce Poisson noise of N -body discrete particle distribution.

However, other limitation is N -body resolution.



From (Hilbert et al. 2009).

- From Cartesian flat-sky simulations, lens planes are by construction **parallel**:
 - Neglects sky curvature.
 - Gradient of potential not orthogonal to light ray

This limits simulated field of view to a few degrees.

With convergence maps created on say grids of 1024^2 pixels \rightarrow resolution of around 0.2 arcmin.

- Newtonian** physics, neglects GR effects. Also, MoG simulations not possible under Newtonian approximation.

Further methods

- Compute lensing Jacobian on the fly while running N -body simulation (White & Hu 2000).

Circumvents lens plane projections, allows for slightly higher time resolution.

Easy for ray-shooting where photon tractories are known before hand, more difficult for ray tracing (Li et al. 2011).

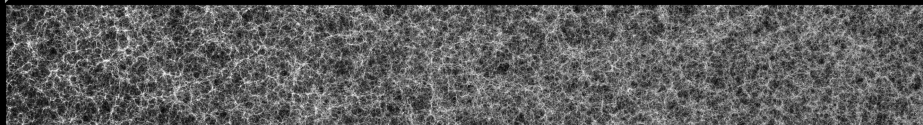
- Store density field at different time steps on surface moving towards box center (= observer) with speed of light, use those after run ends for lensing projections (Teyssier et al. 2009).
- Full-sky simulations, for large upcoming surveys, CMB lensing. Create spherical concentric shells around observer on the fly, project onto *lens spheres*. (Fosalba et al. 2008, Das & Bode 2008, Teyssier et al. 2009, Becker 2013).
- General-relativity simulations.
- Modified gravity simulations.
Take ~ 5 times compared to Newtonian ones.

Flagship mock galaxy catalog

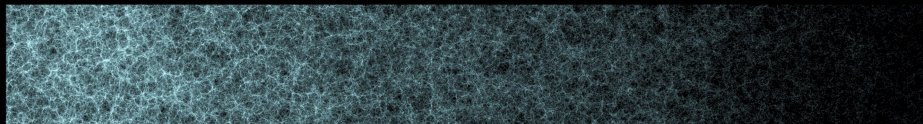
$z = 0$

All galaxies

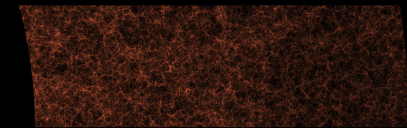
$z =$



$VIS < 24.5$



$NISP\ H\alpha > 1.e-16$



$z = 0.8$

$z = 1.9$

Hydro-dynamical simulations I

Important processes to simulate:

- Gas pressure, $R \sim 1 - 0.1$ Mpc, suppression of structure formation, gas distribution is more diffuse than dark matter
- Baryonic cooling, $R < 0.1$ Mpc ($k > 10/\text{Mpc}$), gas condenses into stars and galaxies, more strongly clustered than dark matter
- AGN and SN feedback

Simulation methods

Dark matter usually simulated as (very massive) particles.

Hydrodynamic physics often simulated in cells on a grid (adaptive).

Non-resolved physical processes, effective treatment within cell (“sub-grid physics”).

Hydrodynamical simulation can often not reproduce observational results, e.g. on AGN feedback. Need to calibrate simulations with observations.

Hydro-dynamical simulations II

Influence on WL

- Need to know total (dark + baryonic) power spectrum to 1-2% at k up to $10h/\text{Mpc}$.
- Baryons (15% of total matter) behave differently than dark matter, but dark matter is influenced by this, e.g. slightly follows distribution of baryons
- P_κ strongly influenced for $\ell \geq 1000$ to 3000 (depending on statistical errors).

Mitigation of baryonic effects

- Removing small scales from survey analysis.
- Model baryonic effects e.g. with halo model. Fit to simulations, marginalise over nuisance parameters or different models.
- Self-calibration using combination of observations. E.g. additional observations of halo structure (Zentner et al. 2008), power spectrum and bi-spectrum (Semboloni et al. 2013).

