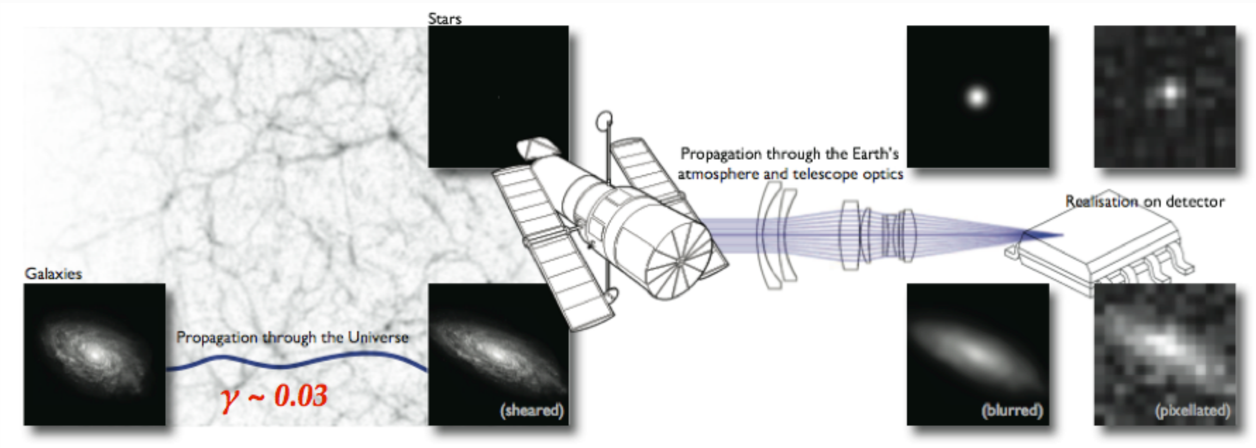
Part I day 2. Reminder: Overview

The shape measurement challenge

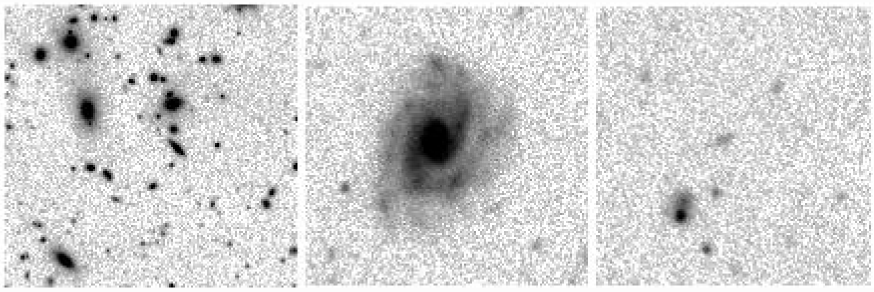


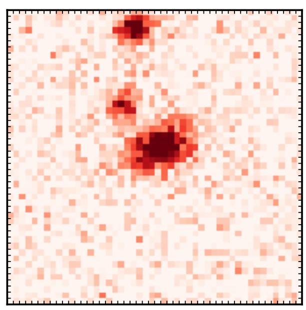
Bridle et al. 2008, great08 handbook

- Cosmological shear $\gamma \ll \varepsilon$ intrinsic ellipticity
- Galaxy images corrupted by PSF
- Measured shapes are biased

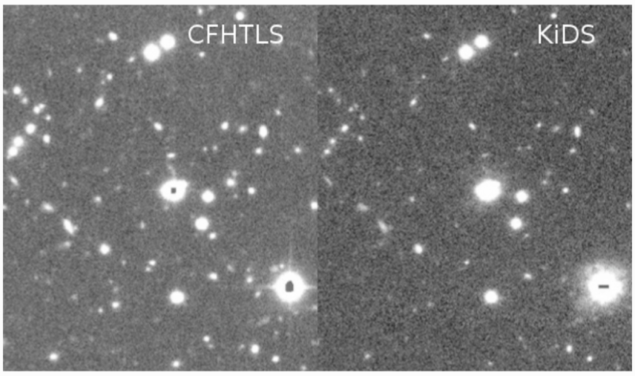
The shape measurement challenge

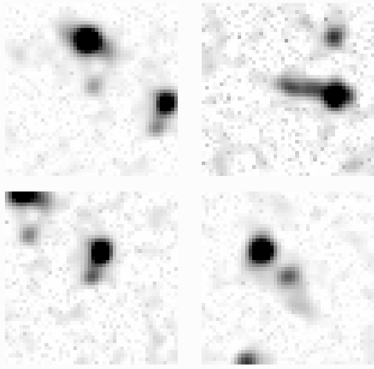
How do we measure ”ellipticity” for irregular, faint, noisy objects?





