

# Weak-lensing statistics on a CFHTLenS field I

This exercise will show you how to estimate second-order shear statistics (2PCF, aperture-mass dispersion, band-power spectrum) and their errors, and how to compare these estimates with theoretical predictions.

## 1. Download shear catalogue

**Note:** The catalogue size is 158 Mb, so this might take a while. So do this step well before the start of the TD, or use the downloaded catalogue on the common disk.

Go to <http://cfhtlens.org> → Fellow astronomers → **Quick link:** Access the CFHTLenS Shear and Photometric Redshift catalogues.

This brings you to the catalogue query page on CADC

<http://www.cadc-ccda.hia-ihp.nrc-cnrc.gc.ca/en/community/CFHTLenS/query.html>.

We will download the shear data from the W1 field (but feel free to use another field — check the coordinates in (Erben et al. 2013)). The following steps are advised (for some of these you have to edit the string in the query field):

## Weak-lensing statistics on a CFHTLenS field II

- Un-select id
- Select ALPHA\_J2000, DELTA\_J2000, e1, e2, weight. These are the  $x$ - and  $y$ -coordinates, the two ellipticity components, and the galaxy weight.
- Choose the ranges  $\text{ALPHA\_J2000} \geq 25$ ,  $\text{ALPHA\_J2000} \leq 45$ ,  $\text{DELTA\_J2000} \geq -20$ , and  $\text{DELTA\_J2000} \leq 0$ . This selects coordinates in the W1 field (you can double-check in (Erben et al. 2013)).
- Choose  $\text{weight} > 0$ , the code `athena` that computes the 2PCF does not like objects with zero weight.
- Choose the range  $\geq 0.0$  and  $\leq 0.0$ , but *do not* select `fitclass`. This flag is zero for galaxies, one for stars, and negative for other detections. We only want galaxies, but do not need this flag in our catalogue.
- If you like you can do a test by clicking on "submit query" to see the first 10 objects. If you are happy with the result, choose "Asynchronous" as submission method, "Tab Separated Values", and delete "top 10" from the query field (we don't only want 10 objects)

## Weak-lensing statistics on a CFHTLenS field III

The text in the query field should now look something like the following:

```
SELECT
ALPHA_J2000, DELTA_J2000, e1, -1*e2, weight
FROM
cfht.clens
WHERE
ALPHA_J2000>=25
AND ALPHA_J2000<=45
AND DELTA_J2000>=-20
AND DELTA_J2000<=0
AND fitclass>=0.0
AND fitclass<=0.0
AND weight>0
```

I recommend to flip the  $\varepsilon_2$ -coordinate, by placing a minus sign in front of e2 in the second line. The original coordinates have North and East defined such that  $(x, y)$  have a left-handed orientation. The  $\varepsilon_2$  flip accounts for that (why?).

## Weak-lensing statistics on a CFHTLenS field IV

- Submit query and wait. The processing of the query can take a few minutes up to **several hours!** After it is done the web page will show a link, from where you can download the catalogue.

Once the catalogue is downloaded, check whether it contains five columns that make sense. You can for example make a scatter plot of  $\alpha$  and  $\delta$  to see whether the selected galaxy coordinates are as desired.

**Before proceeding, remove the first (header) line.**

2. Use `athena` to get the 2PCF  $\xi_+$  and  $\xi_-$ .

If `athena` is not installed, download version 1.7 from `www.cosmostat/athena.html`, and compile.

First, create a config file. The easiest is to copy the example file from `/path/to/athena/test/test_xi/config_tree` and modify it to set the following entries:

- `GALCAT1` to the name of the catalogue you downloaded, `GALCAT2` either the same or “-”.
- `COORD_INPUT` to “deg”

## Weak-lensing statistics on a CFHTLenS field V

- THMIN, THMAX to whatever you like; Note that the shear correlation function is very noisy on scales smaller than 0.1 arcmin due to a very small number of galaxy pairs at such small distances; on scales larger than a few degrees it is more or less consistent with zero for the survey area in consideration.
- RADEC to 1
- OATH: the smaller, the more precise but also the slower the calculation. For testing you can put it to 0.2; for serious calculations it should be 0.05 or smaller.

We will perform two runs: (A) to compute and plot the 2PCF in a few coarse bins; (B) to compute the 2PCF in many narrow bins that then will be integrated to get aperture-mass and band-power spectrum.