WL covariance

General definition

Covariance of data vector $\mathbf{d} = \{d_i\}, i = 1 \dots m$:

$$C_{ij} = \langle \Delta d_i \Delta d_j \rangle = \langle d_i d_j \rangle - \langle d_i \rangle \langle d_j \rangle,$$

Examples of \mathbf{d} :

$$d_i = P_\kappa(\ell_i); \quad d_i = \xi_+(\vartheta_i); \quad d_i = \langle M_{\text{ap}}(\theta_i) \rangle.$$

Case of data vector = $\hat{\xi}_\pm$

Recall the estimator for ξ_\pm :

$$\hat{\xi}_\pm(\theta) = \frac{\sum_{ij} w_i w_j (\varepsilon_{t,i} \varepsilon_{t,j} \pm \varepsilon_{\times,i} \varepsilon_{\times,j})}{\sum_{ij} w_i w_j}.$$

Very roughly:

$$C \sim \langle \xi_\pm \xi_\pm \rangle \sim \langle \varepsilon \varepsilon \varepsilon \varepsilon \rangle.$$

With weak-lensing relation $\varepsilon = \varepsilon^s + \gamma$:

$$C \sim \langle (\varepsilon^s)^4 \rangle + \langle (\varepsilon^s \gamma)^2 \rangle + \langle \gamma^4 \rangle \equiv D + M + V$$

WL covariance components

- $D = \sigma_\varepsilon^4$: Poisson noise from intrinsic ellipticities, shape noise
- M : mixed term
- V : shear covariance, cosmic variance, if shear field approximated having Gaussian distribution (which it does not):
 $V \sim 3\langle\gamma^2\rangle$.

Otherwise, need to account for connected 4-pt (tri-spectrum) term.

Gaussian covariance of power spectrum P_κ

$$\langle(\Delta P_\kappa)^2\rangle(\ell) = \frac{1}{f_{\text{sky}}(2\ell+1)} \left(\frac{\sigma_\varepsilon^2}{2\bar{n}} + P_\kappa(\ell) \right)^2.$$

Non-Gaussian covariance

Mode coupling

- Couples different ℓ -modes, leads to saturation of information content on small scales
- Tri-spectrum coupling on small scales
- Coupling of small with large scales: *halo sample variance* (HSV), *beat coupling*, *super-survey covariance* (SSC) (Takada & Hu 2013).
SSC decreases faster with f_{sky} than other terms \rightarrow sub-dominant for large surveys.

Modelling

- Tri-spectrum from halo model (+ PT)
- N -body simulations, **but**: difficult to include SSC
- From data, by spatial averaging over sub-fields, or Jackknife.

Spatial averaging: number of independent lines of sight n

For non-singular covariance of data vector with length m , need $n > m$.

For precision covariance (error bars on cosmo parameters of $< 5\%$, need $n > 10m$ (Taylor et al. 2013).

Likelihood function

Gaussian likelihood

$$L(\mathbf{d}|\mathbf{p}, M) = (2\pi)^{-m/2} |C(\mathbf{p}, M)|^{-1/2} \\ \times \exp \left[-\frac{1}{2} (\mathbf{d} - \mathbf{y}(\mathbf{p}, M))^{\dagger} C^{-1}(\mathbf{p}, M) (\mathbf{d} - \mathbf{y}(\mathbf{p}, M)) \right].$$

with \mathbf{d} = data vector, C = covariance matrix, \mathbf{y} = model, \mathbf{p} = (cosmo) parameter vector, M = cosmological model.

But: True likelihood is non-Gaussian.

Model non-Gaussianity of observables:

- N -body simulations (very time-consuming)
- Transform data to be more Gaussian
- Approximate Bayesian Computation (ABC) sampling

Bayesian parameter inference

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

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

For example:

$$h(\boldsymbol{\pi}) = \boldsymbol{\pi} : \quad \text{mean}$$

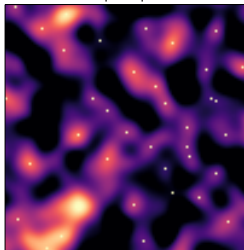
$$h(\boldsymbol{\pi}) = 1_{68\%} : \quad 68\% \text{ credible region}$$

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

WL peak counts: Why do we want to study peaks?