[Y. Mellier/CFHT(?)] — (Jarvis et al. 2016)



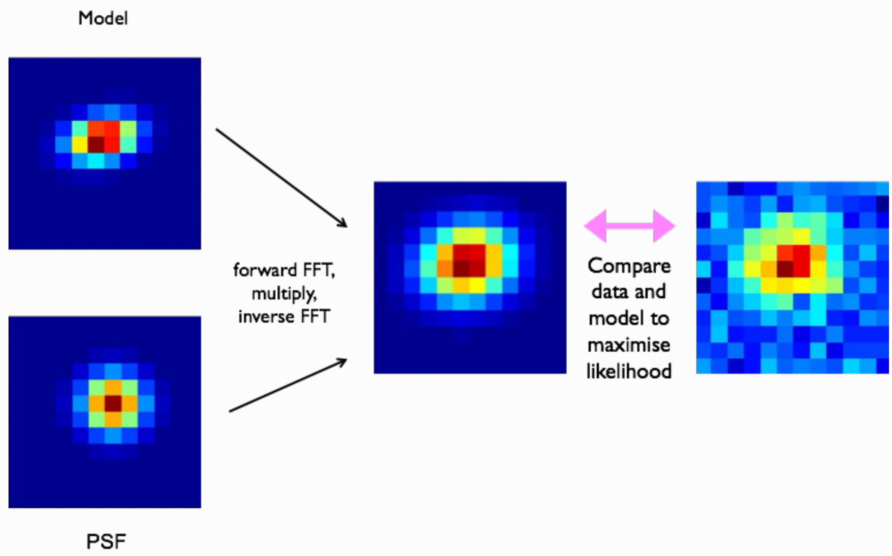


[CFHTLenS/KiDS image — CFHTlenS postage stamps]

Shape measurement methods

- Parametric: model fitting.
  (Kuijken 1999), *lens*fit (Miller et al. 2007)), *gfit* (Gentile et al. 2012), *im3shape* (Zuntz et al. 2013) and many more.
- Non-parametric: direct estimation.
  - Perturbative: weighted moments.
    KSB — (Kaiser et al. 1995) + many improvements
    DEIMOS — (Melchior et al. 2011) (PSF correction in moment space)
    HOLICs — (Okura & Futamase 2009) — Higher-order moments
  - Non-perturbative: Decomposition into basis functions.
    shapelets — (Refregier 2003) + many improvements

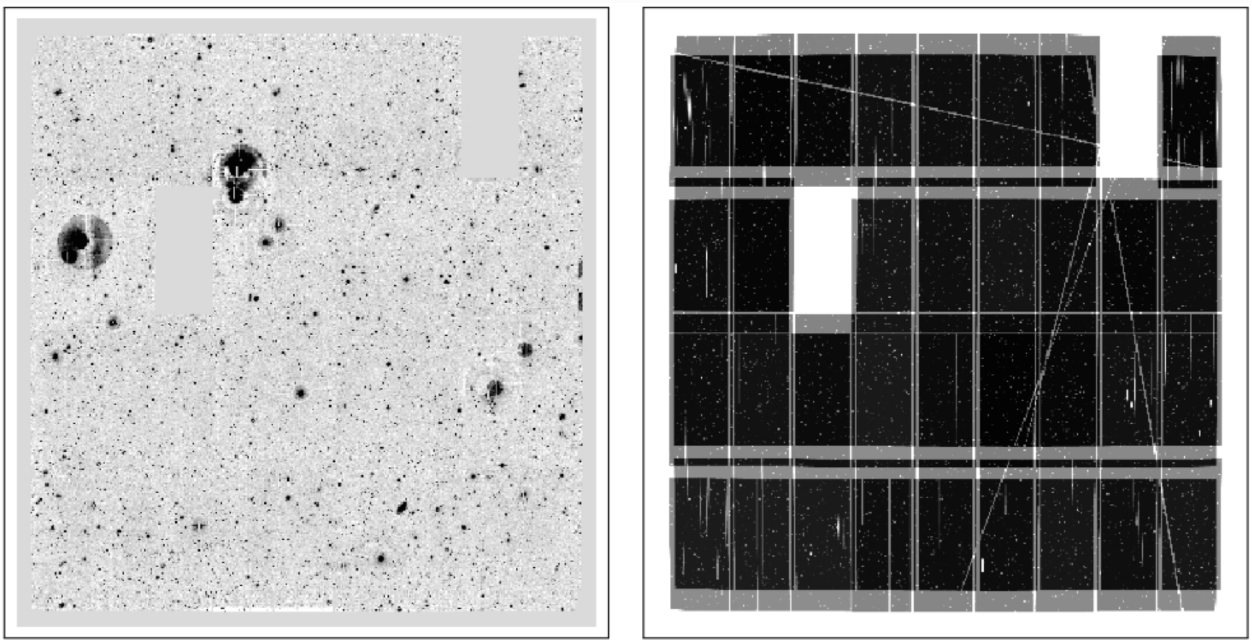
Model fitting methods



Forward model-fitting (example *lens*fit)

- Convolution of model with PSF instead of devonvolution of image
- Combine multiple exposures (in Bayesian way, multiply posterior density), avoiding co-adding of (dithered) images

Dithering



Left: Co-add of two r-band exposures of CFHTLenS.
Right: Weight map.