The settings in the config file for the two cases:

	A	B
NTH	~ 20	~ 1000 or more
BINTYPE	LOG	LIN (recommended) or LOG

## Weak-lensing statistics on a CFHTLenS field VI

Run `athena` with the correct config file for case A or B. You can make sure that the output from another case is not overwritten by running them in different subdirectories, or by using specific suffixes with option (`--out_suf`).

```
> /path/to/athena/bin/athena -c config_tree_A --out_suf _A
```

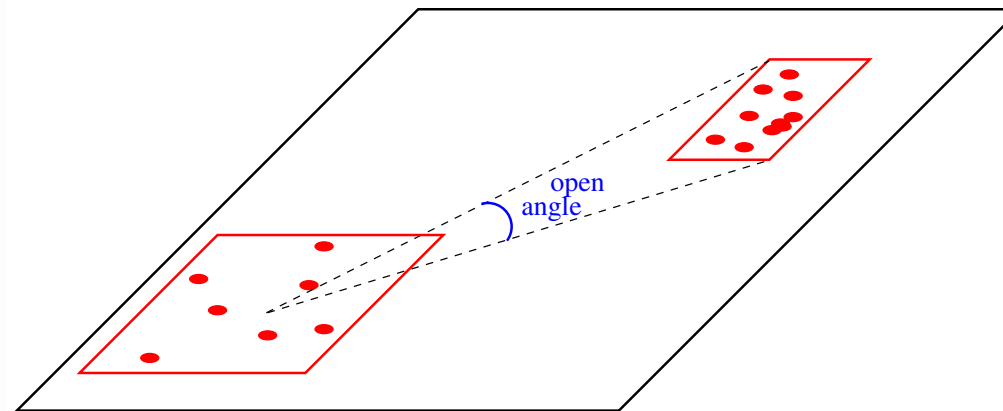
The run will take around 10-20 minutes.

`athena` implements the pairwise galaxy sum estimator of the 2PCF, see Part I:

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

with a tree code algorithm.

# Weak-lensing statistics on a CFHTLenS field VII



[The official Euclid OU-LE3 PF WL-2PCF software is the C++ version of **athena** (written by Bertrand Morin, Florent Sureau).]

### 3. From run (A) plot $\xi_+, \xi_-$ .

The file `xi<suffix>` contains as columns:

angular bin center  $\vartheta$ ,  $\xi_+$ ,  $\xi_-$ ,  $\xi_\times$ , weight  $w$ , raw Poisson error  $\sqrt{D}$ , corrected Poisson error  $\sqrt{D_{\text{cor}}}$ .

Plot  $\xi_+, \xi_-$  with error bars  $\sqrt{D_{\text{cor}}}$  versus  $\theta$ .

The file `xi.resample<suffix>` contain mean and rms of the resampled  $\xi_+$  and  $\xi_-$ , in our case (default config file) from the Jackknife method. Plot the  $\xi_+$  and  $\xi_-$  with resampled error bars, and compare mean and error bars.

## Weak-lensing statistics on a CFHTLenS field VIII

4. Use `pallas.py` to get derived second-order statistics: aperture-mass dispersion and band-power spectrum

```
> /path/to/athena/bin/pallas.py -i xi<suffix>
```

The resulting important output files are:

- `output_map2_poly.txt`, columns: smoothing scale/circle radius  $\theta$ ,  $\langle M_{\text{ap}}^2 \rangle(\theta)$ ,  $\langle M_{\times}^2 \rangle(\theta)$ ,  $\langle M_{\text{ap}} M_{\times} \rangle(\theta)$ .
- `output_pkappa_band.txt`, columns: 2D Fourier mode bin center  $\ell$ ,  $P_{\kappa}^{\text{E}}(\ell)$ ,  $P_{\kappa}^{\text{B}}(\ell)$ ,  $P_{\kappa}^{\text{EB}}(\ell)$ , lower bin limit  $\ell_{\text{lo}}$ , upper bin limit  $\ell_{\text{hi}}$ .

Plot E-, B-, and mixed EB-modes of both quantities in separate plots.

Note: If you find B-mode amplitudes comparable to the E-mode, you might have done the  $\varepsilon_2$ -flip incorrectly.



# Weak-lensing statistics on a CFHTLenS field IX

5. Use the Limber code from cycle I, or `nicaea`, or `CLASS` to create theoretical prediction of the power spectrum.

Use  $\bar{z} = 0.75$  (Kilbinger et al. 2013) for the mean redshift. Add the resulting convergence power spectrum  $P_\kappa$  to the previous plot, by plotting on the  $y$ -axis  $\ell(\ell + 1)/(2\pi)P_\kappa(\ell)$ .

Note that the shear catalogue is not calibrated. The calibration for the multiplicative shear bias  $m$  is around 6% on average, that makes around 12% in amplitude for the 2PCF.

## Additional bonus exercises

1. Download the catalogue again with additional fields. Note that if you want to re-run `athena`, you have to create a copy of the catalogue without those additional fields, since for ascii catalogues only 5 input columns are accepted.

Extra catalogue on USB sticks.