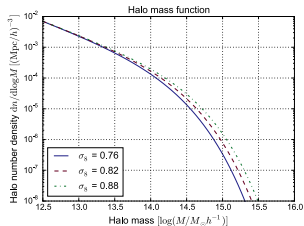
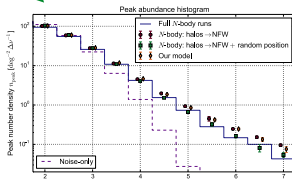
- WL peaks probe high-density regions \leftrightarrow **non-Gaussian** tail of LSS
- **First-order** in observed shear: less sensitive to systematics, circular average!
- High-density regions \leftrightarrow **halo mass function**, but **indirect** probe:
 - Intrinsic ellipticity **shape noise**, creating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

κ map and peaks



modelling

counting



interpretation ?

WL peak counts. What are peaks good for?

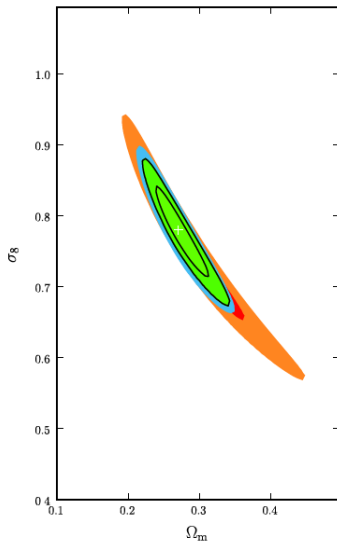
What do we gain from peak counting?

- Additional and complementary information and constraints compared to 2nd order shear
- Non-Gaussian information