Moment-based methods I

Moments and ellipticity

How are moments connected to ellipticity?
Q: Simple case: qualitatively, what are the $0^{\text{th}}, 1^{\text{st}}, 2^{\text{nd}}$ moments of a 1D distribution? Of a 2D distribution?
Quadrupole moment of weighted light distribution $I(\boldsymbol{\theta})$:

$$Q_{ij} = \frac{\int \mathrm{d}^2\theta \, q[I(\boldsymbol{\theta})] \, (\theta_i - \bar{\theta}_i)(\theta_j - \bar{\theta}_j)}{\int \mathrm{d}^2 \, \theta \, q[I(\boldsymbol{\theta})]}, \quad i,j = 1,2$$

q : weight function

$$\bar{\boldsymbol{\theta}} = \frac{\int \mathrm{d}^2\theta \, q_I[I(\boldsymbol{\theta})] \, \boldsymbol{\theta}}{\int \mathrm{d}^2\theta \, q_I[I(\boldsymbol{\theta})]} : \quad \text{barycenter (first moment!)}$$

Ellipticity

$$\varepsilon = \frac{Q_{11} - Q_{22} + 2\mathrm{i}Q_{12}}{Q_{11} + Q_{22} + 2(Q_{11}Q_{22} - Q_{12}^2)^{1/2}}$$

Circular object $Q_{11} = Q_{22}, Q_{12} = Q_{21} = 0$

Moment-based methods II

KSB PSF correction

Perturbative ansatz for PSF effects

$$\varepsilon^{\text{obs}} = \varepsilon^{\text{s}} + P^{\text{sm}}\varepsilon^* + P^{\text{sh}}\gamma$$

[c.f. $\varepsilon^{\text{obs}} = \varepsilon^{\text{s}} + \gamma$ from before]

P^{sm}	smear polarisability, (linear) response of to ellipticity to PSF anisotropy
e^*	PSF anisotropy
P^{sh}	shear polarisability, isotropic seeing correction
γ	shear

$P^{\text{sm}}, P^{\text{sh}}$ are functions (2×2 tensors) of galaxy brightness distribution.

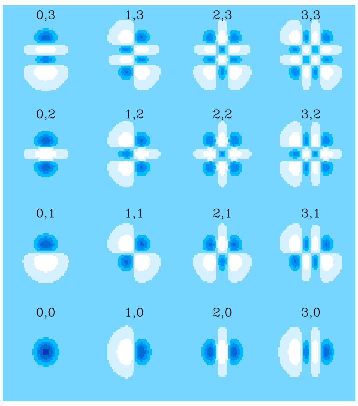
Problematic: Strongly anisotropic PSF, error estimation, combining multiple exposures.

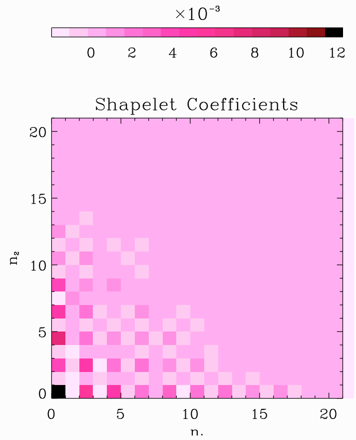
Non-perturbative methods

Shapelets

(Refregier 2003, Massey & Refregier 2005, Kuijken 2006)

- Decompose galaxies and stars into basis functions.





- PSF correction, convergence and shear acts on shapelet coefficients, deconvolution feasible
- Problems: series truncation, basis functions not representative, need to set size parameter

Further methods and techniques

- Machine-Learning, e.g. LUT by supervised learning, (Tewes et al. 2012)
- Self-calibration
- Further Bayesian methods
  - Hierarchical Multi-level Bayesian Inference (MBI), (Schneider et al. 2014). Joint posterior of shear, galaxy properties, PSF, nuisance parameters given pixel data.
  - (Bernstein & Armstrong 2014). Does not measure ellipticity of individual galaxies, direct posterior estimation of shear for population. Needs prior from deep images.

Shear measurement biases I

Origins

- **Noise bias**
  In general, ellipticity is non-linear in pixel data (e.g. normalization by flux). Pixel noise → biased estimators.
- **Model bias**
  Assumption about galaxy light distribution is in general wrong.
  - Model-fitting method: wrong model
  - Perturbative methods (*KSB, DEIMOS, HOLICS*): weight function not appropriate
  - Non-perturbative methods (*shapelets*): truncated expansion, bad eigenfunction representation
  - Color gradients
  - Non-elliptical isophotes
- **Other**
  - Imperfect PSF correction
  - Detector effects (CTI — charge transfer inefficiency)

Shear measurement biases II

  - Selection effects (probab. of detection/sucessful ε measurement depends on ε and PSF)

Characterisation

Bias can be multiplicative ($\boldsymbol{m}$) and additive ($\boldsymbol{c}$):

$$\gamma_i^{\rm obs} = (1 + m_i)\gamma_i^{\rm true} + c_i; \quad i = 1, 2.$$

Biases $\boldsymbol{m}$, $\boldsymbol{c}$ are typically complicated functions of galaxy properties (e.g. size, magnitude, ellipticity), redshift, PSF, ... They can be scale-dependent.