# Weak-lensing statistics on a CFHTLenS field X

- **Redshift distribution.** Select `Z_B`, photometric redshift. From this, create a histogram that you can use as redshift distribution  $n(z)$  for the theoretical prediction, instead of placing all galaxies at one single redshift.  
Extra-bonus if you make a weighted histogram using the weights  $w$ .  
Special extra bonus if you download the full pdf information, `PZ_full`, and create the  $n(z)$  from the sum of weighted pdf's.
- **Shear calibration.** Select `SNratio`, signal-to-noise ratio, and `scalelength` for galaxy size.
  - **Additive shear bias  $c$ :** The correction for  $c$  can be done for each galaxy. Use eq. (19) of (Heymans et al. 2012) for  $c_2$ ; note that `scalelength` is in pixels, with one pixel being 0.187 arc seconds. On average,  $c_2$  should be of order 0.002.  
The 1-component of the additive bias,  $c_1$ , was measured to be consistent with zero, and no calibration is required.  
Subtract  $c_2$  from  $\varepsilon_2$  for each galaxy. This should be done before the  $\varepsilon_2$  flip.  
Note that a constant additive bias shows up in the 2PCF, but not the aperture-mass dispersion. (Why?)

# Weak-lensing statistics on a CFHTLenS field XI

- **Multiplicative shear bias  $m$ :** Correction for  $m$  should not be done on individual galaxies. This might introduce correlations between their weights  $w$  and  $m$ , and could up-weight badly measured galaxies if they have a large  $|m|$ . Instead, we need to compute a total calibration correction for the entire galaxy sample. For the 2PCF, this is the expression (16) in (Miller et al. 2013), or (14) in (Kilbinger et al. 2013),

The 2PCF is then globally calibrated by

$$\xi_{\pm}^{\text{cal}}(\vartheta) = \frac{\xi_{\pm}(\vartheta)}{1 + K(\vartheta)}.$$

We can use `athena` to compute the two-point correlation function of  $m$ ,  $1 + K(\vartheta)$ . Since  $m$  is a scalar and not a spin-2 quantity like ellipticity, we can do the following trick: In the original shear catalogue, we replace  $\varepsilon_1$  with  $1 + m$ , and  $\varepsilon_2$  with 0. The output  $\xi_+ = \langle \varepsilon_1 \varepsilon_1 \rangle + \langle \varepsilon_2 \varepsilon_2 \rangle$  of `athena` then results in  $\langle (1 + m)(1 + m) \rangle$ , which corresponds to  $1 + K$ . ( $\xi_-$  will not be a meaningful output.)

Use (17) of (Heymans et al. 2012) to compute  $m$  for each galaxy. In this equation,  $\log = \log_{10}$ , and  $\alpha$  is in inverse pixel. Do the replacement in the catalogue as described above. The modified catalogue should now contain

# Weak-lensing statistics on a CFHTLenS field XII

the 5 columns ALPHA\_J2000, DELTA\_J2000,  $1 + m$ , 0, w. and run `athena` with the modified catalogue.

Plot the calibrated 2PCF  $\xi^{\text{cal}}$  and compare to the previous result.

- Code up the Hankel transform to obtain  $\xi_+$  and  $\xi_-$  from the theoretical model,

$$\xi_+(\vartheta) = \frac{1}{2\pi} \int_0^\infty dl \, l J_0(l\vartheta) P_\kappa(l)$$

$$\xi_-(\vartheta) = \frac{1}{2\pi} \int_0^\infty dl \, l J_4(l\vartheta) P_\kappa(l),$$

Plot together with the data.

- Plot theoretical power spectrum for different values of  $\sigma_8$ .

# Weak-lensing statistics on a CFHTLenS field XIII

## 4. Error bars for $\langle M_{\text{ap}} \rangle$ and $P_{\kappa}(\ell)$ .

Re-run `athena` with the options `--out_ALL_xip_resample XIP_NAME`  
`--out_ALL_xim_resample XIM_name`. This outputs all resampled realisations of  $\xi_+$  and  $\xi_-$  into the files `XIP_NAME` and `XIM_name`, respectively.

There are two options to proceed:

4.1 Bring these files into the format of `athena` output file `xi`. Use dummy values for columns `xi_x`, `w`, `sqrt_D`, `sqrt_Dcor`, `n_pair` (for example copy the ones from `xi`). Create a different new `xi` file for each of the `NRESAMPLE` resample realisation.

Run `pallas.py` with each of the resample `xi` files. This should provide `NRESAMPLE` output files; make sure they have unique names or are stored in different sub-directories.

The errors bars on  $\langle M_{\text{ap}} \rangle$  and  $P_{\kappa}(\ell)$  are then simply the rms between the different realizations (the errors on  $\xi_{\pm}$  have been properly propagated to the derive quantities).

# Weak-lensing statistics on a CFHTLenS field XIV

4.2 Reading resampled input and computing resample errors bars for  $\langle M_{\text{ap}} \rangle$  and  $P_{\kappa}(\ell)$  is already implemented for FITS format, in a new version of `pallas.py`. See function `read_xi_resample`.

Download this new version and implement resampling for ASCII format.

Compare to option 1.

5. Extend the computation of the jackknife variance (see previous point) to the co-variance.

Code up a simple Gaussian likelihood function with the inverse of this covariance.

Compute the likelihood for various values of  $\sigma_8$ , and make a plot.

**Special extra super bonus:** Use this likelihood in a sampler, e.g. `MontePython`. Do an MCMC and plot parameter constraints.