Figure from Dietrich & Hartlap 2010

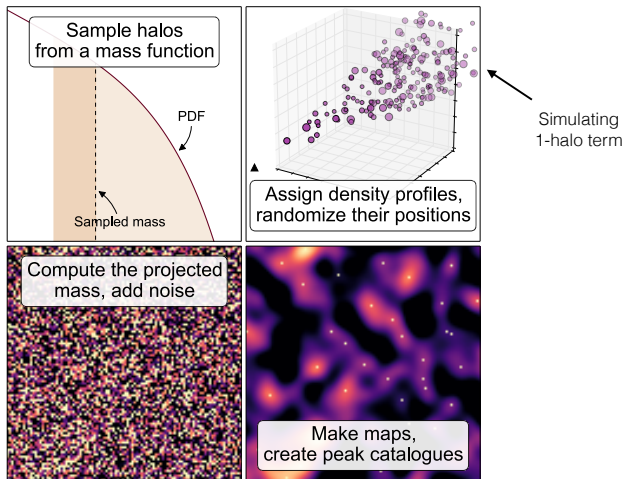
red/orange: cosmic shear

green: shear & peak



WL peaks: A fast stochastic model

Replace N-body simulations by Poisson distribution of halos



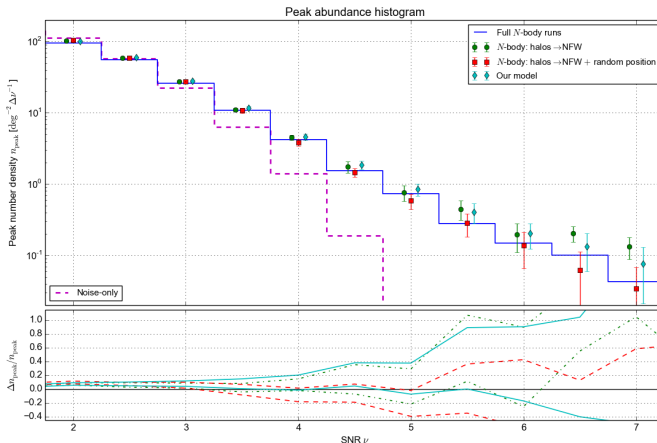
Lin, MK & Pires 2016

WL peaks: histograms

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

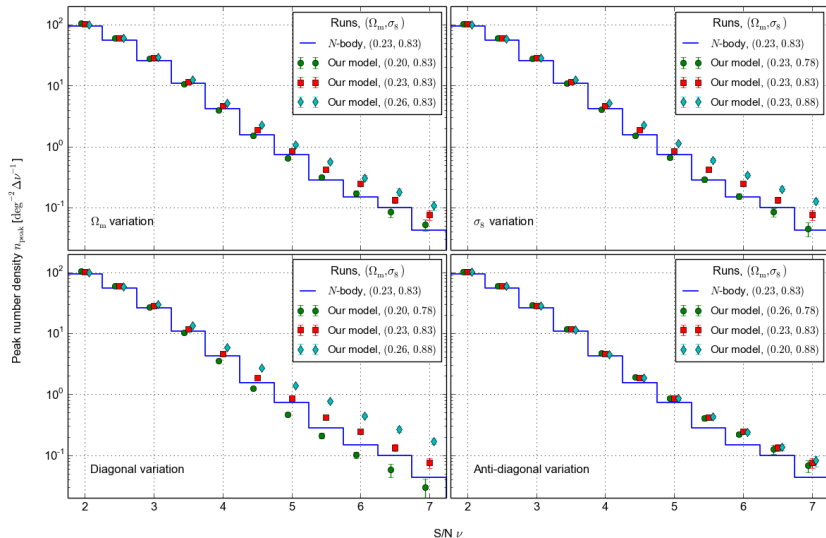
Test:



Field of view = 54 deg²; 10 halo redshift bins from $z = 0$ to 1; galaxies on regular grid, $z_{\text{g}} = 1.0$

WL peaks: cosmological parameters

Peak abundance histogram



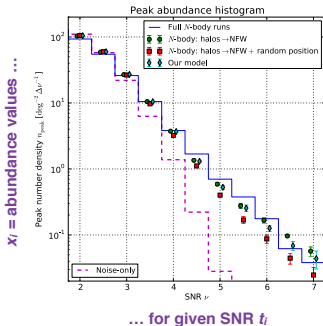
Lin & Kilbinger (2015a)

WL peaks: data vector choices

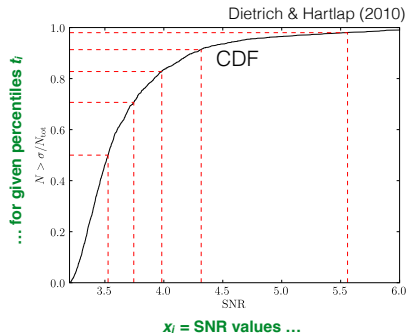
Data vector $\mathbf{x} = \mathbf{x}(t_i)$. Different cases:

- **Abundance** of peaks n_i as fct. of SNR \mathbf{v} (PDF; binned histogram) **or**
- **SNR values** \mathbf{v}_i at some percentile values of peak CDF)
 - with or without lower cut \mathbf{v}_{\min} .