Current methods: $|m| = 1\% - 10\%, |c| = 10^{-3} - 10^{-2}$.

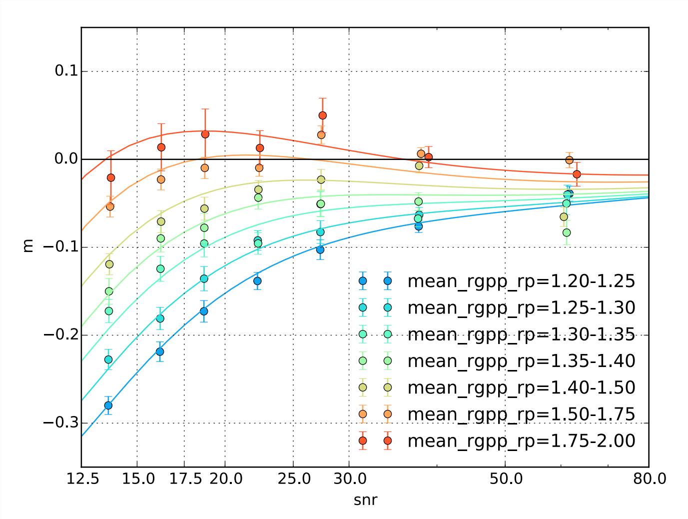
Challenges such as STEP1, STEP2, great08, great10, great3 quantified these biases with blind simulationes.

Calibration

Usually biases are calibrated using simulated or emulated data, or self-calibration.
Current surveys produce their own image simulations with properties of galaxy sample and PSF matching to data.

Shear measurement biases III

Functional dependence of m on observables must not be too complicated (e.g. not smooth, many variables, large parameter space), or else measurement is *not calibratable*!

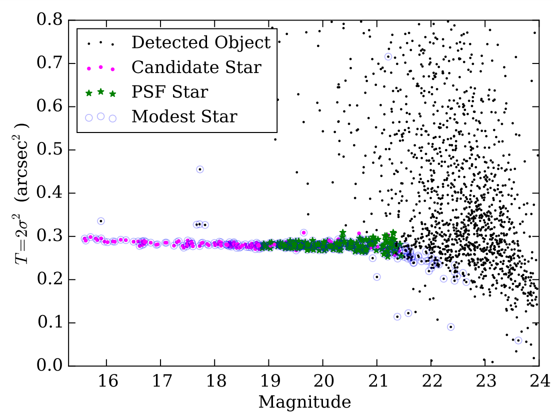


(Jarvis et al. 2016)

Requirements

Normalisation $\sigma_8 \propto m$!
Necessary knowledge of residual biases $\Delta|m|, \Delta|c|$ (after calibration):
Current surveys 1%.
Future large missions (Euclid, LSST, ...) $10^{-4} = 0.1\%$!

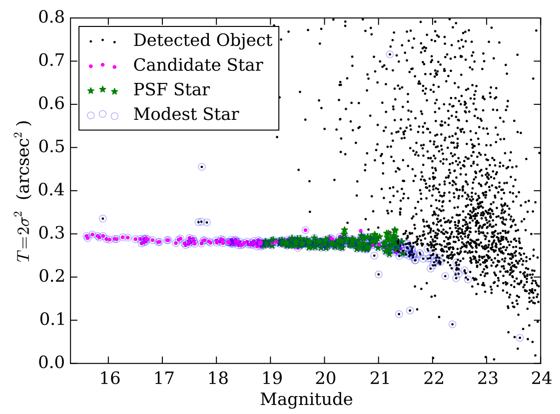
PSF correction



(Jarvis et al. 2016)

- Select clean sample of stars
- Measure star shapes
- Create PSF model and interpolate (pixel values, ellipticity, PCA coefficients, ...) to galaxy positions. Space-based observations: global PSF model from many exposures possible
- Correct for PSF: galaxy image devonvolution or other (e.g. linearized) correction, or convolve model

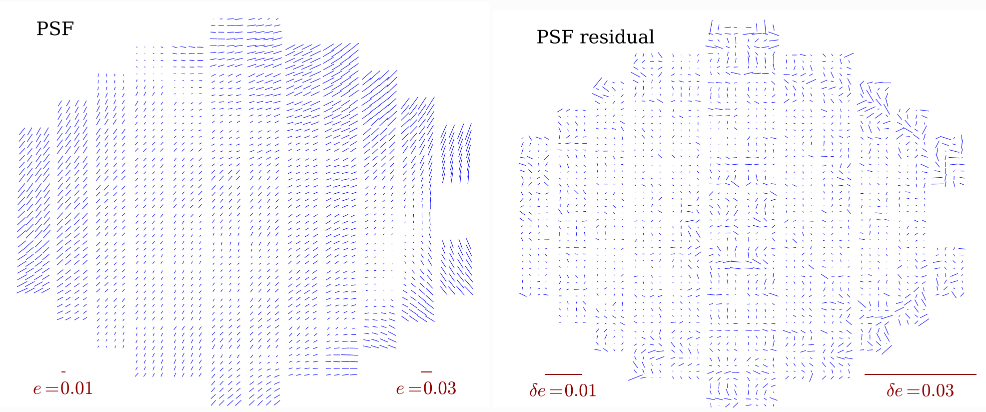
PSF correction



(Jarvis et al. 2016)