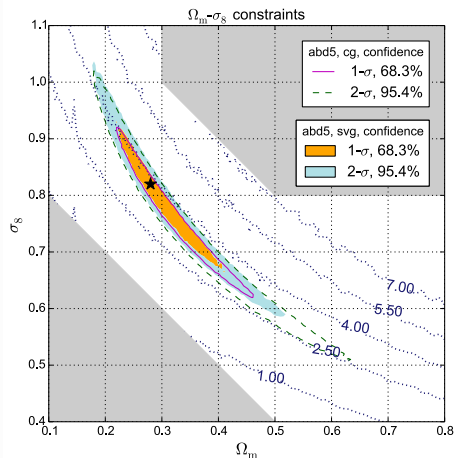
Lin & MK 2015a



or



WL peaks: Gaussian likelihood



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

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

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

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

ABC: Approximate Bayesian Computation I

Likelihood:
probability of data given
parameters and model

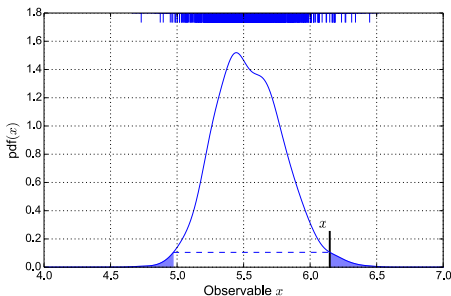
$$p(\boldsymbol{\pi}|\boldsymbol{x}, m) = \frac{L(\boldsymbol{x}|\boldsymbol{\pi}, m)P(\boldsymbol{\pi}|m)}{E(\boldsymbol{x}|m)}$$

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

Likelihood: how likely is it that model prediction $\boldsymbol{x}^{\text{mod}}(\boldsymbol{\pi})$ 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} .



ABC: Approximate Bayesian Computation II

Probability = p/N in frequentist sense.

Magic: Don't need to sample N models.

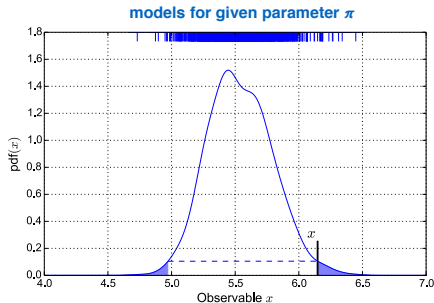
One per parameter π is sufficient with accept-reject algorithm.

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



ABC: Approximate Bayesian Computation III

Example: let's make soup.



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

ABC: Approximate Bayesian Computation IV

Example: let's make soup.



Goal: Determine ingredients from final result.

Model physical processes? Complicated.

Easier: Make lots of soups with different ingredients, compare.

ABC: Approximate Bayesian Computation V

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**)

ABC: Approximate Bayesian Computation VI

Parameter constraints: ABC

- Summary statistic

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

- Metric D : two cases

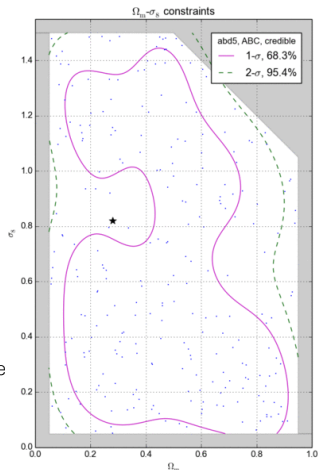
$$D_1(\mathbf{x}, \mathbf{x}^{\text{obs}}) \equiv \sqrt{\sum_i \frac{(x_i - x_i^{\text{obs}})^2}{C_{ii}}},$$

$$D_2(\mathbf{x}, \mathbf{x}^{\text{obs}}) \equiv \sqrt{(\mathbf{x} - \mathbf{x}^{\text{obs}})^T \mathbf{C}^{-1} (\mathbf{x} - \mathbf{x}^{\text{obs}})},$$

D_1 in Lin & MK 2015b

$D_1 + D_2$ in Lin, MK & Pires 2016

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



ABC: Approximate Bayesian Computation VII

ABC's accept-reject process is actually a sampling under P_ϵ (green curve):

$$P_\epsilon(\pi|x^{\text{obs}}) = A_\epsilon(\pi)P(\pi),$$

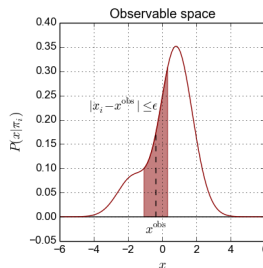
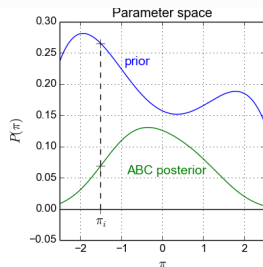
where $P(\pi)$ stands for the prior (blue curve) and

$$A_\epsilon(\pi) \equiv \int dx P(x|\pi) \mathbb{1}_{|x-x^{\text{obs}}| \leq \epsilon}(x),$$

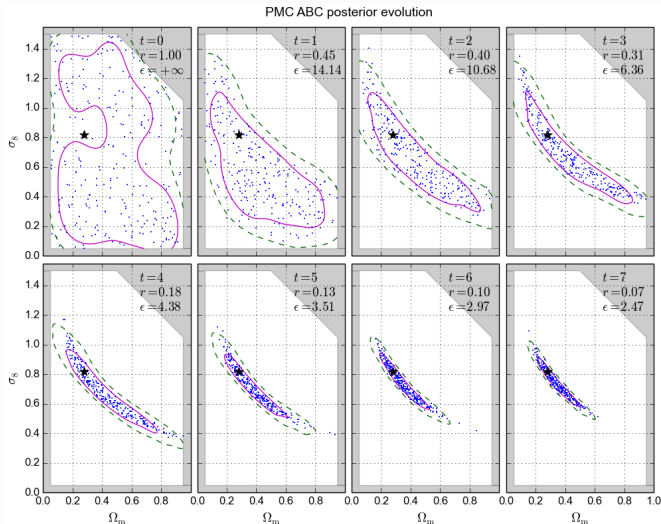
is the accept probability under π (red area). One can see that

$$\lim_{\epsilon \rightarrow 0} A_\epsilon(\pi_0)/\epsilon = P(x^{\text{obs}}|\pi_0) = \mathcal{L}(\pi_0),$$

so P_ϵ is proportional to the true posterior when $\epsilon \rightarrow 0$.

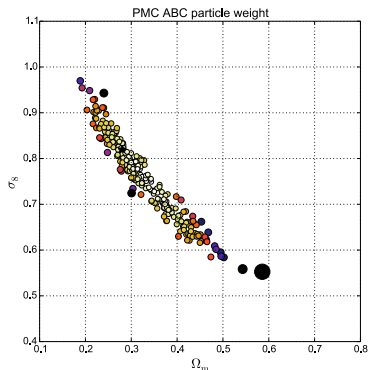
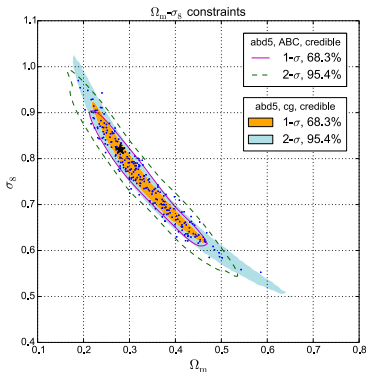


ABC: Approximate Bayesian Computation VIII




Lin & Kilbinger (2015b)

ABC: Approximate Bayesian Computation IX




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

Bibliography I


 Agari M, Schneider P & Simon P 2012 *A&A* **542**, A122.

 Baldauf T, Smith R E, Seljak U & Mandelbaum R 2010 *Phys. Rev. D* **81**(6), 063531.


 Baulieu J P, Bennett D P, Fouqué P, Williams A, Dominik M & al. 2006 *Nature* **439**, 437–440.

 Becker M R 2013 *MNRAS* **435**, 115–132.

 Bernardeau F, Bonvin C, Van de Rijt N & Vernizzi F 2012 *Phys. Rev. D* **86**(2), 023001.

 Blake C, Joudaki S, Heymans C, Choi A, Erben T & al. 2016 *MNRAS* **456**, 2806–2828.











 Bandford R & Narayan R 1986 *ApJ* **310**, 568–582.

 Castro P G, Heavens A F & Kitching T D 2005 *Phys. Rev. D* **72**(2), 023516.

 Chon G, Challinor A, Prunet S, Hivon E & Szapudi I 2004 *MNRAS* **350**, 914–926.

 Lowe D, Bradač M, Gonzalez A H, Markevitch M, Randall S W & al. 2006 *ApJ* **648**, L109–L113.