- Select clean sample of stars
- Measure star shapes
- Create PSF model and interpolate (pixel values, ellipticity, PCA coefficients, . . .) to galaxy positions. Space-based observations: global PSF model from many exposures possible
- Correct for PSF: galaxy image devonvolution or other (e.g. linearized) correction, or convolve model

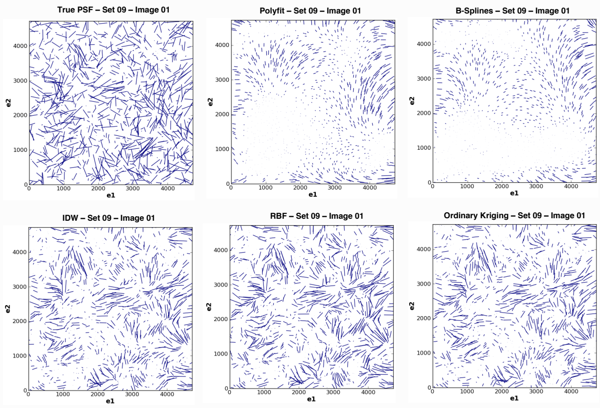
PSF correction



(Jarvis et al. 2016)

- Select clean sample of stars
- Measure star shapes
- Create PSF model and interpolate (pixel values, ellipticity, PCA coefficients, . . .) to galaxy positions. Space-based observations: global PSF model from many exposures possible
- Correct for PSF: galaxy image devonvolution or other (e.g. linearized) correction, or convolve model

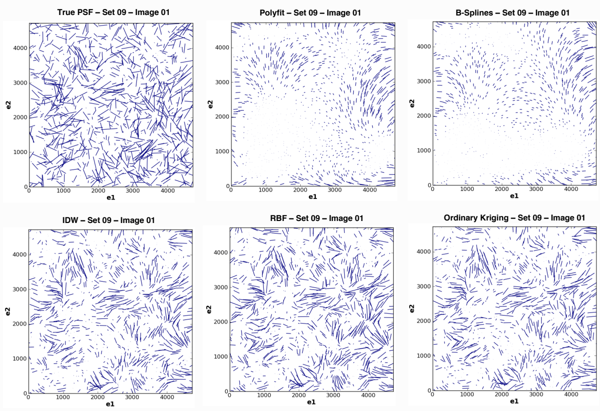
PSF correction



(Gentile et al. 2013)

- Select clean sample of stars
- Measure star shapes
- Create PSF model and interpolate (pixel values, ellipticity, PCA coefficients, ...) to galaxy positions. Space-based observations: global PSF model from many exposures possible
- Correct for PSF: galaxy image devonvolution or other (e.g. linearized) correction, or convolve model

PSF correction



(Gentile et al. 2013)

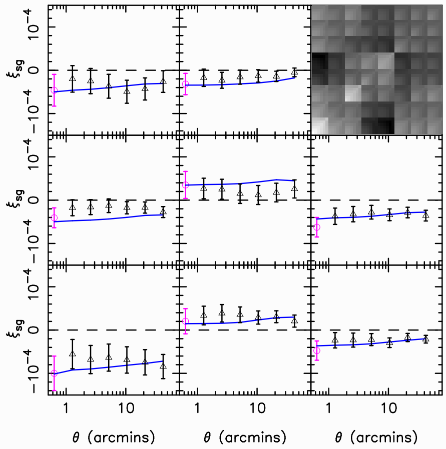
- Select clean sample of stars
- Measure star shapes
- Create PSF model and interpolate (pixel values, ellipticity, PCA coefficients, ...) to galaxy positions. Space-based observations: global PSF model from many exposures possible
- Correct for PSF: galaxy image devonvolution or other (e.g. linearized) correction, or convolve model

Quantifying PSF residuals I

Null test: $\xi_{\rm sys}$ correlation between star and galaxy shapes expected to vanish, unless PSF correction (using stars to correct galaxy shapes) is not perfect.