Bibliography II


-  Coupon J, Kilbinger M, McCracken H J, Ilbert O, Arnouts S & al. 2012 *A&A* **542**, A5.
-  Das S & Bode P 2008 *ApJ* **682**, 1–13.
-  Eifler T, Schneider P & Krause E 2010 *A&A* **510**, A7.
-  Fosalba P, Gaztañaga E, Castander F J & Manera M 2008 *MNRAS* **391**, 435–446.
-  Fu L & Kilbinger M 2010 *MNRAS* **401**, 1264–1274.
-  Fu L, Semboloni E, Hoekstra H, Kilbinger M, van Waerbeke L & al. 2008 *A&A* **479**, 9–25.
-  Heymans C, Van Waerbeke L, Bacon D, Berge J, Bernstein G & al. 2006 *MNRAS* **368**, 1323–1339.
-  Ilbert S, Hartlap J, White S D M & Schneider P 2009 *A&A* **499**, 31–43.
-  Hildebrandt H, Viola M, Heymans C, Joudaki S, Kuijken K & al. 2017 *MNRAS* **465**, 1454–1498.
-  Høg A, Bradac M, Trenti M, Treu T, Schmidt K B & al. 2017 *Nature Astronomy* **1**, 0091.

Bibliography III

 Heekstra H, Viola M & Herbonnet R 2017 *MNRAS* **468**, 3295–3311.

 Heekstra H, Yee H K C & Gladders M D 2001 *ApJ* **558**, L11–L14.

 Hu W 2000 *Phys. Rev. D* **62**(4), 043007.


 Kaiser N, Squires G, Fahlman G & Woods D 1994 *in* ‘Clusters of galaxies, Proceedings of the XIVth Moriond Astrophysics Meeting, Méribel, France’ p. 269.

 Kilbinger M, Fu L, Heymans C, Simpson F, Benjamin J & al. 2013 *MNRAS* **430**, 2200–2220.











 Kilbinger M, Heymans C, Asgari M, Joudaki S, Schneider P & al. 2017 *MNRAS* **472**, 2126–2141.

 Leonard A, Pires S & Starck J L 2012 *MNRAS* **423**, 3405–3412.




 Li B, King L J, Zhao G B & Zhao H 2011 *MNRAS* **415**, 881–892.

 Mandelbaum R, Rowe B, Armstrong R, Bard D, Bertin E & al. 2015 *MNRAS* **450**, 2963–3007.

Bibliography IV

-  Reyes R, Mandelbaum R, Seljak U, Baldauf T, Gunn J E & al. 2010 *Nature* **464**, 256–258.
-  Schneider P 1996 *MNRAS* **283**, 837.
-  Schneider P, Eifler T & Krause E 2010 *A&A* **520**, A116.
-  Schneider P & Kilbinger M 2007 *A&A* **462**, 841–849.
-  Schneider P, Van Waerbeke L, Jain B & Kruse G 1998 *MNRAS* **296**, 873–892.
-  Schneider P, Van Waerbeke L, Kilbinger M & Mellier Y 2002 *A&A* **396**, 1–19.
-  Symboloni E, Hoekstra H & Schaye J 2013 *MNRAS* **434**, 148–162.
-  Simon P, Hetterscheidt M, Schirmer M, Erben T, Schneider P & al. 2007 *A&A* **461**, 861–879.
-  Starck J L, Pires S & Réfrégier A 2006 *A&A* **451**, 1139–1150.
-  Takada M & Hu W 2013 *Phys. Rev. D* **87**(12), 123504.
-  Taylor A, Joachimi B & Kitching T 2013 *MNRAS* **432**, 1928–1946.
-  Teysier R, Pires S, Prunet S, Aubert D, Pichon C & al. 2009 *A&A* **497**, 335–341.

Bibliography V

-  vander M, van Uitert E, Hoekstra H, Coupon J, Erben T & al. 2014 *MNRAS* **437**, 2111–2136.
-  White M & Hu W 2000 *ApJ* **537**, 1–11.
-  Zentner A R, Rudd D H & Hu W 2008 *Phys. Rev. D* **77**(4), 043507–+.