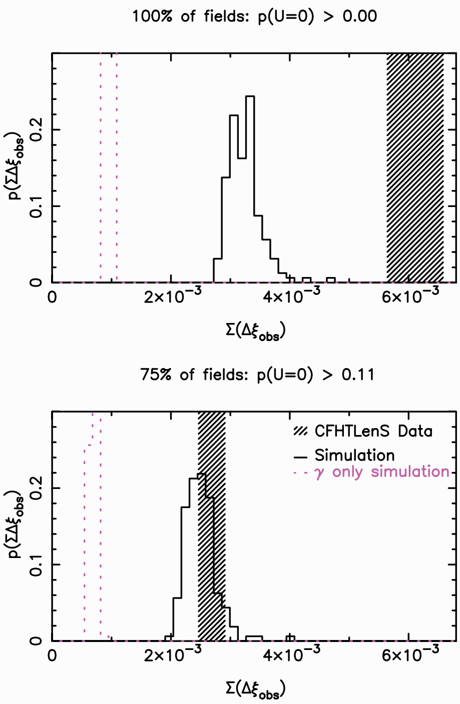
$$\xi_{\rm sys} = \langle \varepsilon^* \varepsilon \rangle$$

This measures residual PSF pattern leakage onto galaxy field.
Caveat: LSS can show chance alignments with PSF pattern. Sample or *cosmic variance* has to be accunted for → N-body simulations!



(Heymans et al. 2012)

Quantifying PSF residuals II

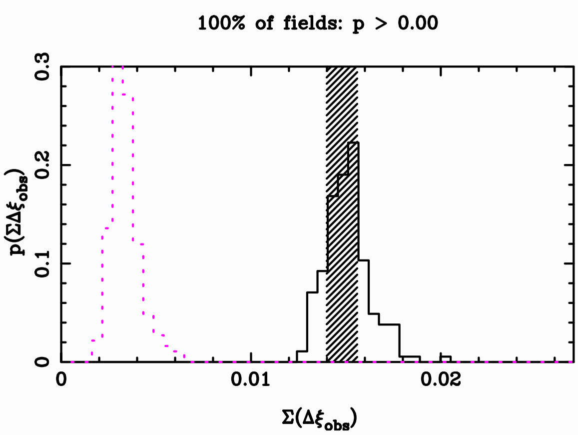


[Heymans et al. 2012, CFHTLenS]

Histogram of probability p that $\Sigma\xi_{\rm obs} \sim \Sigma|\xi_{\rm sys}|$ is not zero (sum over all pointings), from simulations.

Shaded region = data.

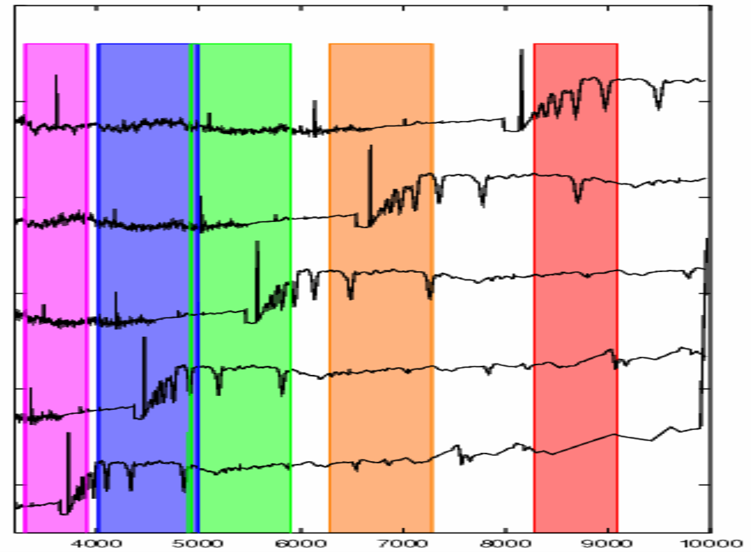
Magenta: simulations without LSS.



[Hildebrandt et al. 2016, KiDS-450]

Redshift estimation I

Redshifted galaxy spectra have different colors.
Photometric redshifts = very low resolution spectra.
#bands between 3 (RCS) and 30 (COSMOS). Typical are 4-5 optical filters (g, r, i, y, z), maybe with UV (u) and IR (I, J, K).



[from Y. Mellier]

4000 Å-break strongest feature → ellipticals (old stellar population) best, spirals ok, irregular/star-burst (emission lines) more unreliable.

Redshift estimation II

Properties

- Redshift desert $z \approx 1.5 - 2.5$, neither 4000 Å-break nor Ly-break in visible range, very hard to access from ground.
- Confusion between low-z dwarf ellipticals and high-z galaxies. Confusion between Balmer and Lyman break. Catastrophic outliers, typically a few to a few 10
- Need UV band and IR for high redshifts! But: UV very inefficient, IR absorbed by atmosphere, have go to space.
- Need spectroscopic galaxy sample for comparison, calibration, or cross-correlation. In general $N_{\rm spec} \ll N_{\rm WL}$.
- Typical accuracy of photo-z's $\sigma/(1+z) \sim 0.05$ (depending on filters).

Redshift estimation III

Redshift accuracy and cosmology

To interpret weak lensing correlations in cosmological context, the redshift distribution needs to be known accurately!
To first order:

$$P_\kappa(\ell \sim 1000) \propto \Omega_{\rm de}^{-3.5} \sigma_8^{2.9} \bar{z}^{1.6} |w|^{0.31} \quad \text{(Huterer et al. 2006)}$$

Methods

- Template fitting.
  Redshifted synthetic or observed templates of various types are fitted to flux in observed bands.
  Examples *LePhare* (Ilbert et al. 2006)), *BPZ* (Benítez 2000), *HyperZ* (Bolzonella et al. 2000).
  Spectroscopic sample for calibration, priors.
- Machine-learning.
  Learn data using training set (of spectroscopic sample).
  Examples: *ANNz* (Collister & Lahav 2004).

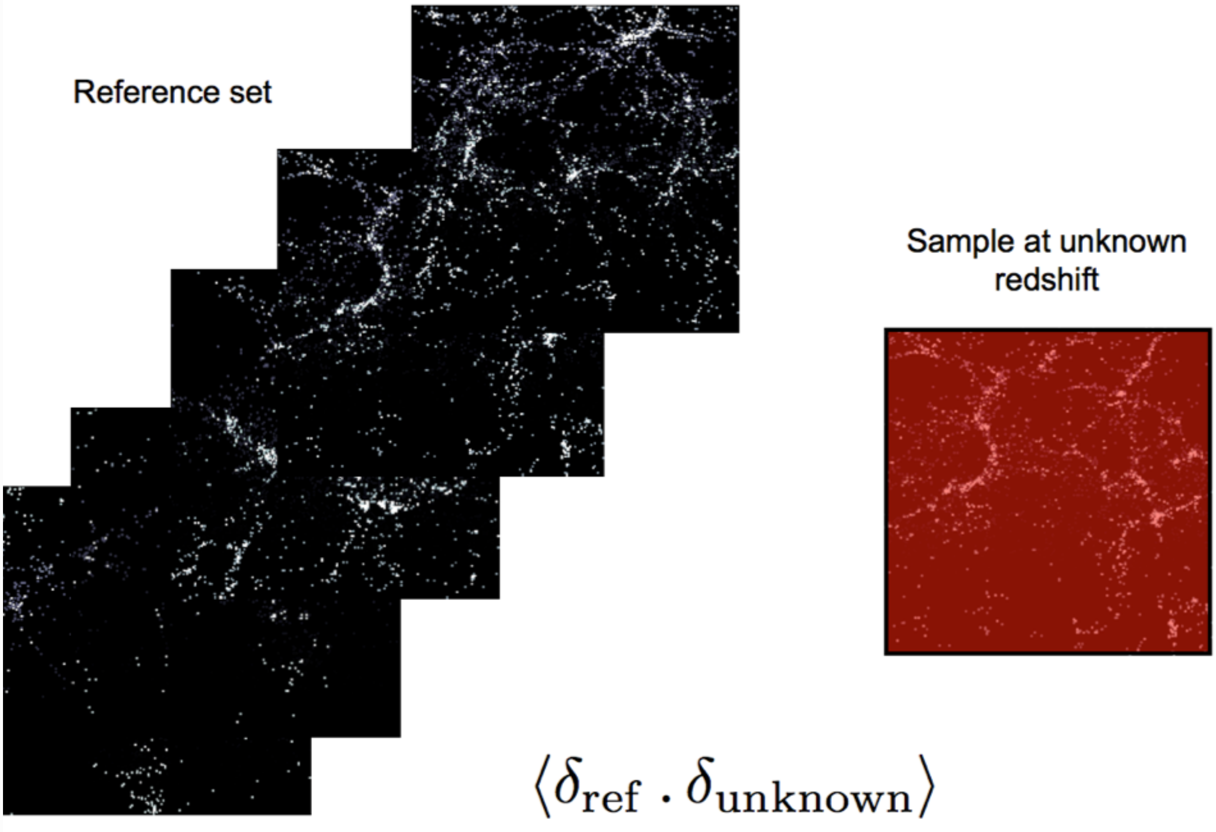
Redshift estimation IV

- Matching photometric properties to spectroscopic sample (?) (direct calibration).
- Spatial cross-correlation with spectroscopic survey (clustering redshifts)

Spectroscopic sample has to be representative in some properties, depending on the method:

- Template fitting: Same magnitude limit as photometric sample
- Neural networks: Cover redshift range, properties (colors)
- Matching: Cover (color) parameter space
- Clustering: Cover redshift range, sky overlap

Clustering redshifts (slide from Vivien Scottez)



Estimator of second-order functions I

Remember the shear two-point correlation function (2PCF)?

$$\xi_\pm(\vartheta) = \langle \gamma_{\rm t} \gamma_{\rm t} \rangle (\vartheta) \pm \langle \gamma_\times \gamma_\times \rangle (\vartheta)$$

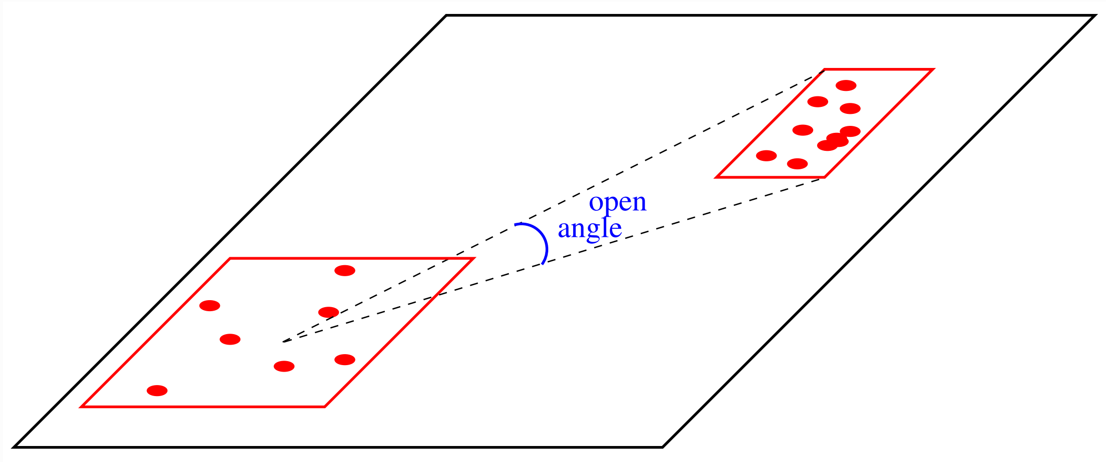
Unbiased estimator of $\xi_\pm$ just involves sums over galaxy pairs:

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

Sum over galaxy pairs with angular distance within bin of θ.

- Unbiased estimator (for bin size $\to 0$, and in absence of intrinsic alignment)
- No need for random catalogue, or mask geometry, since $\xi = 0$ in absence of lensing.
- No need to pixellise data, can use brute-force or tree codes/linked lists (adaptive pixellisation, effective smoothing)

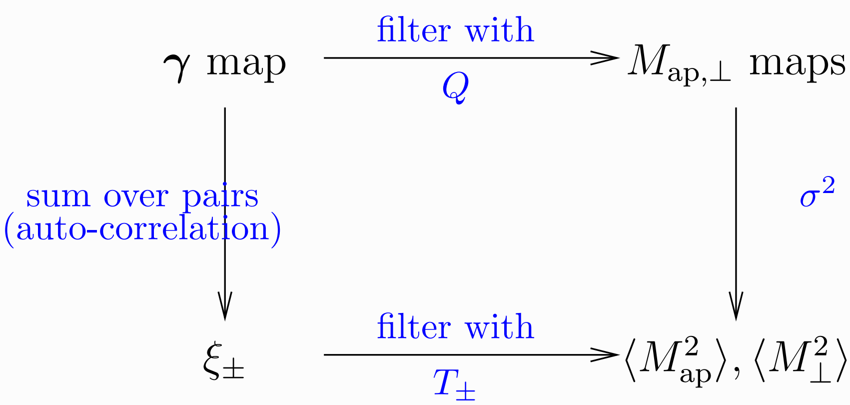
Estimator of second-order functions II



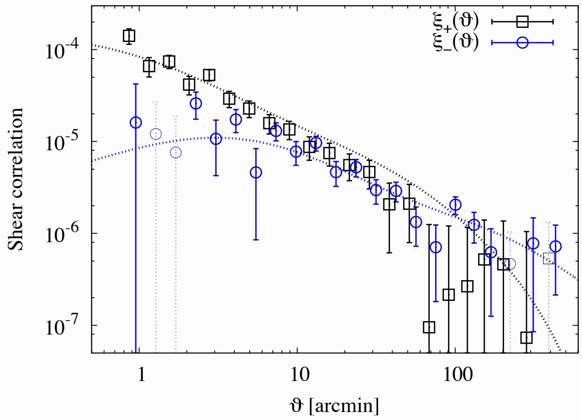
Tree code: correlating two 'nodes' (2D regions).

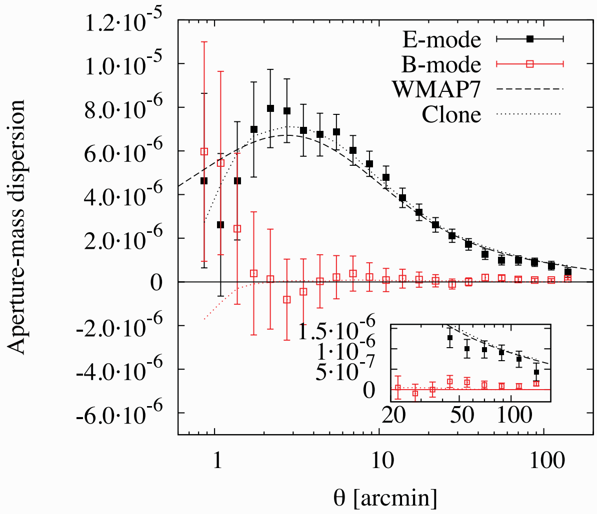
Estimator of second-order functions III

From the 2PCF estimator, the aperture-mass dispersion and other second-order functions can be derived:



Estimator of second-order functions IV





(Kilbinger et al. 2013)

End of day 2.